Applied Geography 53 (2014) 172–186

Contents lists available at ScienceDirect

Applied Geography

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

# Modeling urban vertical growth using cellular automata—Guangzhou as a case study

Jian Lin [a, b], Bo Huang [b, **], Min Chen [c, d, *], Zhen Huang [c]

[a] Graduate Program in Environmental Science, State University of New York College of Environmental Science and Forestry, Syracuse, United States
[b] Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, Hong Kong
[c] Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, Hong Kong
[d] Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, 518057, PR China

*Keywords:*
Vertical growth
Cellular automata
Guangzhou
Urban

A B S T R A C T

Urban development is a complex spatio-temporal process that involves both horizontal and vertical growth. Despite growing recognition of the significance of horizontal development, models of urban vertical growth remain limited. This study aims to develop a GIS-based cellular automata model for exploring the vertical complexities of urban growth. Taking into account a series of variables, including accessibility, population density and building density and height, an "IF-THEN" rule base is designed and employed to simulate different height states of building growth. The model is validated through application to a case study of Guangzhou city for the period of 2001–2010. The results of the proposed model are compared with Guangzhou Urban Planning Bureau reference data for newly authorized construction buildings and then tested using an error matrix for 2001–2005 (overall accuracy 81.2% and Kappa coefficient 74.2%) and a fractal dimension for 2006–2010. Several conclusions are made based on the fractal analysis: (1) low-rise buildings tend to "spread outward," while high-rise buildings exhibit a trend of "compact development"; (2) a "hot zone" of vertical growth in Guangzhou demonstrates that the city is now undergoing a "phase transition" from a mono-center to a bi-center; and (3) low-, moderate-, and high-state buildings are being co-developed and are thus beginning to constitute an important feature of the urban and smart growth landscape.



## Introduction

Human civilization is currently entering an Urban Century (Kourtit, Nijkamp, & Reid, 2014). Urban areas are experiencing rapid growth mainly as a result of growing populations, rising incomes, and declining commuting costs (Bruechner, 2000; Bruechner & Largey, 2008). In developing countries in particular, cities are sprawling at rapid rates, as the number of metropolitan areas has increased considerably (Schneider & Woodcock, 2008). Such processes have had a number of negative influences on citizens with respect to their health and health-related behaviors (Ewing, Schieber, & Zegeer, 2003; Ewing, Schmid, Killingsworth, Zlot, & Raudenbush, 2003). Consequently, modern city expansion monitoring, analysis and simulation have become areas of major interest (*e.g.*, Lavalle, Demicheli, Turchini, Casals-Carrasco, & Niederhuber, 2001; Torrens, 2006; Torrens & Alberti, 2000). One notable byproduct of urban sprawl is vertical development, which is reflected in the growth of buildings of various functions (*e.g.*, commercial, residential, and industrial). As this phenomenon can transform the morphology and functioning of a city, it represents one of the most important aspects of smart growth and sustainable development. However, research in this field currently focuses more heavily on horizontal development, in relation to two phenomena in particular. The first focus involves the study of urban sprawl processes that occur outside of major urban centers, particularly along the fringes, edges and peripheries of cities, which transform landscapes from agricultural land into either built-up or less dense suburban areas. This field of research is not only concerned with conversion analyses of rural-urban land along the perimeter of the city (Huang, Zhang, & Wu, 2009) but also examines the urban expansion of downtown areas and major satellite cities (Li, Zhang, & Liang, 2010). The

* Corresponding author. University of China, Hong Kong.
** Corresponding author.
*E-mail addresses:* bohuang@cuhk.edu.hk (B. Huang), chenmin0902@cuhk.edu.hk (M. Chen).

http://dx.doi.org/10.1016/j.apgeog.2014.06.007
0143-6228/

other major area of focus involves the simulation of land-use changes across entire city regions, which involves the examination of transitions between several interdependent activities. In-depth studies involving land-use change modeling through the use of unbalanced support-vector machines (Huang, Xie, & Tay, 2009) and analyses of land-use change effects on life cycles (Batty, Xie, & Sun, 1999) have been conducted in the field of land-use-change modeling in recent years.

As noted above, these studies are largely based on two-dimensional spatial patterns that do not account for vertical development. However, with a decline in available land and the recent adoption of policies that protect high-quality arable land, the costs of horizontal growth are increasing significantly, necessitating the exploitation of high-level space. Vertical development is defined as vertical building growth in this study. Although vertical development represents an inevitable response to urban development and is therefore a marker of economic vitality, it can also have negative effects. Representing a large proportion of impervious surfaces in urban environments, buildings can heavily influence urban runoff levels (Weng, 2001) and are closely related to non-point pollutant sources (Hurd & Civco, 2004). In addition, building distribution patterns can significantly influence the nature of the urban heat island effect (Yang, Huang, & Homer, 2003) in addition to amplifying road traffic noise (Ko, Chang, & Lee, 2011). Moreover, the vertical growth of urban landscapes can affect urban efficiency levels as well as resident living habits and lifestyles. For example, residents are more likely to suffer from "traffic congestion" resulting from inefficient commercial and residential building distribution and are more likely to be accustomed to a lifestyle characterized by frequent interaction with others and high population densities. There is thus an urgent need to explore spatial patterns and temporal processes of urban vertical development. It is also especially important that the driving forces behind vertical growth be understood in relation to interactions between local actions and global patterns and, based on this knowledge, that 'what-if' decision-making be simulated based on varying combinations of driving forces. Although numerous statistical methods have been employed to capture the complexities of horizontal and vertical urban growth, these methods alone cannot accurately and efficiently quantify dynamic processes of urban growth due to complications arising from spatial heterogeneity and the existence of modifiable areal units (Paez & Scott, 2004). However, computer simulation approaches, such as cellular automata (CA), are frequently and successfully applied in the field of urban analysis and modeling (*e.g.*, issues of urban sprawl, socio-spatial dynamics, segregation, and gentrification have been analyzed) (Batty, 2001; Lagarias, 2012; Linard, Tatem, & Gilbert, 2013; Moghadam & Helbich, 2013). The CA model appears to play a crucial role in generating more comprehensive information on urban patterns and development (He, Okada, Zhang, Shi, & Zhang, 2006). In this paper, we present a GIS-based cellular automata model of urban vertical growth. Importantly, CA offers a bottom-up approach to vertical growth modeling through the use of complex and realistic configurations. A general background on the CA model is provided in the following section. The model's performance is then validated based on a case study of a built-up area of Guangzhou (GZ), a rapidly growing city in southern China.

## Cellular automata modeling

The CA model is composed of five principal elements: a lattice, a set of allowed states, neighborhoods defined by the lattice, transition rules and a temporal component (Batty, 2001). Combining these elements, the following simple and fundamental equation for CA is obtained:

$$S^{t+1} = f(S^t, \Omega^t, T) \tag{1}$$

where $S^{t+1}$ and $S^t$ represent the states of the cell under study at times $t$+1 and $t$, respectively; $\Omega^t$ is the configuration adjacent to the cell under study; $f$ represents a set of transition rules, and $T$ denotes the relevant parameters. Depending on modifications of the aforementioned elements, CA model forms and functions can vary from simple, standard and regular to complex, constrained and irregular. For example, by examining different interpretations of the five components and applying relaxation to the resulting characteristics, Liu developed a systematic scheme for classifying CA models (Liu, 2009). Several advanced forms of cellular models have been developed to distinguish generic from particular cities and realistic from optimal cities. Among these, the following initiatives have been extremely successful in developing cellular theories and useful applications of CA models: the 'Gigalopolis project,' jointly designed and developed by the University of California, Santa Barbara (UCSB) and the United States Geological Survey (USGS) (Project Gigalopolis), the 'GeoSOS group' at Sun Yat-sen University (Project GeoSOS), the 'METRONAMICA group' at the Research Institute for Knowledge Systems (RIKS BV) and the 'DUEM group' at the Bartlett Centre for Advanced Spatial Analysis (CASA). Applying five growth coefficients (*e.g.*, diffusion, breed, spread, road gravity, and slope) to six input data layers (*e.g.*, land cover, exclusion, urbanization, transportation, and hill-shade), the first research group divides land-use changes into four categories: spontaneous growth, new spreading center growth, edge growth, and road-influenced growth (Clarke & Gaydos, 1998) and has successfully applied this model to several regions (Clarke & Gaydos, 1998; Clarke, Hoppen, & Gaydos, 1997; Dietzel & Clarke, 2004; Silva & Clarke, 2002; Syphard, Clarke, & Franklin, 2005). Drawing from the preliminary work of Yeh and Li at the University of Hong Kong, the second research group has developed a computer-based system that is designed based on an Object-Oriented Programming (OOP) paradigm. This system notably situates the cellular model with theories from other fields and thus forms an integrated organism. The approach applies methods such as logistic-based CA, multi-criteria evaluation (MCE)-based CA (Wu & Webster, 1998), principal components analysis (PCA)-based CA (Li & Yeh, 2001b, 2002b), artificial neural network (ANN)-based CA (Li, Lao, Liu, & Chen, 2011; Li & Yeh, 2001a, 2002a) and ant colony optimization (ACO)-based CA (Liu & Li 2008; Li et al., 2011; Liu, Li, & Yeh, 2007). The latter two research groups have also produced significant findings in this field (*e.g.*RIKS BV, 2005; Van Delden, Escudero, Uljee, & Engelen, 2005; Xie, 2002). Due to the recent development of CA theories and their successful application in various studies, CA is employed in this study to simulate urban vertical development for examining building growth height states.

