



OPEN

Detecting the interaction between urban elements evolution with population dynamics model

Min Jin^{1,2}, Lizhe Wang^{1,2}✉, Fudong Ge¹ & Jining Yan^{1,2}

Exploring the evolution of urban elements can improve understanding of the developmental process of city and drive such development into a better direction. However, the non-linearity and complexity of changes in urban elements have brought great challenges to understanding this process. In this paper, we propose a cross-diffusion partial differential equation based on ecological dynamics to simulate the evolutionary process of urban elements from the microscopic viewpoint. The interaction between urban elements is simulated by constructing a non-linear and spatiotemporal change equation, and the main influence between elements is evaluated by the key parameters in the discussed equation. Our model is first experimented to time-series data on population density and housing prices to analyze the interaction of these two elements in the evolution process. We then extend the model to label data, land cover data, to obtain a quantitative expression of the interaction between different land types in the process of urban land cover change.

Researches on the development of urban elements have been a fascinating field in recent decades^{1,2}. Cities play a crucial role in accommodating a significant proportion of the global populace, and this trend is expected to persist and intensify in the foreseeable future^{3,4}. The burgeoning interest in this domain can be attributed, in part, to the intricate and multifaceted nature of urban phenomena, which continues to pose a myriad of unresolved questions and knowledge gaps that warrant further investigation and inquiry⁵. An overarching objective in the study of urbanization is to enhance our understanding of city development processes and utilize this knowledge to facilitate a shift towards a more sustainable path of urban development^{6,7}.

With the efforts of generations of outstanding scientific researchers, a number of impactful works have been witnessed on dealing with the development of urban elements. Research on the relationship between urban elements can be divided into two categories. One is concerned with filtering out those factors that play a leading role in changing a specific event when there are many factors at play, analyzing the driving force, and simulating the driving process based on the leading factors^{8,9}. A typical representative of this type of research is the analysis of the drivers of the land-use change process^{10,11}. Another category seeks to model the relationship between factors directly, without weeding out the dominant factors, and to analyze the evolution of elements based on the results obtained. A typical representative of this type of research is the analysis of the interaction between elements of panel data¹².

The main process of interest for the research method that focuses on driving forces is land-use change, as the core objective is to first identify the driving factors that play a leading role in land-use change. Typical used methods include the logistic regression¹³, the analytic hierarchy process (AHP)¹⁴, and the principal components analysis (PCA)¹⁵, etc. Based on the identification of the dominant driving factors, the driving force is further investigated, and the driving process is simulated. In this stage, commonly used models include cellular automata (CA)¹⁶ and a number of improved algorithms that have been extended to CA^{17,18}. For example, using geographically weighted regression (GWR) causes CA's transformation rules to be spatially non-stationary¹⁹. The combination of the agent-based modeling (ABM) model and the CA model reflects the multi-agent impact process of human activities on land-use change²⁰. In addition, machine learning methods have been used to improve the uncertainty generated by artificially setting rules in the CA model²¹ and the improved FLUS model, which has been combined with artificial neural network (ANN), system dynamic (SD), and CA models²². Another type of algorithm is the CLUE-S model²³, which includes the non-required space module, the space allocation module, and the extension of this model. The Markov, the SD, and the random forest algorithm are used to calculate the total land requirements of the non-spatial modules in the CLUE-S model.

¹School of Computer Science, China University of Geosciences, Wuhan 430074, China. ²Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan 430074, China. ✉email: lizhe.wang@gmail.com

When examining the element relationship with panel data, surveys can be classified according to whether a statistical method is used. In the statistical relationship-based model, scientists usually pre-examine some factors and use statistical analysis methods, such as correlation analysis or regression analysis, to examine the relationship between these factors. Ultimately, a statistical model was used to express the relationship between the factors. Models that do not involve spatial heterogeneity typically use linear regression models, such as piecewise linear regression²⁴, among other options. With an emphasis on the spatial relationship between elements, a number of spatial econometric models have emerged. The first to be proposed was the spatial autoregressive (SAR) model²⁵. The SAR is the modeling of Tobler's first geographic law, which describes spatial autocorrelation. Some of the more commonly used spatial econometric models include the spatial lag model (SLM, also known as the SAR), the spatial error model (SEM)²⁶, and the spatial Durbin model (SDM)²⁷. The SDM is a combined and augmented form of the SLM and the SEM that is constructed by adding appropriate constraints.

