International Journal of
*Geo-Information*

MDPI

*Article*

# Modeling Patterns of Land Use in Chinese Cities Using an Integrated Cellular Automata Model

**Yanlei Feng [1,*] and Yi Qi [2]**

1 Department of Geography, University of California, Berkeley, CA 94720, USA
2 School of Architecture and Urban Planning, Nanjing University, Nanjing 210023, Jiangsu, China; jnjn@163.com
* Correspondence: ylfeng@berkeley.edu

Received: 22 August 2018; Accepted: 9 October 2018; Published: 12 October 2018

check for updates

**Abstract:** This paper introduces an urban growth simulation model applied to the full scope of China. The model uses a multicriteria decision analysis to calculate the land conversion probability and then integrates it with a cellular automata model. A nonlinear relationship is incorporated in to the model to interpret the impacts of different Land Use and Cover Change driving forces. The Analytical Hierarchical Process is also implemented to compute the variance between weights of different factors. Multiple sizes of neighborhood and different urban ratios in the model rules are tested, and a $5 \times 5$ neighborhood and an urban threshold of 0.33 are chosen. The study demonstrates the importance of spatial analysis on socioeconomic factors, population, and Gross Domestic Product in land use change simulation modeling. The model fills the gap between the purely economic theory simulation model and the geographic simulation model. The nationwide urban simulation is an example that addresses the lack of urban simulation studies in China and among large-scale simulation models.

**Keywords:** urban growth modeling; multicriteria simulation; cellular automata; China

## 1. Introduction

There has been long-term steady urban growth worldwide. Developing countries have been the major contributors to urbanization. In 2018, nine out of the top 10 largest cities in the world were located in Asia. It is projected that the number of urban populations will double and the urban area will triple in size by 2030 [1]. Urban expansion has increased at a higher rate than the urban population, and rapid expansion has led cities to face enormous challenges [1], such as uncontrolled informal settlement, insufficient urban service [2], climate change and global warming [3,4], negative effects on social-environmental responses [1,5], and consumption of agriculture and natural land [6].

Research from multiple disciplines has addressed the importance of studying urban growth to better understand the occurrence and consequences of urban expansion and to explore urban sprawl across space and over time [2,6,7]. Simulations of the future urbanized area can assist the government and urban planners in policy making, land use, and land management in response to fast economic development and rapid population growth. Such simulations are typically generated using urban growth models and mapped with GIS and remote sensing techniques.

Cellular automata (CA) models for urban growth simulation have proliferated over the past 20 years due to their simplicity, flexibility, and intuitiveness [8]. CA was first developed by S. Ulan and J. von Neumann in the 1940s and applied as a theoretical approach for the simulation of urban expansions in the 1980s [9–11]. The assumption of CA models is that future patterns of land use will be affected by past urban development through local interactions [8], so the development of computing power led directly to the emergence of operational CA models. CA models are effective and reliable in spatial and temporal simulation research [12]. CA models have four elements: a discrete cell space, states of cells, rules of neighborhoods, and rules of transition. The state of a cell is determined by its previous state and the states of its neighboring cells according to transition rules (Betty 2007). Though simple, CA models are capable of incorporating the spatial and temporal dimensions of processes and modeling complex dynamic systems. Recently, advances in computers have increased the number of CA models used in real-world urban simulations [8,12]. Moreover, the urban CA model can be easily integrated with the GIS environment, which can produce high spatial resolution maps [2].

Compared to a general CA model, the Land Transformation model, although demonstrating a high capacity for prediction with high resolution, has complex operational steps, which make it one of the least popular applications [13]. The Weights of Evidence model (WE) requires rich data and detailed maps, making it very hard to collect all the needed data [14]. An integrated model combining the Frequency Ratio (FR), Analytical Hierarchical Process (AHP), Logistic Regression (LR), and Artificial Neural Network (ANN) was designed to predict and compare urban growth [15], but the model had difficulty identifying the best method due to the differences among the requirements, needs, and the means of each method. Therefore, CA is one of the most popular approaches in urban growth modeling. Agent-based model (ABM) is also one of the most popular models in land use and land cover change studies [16,17] which include individual agents. By following rules, those agents steer land use development by planning, permitting, or restricting land use changes [18]. ABMs are applicable to land use change studies with individual decision-making processes, while CA models are mainly used in studies based on historical land use change patterns.

Early applications of CA to urban dynamic modeling were theoretical models, which allowed modelers to test hypotheses of urban theory and simulate simple urban structures [19]. Theoretical CA models in their simple forms are not strong enough to create realistic simulations due to their simplicity and inability to consider social, economic, and demographic factors when simulating urban dynamics. Hence, these conventional CA models were integrated with quantitative systems to simulate real-world urban development processes [20,21]. Integrating CA models with the Markov Chain [22,23], Analytical Hierarchical Process [24], Logistic Regression [25,26], Multicriteria Evaluation [27,28], Support Vector Machine [29], and ANNs [30] overcomes the limitation of conventional CA [31]. The SLEUTH model was built with multiple data layers, including slope, land use, exclusion, urban extent, transportation, and hill shade [20], so the possibility that the transition rules would affect the state of cells is tuned and weighted through those layers, making it a widespread model [6,20,32]. Integrated CA-AHP and CA-LR models are the strongest in terms of their ability to deal with the most factors; their strong validity in simulations; their effectiveness in explaining the results; and their ability to generate different scenarios using both environmental and social-economic factors [12]. Moreover, previous studies have shown the effectiveness of AHP in using weights of factors on quantitative models and spatial-temporal factors to create realistic simulations. Moreover, AHP has the ability to combine mathematical methods and the experience of experts in the field of urban studies.

Based on Santé et al.'s review (2010) [8], there are relaxations in the original structure of CA that allow the introduction of more complexity to the models.