However, the model-building process is complicated by the presence of various building representations. For example, the basal areas for buildings with different functions range from $10^2$ $m^2$ (for residential buildings) to $10^5$ $m^2$ (for warehouse buildings). Thus, it is in practice either difficult or even impossible to associate cells with buildings. In addition, as several distinct activities may occur within one building, it is inaccurate to refer to buildings as cells. In consideration of these particularities, we have developed a GIS-based CA model that uses a linguistic approach. This model focuses on the simulation of an actual city (*i.e.*, Guangzhou, particularly in terms of vertical growth) from 2001 to 2010. By combining a series of spatial variables, the model is designed to capture building distribution patterns across space and time.

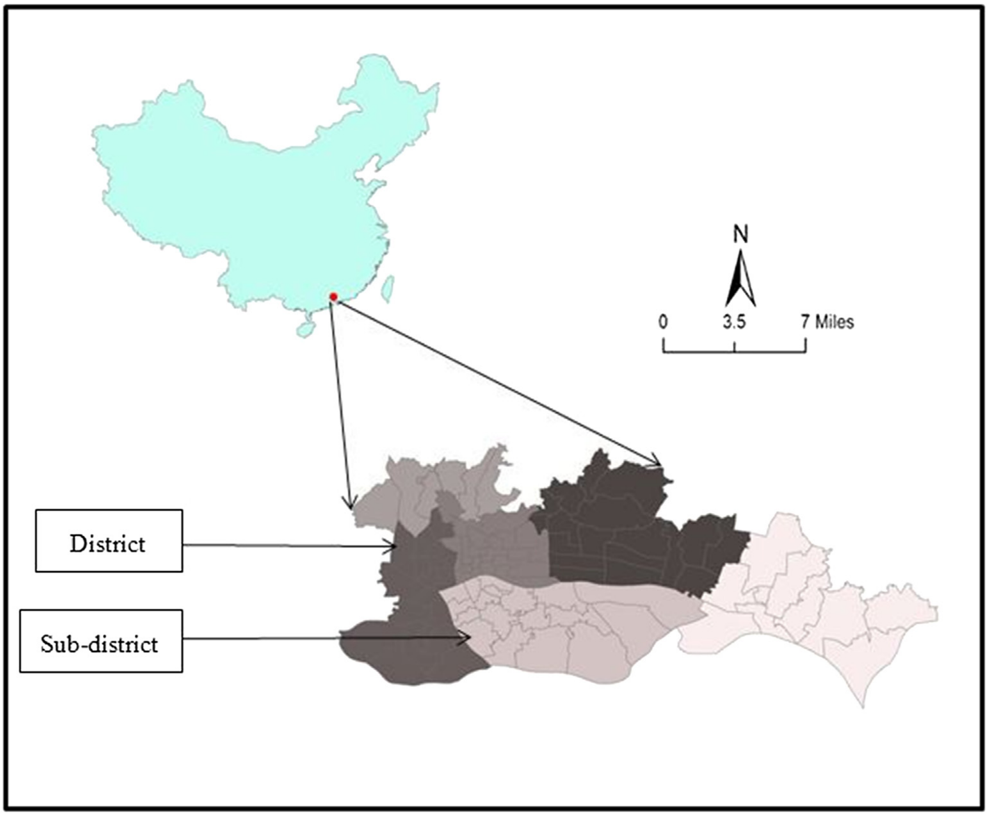



**Fig. 1.** Map of the case study area.

## Study area and data processing specifications

*Study area.*

The study area, spanning approximately 436 km$^2$, covers a developed area of the GZ megacity (Fig. 1). Although district accounts for only 20.86% of the entire megacity, the area is already highly developed and densely populated. The site represents the economic and political center of GZ and houses approximately 70% of the GZ population. The area became highly urbanized by 2000, emerging as the urban center of the Pearl River Delta in southern China. Building growth volumes from 2001 to 2010 show that the area experienced rapid development during this period. It thus became urgently necessary for such development to be coordinated through an optimized approach to spatial structure design.

### Data collection

The data used in this study are shown in Table 1. By combining the building distribution map for 2003 with newly authorized construction data, we can deduce the structure of the building map for the year 2000. To simplify the analysis, we merged building units with the same number of floors that were situated close to one another and deleted buildings with basal areas (S) of less than 100 m$^2$. Although this may have affected the accuracy of our model, this impact is assumed to be negligible because the excluded buildings ($S$ < 100 m$^2$) typically represent affiliated or extended sections of original buildings. Additionally, the purpose of our research is to simulate general patterns and trends of vertical development rather than exact building growth statistics. After performing this operation and removing invalid data (*e.g.*, buildings with a floor number of 0), the number of buildings was reduced from 259,669 to 133,257. The building map for the year 2000 was used as the base for the initial status representation.

Building uses and functions are classified into four types: commercial (C), residential (R), industrial (I), and mixed commercial and residential (C/R). The 'industry' category includes factories, warehouses and workshops. Because the exact ratio of commercial to residential activities within the C/R category is unknown, we consider these forms together as one type. The newly authorized construction data (new buildings) from 2001 to 2010 were categorized using the above classification scheme. The reference data were then used for our simulation.

Road network data for 2000 were used to construct the model (Fig. 2). The road system is categorized into five types: highway,

**Table 1**
Summary of data used in the vertical growth simulation.

| Data | Attribute | Source |
|---|---|---|
| Sub-district based map | Area, perimeter | Guangzhou Urban Planning Automation Center |
| New authorized construction data (2000–2010) | Floor number, basal area, location | Guangzhou Urban Planning Bureau |
| The buildings distribution map of 2003 | function (only new authorized buildings have) | |
| Demographic data (2000–2010) | Population base on sub-district | Guangzhou Statistics Bureau |
| Transportation network data (2000 and 2010) | The level of road, subway stations | Guangzhou Urban Planning Bureau |

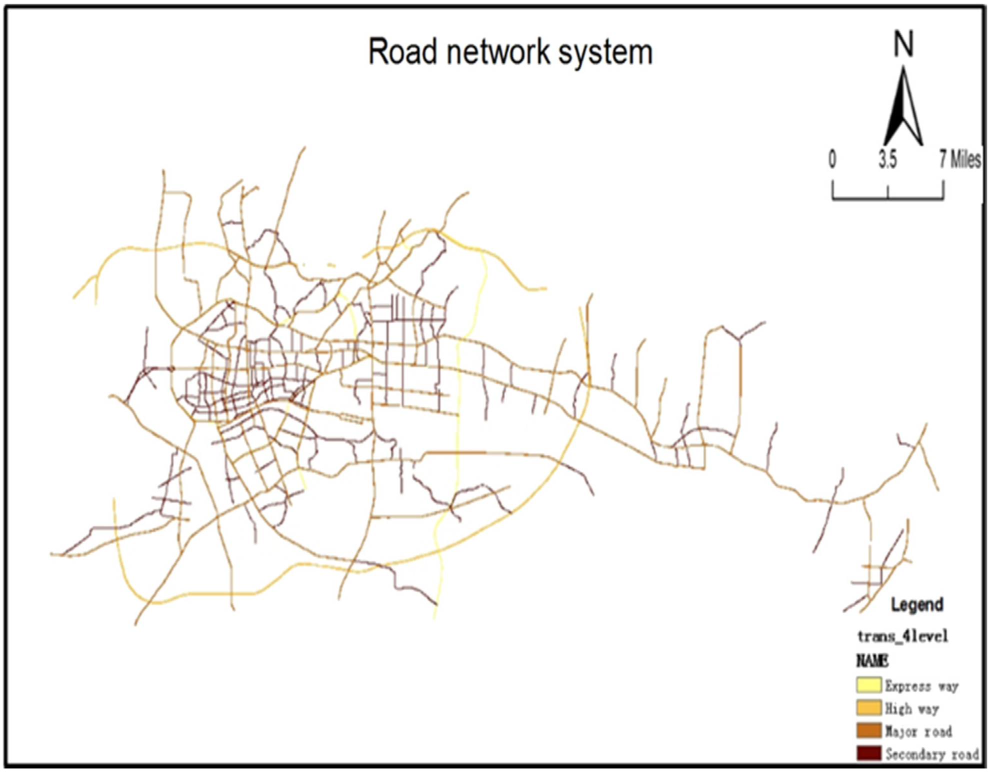



**Fig. 2.** The road network system of the study area.

expressway, major road, secondary road and neighborhood road. Comparing the road network data for 2000 and 2010, which constitute the only network data available, major differences can be observed with respect to the expressway and neighborhood road categories. However, neighborhood road status is not our main concern, and the GZ government has prioritized upgrading existing road systems over constructing new roads since 2000. Hence, the effect of this variable is assumed to be negligible. In addition to linear transportation data, subway station data were also incorporated throughout the modeling process. These data were updated at each loop location according to the time of initial operation. In summary, the transportation data used in the model include transport linear data (*e.g.*, major and secondary roads) and transport point data (*e.g.*, subway stations and expressway entrances).

*Building attributes and linguistic approach*

*Fuzzy set logic and linguistic approach*

Fuzzy set logic, in contrast to Boolean logic, has frequently and successfully been applied in recent studies on uncertainty representation, rule formulation and decision support. Leung provided the following systematic summary of the advantages of fuzzy theory: (1) several geographical phenomena are spatially continuous; (2) our daily languages are essentially imprecise; and (3) differences between objects represent differences "in degree rather than kind" (Leung & Leung, 1993). In this study, we use a linguistic approach in identifying building attributes.

*Building attributes*

The three building attributes—basal area, height (floor number), and function—are undoubtedly related. For example, the height of an industrial building is limited depending on the nature of the industry it is used for and the activity cost; when the basal area of a building is very large, the building cannot include a large number of floors due to associated construction costs. It is therefore important to perform a systematic classification that reflects the interrelationships between these attributes.