Besides being involved with the land-use drivers, the GWR model also can be used to study factor relationships based on panel data²⁸. GWR illuminates the spatial variation and associated driving factors of the research object at a given scale by establishing a local regression equation at each point in the spatial domain that can be used to predict future outcomes. It has the advantage of better performance, since it takes into account the local effects of spatial objects. And thus derived semiparametric geographically weighted regression (SGWR)²⁹, Fast geographically weighted regression (Fast GWR)³⁰. Wang et al. also proposed a statistical method called GeoDetector³¹. The core idea of GeoDetector is based on the assumption that if an independent variable has a significant impact on a dependent variable, then the spatial distributions of the independent and dependent variables should be similar.

Among these methodological models, scholars have invariably focused on the spatial heterogeneity between time-series data items. Describing the spatial heterogeneity of elements in the process of time change has become of special interest to scientists. In addition to these models, one particular group of scientists is trying to understand the urban development process from a different angle. For example, researchers have carefully explored the laws and models of urban development, including some physicists who have made brilliant contributions^{32,33}. Specifically, these researchers model city growth from a physical perspective by applying models from typical research areas of physics to questions of city growth and urban development^{34,35}. Among them, Juste Raimbault et al. used the reaction–diffusion equation to study the growth of urban population. Three parameters, namely aggregation, diffusion and growth speed parameters, were used to reflect the change rule of urban population density, which reflected the effectiveness of the simple model^{36,37}. In addition to the reaction–diffusion model, a gravity-based model and correlated percolation have also been considered in the exploration of the law of urban development^{38,39}. These models can provide valuable perspectives on urban development theories and generate models that can be used to predict urban growth.

Combining the various models of the development process of urban elements that have been advanced by researchers⁴⁰, in this article, we adopt the cross-diffusion model in pattern dynamics to simulate the mutual evolution process between different elements. The self-diffusion element indicates the spatial overflow or propagation process of the elements, and the cross-diffusion element characterises the interaction between the elements. At the same time, since the model was created for each spatial grid, it exhibits spatial heterogeneity.

In this paper, we first explored the development process of urban elements from the perspective of reaction–diffusion equation. We extended a well-known model from the field of chemistry and biology to the development of urban systems. Secondly, since the parameters of our model could be identified from the actual data, the model provided an explicit expression of the interaction between urban elements. Finally, considering different data types, we applied the model to three kinds of data: population density, housing price and land cover, and studied the correlation between these three kinds of data.

The remain of this paper is divided into three sections. We first present our experimental results in “**Results**” section, and then comes our “**Discussion**” section. Next, we illustrate the setup and composition of our algorithmic model in “**Methods**” section. Finally, we introduce our research area and data processing in “**Materials**” section.

Results

In the experimental part, we first feed population density data and house price data into our model to obtain the evolutionary relationship between population density and house prices, and then analyze the results to complete the simulation of unlabeled data. In the next step, we perform experiments for different types of land cover data. The process of the experiment consist of treating each land cover type as an element, performing a mutual evolution simulation, and completing the simulation of the label data.