- **Transition rules:** One of the most commonly used rules is transition potential. It is usually calculated as the weighted sum of a number of factors, including road accessibility, distance to urban centers, slope, accessibility to railways and water, planning and environmental factors, suitability for development, population density, and Gross Domestic Production (GDP). GDP is the most commonly used factor in the majority of the research [25,33]. Different techniques can be used to calculate the transition potential, such as logistic regression or multicriteria evaluation [28].
- **Cell space:** The cell space can be irregular and non-uniform, for example, graphs [34,35].
- **Neighborhood:** The neighborhood space can be extended with a distance-decay effect or defined differently according to its state and location [36].
- **Cell states:** Most urban simulation models make transitions between two states: urban and non-urban. Some models extend the transition to multiple land uses using a Markov chain [37].

The integration of CA models with GIS and AHP is widely used to simulate land use changes. However, most of the models are used at regional scales and only a few of them are at a large scale. The aims of this study are (1) to develop a prototype simulation model for national-scale urban simulation studies by integrating GIS and CA and (2) to analyze the sensitivity of the model with respect to:

- **The neighborhood size:** How does the size of the neighborhood affect the simulation results? What is the best fitted neighborhood size in this study?
- **The urban ratio:** How does the urban ratio in a transition rule affect the urban growth speed? What is the most appropriate urban ratio in this study?

We employ the transition probability as a transition rule in the model. AHP is used to determine the weights of multiple land use change drivers. China is used as a case study to test the model.

## 2. Study Area

China has the fourth largest land area in the world. Its urbanized area has experienced a fast-expanding period over the past 40 years. The urban population in China is anticipated to reach a level higher than 70% by 2035 [38]. The number of cities increased from 193 to 658 between 1978 and 2013 in China [39].

China also has a very diverse urban area across more than 650 cities. The rate of urban expansion across cities has varied from 0 to over 90%, showing a broad spatial and temporal variability [33]. In Deng et al.'s (2008) econometric model, GDP (or income) is a positive and highly significant coefficient, and it was shown to explain nearly 40% of the variability of the urban core expansion in the model [33]. By comparing the model with both economic (such as GDP, population, agricultural investment) and geographical factors (such as existing urban area, highway density, distance to port cities and provincial capital, rainfall, slope, temperature, elevation) with a model that includes only geographical factors, the authors concluded that geographical variables play the major role in explaining the diversity across space in urban core expansion, while economic factors measure their impacts more precisely [33].

Combining urban growth drivers from past urban simulation studies and Deng et al.'s research of urban expansion in China, we chose seven factors as the urban growth drivers for generating land conversion probability, which are displayed in the Data section.

## 3. Data

All spatial datasets were projected into the same coordinate system, China Albers Equal Area Conic, in meters. Euclidean distances were generated from the city center, railway, major roads, and rivers in ArcMap, and then clipped to China's boundary files (Table 1). Missing data values in the layer of the slope, land use in 2000, or population per hectare were converted to 0 for better calculation.

**Table 1.** Data content and source.

| Class | Data Type | Year | Data Source | Resolution |
|---|---|---|---|---|
| Impervious Surface Area | Nighttime images DMSP-OLS and MODIS | 2000 | Beijing City Lab | Downscale to 1 km from 500 m resolution |
| City centers | shapefile (point) | 2000 | China Data Center University of Michigan | |
| Railroads | shapefile (lines) | | Single, multiple, and light track rails and connectors | |
| Major roads | shapefile (lines) | | Highways, trails, and footpaths | |

**Table 1.** *Cont.*

| Class | Data Type | Year | Data Source | Resolution |
|---|---|---|---|---|
| River | shapefile (lines) | | Rivers stretching for thousand kilometers | |
| Slope | raster | | | Convert from DEM |
| Population density | raster | 2000 | China Temporal Dataset in Harvard Dataverse | Convert to 1 km grid |
| GDP | raster | 2000 * | National Earth System Science Data Sharing Infrastructure | 1 km grid |

* Estimated GDP in 2000 based on data from 2005 using the World Bank GDP growth rate.

## 4. Methods

### 4.1. Algorithm

In a cellular automata model, the state of a cell at time $t+1$ is the result of its own state at time $t$ and the state of its neighbors under certain conditions. In this urban growth research, the state of the cell at $t+1$ (1) is a function of (a) its own state at time $t$; (b) the land suitability score generated by the multicriteria analysis on independent drivers; and (c) its neighbors' states under specific rules, which is explained in the next section.

$$S_{t+1} = f(S_t, P_t, N_t) \tag{1}$$

$$P_t = \sum_{k=0}^{n} X_k w_k \tag{2}$$

$$X_{k1} = e^{-aD_{k1}} \tag{3}$$

$$X_{k2} = b * ln(D_{k2}) \tag{4}$$

$$a = 1/(V_{max}/10) \tag{5}$$

$$b = 1/ln(V_{max}) \tag{6}$$

where $S_{t+1}$ is the state of the cell at time $t+1$, and $p_t$ is the land conversion probability. Equation (1) used random numbers to decide whether to change the state of the cell based on $p_t$. In this research, Moore's eight-cell neighborhood was chosen initially. $N_t$ represents the neighborhood Land Use and Cover Change (LUCC) change rules:

1. Calculate the urban ratio of the cell: The number of urban neighbors in a total neighborhood of N (e.g., Moore Neighborhood is 8), which indicates an $X/N$ urban ratio. The urban ratio should be equal to a user-defined urban ratio.
2. Compare a randomly generated number between 0 and 1. The random number should be smaller than the land suitability of the cell.

Stochasticity is included in the LUCC process. All the non-urban cells meeting the requirements at the same time are converted from non-urban to urban. If the random number is lower than $p_t$ and the urban ratio is equal to a defined one, the cell meets requirements and is input into the next run of calculations. Otherwise, the cell remains the same and waits for the next iteration of calculations. Therefore, cells with high land suitability scores have a greater chance of conversion to urbanized land (Figure 1). The process of randomness simulates real-world land use changes that are caused by subjective factors.

