Modeling spatial variations of urban growth patterns in Chinese cities: The case of Nanjing

Jun Luo a,1, Y.H. Dennis Wei b,*

a Department of Geography, Geology and Planning, Missouri State University, 901 S National Ave., Springfield, MO 65897, United States
b Department of Geography & Institute of Public and International Affairs, University of Utah, 260 S. Central Campus Dr. Rm. 270, Salt Lake City, UT 84112-9155, United States

A R T I C L E   I N F O

Article history:
Received 14 November 2007
Received in revised form 12 November 2008
Accepted 30 November 2008
Available online 9 January 2009

Keywords:
Urban growth
GIS
Remote sensing
GWR
Nanjing
China

A B S T R A C T

Revealing spatially varying relationships between urban growth patterns and underlying determinants is important for better understanding local dimensions of urban development. Through a case study of Nanjing, China, we employ both global and local logistic regressions to model the probability of urban land expansion against a set of spatial variables. We found that compared with other fast growing coastal cities, Nanjing remains a relatively compact city. The orthodox logistic regression found the significance of proximity, neighborhood conditions, and urban agglomeration in urban land change. The logistic GWR significantly improves the global logistic regression model in terms of better model goodness-of-fit and lower level of spatial autocorrelation of residuals. More importantly, the local estimates of parameters of spatial variables enable us to investigate spatial variations of the influences of spatial variables on urban growth. We have found distinctive local patterns and effects of urban growth in Nanjing, shaped by local urban spatial and institutional structures. A probability surface of urban growth, which is generated from raster calculations among the parameter and variable surfaces, provides a clear scenario of urban growth patterns and can be useful for decision making. This study also shows the importance of policy studies and fieldwork in the interpretation of results generated from statistical and GIS modeling.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Since the launch of economic reforms in the late 1970s, and even more so in the mid-1980s urban reforms and open door policies, China has experienced unprecedented urbanization. While urbanization on such a massive scale as that of China’s is hardly seen in human history, urbanization has long been the subject of academic inquiry, and modeling urban growth patterns even draws the attention of premiere scientific journals such as Nature and Science (e.g., Makse et al., 1995). The efforts to understand patterns, mechanisms, and effects of urbanization and urban growth have been heightened recently, especially in China, due to the intensified consequences of human activities on resources, open spaces, and environment. Being able to understand complex issues of this sort has been made easier with advances in GIS and remote sensing. While reform and growth have pulled millions of people in the countryside out of poverty and made China a growth engine of the global economy, the rise of China and rapid urbanization are accompanied by the disappearance of rural agricultural land, spatial fragmentation, and sustainability challenges (Xie et al., 2006; Yeh and Li, 1999; Wei, 2007). The United Nations projected that more than half of China’s population will be urban by 2030 (United Nations, 2001), and there is no doubt that the demand for urban land and the pressure for sustainable development will continue to increase. Better understanding and managing of urban growth are critical to the development and sustainability in China.

Urban growth and land expansion in China have been studied from various perspectives. Scholars have made efforts to understand the process and mechanisms of urban growth in China from institutional and political economy perspectives (Lin and Wei, 2002; Li, 2005; Ma, 2002), and argued that urban growth in China is driven by reform and globalization and led by the state, state-centered development alliances, and multinational enterprises (e.g., Deng and Huang, 2004; Ding, 2003; Lin and Ho, 2005; Wei and Li, 2002). By using spatial modeling, GIS and remote sensing technologies, a lot of works have been done to understand the patterns, mechanisms and effects of urban growth in China (e.g, Li, 1998; Schneider et al., 2005; Sui and Zeng, 2001; Weng, 2002; Wu and Yeh, 1997; Xie et al., 2005; Yeh and Li, 1998, 1999; Zhai and Ikeda, 2000). Various models have been developed to analyze and simulate urban growth patterns (e.g., Arai and Akiyama, 2004; Clarke et al., 1997; Landis and Zhang, 1998), including in the context of China (Chen et al., 2002; Cheng and Masser, 2003a; Liu and Zhou, 2005; Xiao et al., 2006; Wu et al., 2006). Among them, Cellular Automata
Previous land use models, whether CA or statistical models, only reveal urban growth patterns from a global or whole map view. The transition rules of CA models or the parameters of explanatory variables of statistical models, which actually indicate the influences of various factors on urban land expansion, are applied uniformly for the whole study area without considering the variation across space. Recently, scholars have attempted to localize statistical models by developing and using local indicators of spatial analysis (LISA) and geographically weighted regression (GWR) (e.g., Anselin, 1995; Fotheringham et al., 2002), well reflected in the recent research on China (e.g., Cheng and Masser, 2003a; Luo and Wei, 2006; Yu, 2006).