Population density data and house price data. Based on house price data, and considering the impact of population density changes on house price growth, we obtain the following data results. Figure 1a parades the spatial distribution of the diffusion coefficient. It can be seen that the diffusion coefficient of housing prices in Shenzhen exhibits low distribution in the south, high distribution in the north, high in the east, and low in the west, which means that in the north and east of Shenzhen, the higher house price range radiating outward is wider. We compare these findings with the subway lines in Shenzhen, as demonstrated in Fig. 1b. It is found that the diffusion coefficient of housing prices in Shenzhen basically increased outward along subway lines. At the same time, the diffusion coefficient of housing prices in areas close to subway lines is low, which also means that housing prices in Shenzhen radiated outward along subway lines. This also can be seen in Fig. 1c, the proportion of housing price growth ratio in Shenzhen from 2011 to 2019. At the same time, based on the housing price in 2011, along the subway lines that have been built in Shenzhen in 2019, it can be seen that the growth rate of house prices in Shenzhen is generally low. As can be seen in Fig. 1d, along the subway lines planned to be

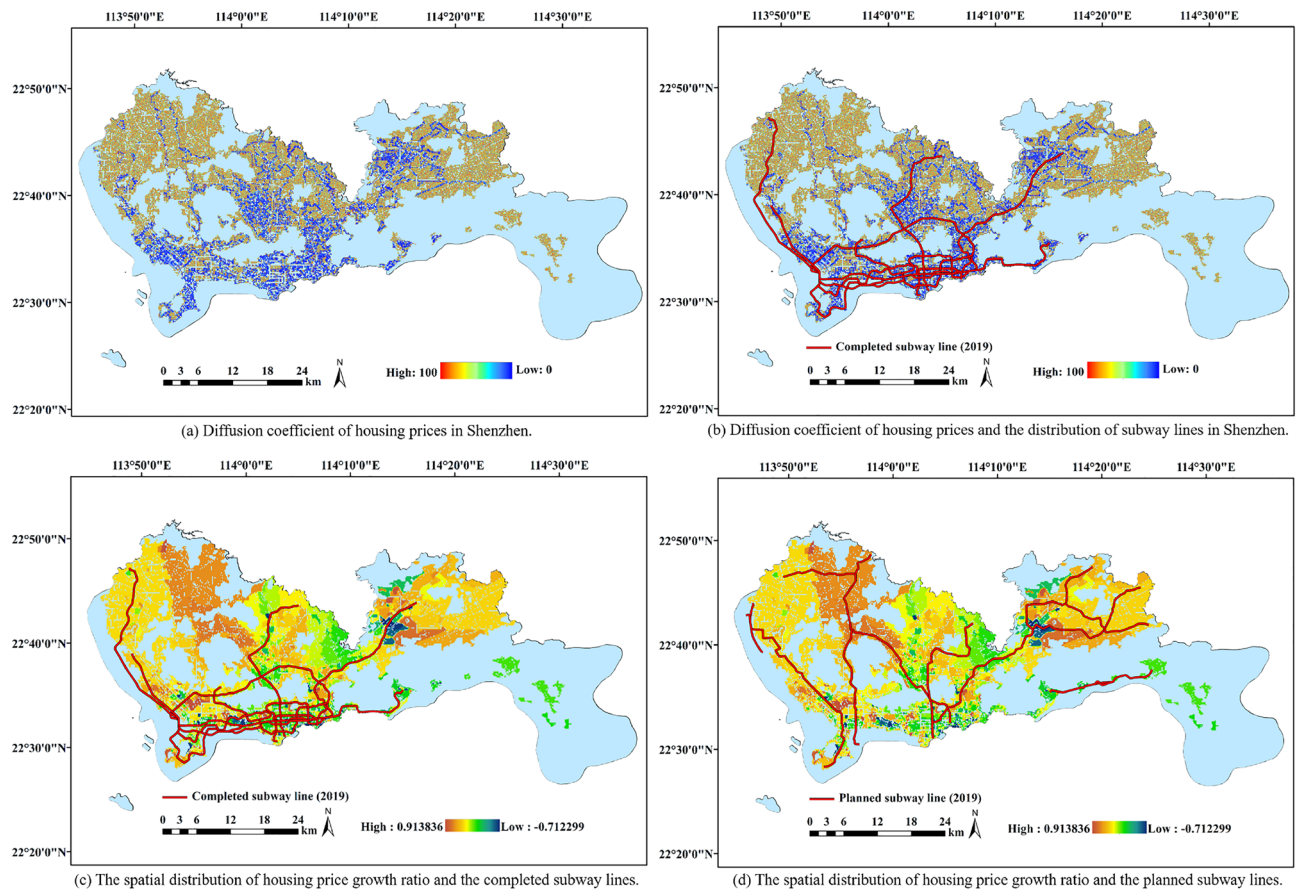


Figure 1. The correlation between the diffusion coefficient of house price and subway lines. The figure was created by QGIS (QGIS Desktop 3.22.4, <https://www.qgis.org>).