As explained above, there are four types of building functions: commercial (C), residential (R), industrial (I), and mixed commercial and residential (C/R). Admittedly, it is unrealistic to disregard vertical mixed-uses and to refer to an entire individual building as a single type. This simplified approach can, however, capture the fundamental features of each function.

The number of floors for the buildings examined ranged from 1 to 83. It is unnecessary and inefficient to develop a single attribute based on this enormous range. Therefore, floor numbers are divided into a range of "states" that are divided across the following series of linguistic variables: "very low," "low," "moderate," "rather high," "high," "very high" and "extremely high." The number of buildings with floor numbers falling within the "very low" and "low" categories was significantly larger than that of buildings occupying the "high" category. To distinguish between each state and to clarify the characteristics of this distribution, we apply the classification shown in Table 2. Floor number levels and associated membership functions can be adjusted throughout the calibration process.

**Table 2**
Floor number classification system.

| Category | Floor number |
|---|---|
| Very low | 1–2 |
| Low | 3–5 |
| Moderate | 5–7 |
| Rather high | 7–9 |
| High | 9–19 |
| Very high | 19–29 |
| Extremely high | >29 |

## Simulation of building growth in GZ

Urban development involves both horizontal and perpendicular growth. Previous studies have placed greater emphasis on horizontal expansion while tending to overlook vertical growth. The proposed CA model was constructed using building growth data. The model enables the simulation of urban complexities in three-dimensional form. Horizontally, a cell is regarded as undergoing development if it transforms from a vacant state to a different building function. Vertically, this cell will reach different building heights.

### *Cell size and neighborhood configuration*

Because buildings take different shapes and sizes, an overly large cell size will not accurately represent the buildings and may introduce errors. However, cell units that are too small may also limit the explanatory power of the model because bordering areas where cells meet may cover part of the same building. After testing a number of size options, a cell size of 20 $m^2$ was chosen for this study so that each building is composed of at least five cells. A neighborhood is denoted by a circular area with a 10 m radius. This approach was chosen due to distortions that can occur when examining neighborhoods following a rectangular configuration (Li & Yeh, 2000).

### *Driving forces*

Urban growth results from a series of driving forces. These forces can be categorized as either internal or external and as spatial or socio-economic. A considerable body of research has been conducted in exploration of these influencing factors. For instance, Huang grouped seven variables into three types: "(1) site specific characteristics; (2) proximity extent; (3) neighboring characteristics" (Huang, Zhang, et al., 2009). According to the major principle of the CA model, namely, the relationship between local actions and global patterns, the nature in which a cell develops depends on its intrinsic suitability and on aggregate processes occurring within its immediate neighborhood. Moreover, additional catalyzing and limiting factors can be identified to develop a more accurate and realistic model. We adopted Huang's classification logic in this study, and hence, three groups of variables are defined and shown in Table 3.

Population plays a pivotal role in the process of building growth. Family members are the "users" of residential buildings, while employees are the "users" of industrial buildings; for commercial buildings, individuals act as both consumers and employees. In this study, population density data for the period of 2001–2010 were used to determine the intrinsic suitability of a cell. Suppose $\rho(x_{ij}^t)$ is the population density of cell $x$ at time $t$. We use equation (2) to standardize $\rho(x_{ij}^t)$ between 0 and 1:

$$\beta\left(x_{ij}^t\right) = \begin{cases} 0 & \rho\left(x_{ij}^t\right) = \rho_{\min} \\ \dfrac{\rho\left(x_{ij}^t\right) - \rho_{\min}}{\rho_{\max} - \rho\left(x_{ij}^t\right)} & \rho_{\min} < \rho\left(x_{ij}^t\right) < \rho_{\max} \\ 1 & \rho\left(x_{ij}^t\right) = \rho_{\max} \end{cases} \tag{2}$$

where the maximum and minimum population densities are represented as $\rho_{\max}$ and $\rho_{\min}$, respectively. The population data were, however, incomplete for a number of sub-districts. For example, 2002 to 2004 population data for the Haizhu district were not available. Although spatially and temporally based interpolation methods can estimate missing population data, these statistical approaches may distort intrinsic patterns and trends in the original data. In this study, we simply labeled missing values as "No data" and removed them from the evaluation criteria.

Another significant factor to consider is the building distribution of the immediate neighborhood surrounding the cell under study. This value acts as an indirect indicator of building height regulations in the local area. A cell is assigned a state value according to the height states of existing buildings in the neighborhood in which it is situated. For example, if the majority of the neighboring cells occupy a "very high" state, then the cell in question is classified as "very high."

Apart from these basic settings, a number of catalyzing and constraining factors (*e.g.*, transportation networks and zoning plans) can also stimulate or prevent urban growth, thus affecting overall patterns and trends. Although highways and expressways represent line transportation data, they are classified as closed or semi-closed road networks. Therefore, it is more suitable to regard these features as transport point data at the entrance and exit points of a road. An effective road system can improve accessibility levels and thus encourage new development. However, the degree of accessibility can change depending on the forms of transport that are most prevalent. A series of weights were thus assigned to different levels and modes of transport (Table 4).

The accessibility measure proposed by White, Engelen, and Uljee, (1997). was adopted in our model, although a number of adjustments were made. While the accessibility coefficient $\delta_j$ is activity oriented in White's equation, the coefficient is transport mode-oriented in our model. The equation is therefore as follows:

$$\alpha_{ij} = \left(1 + \frac{D}{\delta_{ij}}\right)^{-1} \tag{3}$$

where $\alpha_{ij}$ represents the level of transport accessibility from $i$ for building function $j$; D is the Euclidean distance between the cell of concern and the closest transportation point; and $\delta_{ij}$ is a coefficient that reflects the importance of transport from $i$ for building function $j$.

**Table 3**
Summary of model driving forces.

| Category | Variable | Description |
|---|---|---|
| Site-suitability | Population density β | Standardizing between 0 and 1 while representing no data as one category |
| | Zoning | Including core, fringe, periphery and hinterland |
| Neighborhood configuration | Buildings density γ; | Cell number occupied by buildings |
| | Buildings height states μ | Including 7 states from "very low" to "extremely high" |
| Proximity extent | The accessibility α | Transport linear measure: Major and secondary road |
| | | Transport point measure: subway stations and high way & express-way entrances |

**Table 4**
Transport data initial weight categorization.

| Category | Variable | Weight |
|---|---|---|
| Transport linear data | Major road | 4 |
| | Secondary road | 3 |
| Transport point data | Subway stations | 4 |
| | High way entrances | 1 |
| | Express-way entrances | 2 |

*Transition rules*

Transition rules denote the logical connections between all driving forces that are adjusted using different weights and orders. Transition rules are typically specified in such a way that incorporates geographic theories or methodologies (Longley & Batty, 2003). The fundamental principle of our model is based on the concentric zone model developed by Ernest Burgess, which was originally developed to describe the concentric circle structure of urban land use (The Burgess Urban Land Use Model) (Burgess, 2000). Transition rules taking an "IF-THEN" form and including linguistic variables are adopted in this model.

The basic skeleton of the road system in GZ was fully developed by 2000 and included (1) a checkerboard pattern within the old town; (2) an inner and outer ring serving the old town and Tianhe new district; and (3) a strip extension running in a primarily west-east direction. This road network system contributed to the development of a layered structural pattern and a highly dense building distribution within the old town. Furthermore, urban planning schemes applied to GZ strengthened this trend by dividing the developed area into three groups: (1) the old town group, which is the economic and political center; (2) the Tianhe group, which became the central business district (CBD) of GZ city; and (3) the Huangpu group, where industry and warehousing represent the main functions. Although the GZ government has favored eastward and southward growth as the main avenues for urban expansion, the spatial pattern of the mono-center has not changed.

Based on the aforementioned characteristics of the study area, we divided the entire study area into four zones: the core, the fringe, the periphery, and the hinterland (Fig. 3). From the geometric center of the old town, we created a series of buffer areas (*e.g.*, 4 km, 7 km, 10 km and 16 km) to cover each zone. Among these buffering radii, the fringe zone spans two sections (7 km and 10 km) due to the effect of the Zhujiang River. The northern section of Zhujiang was attributed a radius of 10 km, and the southern section was designated a radius of 7 km due to obstacles imposed by the river.

Our model assumes that despite functional differences between the buildings, each block exhibits the same locational preferences. Namely, all entities prefer to occupy "the best or optimal location." While commercial and industrial entities aim to maximize profits, residential entities are attracted by proximate services and job opportunities. For example, while the core zone is highly appealing for C or C/R building development, industrial buildings are more likely to be constructed in hinterland regions due to lower associated costs and preferential policies. This general pattern can be either reinforced or altered by the existing transportation network, the quantity of available land and the nature of historical patterns, and these factors have a cumulative effect on overall development. At the end of each iteration, cell transformations from vacant to building functions were added to the existing building set as part of the simulation base for the next year.

## Building cell growth simulation—from vacant to varying heights

The model is implemented using Python and Arcpy. Arcpy is a site package that uses Python to process geographical data and perform spatial analysis. Although Arcpy is still undergoing development, the program is relatively convenient to implement for the purposes of our study.