Most studies on urban growth in China have been focused on fast growing coastal cities and major interior cities. The fact that urban growth is a non-stationary process over space and that rules governing land use change vary from place to place might be masked (McDonald and Urban, 2006), which implies that the same set of underlying factors may yield various effects in different cities as well. Research on Nanjing, a peripheral city of the Yangtze Delta, remains limited as well, and the effects of exploratory variables might be different from existing findings. The case of Nanjing therefore helps to improve the understanding of patterns and determinants of urban growth and land expansion in China. After the following section on data and methodology, the third section uses global logistic regression to model urban land expansion from 1988 to 2000. We then develop a logistic GWR (geographically weighted regression) model and analyze the spatial variation of urban growth patterns in Nanjing. The last section presents our conclusions.

2. Data and methodology

2.1. The study area

Nanjing, with a history of over 2000 years, is one of the ancient capitals and cultural-historical cities of China. It is currently the capital of Jiangsu province and remains one of the most important political, economic, and cultural centers in the Yangtze Delta (Fig. 1). As a peripheral city to the delta, Nanjing is considered a relatively conservative (Zhao, 2005) and compact city (Luo and Wei, 2006) when compared with other major coastal cities in China. Like other cities in China, Nanjing has recorded fast growth, with the GDP and population of the municipality growing from 1.2 billion RMB and 4.88 million in 1988 to 102.1 billion RMB and 5.45 million in 2000, respectively. However, compared with other leading coastal cities which are the focus of the research on urban expansion in China, such as Shenzhen and Guangzhou in the Pearl River Delta and Hangzhou and Suzhou in the Yangtze Delta (e.g., Li, 1998; Yeh and Li, 1999; Wei and Li, 2002; Xie et al., 2005), Nanjing is less studied and its growth is somewhat slower.

Nanjing Municipality includes the city proper (Nanjing City), five suburban districts, and two remote rural counties, with a total land area of 6597 km². Subdistricts are the lowest administrative level under urban districts and the finest level at which census data is available; it is used by planners to define urban core and suburban areas (Fig. 1). The study area encompasses the majority of the built-up areas in Nanjing Municipality, within which most urban land expansion has occurred since the economic reform. It covers all urban districts and parts of surrounding suburban districts with 67 subdistricts and a total land area of 1128.89 km² (Fig. 1).

2.2. The data

A challenging issue for the study of uneven urban growth patterns within cities is the problem of data, since data at the district level in China is very limited. Fortunately, we were able to gather a substantial amount of data compiled from varied sources in Nanjing and abroad. Land use data employed in this research was derived from Landsat TM remote sensing imagery in 1988 and 2000 (30 m × 30 m resolution, five bands for each), which cover the whole study area. Several steps have been conducted for image processing and classification. The ERDAS IMAGINE 8.7 software package was used for image processing and land use classification (Leica Geosystem GIS & Mapping LLC, Atlanta, GA). Topographic maps of 1988 and land use survey maps of 2000 (scale 1:10,000 and 1:1000, respectively) were used to assist image geometric correction and land use classification. First, all images were geo-registered to the topographic map under a national coordinate system (i.e., Xian 1980, 3 degree zone) with a Transverse Mercator map projection. Eight ground control points were chosen systematically and evenly distributed over the images. A second-order polynomial transformation was used with nearest-neighborhood algorithm for pixel re-sampling. Root-mean-square pixel error was strictly limited to less than 0.5 pixels. Second, a supervised maximum likelihood classifier was used to classify the georeferenced images. Four types of land use are classified for the research: built-up, agriculture, forest, and water body. Third, we did field work in Nanjing, mainly in summer 2002, to systematically investigate the accuracy of the classifications, and interviews with local planners were conducted to enhance the error verification. By presenting the classified images to local planners, we asked them to visually judge the general accuracy of the classifications, and, for some local development hot spots, we asked them to visually determine if the classified images could reflect the direction, size and shape of the non-urban to urban land use conversion.

A spatial overlay operation was performed between the two classified images to extract the conversion between non-urban to urban land use. The other data sets needed for building the land use conversion model were extracted from various data sources. We extracted roads and rail networks from the land survey maps of 2000, and bridges, major centers, suburban centers, and industrial centers from plan scheme maps (Fig. 2). Road networks were divided into two types: inter-city highways and local artery roads.

2.3. Land use data sampling

The classified image for urban growth between 1988 and 2000 has 2684 × 2627 cells with 30 m resolution. Most statistical software cannot handle such a large data set, thus data sampling is needed for the research. However, while spatial sampling can reduce spatial dependence for the classical multivariate models, small sample size can make the maximum-likelihood method less reliable (Cheng and Masser, 2003a), and thus requiring a proper sample size is very important. To ensure that sampled land use data represents the study area systematically and provides enough information on land use change, we used a combined systematic and random scheme for land use data sampling. Systematic sampling (i.e. stratified sampling) is effective in reducing spatial dependence (Longley et al., 2005), and random sampling is efficient in representing population (Cheng and Masser, 2003a). From the non-urban land use areas in 1988, we extracted regularly spaced points with 300 m or 10 pixels internal, which captured the spatial variations of land use change well and avoided data redundancy; from the result we then extracted all 1332 points with non-urban to urban land use conversion and randomly selected 1350 points without land use conversion, which gave us a total of 2682 sample points. Such a sample size can be handled by software packages and is similar to...
the size used by other studies, such as 3002 in Cheng and Masser (2003a).