built in Shenzhen after 2019, the increase in housing prices is even greater. With the completion of the subway, the housing prices in this area gradually spread outward and reached a higher price, so the diffusion coefficient also decrease.

From the result of Fig. 2a, the cross-diffusion coefficient of the corresponding population density is distributed. The increase in population density is exposed in Fig. 2b. Overall, the distribution of the cross-diffusion coefficients is consistent with the increase in population density. In regions where the cross-diffusion coefficient is found to be unstable, negative population growth occurred due to the properties of the diffusion equation. At the same time, a positive cross-diffusion coefficient also means that the housing prices in this area are increasing toward the area with population growth. A negative cross-diffusion coefficient indicated that housing prices are increasing from areas with lower populations to areas with higher populations. This characteristic, reflects by the cross-diffusion coefficient in the diffusion equation, is consistent with the actual situation.

From the spatial distribution of the competition coefficient, it can be found that, as displayed in Fig. 3a, the internal competition of housing prices is also more intense in areas with larger growth coefficients. At the same time, compared with the overall growth rate of house prices, on view of Fig. 3b, the competition within the region with more house price appreciation is more intense. This means that in these areas, the self-growth potential of housing prices is still relatively large. At the same time, in these areas, the phenomenon of housing price competition between sub-regions, that is, real estate speculation, may be more serious.

Land cover data. In the simulation experiment for land cover data, we first conduct a preliminary analysis of the land cover change in Shenzhen, as shown in Table 1 below. From the figure, we can see that Baoan District had the most newly-added impervious areas, and Yantian District had the least. From the perspective of the proportion of new impervious areas, Guangming District and Dapeng District have the highest and lowest proportions of new areas, respectively. Therefore, we select those four districts for the simulation experiments.

The experiment in Baoan District. The first sub-region of our experiment is the Baoan District. From the distribution of the diffusion coefficient in Fig. 4, within the time frame of the study, the impact of the spatial neighborhood on the diffusion process of cropland, water, and impervious in Baoan District is still relatively large. During the diffusion process of forest and barren, the influence of neighborhoods on this process is relatively small. Since there are no shrubs or snow in the Baoan District, their diffusion coefficients are 0. At the same time, the areas with large fluctuations in diffusion coefficient are concentrated in the western coastal area of Baoan,