*First stage simulation for the period of 2001–2005*

*Rule setting and model implementation*

When applying an "IF-THEN" rule, it is critical to confirm that the input data satisfy the "IF conditions" and that these conditions thus trigger the rule. Several pre-specific thresholds were established to fulfill this task. Table 5 shows the classification schemes for the different driving forces. We classified accessibility into the following three categories: 0, 1 and 2, signifying basic support,

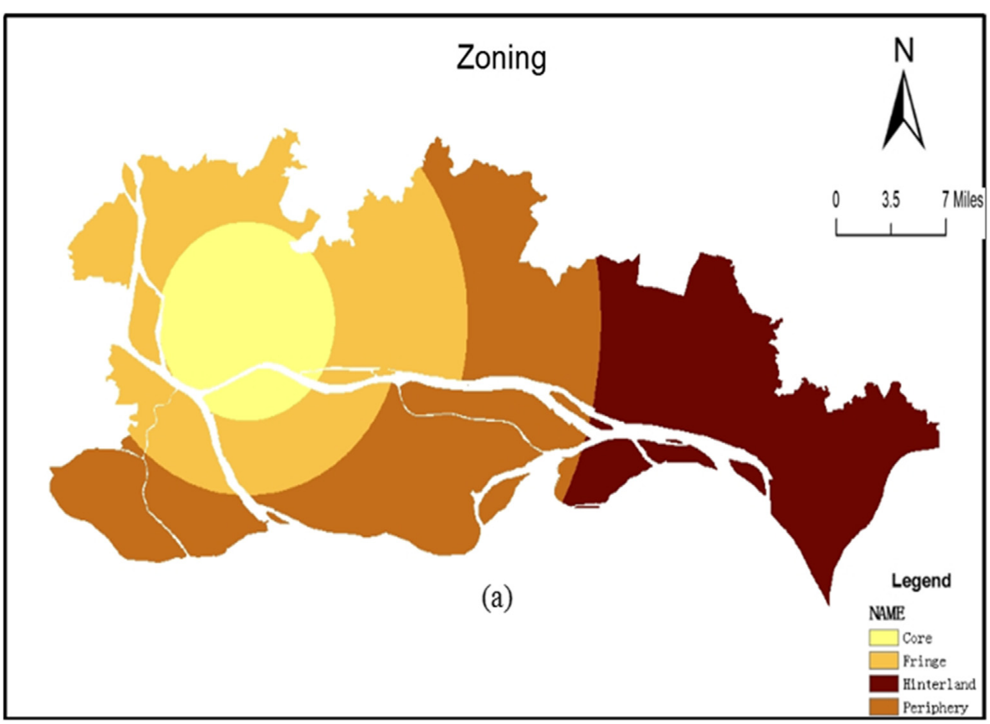



**Fig. 3.** The zoning system of the study area.

**Table 5**
Model driving factor interval ranges.

| Category | Population density β | Building density γ | Accessibility α (old) | Accessibility α (new) |
|---|---|---|---|---|
| 0 | 0–0.005 | 0–15 | 0–0.005 | 0–0.01 |
| 1 | 0.005–0.01 | 15–50 | 0.005–0.015 | 0.01–0.02 |
| 2 | 0.01–0.02 | 50–100 | >0.015 | 0.02–0.05 |
| 3 | 0.02–0.05 | >100 | | 0.05–0.2 |
| 4 | 0.05–0.1 | | | >0.2 |
| 5 | 0.1–0.2 | | | |
| 6 | 0.2–0.5 | | | |
| 7 | 0.5–0.8 | | | |
| 8 | 0.8–1 | | | |
| 9 | No data | | | |

weak support and strong support, respectively. We assumed that cells could only transform when provided with weak support or higher. Other initial parameter settings included population and building densities larger than 0.05 and 50, respectively.

For the cells that are excluded from building growth (*e.g.*, water bodies and road networks) and for those cells that are already occupied by existing buildings, no transition rules apply. Although these cells are excluded from the conversion, they may affect the transition probabilities of vacant cells. As illustrated in Fig. 4, the model is built over three stages.

*Results and cell-based accuracy assessment*

Fig. 5 (a) shows the initial building height states in GZ used in our study, and the actual versus simulated height states in the year 2005 are displayed in Fig. 5 (b) and (c), respectively.

The cell-based error matrix was used to estimate the simulation results for 2005. As shown in Table 6, the overall accuracy of the simulation reached 81.2%, and the main error value fell within the user accuracy range, which varied from 68.8% to 96.5%. The user accuracy range signifies the level of commission error, or the number of cells that appear in the wrong categories.

These phenomena are clearly demonstrated in the enlarged display of height states in the Haizhu district (Fig. 6), where several vacant cells in the reference data are instead occupied by buildings in the simulation results for the same period.

The overall accuracy of the results is relatively high for the following two reasons: (1) there were a large number of building cells that occupied the initial state, which led to an overwhelming predominance of unchanged cells; and (2) the error matrix itself only