2.4. Dependent and exploratory variable

Classic multivariate statistical analysis has been typically used to model the transition from non-urban land-use to urban land-use. For example, Wu and Yeh (1997), and Cheng and Masser (2003b), used logistic regression models to detect the determinants of land use change in Guangzhou and Wuhan, respectively; Arai and Akiyama (2004) used linear regression and discriminant analyses to identify significant explanatory factors of land use transition potentials in Tokyo. We chose the probability of non-urban to urban land conversion during 1988 and 2000 as the dependent variable for the proposed logistic models, with values of 0 (no conversion) and 1 (with conversion). All variables used in the proposed land use models are listed in Table 1.

Table 1

<table>
<thead>
<tr>
<th>Variables</th>
<th>Type</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable</td>
<td>ChangeProb</td>
<td>Probability of land use conversion</td>
</tr>
<tr>
<td>Explanatory variable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dis2Hwy</td>
<td>Continuous</td>
<td>Distance to inter-city highway</td>
</tr>
<tr>
<td>Dis2Lard</td>
<td>Continuous</td>
<td>Distance to local artery roads</td>
</tr>
<tr>
<td>Dis2Rail</td>
<td>Continuous</td>
<td>Distance to railways</td>
</tr>
<tr>
<td>Dis2River</td>
<td>Continuous</td>
<td>Distance to the Yangtze River</td>
</tr>
<tr>
<td>Dis2YBrid</td>
<td>Continuous</td>
<td>Distance to the Yangtze bridge</td>
</tr>
<tr>
<td>Dis2Mcen</td>
<td>Continuous</td>
<td>Distance to major city centers</td>
</tr>
<tr>
<td>Dis2MNCen</td>
<td>Continuous</td>
<td>Distance to suburban centers</td>
</tr>
<tr>
<td>Dis2Induc</td>
<td>Continuous</td>
<td>Distance to industrial centers</td>
</tr>
<tr>
<td>Neighborhood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AgriDen</td>
<td>Continuous</td>
<td>Density of agriculture land</td>
</tr>
<tr>
<td>BuiltDen</td>
<td>Continuous</td>
<td>Density of built-up land</td>
</tr>
<tr>
<td>WaterDen</td>
<td>Continuous</td>
<td>Density of water body</td>
</tr>
<tr>
<td>FrostDen</td>
<td>Continuous</td>
<td>Density of forest land</td>
</tr>
</tbody>
</table>
Different types of explanatory factors have been identified in land use models, such as proximity to road infrastructure, attributes of specific land use site, attributes of neighborhoods of the land use site, etc. First, proximity factors have been extensively used in previous land use models and proven to be important in explaining land use transition (e.g., Cheng and Masser, 2003a; White and Engelen, 1997; Wu and Yeh, 1997). Nanjing has set up a transportation-oriented urban development plan (Nanjing City, 1993) and the last two decades have seen a dramatic increase in road construction and renovations, which have apparently stimulated the urban expansion. Considering that the inter-city highways, local artery roads, and railways form the cadre of transportation networks in the study area, we selected three variables of accessibility to transportation networks: distance to inter-city highways, distance to local artery roads, and distance to railways. The Yangtze River plays an important role in Nanjing’s urban development since it physically divides the study area into two parts and provides aquatic transportation access. We selected two variables related to the Yangtze River: distance to the Yangtze River and distance to the bridge over the Yangtze River. To obtain values of proximity variables for the sample points, we first generated a set of distance raster surfaces (30 m × 30 m cell size) using Euclidean Distance geoprocessing tool in ArcGIS™, then we extracted variable values for each point from the generated distance raster surfaces. These variables are expected to negatively affect the land transition from rural to urban uses.

Second, land use development is largely dependent on the neighborhood land-use conditions, and neighborhood variables are usually density-oriented (Cheng and Masser, 2003a; White and Engelen, 1997; Wu and Yeh, 1997). This study selected several neighborhood land-use attributes, including density of agriculture land, density of built-up land, density of water body, and density of forest land. They indicate the availability of development land (density of agriculture land: positive), or constraints (density of water body and forest land: negative), or neighborhood promotion (den-
Fig. 3. Spatial distribution of agriculture land (a), water body (b) and forest land (c).