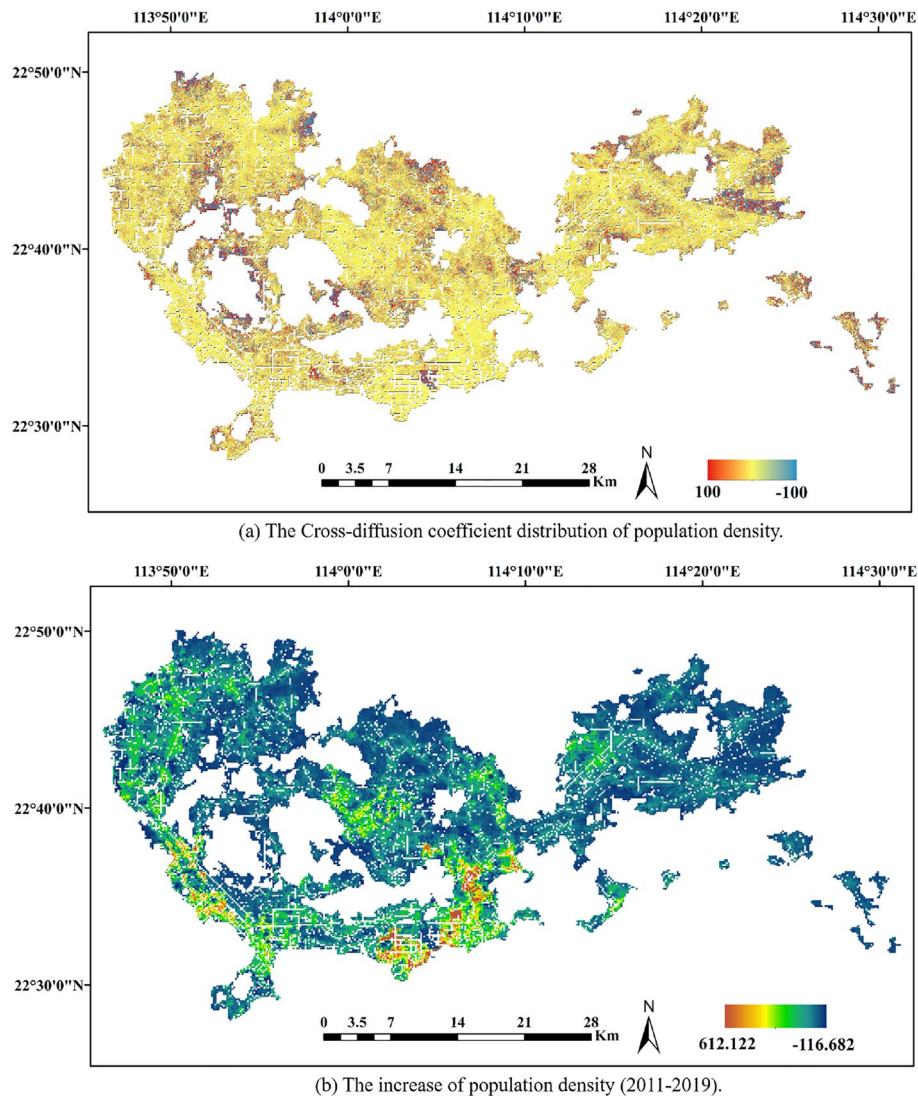


Figure 2. The correlation between diffusion coefficient and subway lines. The figure was created by QGIS (QGIS Desktop 3.22.4, <https://www.qgis.org>).

which means that the diffusion trend of impervious surfaces in the western coastal area is more obvious during the study period.

From the distribution of the growth coefficients, it is found that the impervious in the Baoan District had obvious self-growth trends while being affected by the neighborhood. Which means that during impervious expansion, the larger the impervious in the spatial neighborhood, the easier it is for the central grid to become new impervious. Similarly, the aforementioned areas of expansion are predominantly clustered within the western coastal region of Baoan.

For the competition coefficient, similarly, cropland and impervious are accompanied by a strong competition coefficient. In addition, the water in Baoan District show a strong competition coefficient, which indicates that, even though the water in Baoan District has an increasing trend from the historical time series, the development of other land cover types often limits the spread of water area.

The experiment in Dapeng District. As shown in Fig. 5, Dapeng District is similar to the Baoan District, in that cropland, forest, and impervious are shown to have a strong tendency to be influenced by spatial neighborhood, but the degree of influence is relatively small on the whole. It is reflected in the spatial distribution that only the local values change drastically.

The results of growth coefficient show that in addition to cropland and impervious, forest and water in Dapeng also exhibit large growth coefficients. This indicates that the forest and grassland in the Dapeng District have increasingly more possessions during the study period. Correspondingly, in the spatial space, the growth coefficient of the region where the above-mentioned types of changes have occurred is more drastic.