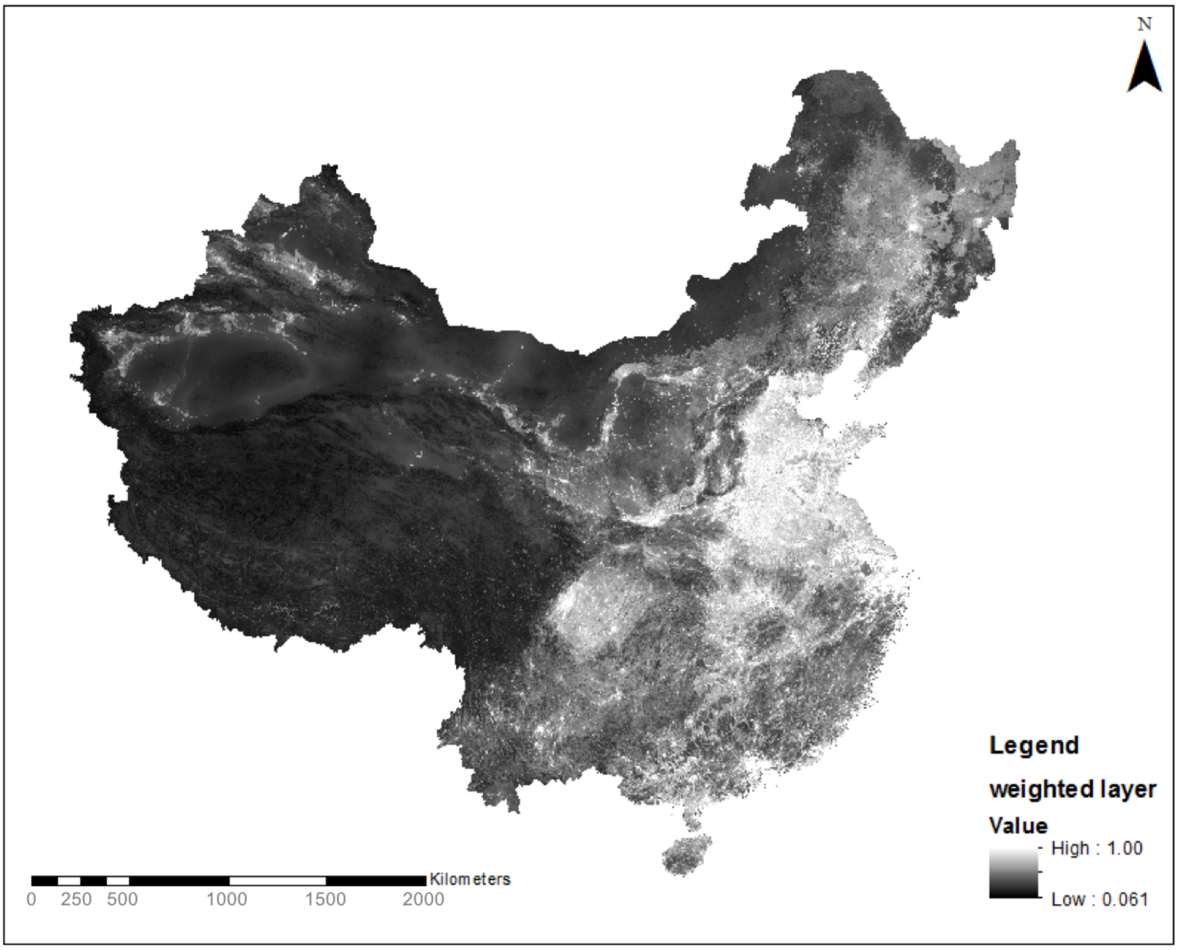


**Figure 1.** Land suitability map.

Equation (2) explains the process for calculating land suitability scores. $X$ represents multiple factors and $w$ is the weight assigned to each factor. There are seven independent factors: distance to city center, distance to highway, distance to railway, distance to river, slope, GDP, and population. The AHP method was employed to determine the weights of various factors and reach a balanced combination. The land conversion probability was generated by the weighted sum of factors. Equations (3) and (4) express the relationship between each independent geographical factor and the land use change probability. $D$ is a factor and '$a$' is a parameter that controls the growth speed. $D_1$ represents factors that are negatively related to urban expansion. For example, the land conversion probability becomes larger when moving close to city and town centers. The land conversion probability increases when building in a flatter area. Therefore, '$a$' is a positive number for normalizing the scale of multiple factors and maintaining the negative relationship. However, the GDP and population factors play different roles. GDP is a highly significant coefficient in an ordinary least-squares estimator of the expansion of the spatial size of Chinese cities which is positively related to the size of the urban area [33]. Clusters of population attract new land development. Therefore, different from Equation (3), Equation (4) was used to fit the positive relationship between the GDP, population, and urban area size. The formula showed an increasing growth rate at first, and then the growth rate decreased, which simulates the correlation between GDP and urban growth in China. There is also a parameter '$b$' in the formula that normalizes the scales of factors.

The AHP was finished in three consecutive steps. The first step was to calculate the vector of criteria weights based on a pairwise matrix. Saaty (1987) introduced a table of relative scores when making the pairwise comparison: 1:1 means that the two factors are equally important, 3:1 represents slightly more importance, 5:1 means moderate importance, 7:1 demonstrates stronger importance, and 9:1 indicates that one in the pair is much greater than the other [40]. In the matrix, $a_{jk}$ represents the relative importance of the $j$th criterion to the $k$th criterion. The relative importance of those parameters is achieved by consulting and surveying the opinions of 10 experts in city planning and urban growth.

Based on Saaty's theory, the criteria weight vector $w$ was built (Table 2). The final result gave GDP a very significant weight of 0.43 and Population Per Hectare (PPH) was given a weight of 0.32, while the weights of the remaining physical geography factors were below 0.1.

Due to the variety in their scales, each geographic factor has a distinct parameter a, and each socioeconomic factor has a parameter b to standardize the scales of the factors. Equations (5) and (6) were used to standardize the scales of both geographic factors and socioeconomic factors on a scale of 0–1.