Fig. 4 shows urban growth patterns in the study area of Nanjing from 1988 to 2000. We found that built-up areas increased by 30.61% from 392 km² in 1988 to 512 km² in 2000, which was modest compared with the fastest growing cities in China. Four broad areas of growth can be identified: the southeast, the northeast, the southwest, and areas along the northwest side of the Yangtze River. Larger and more concentrated development is related to development zones: Jiangning Economic and Technological Development Zone (ETDZ) (southeast), Nanjing ETDZ (northeast), Nanjing High-Tech District (northwest of the river), and Nanjing Hexi New Urban District (including Nanjing Olympic Sports Center) (southwest). The growth, however, has not changed the concentrated patterns of the city, where major centers are still located in the traditional urban core (Fig. 2) and not challenged significantly by any new centers. Such an urban structure is quite different from other major cities in coastal China, such as Suzhou where the dominance of the old city district has been replaced by Suzhou Industrial Park and Suzhou New District, and Shanghai where Pudong District has become a new urban center as important as the old city district (Wei et al., 2006).

While typically used to model the rural–urban land transition, classic statistical analysis has potential problems when modeling land use change at the local urban level. Land use transition as a
The spatial process usually presents high spatial dependency among observations, and ignoring such spatial dependency might lead to model misspecifications. Moreover, the conventional statistical analysis of urban land expansion implicitly assumes that relationships between explanatory factors and land use transition are spatially stationary, which is often untenable (Anselin, 1995; Fotheringham et al., 2002), especially for urban growth studies which usually involve large spatial data sets from remote sensing at relatively small urban scales.

This paper uses both the classic multivariate statistic analysis and GWR in modeling urban land expansion. To quantify the influences of explanatory variables on probability of urban land expansion, we first tested the classic logistic regression, which takes the following forms:

\[
\text{ChangeProb}_i = \frac{e^{(C + \sum \beta_k X_{ki})}}{1 + e^{(C + \sum \beta_k X_{ki})}}
\]

in which ChangeProb is the probability of land-use change to be regressed, \( C \) is constant, and \( \beta_k \) is the parameter for individual explanatory variable \( X_{ki} \) \((k = 1, 2, 3, \ldots, n)\). The global logistic regression model is estimated by the maximum-likelihood algorithm and the results are presented in Table 2. The model is significant at the 0.01 level. The \(-2\) Log likelihood value and percentage correctly predicted (PCP) are 2783.3 and 70.1%, respectively, which indicate a moderate goodness of fit of the model and a moderate level of prediction accuracy.

All explanatory variables are significant: Dis2Rail at the 0.05 level and the rest of the variables at the 0.01 level. Among the proximity variables, distance to local artery roads (Dis2Lard) has the strongest negative effect on land conversion probability, followed by distance to highways (Dis2Hwy). The significance of distance to
minor and major roads in land use transition is exactly the same as the findings of Cheng and Masser (2003a). Wu and Yeh (1997) also found the rise of proximity to city streets in land development to be significant. These findings suggest that urban growth in Nanjing, as well as across China, is largely dependent on road infrastructure development and that local roads are a more important determinant than highways.

Distance to the Yangtze River (Dis2YRiver) also has a negative effect on urban growth, while distance to the Yangtze Bridge (Dis2YBrid) has a positive effect. These findings are contradictory to Cheng and Masser’s (2003a) findings that in Wuhan the distance to the Yangtze River is not effective for probability of change. This difference reflects the policy change in Nanjing, which has increasingly recognized the importance of the Yangtze River to urban growth, as exemplified by the development of the NJ ETDZ and the Hexi district. Areas surrounding the Yangtze Bridge were developed much earlier and with little land available for further development, which explains the positive effect. It can be seen that global urban growth patterns in Nanjing have not relied on railways because distance to railways (Dis2Rail) has the weakest significance and the lowest positive parameter value. This makes sense since railways serve long-distance interurban travelers, rather than movement within the city.

All four neighborhood density variables except the density of built-up land (BuiltDen) have negative effects on urban land expansion. The density of forest land distribution has the strongest restriction on urban land expansion, followed by the density of water body distribution. This results from the fact that major forests and water bodies, characteristics of Nanjing (Zhao, 2005), are highly protected. Density of agriculture land, although providing the available land for urban growth, has negative effect on land use conversion as well, given the increasing protection of major agricultural land. Our findings are somewhat contradictory to Wuhan where the density of agricultural land is not a strong predictor for probability of change (Cheng and Masser, 2003a), suggesting that Nanjing is more protective of major agricultural land areas and water bodies, which explains why Nanjing is a more compact city than many other fast growing cities in China (Luo and Wei, 2006). It is not surprising that the density of built-up land (BuiltDen) has a significant promotion role for urban growth, which is similar to what others have found (e.g., Cheng and Masser, 2003a).