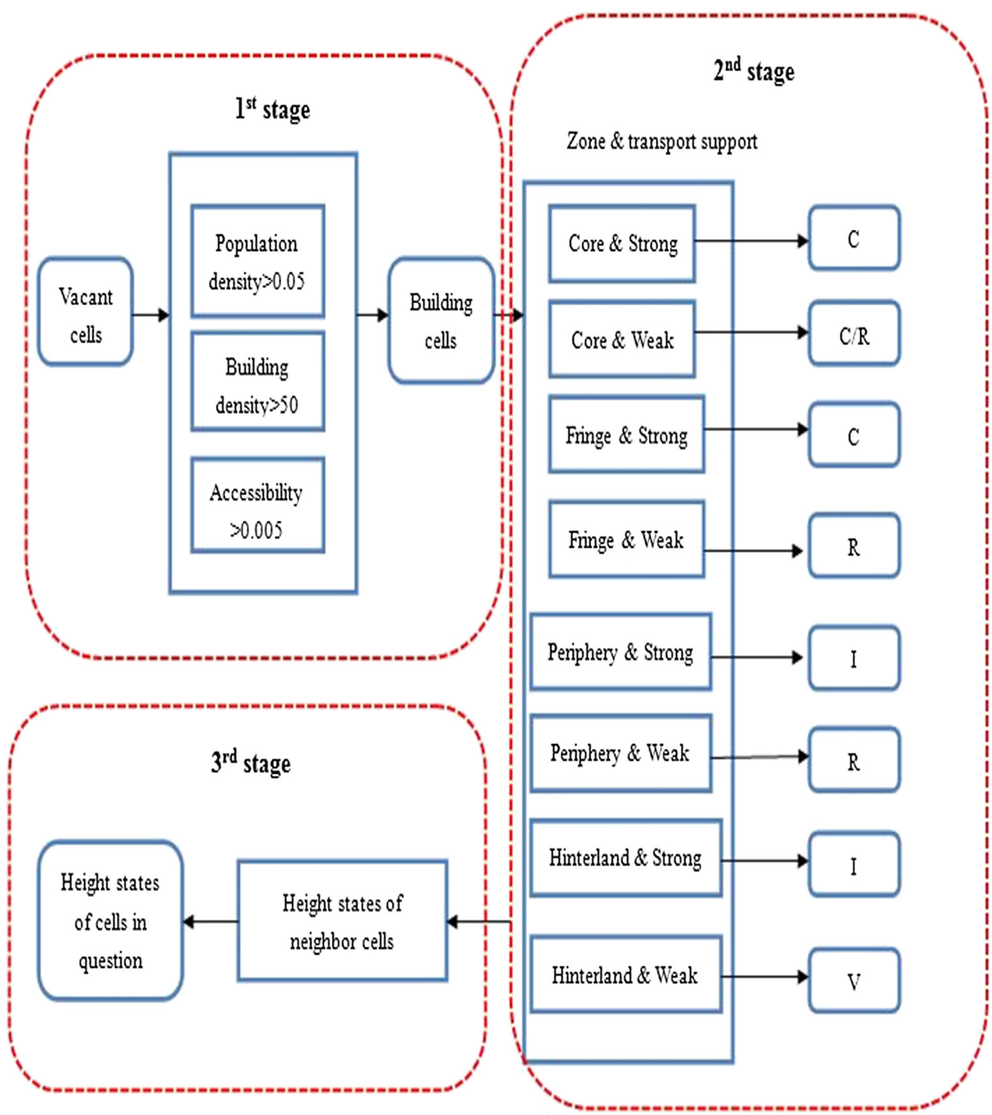



**Fig. 4.** Model-building process flowchart.

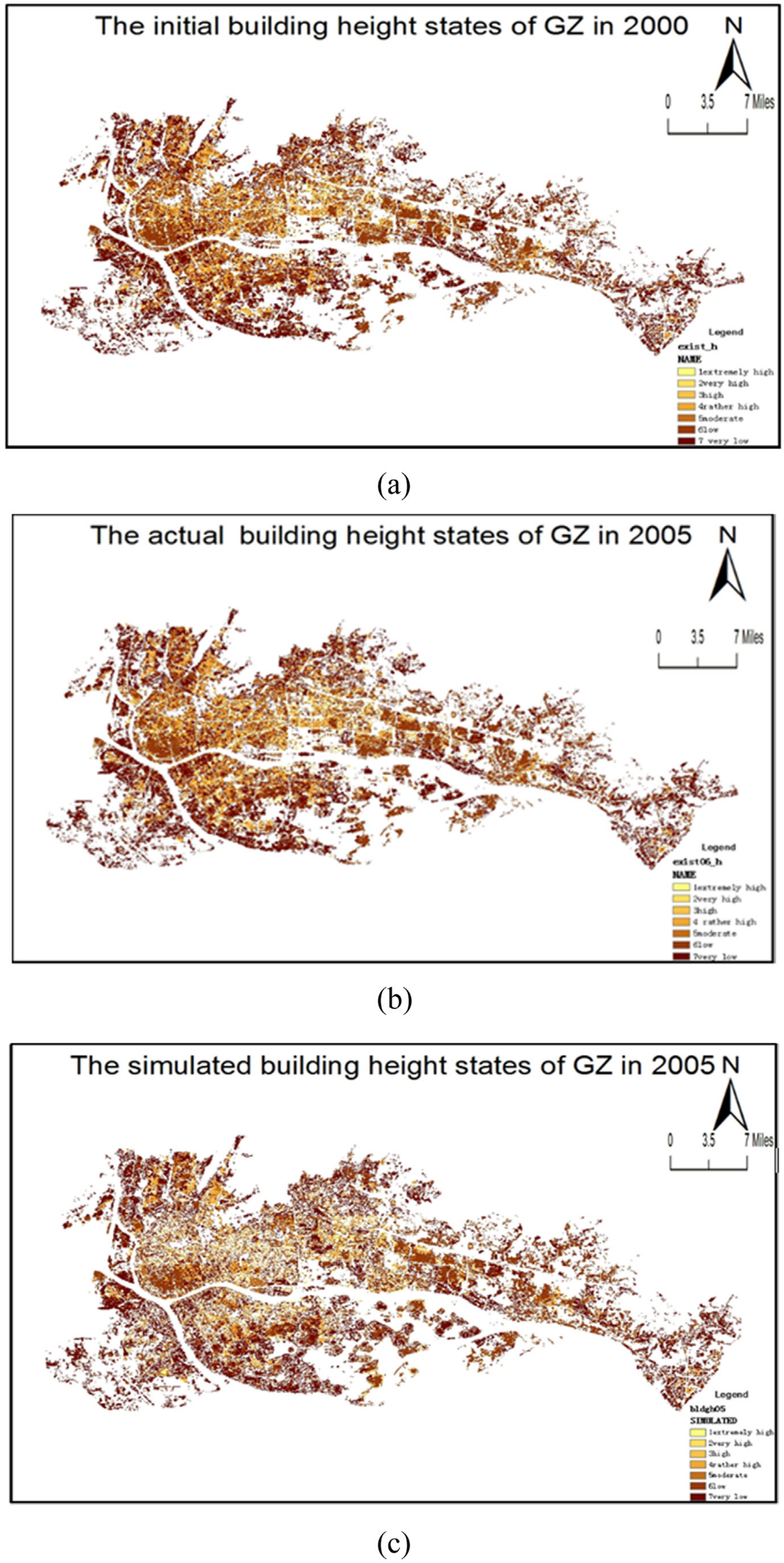



**Fig. 5.** The spatial distribution of building height states: (a) initial distribution for 2000; (b) actual distribution for 2005; (c) simulated result for 2005.

**Table 6**
Error matrix for the first simulation stage (2001–2005).

| | | The actual building height states of GZ in 2005 | | | | | | | Row total |
|---|---|---|---|---|---|---|---|---|---|
| | | Very low | Low | Moderate | Rather high | High | Very high | Extremely high | |
| The simulated building height states of GZ in 2005 | Very low | 140,736 | 207 | 183 | 47 | 408 | 208 | 211 | 142,000 |
| | Low | 76 | 107,807 | 27 | 0 | 97 | 33 | 33 | 108,073 |
| | Moderate | 27 | 7 | 19,730 | 8 | 140 | 70 | 95 | 20,077 |
| | Rather high | 16 | 11 | 6 | 31,660 | 66 | 32 | 12 | 31,803 |
| | High | 6 | 3 | 0 | 5 | 9507 | 7 | 15 | 9543 |
| | Very high | 0 | 0 | 0 | 0 | 0 | 3235 | 0 | 3235 |
| | Extremely high | 0 | 0 | 0 | 0 | 0 | 0 | 2143 | 2143 |
| | Column total | 140,861 | 108,035 | 19,946 | 31,720 | 10,218 | 3585 | 2509 | 316,874 |
| | Producer's accuracy | | | | | | | | |
| | all the accuracies of the individual category are larger than 90% (except that "extremely high: 85%") | | | | | | | | |
| | User's accuracy | | | | | | | | |
| | Very low | 75.30% | Moderate | 68.80% | High | 75.20% | | | |
| | Low | 91.60% | Rather high | 87.50% | Very high | 96.50% | | | |
| | Overall accuracy | | 81.20% | | | | | | |

considered major diagonal data while neglecting all off-diagonal cells. An enhanced approach—"Kappa analysis"—is often used to compensate for error matrix shortcomings. When off-diagonal data are combined, the Kappa coefficient reaches a mere 74.2%, which is significantly lower than the overall accuracy value of 81.2%.

### *Second simulation stage for the period of 2006–2010*

#### *Calibration based on the first stage*

The low level of user accuracy (commission error) can be attributed to the rule setting that regulated only the lower threshold while neglecting the upper limit. Theoretically, larger population density values should correspond to a higher probability for cells to transform from vacant into building cells. However, in practice, dense populations always correspond with high levels of development or competition and hence high construction costs. Low building density causes "death by isolation," while high density leads to "death by crowding." As high accessibility levels can also result in negative repercussions, such as "noise pollution," building cell location selections must achieve a balance between each element. Using the actual cell distribution for each interval range, we performed a new classification for accessibility (Table 4) and adjusted the threshold ranges for compatibility with the factors. Accessibility, $\alpha$, ranges from 0.02 to 0.2 (namely, corresponding to categories two and three); for building density $\gamma$, the range varies from 50 to 100 (category three); with decreasing land availability, values can extend between categories two and four; and population density $\beta$ values range between categories two and eight as a result of the differing functions of building cells. For example, commercial cells can afford high land rents, while industrial cells do not require "high population density" levels.

#### *Results and pattern-based accuracy assessment*

Based on the aforementioned settings and parameter combinations, the simulation for GZ for 2006 to 2010 was implemented (Fig. 7).

The main objective of this study is to explore building cell distribution patterns and trends to gain insight into urban landscape and vertical growth attributes. It is not the cells themselves but rather the patterns that cells develop that are being measured through the simulation. In addition, traditional statistically based accuracy assessment methods are non-site-specific. A more precise approach, namely, fractal dimension-based area–radius relationship measurement, is used to evaluate simulation patterns across space.

Although the simulated and actual building cell results do not coincide completely, they exhibit similar patterns, as shown by the accumulated growth figures for new building cells (Fig. 8). This demonstrates that the model accurately simulates building patterns and trends. Overall, the building cells are initially small within the city center (at a distance of <2 km on the $x$ axis) while increasing significantly outward towards the hinterland. The highest building cell state values from "moderate" to "extremely high" appear close to the fringe edge (namely, at a distance of 10 km on the $x$ axis), while "very low" and "low" categories are most prevalent in the hinterland.

(1) For the "very low" and "low" categories, both the actual and simulated plots show rapid increases as a function of distance, indicating that these two categories are sensitive to distance changes. Although the simulated plots produce noticeably higher values than the actual plots, especially beyond 10 km, they exhibit the following similar features: (a) the most active region of the building distribution ranges between 5 and 10 km for the "very low" category and spans the entire study area for the "low" category; (b) both categories show an even higher level of increase beyond 20 km, indicating a clear intensification of the two classes in the hinterland. This pattern was expected because in regions farther than 20 km from the city center (namely, less-dense suburban areas), scattered and less active building growth patterns are typical.

(2) For the "moderate" and "rather high" categories, areas within 10 km of the urban center experienced a rapid increase in building cells followed by a slow upward trend for the region beyond 10 km from the urban center. Although peak values decreased from $4.82 * 10^4$ (for "rather high") to $3.14 * 10^4$ (for "moderate"), these two classes exhibited virtually the same trend. Growth nearly ceases altogether beyond the fringe zone border, illustrating the compact nature of these two categories.

(3) In contrast to the previous five categories, a unique phenomenon was observed for the "very high" and "extremely high" categories wherein simulated building cells were found to be smaller than actual building cells (Fig. 8). This may be attributable to our rule setting: we determined the height state of the cell in question based on the states of the majority of adjacent cells. Our model therefore makes calculations only on the basis of existing states, and thus, "mutated" building height states cannot be adequately predicted. In practice, the metropolitan area witnessed a higher degree of skyscraper development, especially in the central business area (CBD).