### 4.2. Model Implementation

The model was attached to the python module in ArcMap, which used ArcPy to link between the GIS software and Python. Through the connection, additional functions were implemented in ArcMap. Parameters '*a*' and '*b*' were calculated to ensure all the factors were on the same scale, from 0 to 1. Then, Formulas (3) and (4) were applied. A weighted sum was calculated. Then, a weighted layer was generated and saved for use in the CA model. The detailed process is shown in (Figure 2).

The model ran from the left upper corner of the NumPy array. New land use changes were generated in separate layers. The model iterated by using the previous output LUCC layer as its input and repeating the same process.

The neighborhood definition and LUCC rules can be easily modified through programming. Therefore, different scenarios were used to test the model's stability and to validate the model, as discussed in Section 5.

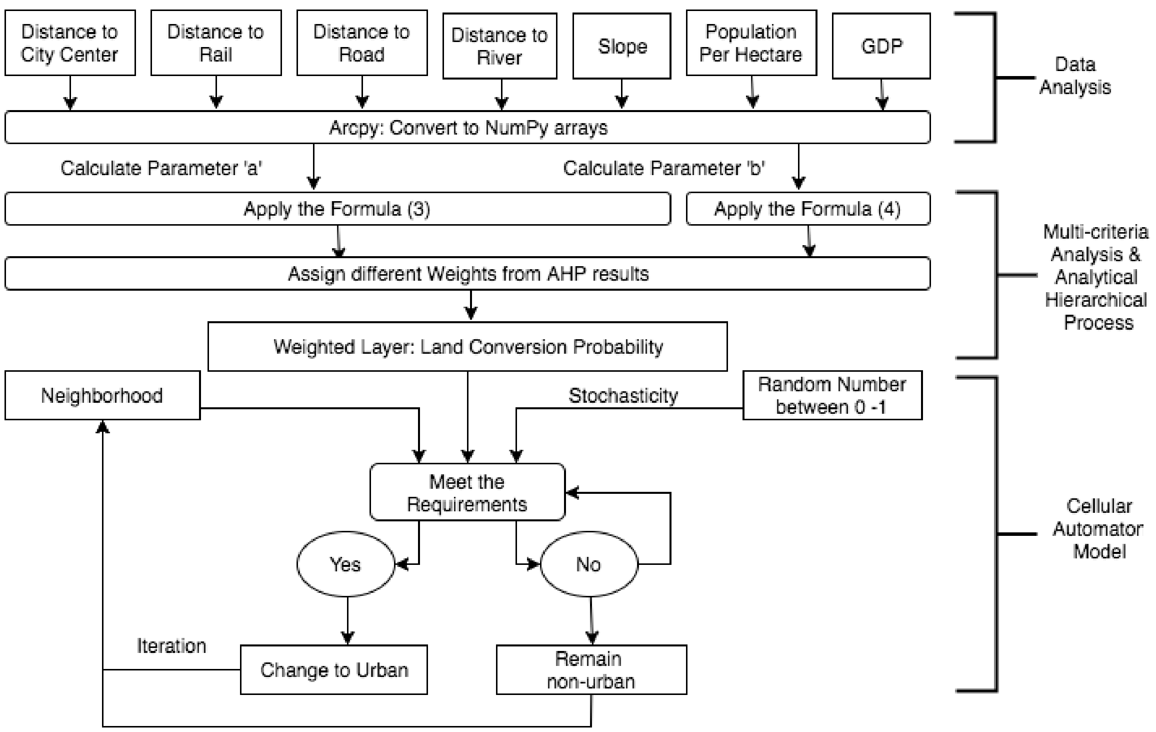


**Figure 2.** Methodology flowchart.

**Table 2.** Analytical Hierarchical Process (AHP) pairwise matrix.

| Class | Dist_CityCenter | Dist_Rail | Dist_Road | Dist_River | Slope | PPH | GDP |
|---|---|---|---|---|---|---|---|
| Pairwise Matrix * | 1 | 3 | 3 | 3 | 5 | 9 | 9 |
| Results | 0.02 | 0.04 | 0.06 | 0.05 | 0.09 | 0.32 | 0.43 |

* The numbers represent the relative scores of each factor compared to the distance to the city center (Dist_CityCenter). Dist_CityCenter = Distance to City Center, Dist_Road = Distance to Road, Dist_River = Distance to River. The numbers represent the relative scores of each factor compared to the distance to the city center (Dist_CityCenter).

### 4.3. Edge Pixels

The edge cells of the map do not have the same number of neighborhoods as the center cells. In a $3 \times 3$ neighborhood simulation, a simple model starts from the second cell of the second line to the second last cell of the second last line (black window in Figure 3), leaving the outermost boundary cells out of the simulation. There are no urban cells in the outermost boundary, so this is an easy way to solve the edge problem without edges. However, if the number of neighborhoods keeps increasing, there will be an increasing number of errors and bias using the previous method. Therefore, new rules had to be defined to include the edge cells in the model.

We expanded the original land use map with an outer boundary filled with the number 2 to solve the issues of neighbor overflow (Figure 3).

$$W_O = 1/2 * (L_n - 1) \tag{7}$$

$$R_e = N_u / N_{s<2} \tag{8}$$

where $W_O$ represents the width of the added outer boundary, and $L_n$ represents the length of the neighborhood. Equation (7) ensures that the original edge cells have the same number of neighborhoods as the center cells. We started the simulation from the first cell in the first line of the original land use map and calculated the ratio of urban cells by Equation (8). $R_e$ represents the urban ratio of edge pixels, $N_u$ represents the number of urban pixels whose state is 1, and $N_s$ represents the number of pixels whose state is less than 2. Equation (8) is a convenient way to separate necessary pixels, so the number of the neighborhoods can easily be modified.

| | | | | | |
|---|---|---|---|---|---|
| 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 0 | 0 | 0 | 0 | 2 |
| 2 | 0 | 1 | 0 | 1 | 2 |
| 2 | 1 | 0 | 1 | 1 | 2 |
| 2 | 0 | 0 | 0 | 1 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 |

**Figure 3.** The urban ratio of the first non-urban cell in the upper left window was 1/3 = 0.334 (cells with 0 are non-urban, those with 1 are urban, and those with 2 represent the added outer boundary, which makes the computation easier).