We can also see that distance to major city centers (Dis2MCen) has greater influence than distance to suburban centers (Dis2MCNcen), which contrasts with Cheng and Masser’s (2003b) finding that in Wuhan subcenters were making more effective impacts than major centers. This difference again indicates the relatively more compact urban growth patterns in Nanjing where traditional urban centers still exert great impacts on urban expansion. Distance to industrial centers (Dis2Induc) has positive effect on urban growth, which implies that large scale urban land development is not dependent on existing industrial centers. In fact, most of the existing industrial centers were established based on socialist principles of industrial allocation. They are either located in city districts which have largely been developed or in more the remote north and northwest which no longer serve as major areas of urban development in Nanjing.

Our classic logistic regression model effectively explains the determinants of probability of urban land expansion from the global view. We have found that major determinants of urban expansion in Nanjing were distance to roads, distance to existing major city centers, and the land use conditions of neighborhood area. The more conservative and compact nature of Nanjing is well reflected from our logistic model. However, the potential spatial non-stationarity of urban growth patterns still remains unknown. GWR is an appropriate technique for the detection of spatial non-stationarity, which allows the regression parameters to vary across the space, and can therefore expose local spatial patterns of urban growth in Nanjing.

### 4. Geographically weighted logistic regression model

Built on the spatial expansion method, GWR is a local regression technique for investigating the spatial non-stationarity, which aims at estimating parameters of a global regression model with a function of some other attributes representing spatial variation. The expansion method can only represent the broad spatial trends and may mask significant local variation (Fotheringham et al., 2002). In contrast, GWR is suitable for modeling the complex local variation of regression parameters and has been recently applied in various studies (e.g., Fotheringham et al., 2001; Lloyd and Shuttleworth, 2005; Longley and Tobón, 2004; Mennis and Jordan, 2005). In its most basic form, GWR model takes the following equation (Fotheringham et al., 2002):

$$Y_i = C_i + \sum_k \beta_k X_{ki} + e_i$$

(3)

$C_i$ is the constant parameter which is specific to location $i$; $\beta_k$ is the parameter of independent variable $X_k$ at location $i$. Based on Eqs. (1)–(3) can be modified to the following forms, which represent the logistic GWR:

$$\text{ChangeProb}_i = \frac{e^{(C_i + \sum_k \beta_k X_{ki})}}{1 + e^{(C_i + \sum_k \beta_k X_{ki})}}$$

(4)

$$\ln \left( \frac{\text{ChangeProb}_i}{1 - \text{ChangeProb}_i} \right) = C_i + \sum_k \beta_k X_{ki}$$

(5)

GWR estimates the parameters for each observation at location $i$ using all observations with assigned weights through a weighting scheme to data at other locations according to their spatial proximity, which is represented by Euclidean distance in this study. Nearer locations gain higher weights and vice versa. Two types of functions are usually used to obtain weights: fixed and adaptive kernels. The fixed kernel function applies an optimal spatial kernel (bandwidth) over the space, which is usually less computationally intensive. However, this approach can produce large local estimation variance in areas where data are sparse, and may mask subtle local variations in areas where data are dense (Fotheringham et al., 2002; Páez et al., 2002a,b). The adaptive kernel function, on the other hand, ensures a certain number of nearest neighbors as local samples, and better represents the degree of spatial heterogeneity. In this study, the adaptive kernel function was used, which is based...
5. Spatial variations of urban growth patterns

The logistic GWR generates a set of parameter estimates of explanatory variables for each land use sample point, which can be used to analyze spatial variations of the effects of urban growth determinants. In addition, a pseudo t-statistic is also calculated to indicate the significance of the parameters, which can be obtained by dividing a parameter estimate by its standard error (Fotheringham et al., 2001). Based on the sample points with parameter estimates and corresponding t-statistics, we generated a set of parameter and t-statistic surfaces to reveal the spatial variations of urban growth patterns. We employed an inverse distance weighted (IDW) interpolation algorithm to generate parameter and t-statistic surfaces. IDW assumes that the surface is being driven through the neighborhood (Watson and Philip, 1985), and hence is an appropriate approach in the context of this study. Figs. 5–12 presents the generated parameter and t-statistic surfaces with cell size of 30 m × 30 m.

Unlike the global logistic model which has unified parameters across space, it is apparent from Figs. 5–12 that all parameter vary across the study area with generally regular spatial patterns. The spatial distributions of corresponding t-statistics generally coincide with the spatial distributions of parameters. In terms of parameter significance, the parameters of Dis2Lard are highly significant over the entire study area, while all other parameters have certain parts in the study area where they are non-significant (at the 0.05 level).

Figs. 5 and 6 present the parameter and t-statistic surfaces of the three variables of proximity to transportation networks. One can see that the distance to highway has a more negative effect on land development in the northern part of the study area than the southern part (Fig. 5a). And the areas where the distance to highway is not significant are southeast and southwest of the study area, as indicated by t-statistics (Fig. 6a). The distance to local artery roads has even larger negative effect on land development, given the parameter has a distribution of relatively larger values in the study area (Fig. 5b) and is significant over the entire study area (Fig. 6b). However, it also varies across the study area with the influence decreasing from north to south in the vicinity of the northern Yangtze River. Near the southern Yangtze River, the parameter varies less with stronger influence to the northeast and southwest of the urban core. While the global logistic regression model has proven that the distance to railways has positive influence on land development, the parameter surface of the distance to railways suggests that there are large portions of the southern of the study area where distance to railways has negative influence on land development. In summary, the two variables of distance to highway and distance to local artery roads, which are two of the major determinant factors for land use conversions found in the global logistic model, played more important roles in land development north of Yangtze River than in the south.