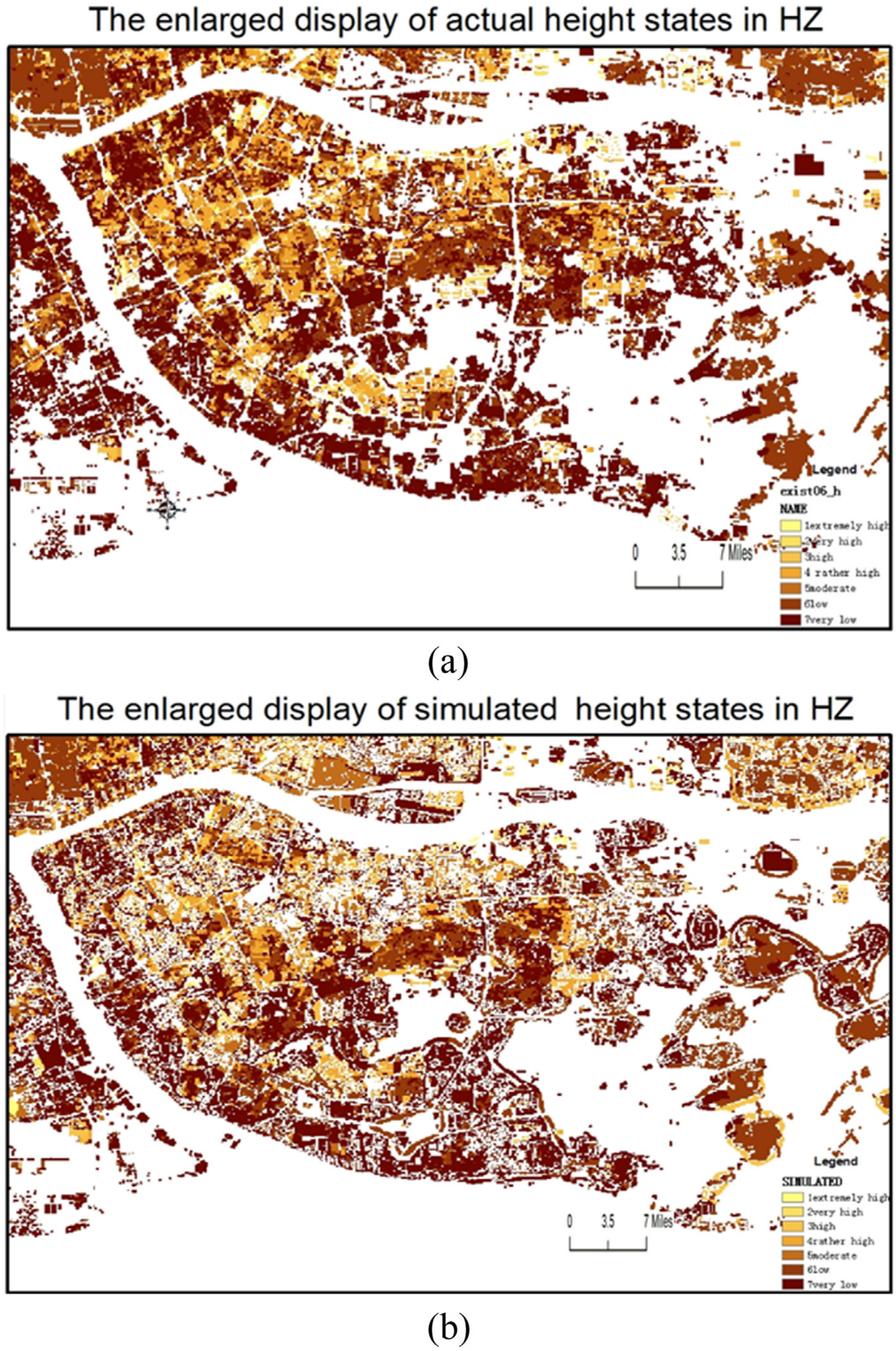
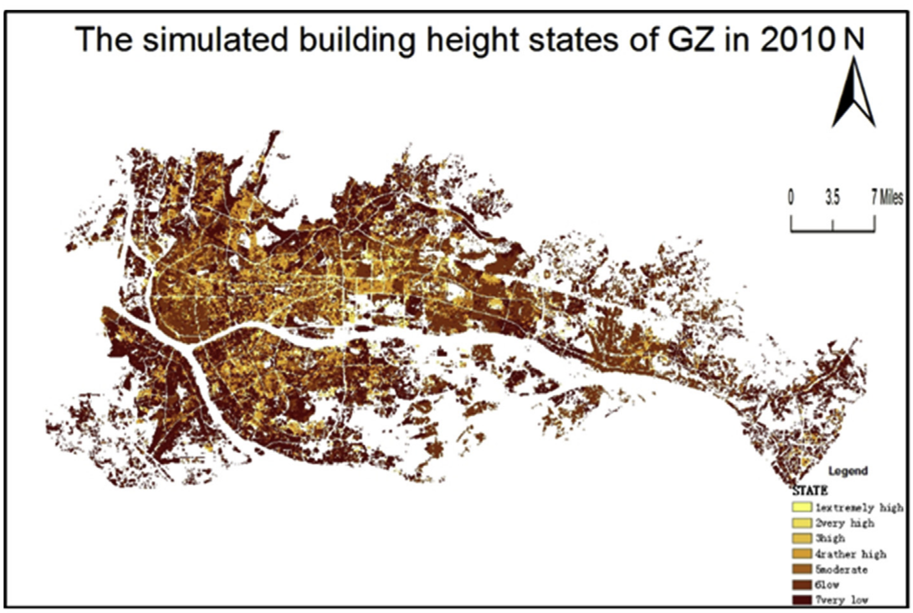



**Fig. 6.** Enlarged display of building distribution in HZ: (a) actual; (b) simulated.

In summary, for the “moderate” to “extremely high” categories, building distribution curves exhibited an initial rise in combination with relatively stable or slightly upward trends throughout the periphery and hinterland areas. The overall distribution patterns and trends for the “very low” and “low” categories were found to be quite different. This contrast shows that moderate to high buildings are more characteristic of “compact development,” while relatively low buildings tend to sprawl outward in a dispersed manner. In addition, the peak values for the “very low” and “low” categories were almost a full order of magnitude larger than those for the other categories. These two features of “very low” and “low” groups demonstrate that these cell categories are undergoing the most rapid development of the whole region and at almost unrestricted levels.

Despite the aforementioned notable differences, all the groups experienced significant degrees of development within the range of 5–10 km from the city center, which denotes this area as a “hot zone” of vertical growth. Although this region lies within the scope of the fringe zone rather than the core zone, the results are reasonable for the following two reasons: (1) due to long-term historical development, available land in the core zone is limited, and this is a significant constraining factor influencing building growth; (2) the 5–10 km zone falls primarily within the Tianhe district, which is the new CBD of GZ city. As the core is already


(a)




(b)

**Fig. 7.** Spatial distribution of building height states: (a) actual distribution for 2010; (b) simulated result for 2010.

highly developed, the fringe zone is now developing into a new section of the city center. The relationship between these two zones illustrates the outward movement of vertical city growth.

*Function and height*

Because the "industrial" function fell almost entirely under the "very low" and "low" categories, the relationship between function and height was not as significant as previously thought. Although each height state includes varying degrees of the four building functions, each function is generally associated with a different height state ranging from "very low" to "very high," indicating the correlation between function and height.

*Discussion*

*Relationships between building heights*

For ease of analysis, we reclassified the height states into three categories: low (for very low and low), moderate (for moderate and rather high), and high (for high, very high and extremely high). According to the above analysis, the low category tends toward a pattern of "outward dispersion," the high category follows a trend of "compact development" and the moderate category exhibits characteristics of both. Although these three categories exhibit different characteristics and magnitudes, they demonstrate a relationship of interaction and co-development rather than of competition. As shown in Fig. 9, the values for all three categories

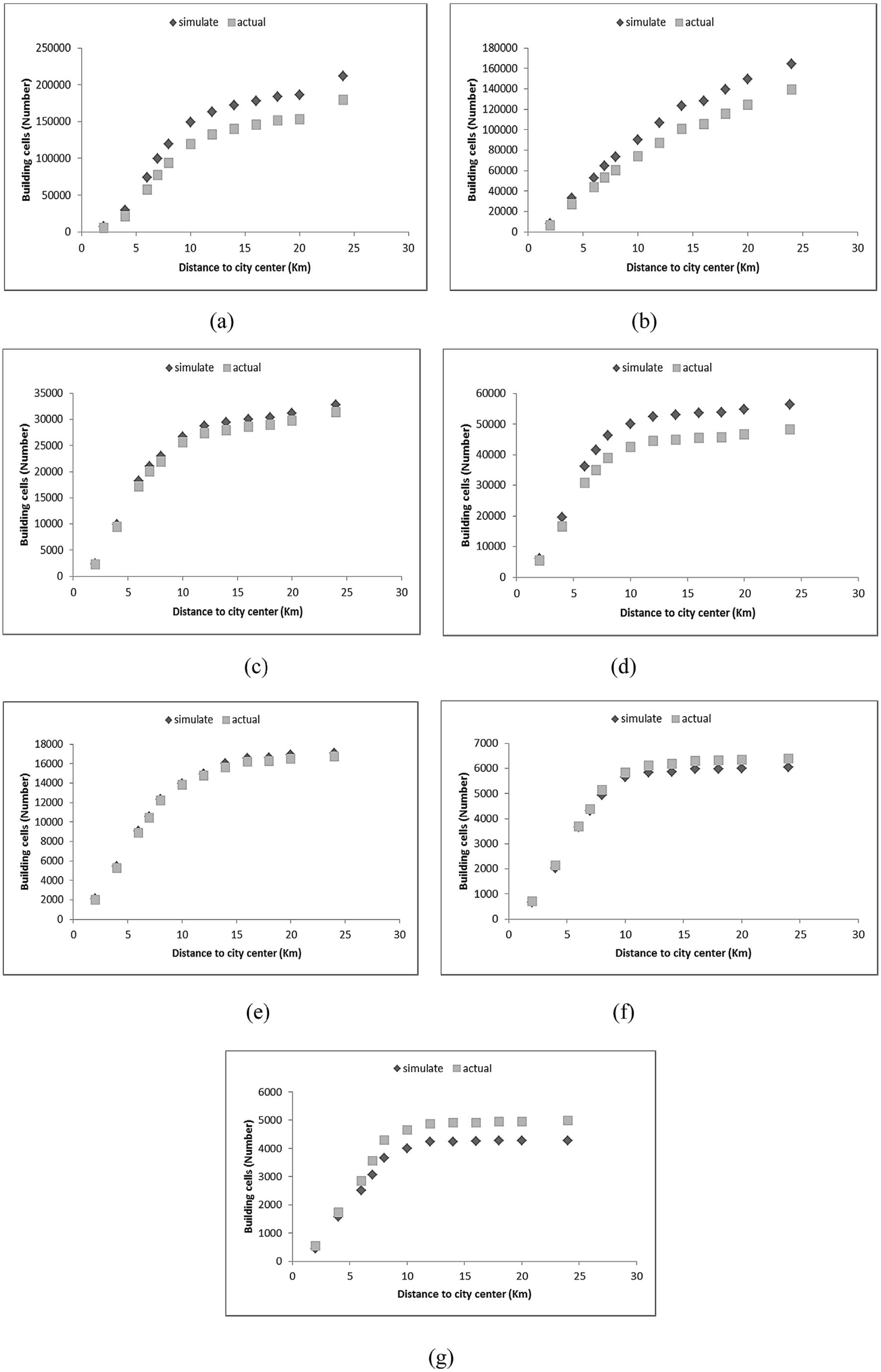



**Fig. 8.** The accumulated growth of new building cell height state changes by distance: (a) very low; (b) low; (c) moderate; (d) rather high; (e) high; (f) very high; (g) extremely high.