## 5. Parameter Tests

To improve the results of the model, besides the modifications to the edge effects, there were still two parameters in the model which could be adjusted: the number of neighborhoods and the urban ratio threshold. We modified the original rules: if the urban ratio of a non-urban cell's neighborhoods was over a certain threshold, and the generated random number was greater than the cell's land suitability at the same time, the cell became an urban cell and was exported to the new layer. In this new rule, two parameters were edited to evaluate their influence on the final results of the model. Then, we performed multiple tests on different sizes of the neighborhood with the same urban ratio, and then changed the urban ratio threshold, expecting to understand the role that these two parameters play in this model.

### *5.1. Neighborhood Size Tests*

We tested the Moore Neighborhood first and began with an urban ratio of 0.375. Then, we increased the number of neighborhoods while keeping the urban ratio stable at 0.375, and performed different tests (Figure 4).

With a neighborhood of 3 × 3, almost 95% of the existing urban cells were surrounded by new urban cells. The smallest urban area with new urban growth had only two original urban cells. The new urban growth was 5201 km$^2$.

With a neighborhood of 5 × 5, the smallest urban area with new urban growth had seven original urban cells. The larger the historical urban area was, the more new urban cells that grew adjacent to it. New urban cells not only appeared next to existing cells but also accumulated near newly grown cells. Therefore, the original large urban areas, such as the city of Beijing and the city of Tianjin, were enlarged by a circle of new urban cells. This simulation with a 5 × 5 neighborhood had the largest new urban growth: 6467 km$^2$.

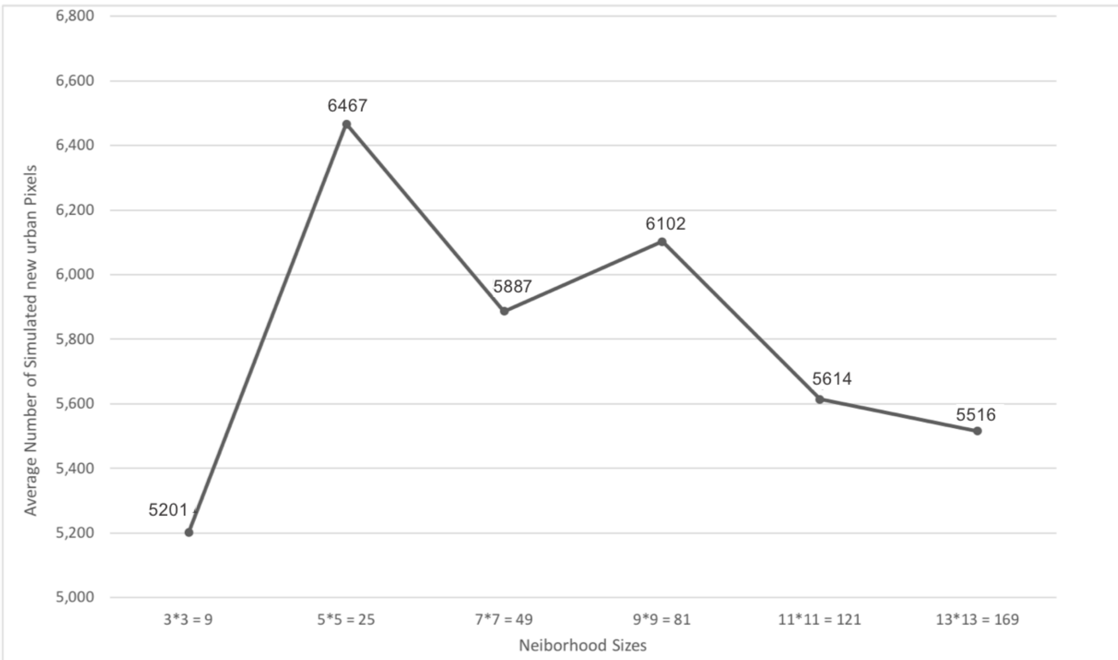


**Figure 4.** Parameter tests on different neighborhood sizes, with a fixed urban ratio of 0.375. Each value represents an average of 10 tests.

With a neighborhood of 7 × 7, a large number of small existing urban clusters remained the same after this simulation. Around large urban clusters, some new urban cells were not tied closely to the edges. They extended close to but not directly adjacent to urban clusters. Starting from this simulation, the computation time of one simulation increased exponentially if we increased the neighborhood size. The new urban growth decreased to 5887 km$^2$.

With larger neighborhoods, only a few metropolitan clusters expanded with new urban cells, which is not true in reality because most urban growth results from suburban sprawl.

Due to the stochasticity in the model, the results were different each time, so we ran each test 10 times and calculated the average number of simulated urban pixels (Figure 4). There was a regular pattern in the results. The simulated urban development was shown to grow when the number of neighborhoods increased and reached a peak at around $5 \times 5$. After that, the number of changes started to decrease with the continuous increase of neighborhoods. The variation among those results was not obvious. In Figure 4, there is a difference of around 1300 pixels between the peak number of changes at $5 \times 5$ and the fewest changes at $3 \times 3$. When we conducted simulations using neighborhoods of $3 \times 3$ and $5 \times 5$, new urban cells occurred adjacent to a small area of original urban land use, which is true in reality. However, when the simulation was performed with an increasing number of neighborhoods, new urban cells only occurred adjacent to a large area of original urban areas (See Supplementary Material).

### *5.2. Urban Ratio Tests*

After testing different sizes, the combination of $3 \times 3$, $5 \times 5$, and $7 \times 7$ neighborhoods stood out and generated the most realistic results. In the following tests, we kept the neighborhood size stable while changing the urban ratios (Table 3).