Similarly, the global logistic regression model has shown that the distance to the Yangtze River has a negative influence on urban land expansion. This, however, does not hold true for all portions of the study area. For all the areas to the south of the river, the dis-

Table 3
Comparison between global logistic regression and logistic GWR.

<table>
<thead>
<tr>
<th></th>
<th>Global logistic regression model</th>
<th>Logistic GWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 Log likelihood</td>
<td>2783.3</td>
<td>1873.5**</td>
</tr>
<tr>
<td>PCT</td>
<td>70.1</td>
<td>85.6</td>
</tr>
<tr>
<td>Residual sum of squares</td>
<td>450.5</td>
<td>297.6</td>
</tr>
<tr>
<td>Moran’s I of residuals</td>
<td>0.744</td>
<td>0.48</td>
</tr>
</tbody>
</table>

* Indicates significance at 0.01 level.
** Indicates significance at 0.01 level.

\[
W_{ij} = \left[1 - \left(\frac{d_{ij}}{D}\right)^2\right]^2 \quad \text{if } j \text{ is } N \text{ nearest neighbour points}
\]

\[
d_{ij} = \text{the distance from } j \text{ to } i
\]

\[
b = \text{the distance from } N \text{ nearest neighbour to } i
\]

in which the number of nearest neighbor points, 138, is chosen by minimizing an Akaika Information Criterion (Fotheringham et al., 2002).

We used the same set of 2682 sample point data to construct the logistic GWR model. Table 3 presents a comparison between the global logistic and logistic GWR models. The logistic GWR model shows significant improvement over the global logistic model. First, the significant decrease of -2 Log likelihood indicates that the logistic GWR model has a much better goodness-of-fit than the global logistic model. Second, the increase of PCT from 70.1% to 85.6% and the significant decrease in residual sum of squares from 450.5 to 297.6 suggests the logistic GWR model has much better performance in exploring the relationships between land use change and explanatory variables than the global logistic model.

Third, residuals in the global logistic model have significant spatial dependence as indicated by a Moran’s I index significant at 0.01 level, while residuals in the logistic GWR model show non-spatial dependence indicated by a non-significant Moran’s I index. Furthermore, Monte Carlo significance tests for the parameter estimates show that all of the parameters have significant spatial variability with p-value of each below 0.01, which further justify the use of GWR in the study. Logistic GWR is calibrated using GWR 3 package and the calibration time is about 40 min, much longer than that of conventional logistic model, which is only about 10 s, using the same Windows XP desktop computer. We can expect that as the size of samples increase the calibration time of logistic GWR will increase significantly. Modeling spatial non-stationarity like logistic GWR of this research requires a high-profile computer hardware configuration and optimized codes of modeling package, which may enable scholars to deal with large size of samples and view the results sooner.

Compared with the global logistic regression model, the values of parameter estimates of GWR show significant variations. Table 4 presents the summary statistics of the GWR parameter estimates for the 2682 sample points. All the variables except the Dis2Lard and ForeDen have both positive and negative parameter values, although with differences in the portions of both values. This suggests that even from the aspatial perspective, the stationarity of parameter estimates in the global logistic regression model might be problematic. The parameter values of Dis2Lard and ForeDen, which have the strongest negative effects on land use change in the global logistic regression model among the distance and density variables, respectively, present single negative signs with relatively high standard deviations. Also we can see that the parameter values of Dis2Rail, Dis2YRiver, Dis2YBrid, Dis2MCen, Dis2MNCCen, and Dis2Induc have apparent divisions of positive and negative results.

Such significant spatial variations are largely ignored by orthodox logistical models, but can be revealed by GWR, as our study found. Though presenting relatively high standard deviations, the parameter values of the other three density variables have relatively low level divisions of positive and negative results, indicating that these three variables have less spatial variations. The following section explains in detail the spatial variations of urban growth patterns found from the GWR model.
Table 4
Summary statistics for GWR parameter estimates.