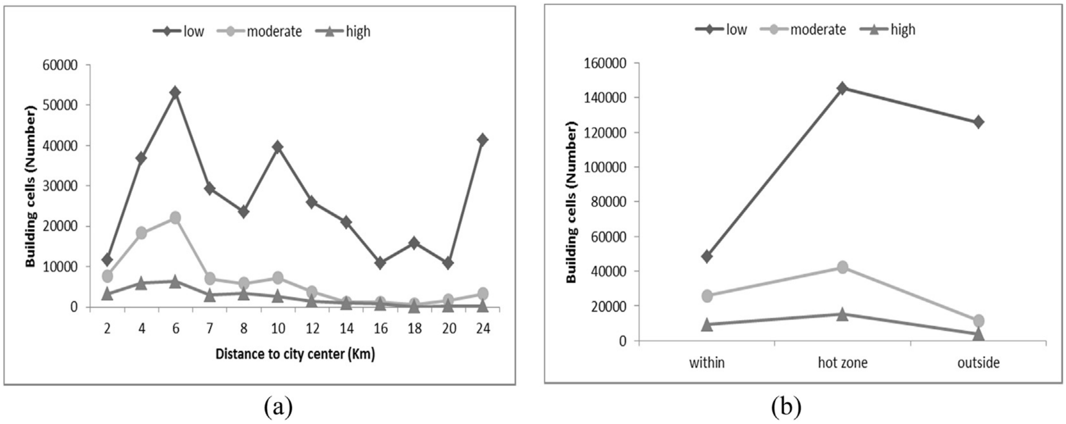



**Fig. 9.** Distribution of building states (low, moderate, high) by distance: (a) detailed results; (b) general pattern extracted from the detailed results.

initially increased and subsequently decreased with increasing distance from the city center.

A scale relationship, also known as a self-similar relationship, was also found among the buildings. As the scale was decreased from the total area to the hot zone, similar obtuse triangles were identified among the different categories of buildings (Fig. 10). Although the exact slopes for the three categories differed, their results were similar to one another and displayed the same scalar relationship. Different building height states can co-exist and can even co-develop. However, it is extremely difficult to determine building height states that are more efficient in terms of urban growth or to identify the most efficient combination of different building states.

### *Zone characteristics and phase transition*

Phase transition denotes a type of change that complex systems strive to avoid. Due to the self-organizing nature of the urban system, the system tends to maintain its current structure. However, phase transition is a critical phenomenon that is occurring in modern cities. The process is described as "unraveling the functional change within existing spatial structures" by Michael (Batty, 2005) and as "the growth and accumulation of disutilities" by Wu (Wu, 1998). As significant developments and consequent qualitative changes occur within the fringe area, the functional structure of the urban system and the relative relationships between different zones evolve in a similar fashion. The hot zone, namely, the fringe area, transitioned from its previous function as a rapidly developing area to that of a new city center.

## Conclusion and future development

### *Conclusion*

Urban vertical growth plays a significant role not only in the transformation of urban landscapes and city skylines but also in the development of urban economies. Moreover, building height states can affect citizens' lifestyles and living environments and thus represent an important facet of urban efficiency and culture. Although low-rise, sprawling buildings dominate the area, the cluster of high-rise buildings within the core and fringe zone creates an organism that causes both positive and negative outcomes. It is difficult to predict whether low- or high-rise buildings are more suitable to sustainable vertical development and thus whether either of these building forms will benefit urban residents to a greater degree. The co-development of varying building heights in appropriate proportions may be the best course of action. However, determining these proportions and effectively integrating low- and high-rise buildings represent major challenges in urban planning, especially with regards to megacities. There is an urgent need to understand urban vertical development patterns and to formulate effective policies that either promote or discourage vertical development based on this understanding.

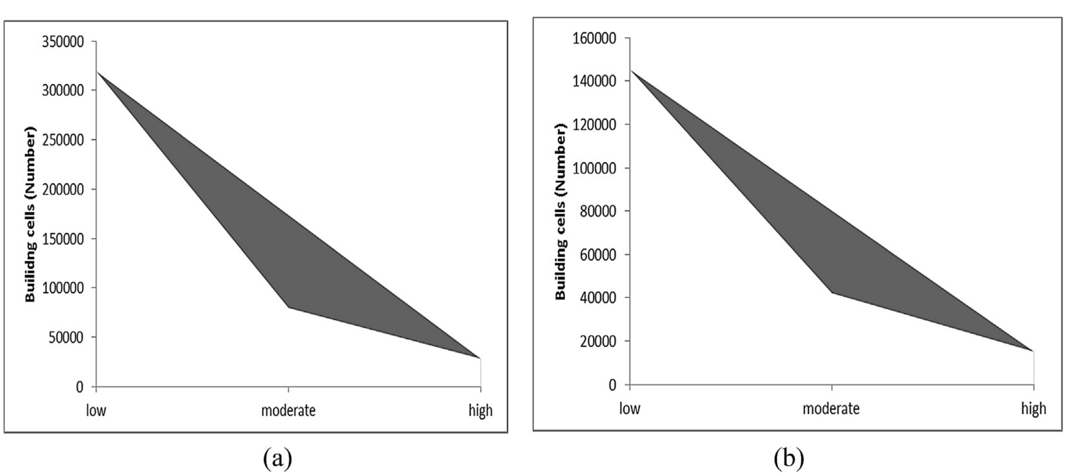



**Fig. 10.** Similar obtuse triangle relationships between low, moderate and high scales: (a) entire study area scale; (b) hot zone scale.

This study developed a CA based approach to effectively model and analyze building distribution patterns. By introducing a series of influencing factors (*e.g.*, accessibility, building and population density), the building height patterns were successfully simulated. Using newly authorized Guangzhou Urban Planning Bureau construction building data for each year, the accuracy of the model was assessed. This model demonstrates that urban vertical growth presents the following pattern of intensification and decline in terms of building floor values across space: the fringe area tends to construct high-rise buildings, while the hinterland zone has witnessed a significant increase in low-rise buildings. However, due to missing data and the variable constraints noted above, the model developed here is limited in capturing exact volumes of building growth. Nonetheless, urban vertical growth patterns across space from the city center, through the fringe, and to the periphery and hinterland zones can be effectively modeled using the technique developed in this study.

*Limitations*

Although we improved the second phase of the simulation for the period of 2006–2010, errors remained in the building growth volume control, which is expressed in terms of low user accuracy. This may be due to several factors, including inconsistencies in the data and variables, such as missing population data for certain years and the introduction of the subway into the model according to its initial operation time. However, the subway has had an effect since the planning stage. Errors may also have been introduced as a result of the nature of building simulation itself. Unlike land-use simulation, which applies a relatively large basal area, each individual building represents a "point" in relation to the entire study area. This complicates building growth simulation, because data integrity and sufficiently sensitive variables are required for a building point to be distinguished from points that surround it. One strategy for remedying this issue involves designing driving force intervals to be sufficiently small. However, this modeling method is far more complex due to numerous possible combinations of parameter intervals. The determination of thresholds and intervals that are sensitive but maintain computing efficiently is an avenue for future research. The model is limited in that it only considers the growth of new buildings while neglecting changes in existing buildings through processes such as destruction or renovation. However, as an early attempt to simulate building growth patterns, this model achieves relatively high levels of accuracy. The resulting data provide us with strong insight into patterns of urban vertical growth.

*Future developments*

This bottom-up characterization of the CA approach has enabled us to avoid the modifiable areal unit (MAUP) (Paez & Scott, 2004) and spatial heterogeneity problems that often complicate traditional mathematical models. However, it was challenging to determine the extent to which the size of a neighborhood is geographically defined. This may introduce the problem of "uncertain geographic context (UGCoP)," as proposed by Kwan (Kwan, 2012). Although methods designed to handle this problem are well established and involve the use of empirical findings by adding the distance-decay effect to the neighborhood configuration and multiscale definition (Batty et al., 1999), these approaches are built on a weak theoretical basis. The optimal scale for delineating contextual units and hence for accurately capturing local–global interactions requires further exploration.

One additional aspect to consider in future research is the model's flexibility and its applicability to other cities. Changes to the model, such as altering the parameter settings or adding or removing driving forces, would be necessary when considering developmental stages of city growth (*e.g.*, rapid growth and highly mature) and city characteristics (*e.g.*, topographical factors and policy planning). The fact that the model developed here is based on the highly mature megacity of GZ must also be noted. Although the task is beyond the scope of the current analysis, future research projects may examine whether the model is sufficiently robust and flexible to evaluate spatio-temporal processes that occur in other types of cities. Another critical area for future exploration would involve determining strategies for systematically selecting intervals for building floor classification and for interpreting results accordingly.

## Acknowledgments

We appreciate the detailed suggestions and comments from the editor and the anonymous reviewers. The work described in this article was supported by the National Natural Science Foundation of China (grant no. 41101439, grant no. 41371417, and grant no.41371424).