Similar to the Baoan District in the coastal area, apart from farmland and forests, the Dapeng District also have a strong competition coefficient in the category of water, indicating that water is more occupied in the

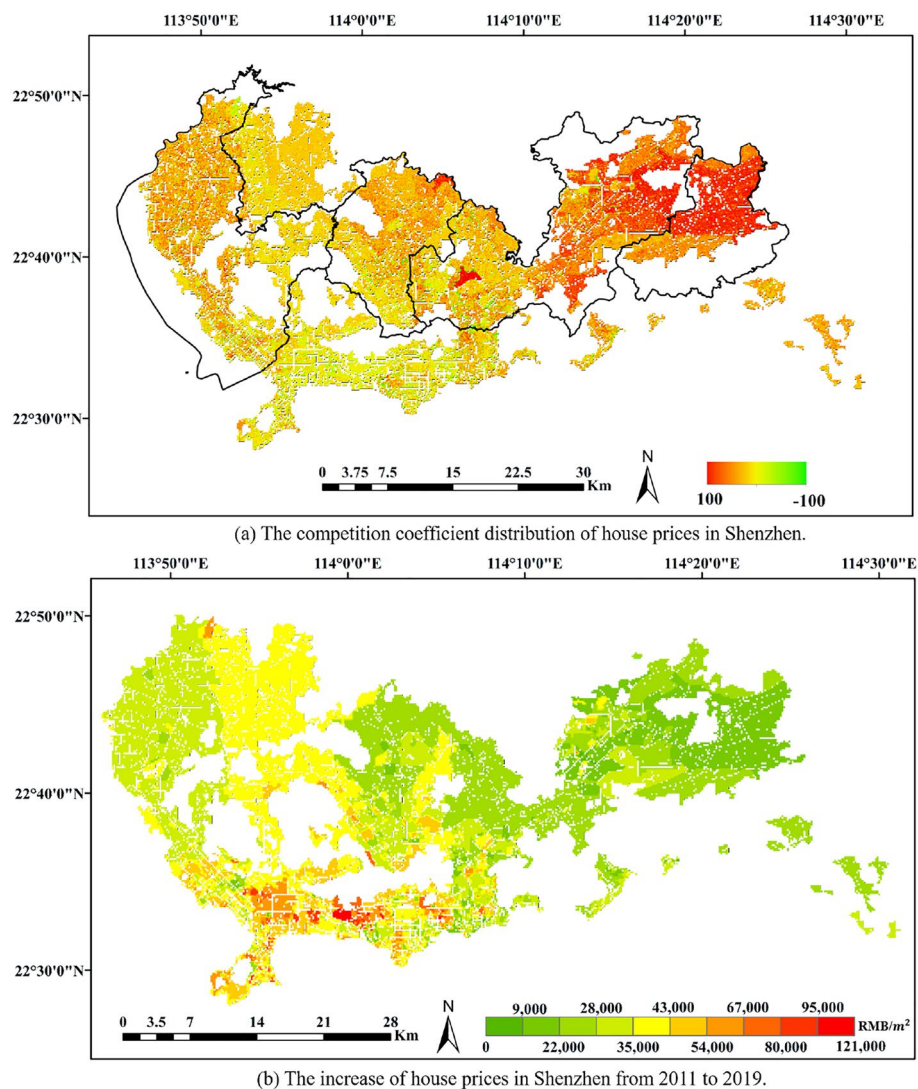


Figure 3. The correlation between competition coefficient and increase in house prices (RMB/m²). The figure was created by QGIS (QGIS Desktop 3.22.4, <https://www.qgis.org>).

District	Baoan	Dapeng	Futian	Guangming	Longgang	Longhua	Luohu	Nanshan	Pingshan	Yantian
New area (km ²)	16.18	1.65	1.09	8.69	11.65	8.42	0.76	4.13	6.97	0.54
New area proportion (%)	4.17	0.53	1.36	5.12	2.76	4.38	0.89	2.57	3.86	0.74

Table 1. The changes of impervious areas in various districts of Shenzhen.

process of urban expansion. This also shows that in the process of land cover change in Dapeng District, water is one of the most impermeable categories in transformation.