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>% of positive</th>
<th>% of negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dis2Hwy</td>
<td>-5.569</td>
<td>1.266</td>
<td>-1.275</td>
<td>1.045</td>
<td>8.24</td>
<td>91.76</td>
</tr>
<tr>
<td>Dis2Lard</td>
<td>-17.231</td>
<td>-3.156</td>
<td>-7.877</td>
<td>2.333</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Dis2Rail</td>
<td>-1.351</td>
<td>4.090</td>
<td>0.428</td>
<td>1.003</td>
<td>61.26</td>
<td>38.74</td>
</tr>
<tr>
<td>Dis2River</td>
<td>-7.135</td>
<td>10.198</td>
<td>-0.165</td>
<td>1.823</td>
<td>36.73</td>
<td>63.27</td>
</tr>
<tr>
<td>Dis2YBrid</td>
<td>-2.263</td>
<td>9.408</td>
<td>1.308</td>
<td>1.736</td>
<td>72.97</td>
<td>27.03</td>
</tr>
<tr>
<td>Dis2MNuc</td>
<td>-8.160</td>
<td>2.056</td>
<td>-1.438</td>
<td>1.570</td>
<td>20.95</td>
<td>79.05</td>
</tr>
<tr>
<td>Dis2Induc</td>
<td>-2.919</td>
<td>10.317</td>
<td>0.223</td>
<td>1.354</td>
<td>59.02</td>
<td>40.98</td>
</tr>
<tr>
<td>AgrDen</td>
<td>-35.034</td>
<td>14.586</td>
<td>-14.529</td>
<td>7.182</td>
<td>2.46</td>
<td>97.54</td>
</tr>
<tr>
<td>BuildDen</td>
<td>-11.746</td>
<td>150.900</td>
<td>24.482</td>
<td>17.812</td>
<td>89.37</td>
<td>10.63</td>
</tr>
<tr>
<td>WaterDen</td>
<td>-94.201</td>
<td>52.989</td>
<td>-26.648</td>
<td>16.928</td>
<td>7.53</td>
<td>92.47</td>
</tr>
<tr>
<td>ForeDen</td>
<td>-77.013</td>
<td>-17.955</td>
<td>-41.671</td>
<td>13.483</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

The distance to the Yangtze River does present negative influence on land development (Fig. 7a). However, positive influence of the distance to the Yangtze River can be found in some parts north of the river. Furthermore, non-significant areas are found in some parts of both the northern and southern sections of the study area (Fig. 8a). The distance to the bridge over the Yangtze River has positive influence on urban growth from the global view. But once again, from the local view, the positive influence is only found in the northern part of the study area (Fig. 7b). Areas in both the northern and southern portions of the study area have some locations with no significant influence relating to the distance to the bridge over the Yangtze River (Fig. 8b).

The distance to major city centers has a stronger influence in both the global and local views, which implies a compact urban land expansion in the study area, similar to what Luo and Wei (2006) found. In particular, the distance to major city centers has a stronger influence north of urban core than in the south (Fig. 9a). This is understandable because the north is a more mature area and the traditional focus of the urban life in Nanjing is more integrated with the major urban center. There are some areas south of the urban core and north of the Yangtze River where the distance to major city centers is not a significant factor of land development (Fig. 10a); most of the aforementioned areas are relatively independent industrial areas or newly developed areas. The distance to suburban centers, though weaker in terms of influence, varies more than the distance to major city centers (Fig. 9a), which suggests that suburban centers tend to have local influence on land development while major city centers tend to have global influence on land development. This reflects the reality of suburban centers which tend to be smaller and are built to serve local communities. Furthermore, regarding the significance, the distance to suburban centers has smaller distribution of non-significant areas than that of the distance to major city centers (Fig. 10b). This indicates that land development physically and locally relies more on the immediate nearby suburban centers, while the apparent global compact urban expansion pattern of Nanjing still shows that distance to urban major centers is a very important factor.

The global logistic regression model has demonstrated that land development tends to be located away from the existing industrial centers. Our logistical GWR model find that this statement is over...
Fig. 6. GWR parameter t-statistic surfaces of: distance to inter-city highway (a); distance to local artery roads (b); distance to railways (c).

Fig. 7. GWR parameter surfaces of: distance to the Yangtze River (a); distance to the Yangtze bridge (b).

Fig. 8. GWR parameter t-statistic surfaces of: distance to the Yangtze River (a); distance to the Yangtze bridge (b).
More specifically, it can be seen that the positive influence of the distance to industrial centers are mainly concentrated in the east of the study area (Fig. 9c), which is a more scenic, mountainous area in Nanjing with restricted development and more upscale residential buildings. What the global logistic model masked is that north of the Yangtze River most land development has been dependent on the existing industrial centers, and such influence is also significant (Fig. 10c). This is a result of government policy in Nanjing allocating major industrial projects to the north of the river, which served as
Fig. 11. GWR parameter surfaces of: density of agriculture land (a); density of built-up land (b); density of water body (c); density of forest land (d).

Fig. 12. GWR parameter t-statistic surfaces of: density of agriculture land (a); density of built-up land (b); density of water body (c); density of forest land (d).
the basis for urban expansion by providing jobs, transportation, and services for the residents.