## Appendix A

First, for cells occupying the "vacant" state, the following rules apply:

IF $\beta > 0.05$ AND $\gamma > 50$ AND$\alpha$>0.005

THEN the cell will transform from a vacant to building cell

OTHERWISE the cell remains vacant

Second, for cells occupying the "building" state, the following rules apply:

IF a cell is in the 'core' zone AND receives strong transport support

THEN it will transform into a "commercial" cell

IF a cell is in the 'core' zone AND receives weak transport support

THEN it will transform into a "commercial/residential" cell

IF a cell is in the 'fringe' zone AND receives strong transport support

THEN it will transform into a "commercial" cell

IF a cell is in the 'fringe' zone AND receives weak transport support

THEN it will transform into a "residential" cell

IF a cell is in the 'periphery' zone AND receives strong transport support

THEN it will transform into an "industrial" cell

IF a cell is in the 'periphery' zone AND receives weak transport support

THEN it will transform into a "residential" cell

IF a cell is in 'hinterland' zone AND receives strong transport support

THEN it will transform into an "industrial" cell

IF a cell is in the 'hinterland' zone AND receives weak transport support

THEN it will return to "vacant" cell status

Third, according to the height states $\mu$ of existing buildings immediately surrounding the neighborhood of the building cells in question, a corresponding state is assigned. For example, if the majority of the neighboring cells exhibit a "very low" state, then the cell in question is labeled "very low."

## References

Batty, M. (2001). Cellular automata and urban simulation, where do we go from? *Environment and Planning B: Planning and Design, 28*, 163–168.

Batty, M. (2005). *Cities and complexity: Understanding cities with cellular automata, agent-based models and fractals*. Cambridge, MA: MIT Press.

Batty, M., Xie, Y., & Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems, 23*, 205–233.
Bruechner, J. K. (2000). Urban sprawl: disgnosis and femedies. *International Regional Science Review, 23*(2), 160–171.
Bruechner, J. K., & Largey, A. G. (2008). Social interaction and urban sprawl. *Journal of Urban Economics, 64*(1), 18–34.
Burgess, P. (2000). *Conservation and development in conflict: Using GIS to improve land management strategies in the Ipswich Watershed* (Master's Project in International Development). Worcester, MA: Clark University.
Clarke, K., & Gaydos, L. J. (1998). Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science, 2*(7), 699–714.
Clarke, K., Hoppen, S., & Gaydos, L. (1997). A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B: Planning and Design, 24*, 247–261.
Dietzel, C., & Clarke, K. (2004). Spatial differences in multi-resolution urban automata modeling. *Transactions in GIS, 8*(4), 479–492.
Ewing, R., Schieber, R. A., & Zegeer, C. V. (2003b). Urban sprawl as a risk factor in motor vehicle occupant and pedestrian fatalities. *American Journal of Public Health, 93*(9), 1541–1545.
Ewing, R., Schmid, T., Killingsworth, R., Zlot, A., & Raudenbush, R. (2003). Relationship between urban sprawl and physical activity, obesity, and morbidity. *American Journal of Health Promotion, 18*(1), 47–57.
He, C., Okada, N., Zhang, Q., Shi, P., & Zhang, J. (2006). Modeling urban expansion scenarios by coupling cellular automata model and system dynamic model in Beijing, China. *Applied Geography, 23*, 323–345.
Huang, B., Xie, C., & Tay, R. (2009). Land-use-change modeling using unbalanced support-vector machines. *Environment and Planning B: Planning and Design, 36*, 398–416.
Huang, B., Zhang, L., & Wu, B. (2009). Spatiotemporal analysis of rural-urban land conversion. *International Journal of Geographical Information Science, 23*(3), 379–398.
Hurd, J. D., & Civco, D. L. (2004). Temporal characterization of impervious surfaces for the State of Connecticut. In *ASPRS Annual Conference Proceedings, Denver, Colorado*.
Ko, J. H., Chang, S. I., & Lee, B. C. (2011). Noise impact assessment by utilizing noise map and GIS,a case study in the city of Chungju, Republic of Korea. *Applied Acoustics, 72*, 544–550.
Kourtit, K., Nijkamp, P., & Reid, N. (2014). The new urban world: challenges and policy. *Applied Geography*, 1–3.
Kwan, M. P. (2012). The uncertain geographic context problem. *Annals of the Association of American Geographers, 00*(00), 1–11.
Lagarias, A. (2012). Urban sprawl simulation linking macro-scale processes to micro-dynamics through cellular automata, an application in Thessaloniki, Greece. *Applied Geography, 34*, 146–160.
Lavalle, C., Demicheli, L., Turchini, M., Casals-Carrasco, P., & Niederhuber, M. (2001). Monitoring megacities: the MURBANDY/MOLAND approach. *Development in Practice, 11*(2–3), 350–357.
Leung, Y., & Leung, K. S. (1993). An intelligent expert system shell for knowledge-based geographical information system: 2. Some applications. *International Journal of Geographical Information Science, 7*(3), 201–213.
Li, X., Lao, C., Liu, X., & Chen, Y. (2011). Coupling urban cellular automata with ant colony optimization for zoning protected natural areas under a changing landscape. *International Journal of Geographical Information Science, 25*(4), 575–593.
Li, X., & Yeh, A. G. O. (2000). Modeling sustainable urban development by the integration of constrained cellular automata and GIS. *International Journal of Geographical Information Science, 14*(2), 131–152.
Li, X., & Yeh, A. G. O. (2001a). Calibration of cellular automata by using neural networks for the simulation of complex urban systems. *Environment and Planning a, 33*, 1445–1462.
Li, X., & Yeh, A. G. O. (2001b). The application of principal component analysis and cellular automata in spatial decision making and city simulation. *Science China Earth Sciences, 31*, 683–690 (in Chinese).
Li, X., & Yeh, A. G. O. (2002a). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science, 16*(4), 323–343.
Li, X., & Yeh, A. G. O. (2002b). Urban simulation using principal components analysis and cellular automata for land-use planning. *Photogrammetric Engineering & Remote Sensing, 68*(4), 341–351.
Li, X., Zhang, L., & Liang, C. (2010). A GIS-based buffer gradient analysis on spatio-temporal dynamics of urban expansion in Shanghai and its major satellite cities. *Procedia Environmental Sciences, 2*, 1139–1156.
Linard, C., Tatem, A. J., & Gilbert, M. (2013). Modelling spatial patterns of urban growth in Africa. *Applied Geography, 44*, 23–32.
Liu, Y. (2009). *Modeling urban development with geographical information systems and cellular automata*. Boca Raton, FL: CRC Press.
Liu, X., & Li, X. (2008). A bottom-up approach to discover transition rules of cellular automata using ant intelligence. *International Journal of Geographical Information Science, 22*(11–12), 1247–1269.
Liu, X., Li, X., & Yeh, A. G. O. (2007). Discovery of transition rules for geographical cellular automata by using ant colony optimization. *Science in China (Series D), 50*(10), 1578–1588 (in Chinese).
Longley, P., & Batty, M. (2003). *Advanced spatial analysis*. Redlands, CA: ESRI Press.
Moghadam, H. S., & Helbich, M. (2013). Spatiotemporal urbanization processes in the megacity of Mumbai, India: a Markov chains-cellular automata urban growth model. *Applied Geography, 40*, 140–149.
Paez, A., & Scott, D. M. (2004). Spatial statistics for urban analysis: a review of techniques with examples. *GeoJournal, 61*, 53–67.
Project GeoSOS. <http://www.geosimulation.cn/index.html>.
Project Gigaplpolis. <http://www.ncgia.ucsb.edu/projects/gig/index.html>.
RIKS BV. (2005). *Metronamica: A dynamic spatial land use model*. http://www.riks.nl/resources/papers/Metronamica_Brochure.pdf.
Schneider, A., & Woodcock, C. E. (2008). Compact, dispersed, fragmented, extensive? A comparison of urban growth in twenty-five global cities using remotely sensed data, pattern metrics and census information. *Urban Studies, 45*(3), 659–692.
Silva, E. A., & Clarke, K. (2002). Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. *Computers, Environment and Urban Systems, 26*, 525–552.
Syphard, A. D., Clarke, K., & Franklin, J. (2005). Using a cellular automaton model to forecast the effects of urban growth on habitat pattern in southern California. *Ecological Complexity, 2*, 185–203.
Torrens, P. M. (2006). Simulating sprawl. *Annals of the Association of American Geographers, 96*(2), 248–275.
Torrens, P. M., & Alberti, M. (2000). *Measuring sprawl*. http://discovery.ucl.ac.uk/1370/1/paper27.pdf.
Van Delden, H., Escudero, J. C., Uljee, I., & Engelen, G. (2005). METRONAMICA: a dynamic spatial land use model applied to Vitoria-Gasteiz. In *Virtual Seminar of the MILES Project*. Vitoria-Gasteiz: Centro de Estudios Ambientales.
Weng, Q. (2001). Modeling urban growth effect on surface runoff with the integration of remote sensing and GIS. *Environmental Management, 28*, 737–748.
White, R., Engelen, G., & Uljee, I. (1997). The use of constrained cellular automata for high-resolution modeling of urban land-use dynamics. *Environment and Planning B: Planning and Design, 24*, 323–343.
Wu, F. (1998). An experiment on the generic polycentricity of urban growth in a cellular automatic city. *Environment and Planning B: Planning and Design, 25*, 731–752.
Wu, F., & Webster, C. J. (1998). Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B: Planning and Design, 25*, 103–126.
Xie, Y. C. (2002). *Projecting urban land use changes: DUEM – Dynamic urban evolution modeling*. http://www.casa.ucl.ac.uk/duem/DuemHelp.pdf.
Yang, L., Huang, C., & Homer, C. G. (2003). An approach for mapping large-scale impervious surfaces: synergistic use of Landsat-7 ETM+ and high spatial resolution imagery. *Canadian Journal of Remote Sensing, 29*, 230–240.