The experiment in Guangming District. From Fig. 6 of the diffusion coefficient, the land cover change in Guangming District is found to be mainly concentrated in the following four types: cropland, forest, water, and impervious. Judging from the distribution of the diffusion coefficients, cropland is more negatively affected by spatial neighbors; that is, the denser the cropland in the surrounding neighborhood, the less favorable it is for outward spread. This may be due to the fact that densely distributed areas of cropland are more likely to be intensively developed as impervious.

Figure 6 exhibits the result of the growth coefficient. From the perspective of growth coefficients, the distribution of the above four types cropland, forest, water and impervious is relatively average, indicating that the impact of historical trends is not easy to evaluate.

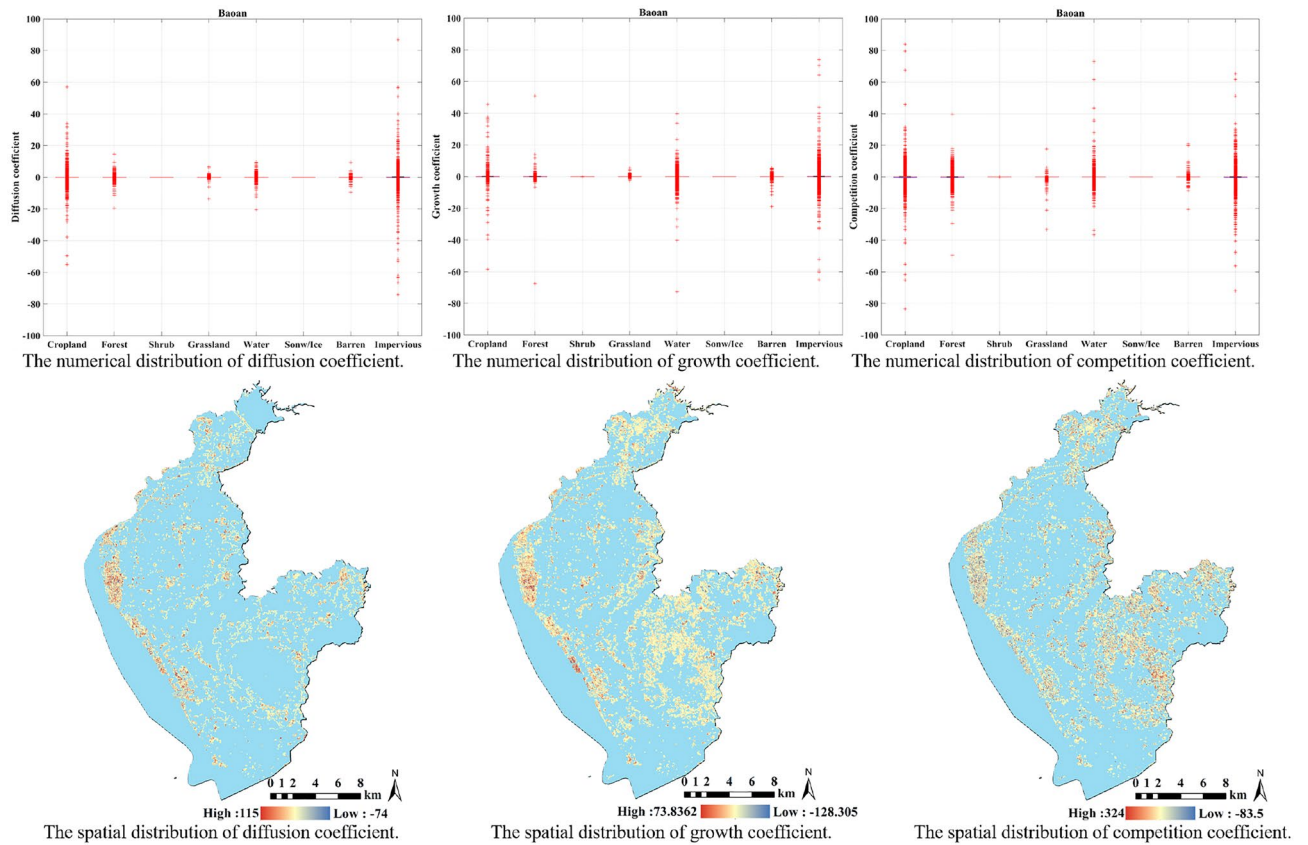


Figure 4. The distribution of different coefficients in Baoan. The figure was created by QGIS (QGIS Desktop 3.22.4, <https://www.qgis.org>).