For the four density variables, the global logistic regression model shows that only the density of built-up land has a positive influence on urban growth. However, the logistical GWR model provides more nuanced analysis. North of the Yangtze River, there are certain parts where the density of built-up land actually has a negative influence (Fig. 11b). But the negative influence in these areas shows non-significance with absolute value of t-statistic less than 1.96 (Fig. 12b). For the other three density variables, the density of forest land, which as previously stated also has a more significant impact in northern Nanjing, has a more significant impact in northern Nanjing. The logistic GWR model confirms that distance to major city centers has a more significant influence than distance to suburban centers, which also shows the relatively compact nature of the city. While some of the findings confirm what Cheng and Masser (2003a,b) found in their case study of Wuhan, others contradict their findings. This indicates disparities in urban growth patterns across China and the importance of case studies.

We found that logistic GWR can significantly improve the global logistic regression. Logistic GWR has a much better goodness-of-fit than global logistic regression model, and has better performance in exploring the relationships between land use change and explanatory variables than the global logistic model. Furthermore, the residuals of the global logistic regression model present significant spatial dependence, which violates the assumption of uncorrelated error, while the residuals in logistic GWR are no longer spatially dependent.

More importantly, the logistic GWR model allows the model parameters to vary across space, which provides deep insights into the spatial variations of the urban growth pattern. It has been demonstrated that the spatial variability of each factor influencing urban land expansion is significant and presents different patterns. We have found that distance to local artery roads has an even larger negative effect on land development, while distance to highways has a more significant impact in northern Nanjing. The logistic GWR model confirms the strong influence of distance to major city centers, with the influence stronger to the north of the urban core, a sign of a more mature urban environment. We also found that suburban centers tend to have local influence on land development. The generated probability surface of urban growth, which is based on the parameter surfaces from logistic GWR estimates, provide us an accurate visualization for urban growth. In general, GWR analysis reveals different effects of growth determinants in more mature urban areas and suburban areas, as well as in different parts of the city.

Although logistic GWR can more efficiently reveal the spatial variations of the influence of spatial variables on urban land expansion, interpretation of such variations should be careful and related to the contextual information of the study area. This paper has shown that policy studies and fieldwork have tremendous values in interpreting the results; otherwise it would be difficult to understand the varied effects of determinant variables found from the logistic regression models. Future studies should incorporate more disaggregate socio-economic variables, which are not currently available for our study area, into the logistic GWR model for investigating the interactions between urban growth and socio-economic conditions.

6. Conclusions

This paper has contributed to recent efforts to improve the understanding of urban growth and expansion in China through a case study of Nanjing. Based on recent developments in GIS and spatial modeling, we have developed global and local logistic regression models, which integrate a set of spatial determining variables to analyze spatial patterns and underlying factors of urban growth in Nanjing from 1988 to 2000. We found that built-up areas in Nanjing increased by 30.61% from 1988 to 2000, which was modest compared with many other fast growing cities in China. Nanjing, a peripheral city of the dynamic Yangtze Delta, remains a relatively compact city.

In our logistic regression model, all of the exploratory variables are statistically significant. We found that among proximity variables, distance to local artery roads has the strongest negative effect on land conversion, followed by distance to highways. Density variables of agriculture, forest, and water bodies all have negative effects, while the density of built-up land tends to promote urban growth. We also found that distance to major city centers has a more significant influence than distance to suburban centers, which also shows the relatively compact nature of the city. While some of the findings confirm what Cheng and Masser (2003a,b) found in their case study of Wuhan, others contradict their findings. This indicates disparities in urban growth patterns across China and the importance of case studies.

The generated probability surface of urban growth, which is based on the parameter surfaces from logistic GWR estimates, provides us an accurate visualization for urban growth. In general, GWR analysis reveals different effects of growth determinants in more mature urban areas and suburban areas, as well as in different parts of the city. Although logistic GWR can more efficiently reveal the spatial variations of the influence of spatial variables on urban land expansion, interpretation of such variations should be careful and related to the contextual information of the study area. This paper has shown that policy studies and fieldwork have tremendous values in interpreting the results; otherwise it would be difficult to understand the varied effects of determinant variables found from the logistic regression models. Future studies should incorporate more disaggregate socio-economic variables, which are not currently available for our study area, into the logistic GWR model for investigating the interactions between urban growth and socio-economic conditions.
Acknowledgements

We wish to thank the anonymous reviewers and editor who provided valuable comments on an earlier version of this paper. This research was financially supported by the Outstanding Young Scientist Award (Overseas) of the National Science Foundation of China (40428003) and the Graduate Student Small Grant from Urban China Research Network.

References

Jin Luo is currently an assistant professor in the Department of Geography, Geology and Planning at Missouri State University. He received a PhD in geography from University of Wisconsin-Milwaukee in 2006. His research focuses on GIS modeling, land use change and China’s urban development.
Yehua Dennis Wei is currently a full professor in the Department of Geography & Institute of Public and International Affairs at University of Utah. His research focuses on China’s economic/urban geography, regional development, and globalization.