With a $3 \times 3$ neighborhood, the urban growth after one simulation had large variance between the simulation using an urban ratio of 0.25 (two urban cells out of eight neighborhoods) and that using an urban ratio of 0.375 (three urban cells out of eight neighborhoods) (Figure S1a). The new urban area generated by the former was 9939 square kilometers, which is almost twice the latter. However, there was less variance between the simulated urban growth using urban ratios of 0.375 and 0.5 (four urban cells out of eight neighborhoods). Their urban growth results were almost the same: around 5200 $km^2$.

The simulation generated by a $5 \times 5$ neighborhood showed a similar trend to the previous simulation (Figure S1b). It had 6506 new urban cells, with an urban ratio of 0.375 (nine urban cells out of 24 neighborhoods) and 32% more urban cells with an urban ratio of 0.33, which means the new urban cells had at least eight urban neighbors. With an urban ratio of 0.5, the simulation generated 32% fewer urban cells compared to that with a 0.375 urban ratio.

The pattern was similar for a $7 \times 7$ neighborhood (Figure S1c). The results showed a new urban area of 7795 square kilometers using a ratio of 0.3125 (15 urban cells in 48 neighborhoods), which is 34% more than the urban growth using a ratio of 0.375. The results generated by a ratio of 0.5 had 25% fewer urban cells than when using 0.375. Moreover, with an urban ratio of 0.5 in the model, urban growth decreased, along with the increase in the size of the neighborhood.

After testing multiple neighborhoods with different urban ratios as the threshold in the model (Figure S1d–f), we picked three tests with results after the first simulation that were very similar to the urban land cover in 2005 and evaluated them. The first chosen test used a $3 \times 3$ neighborhood with an urban ratio of 0.25. The evaluation test produced a kappa coefficient of 0.79, which indicates that there is good agreement between the simulated result and the real urban land cover in the year 2005. The second chosen test used a $5 \times 5$ neighborhood and an urban ratio of 0.33. The kappa coefficient given by this test was 0.78, which is slightly lower than the first one, but it still demonstrates good agreement. The third test used a $7 \times 7$ neighborhood and 0.3125 as the urban ratio, and the kappa coefficient produced by the evaluation test was 0.77. All three tests had an overall accuracy of around 0.89 (Figure S2).

Then, we used simulated land use cover from 2005 as the input layer, and then we generated the land use cover for 2010 and 2015 with three sets of parameters. The results are shown in Tables 4 and 5. The results of the simulation with a $5 \times 5$ neighborhood and an urban threshold of 0.33 were the closest to the observed urban area. The simulated urban growth rate in 2000–2005 was 0.39, 14.7% higher than the observed growth rate, 0.34, and the simulated growth rate between 2005 and 2010 was 0.81, 3.6% less than the observed rate. Therefore, the final results were produced using this set of parameters.

After validation and calibration, new simulations were completed based on the historical land use data in 2010. The model was run four times. Each iteration represented a 5-year urban growth change. The model simulated the urban growth changes until 2030. The results are demonstrated in Figure S3.

The simulated future urban pattern is similar to the historical urban area in 2010. Three major existing large urban clusters, including Beijing-Tianjin, Yangtze River Delta, and Pearl River Delta, were estimated to continue to expand. Capital and large cities in central provinces grow quickly, and Hohhot, Baotou, Xi'an, and Wuhan are expected to have large urban growth compared with a map from 2010.

According to the simulation, there will be an increasing trend in terms of the total urbanized area, and the simulated urban area in 2030 approaches 105,000 square kilometers, which is more than twice the observed urban area in 2010. The simulated results indicate a decreasing trend in urban growth rate, from around 38% in 2015 to 17% in 2030. The new urban growth is predicted to become smaller and smaller in the future (Table 6).

**Table 3.** Urban ratio tests.

| Urban Ratio | 0.25 | 0.3125 | 0.33 | 0.375 | 0.5 |
|---|---|---|---|---|---|
| Neighborhood of 3 × 3 | 9939 | / | / | 5224 | 5202 |
| Neighborhood of 5 × 5 | / | / | 8576 | 6506 | 4910 |
| Neighborhood of 7 × 7 | / | 7795 | / | 5814 | 4354 |

**Table 4.** Comparison between observed and simulated urban growth areas.

| | Urban Area from Remote Sensing Images/km$^2$ | Neighbor 3 × 3 Threshold: 0.25 | Neighbor 5 × 5 Threshold: 0.33 | Neighbor 7 × 7 Threshold: 0.3125 |
|---|---|---|---|---|
| 2005 | 29,251 | 31,787 | 30,424 | 29,643 |
| 2010 | 40,533 | 42,597 | 39,628 | 38,253 |
| 2015 | / | 54,151 | 49,864 | 47,755 |

**Table 5.** Comparison between observed and simulated urban growth rates.

| | Growth Rate from Remote Sensing Images | Neighbor 3 × 3 Threshold: 0.25 | Neighbor 5 × 5 Threshold: 0.33 | Neighbor 7 × 7 Threshold: 0.3125 |
|---|---|---|---|---|
| 2005 | 0.34 | 0.45 | 0.39 | 0.36 |
| 2010 | 0.84 | 0.95 | 0.81 | 0.75 |
| 2015 | / | 1.48 | 1.28 | 1.19 |

**Table 6.** Simulated urban growth until 2030.

| Urban Simulation | Simulated Urban Area (in Square Kilometers) | Simulation Growth Rate % |
|---|---|---|
| 2010 (existing land use) | 40,607 | / |
| 2015 | 55,841 | 38.6 |
| 2020 | 71,273 | 25.2 |
| 2025 | 87,637 | 20.0 |
| 2030 | 105,026 | 17.3 |

## 6. Results

The new land use changes were shown to occur close to the edge of existing urban areas and enlarge the original urban area. Some non-urban areas in the interspace of urban areas were shown to quickly fill with urbanized spaces, which is true in reality: the area between two developed cities rapidly urbanizes and may form a small new city center. Good examples are the satellite cities of Shanghai, such as Jiading, Qingpu, Baoshan, and Songjiang [41].

## 7. Discussion

CA models have inherent challenges associated with the use of parameters and their impacts on model results. Uncertainties in the model generated by those parameters are often ignored or not adequately addressed. This study examined the impacts of the neighborhood size and urban ratio on the generated model outputs using a visual comparison, the coincidence matrix with a KAPPA index. The results indicate that there are significant impacts from a changing neighborhood size and urban ratio on the simulated outputs.

Under the transition rule, a small neighborhood size results in a more dispersed pattern of urban growth. Under the rule of small neighborhood size, only a few urban cells are required to enable urban growth. Therefore, most urban growth occurs adjacent to existing land use, regardless of the patch size of the existing urban area. As the neighborhood size grows, it requires more urban cells to enable urban growth. A lot of small-sized urban patches that are far from large urban centers start to lose the ability to generate urban growth. The majority of urban growth starts to accumulate around the large urban centers. For a given neighborhood size, urban growth decreases as the urban ratio increases. Therefore, the urban ratio also has a significant effect on the model's output. Both neighborhood size and urban ratio require fine-tuning when using the model.

It is worth noting that the significance of urban models does not only lie in the results, but also in their individual behaviors and the settings of behavior rules. Through the process of developing models, we learn how each model functions. For urban simulation models, although the simulated results will never be perfect, planners can still get to know the specific kinds of factors which may influence urban growth and how those possible factors lead to actual land use changes.

This research demonstrates that the cellular automata model is appropriate for simulating urban growth in China. The calibrated model can predict the urban growth rate and spatial location of urban areas, but it still has flaws. The population data used to generate the suitability map is from the year of 2010, and it is stable throughout the whole simulation process. However, there is an obvious increasing trend of urban population in China during these years, which will have contributed to urban growth in terms of of both magnitude and spatial location. Therefore, the stable population embedded in the model will result in a conservative prediction of urban growth.

Although AHP was determined by expertise knowledge, it will still bring human subjective influence to the model outputs. In future research, logistic regression could be used to complement AHP weights. Furthermore, the urban growth in this model was endogenously generated by the CA model. Further growth constraints could be added to control the total area of urban growth. Multiple land use types could be introduced into the model.

## 8. Conclusions

This paper focused on the urban development simulation of China until 2030. GIS, multicriteria analysis, and AHP techniques were adopted, and an integrated simulation model was developed in combination with the traditional cellular automata model. The land suitability score was calculated based on seven factors, including distance to city centers, railroads, major roads, river, slope, population density, and GDP. Two parameter tests were carried out for the urban ratio and neighborhood size. As a result, a $5 \times 5$ neighborhood and an urban ratio of 0.33 were selected to generate the results. The spatial pattern of future urban growth in China will not change a lot, but more satellite cities will be built to connect existing urban centers. The total urban development will still increase and is expected to approach 105,000 square kilometers. However, the simulation indicates a decreasing trend in urban growth rate, from 38% in 2005 to 17% in 2030.

**Supplementary Materials:** The following are available online at http://www.mdpi.com/2220-9964/7/10/403/s1.

**Author Contributions:** Y.F. contributed major parts of modeling, writing and structuring of the contents. Y.Q. initiated the concept of the paper, design the model, structuring the content, critically revised the paper.

**Funding:** This research received no external funding

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Habitat, U. World cities report 2016. In *Urbanization and Development: Emerging Futures*; Pub. United Nations: New York, NY, USA, 2016.
2. Alsharif, A.A.; Pradhan, B. Urban sprawl analysis of Tripoli Metropolitan city (Libya) using remote sensing data and multivariate logistic regression model. *J. Indian Soc. Remote Sens.* **2014**, *42*, 149–163. [CrossRef]
3. López, E.; Bocco, G.; Mendoza, M.; Duhau, E. Predicting land-cover and land-use change in the urban fringe: A case in Morelia city, Mexico. *Landsc. Urban Plan.* **2001**, *55*, 271–285. [CrossRef]
4. Tyson, P.; Steffen, W.; Mitra, A.; Fu, C.; Lebel, L. The earth system: Regional–global linkages. *Reg. Environ. Chang.* **2001**, *2*, 128–140.
5. Lambin, E.F.; Turner, B.L.; Geist, H.J.; Agbola, S.B.; Angelsen, A.; Bruce, J.W.; Coomes, O.T.; Dirzo, R.; Fischer, G.; Folke, C.; et al. The causes of land-use and land-cover change: Moving beyond the myths. *Glob. Environ. Chang.* **2001**, *11*, 261–269. [CrossRef]
6. Martellozzo, F.; Amato, F.; Murgante, B.; Clarke, K. Modelling the impact of urban growth on agriculture and natural land in Italy to 2030. *Appl. Geogr.* **2018**, *91*, 156–167. [CrossRef]
7. Chen, J. Rapid urbanization in China: A real challenge to soil protection and food security. *Catena* **2007**, *69*, 1–15. [CrossRef]
8. Santé, I.; García, A.M.; Miranda, D.; Crecente, R. Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landsc. Urban Plan.* **2010**, *96*, 108–122. [CrossRef]
9. Tobler, W.R. Cellular geography. In *Philosophy in Geography*; Springer: Cham, Switzerland, 1979; pp. 379–386.
10. Batty, M.; Xie, Y. Modelling inside GIS: Part 1. Model structures, exploratory spatial data analysis and aggregation. *Int. J. Geogr. Inf. Syst.* **1994**, *8*, 291–307. [CrossRef]
11. White, R.; Engelen, G. Cellular automata and fractal urban form: A cellular modelling approach to the evolution of urban land-use patterns. *Environ. Plan. A* **1993**, *25*, 1175–1199. [CrossRef]
12. Aburas, M.M.; Ho, Y.M.; Ramli, M.F.; Ash'aari, Z.H. The simulation and prediction of spatio-temporal urban growth trends using cellular automata models: A review. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 380–389. [CrossRef]
13. Pijanowski, B.C.; Brown, D.G.; Shellito, B.A.; Manik, G.A. Using neural networks and GIS to forecast land use changes: A land transformation model. *Comput. Environ. Urban Syst.* **2002**, *26*, 553–575. [CrossRef]
14. Amato, F.; Pontrandolfi, P.; Murgante, B. Using spatiotemporal analysis in urban sprawl assessment and prediction. In *International Conference on Computational Science and Its Applications*; Springer: Cham, Switzerland, 2014; pp. 758–773.
15. Park, S.; Jeon, S.; Kim, S.; Choi, C. Prediction and comparison of urban growth by land suitability index mapping using GIS and RS in South Korea. *Landsc. Urban Plan.* **2011**, *99*, 104–114. [CrossRef]
16. Batty, M. *Cities and Complexity: Understanding Cities With Cellular Automata, Agent-Based Models, and Fractals*; The MIT Press: Cambridge, MA, USA, 2007.
17. D'Aquino, P.; August, P.; Balmann, A.; Berger, T.; Bousquet, F.; Brondízio, E.; Brown, D.G.; Couclelis, H.; Deadman, P.; Goodchild, M.F.; et al. Agent-Based Models of Land-Use and Land-Cover Change. Available online: https://www.zef.de/fileadmin/template/Glowa/Downloads/Berger_et_al_2002.pdf (accessed on 16 September 2018)
18. Haase, D.; Schwarz, N. Simulation models on human-nature interactions in urban landscapes: A review including spatial economics, system dynamics, cellular automata and agent-based approaches. *Living Rev. Landsc. Res.* **2009**, *3*, 1–45. [CrossRef]
19. Batty, M. Agents, cells, and cities: new representational models for simulating multiscale urban dynamics. *Environ. Plan. A* **2005**, *37*, 1373–1394. [CrossRef]
20. Clarke, K.C.; Hoppen, S.; Gaydos, L. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environ. Plan. B Plan. Des.* **1997**, *24*, 247–261. [CrossRef]

21. White, R.; Engelen, G. High-resolution integrated modelling of the spatial dynamics of urban and regional systems. *Comput. Environ. Urban Syst.* **2000**, *24*, 383–400. [CrossRef]
22. Sun, H.; Forsythe, W.; Waters, N. Modeling urban land use change and urban sprawl: Calgary, Alberta, Canada. *Netw. Spat. Econ.* **2007**, *7*, 353–376. [CrossRef]
23. Mondal, B.; Das, D.N.; Bhatta, B. Integrating cellular automata and Markov techniques to generate urban development potential surface: A study on Kolkata agglomeration. *Geocarto Int.* **2017**, *32*, 401–419. [CrossRef]
24. Mohammad, M.; Sahebgharani, A.; Malekipour, E. Urban growth simulation through cellular automata (CA), analytic hierarchy process (AHP) and GIS; case study of 8th and 12th municipal districts of Isfahan. *Geogr. Tech.* **2013**, *8*, 57–70.
25. Liu, J.; Zhan, J.; Deng, X. Spatio-temporal patterns and driving forces of urban land expansion in China during the economic reform era. *AMBIO J. Hum. Environ.* **2005**, *34*, 450–455. [CrossRef]
26. Liao, J.; Tang, L.; Shao, G.; Su, X.; Chen, D.; Xu, T. Incorporation of extended neighborhood mechanisms and its impact on urban land-use cellular automata simulations. *Environ. Model. Softw.* **2016**, *75*, 163–175. [CrossRef]
27. Wu, F.; Webster, C.J. Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environ. Plan. B Plan. Des.* **1998**, *25*, 103–126. [CrossRef]
28. Wu, F. SimLand: A prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. *Int. J. Geogr. Inf. Sci.* **1998**, *12*, 63–82. [CrossRef]
29. Yang, Q.; Li, X.; Shi, X. Cellular automata for simulating land use changes based on support vector machines. *Comput. Geosci.* **2008**, *34*, 592–602. [CrossRef]
30. Li, X.; Yeh, A.G.O. Neural-network-based cellular automata for simulating multiple land use changes using GIS. *Int. J. Geogr. Inf. Sci.* **2002**, *16*, 323–343. [CrossRef]
31. Couclelis, H. From cellular automata to urban models: New principles for model development and implementation. *Environ. Plan. B Plan. Des.* **1997**, *24*, 165–174. [CrossRef]
32. Clarke, K.C.; Gaydos, L.J. Loose-coupling a cellular automaton model and GIS: Long-term urban growth prediction for San Francisco and Washington/Baltimore. *Int. J. Geogr. Inf. Sci.* **1998**, *12*, 699–714. [CrossRef] [PubMed]
33. Deng, X.; Huang, J.; Rozelle, S.; Uchida, E. Growth, population and industrialization, and urban land expansion of China. *J. Urban Econ.* **2008**, *63*, 96–115. [CrossRef]
34. O'Sullivan, D. Exploring spatial process dynamics using irregular cellular automaton models. *Geogr. Anal.* **2001**, *33*, 1–18. [CrossRef]
35. O'Sullivan, D. Graph-cellular automata: a generalised discrete urban and regional model. *Environ. Plan. B Plan. Des.* **2001**, *28*, 687–705. [CrossRef]
36. White, R.; Engelen, G. Cellular Dynamics and GIS: Modelling Spatial Complexity. *Geogr. Syst.* **1993**, *1*, 237–153.
37. Arsanjani, J.J.; Helbich, M.; Kainz, W.; Boloorani, A.D. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 265–275. [CrossRef]
38. Gu, C.; Guan, W.; Liu, H. Chinese urbanization 2050: SD modeling and process simulation. *Sci. China Earth Sci.* **2017**, *60*, 1067–1082. [CrossRef]
39. John, W.; Fook, L.L. *The Challenge of Making Cities Liveable in East Asia*; World Scientific: Singapore, 2016.
40. Saaty, R.W. The analytic hierarchy process—What it is and how it is used. *Math. Model.* **1987**, *9*, 161–176. [CrossRef]
41. Li, X.; Zhang, L.; Liang, C. A GIS-based buffer gradient analysis on spatiotemporal dynamics of urban expansion in Shanghai and its major satellite cities. *Procedia Environ. Sci.* **2010**, *2*, 1139–1156. [CrossRef]