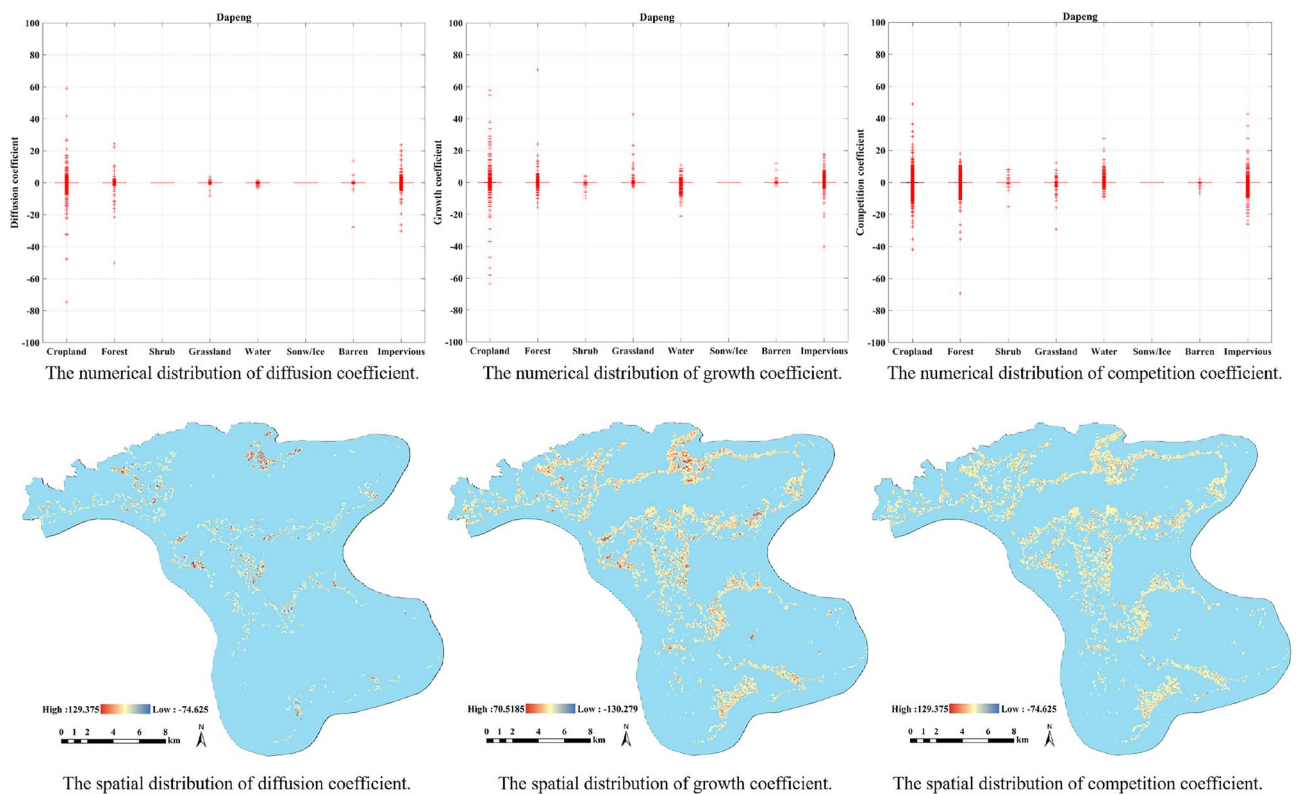


Figure 5. The distribution of different coefficients in Dapeng. The figure was created by QGIS (QGIS Desktop 3.22.4, <https://www.qgis.org>).

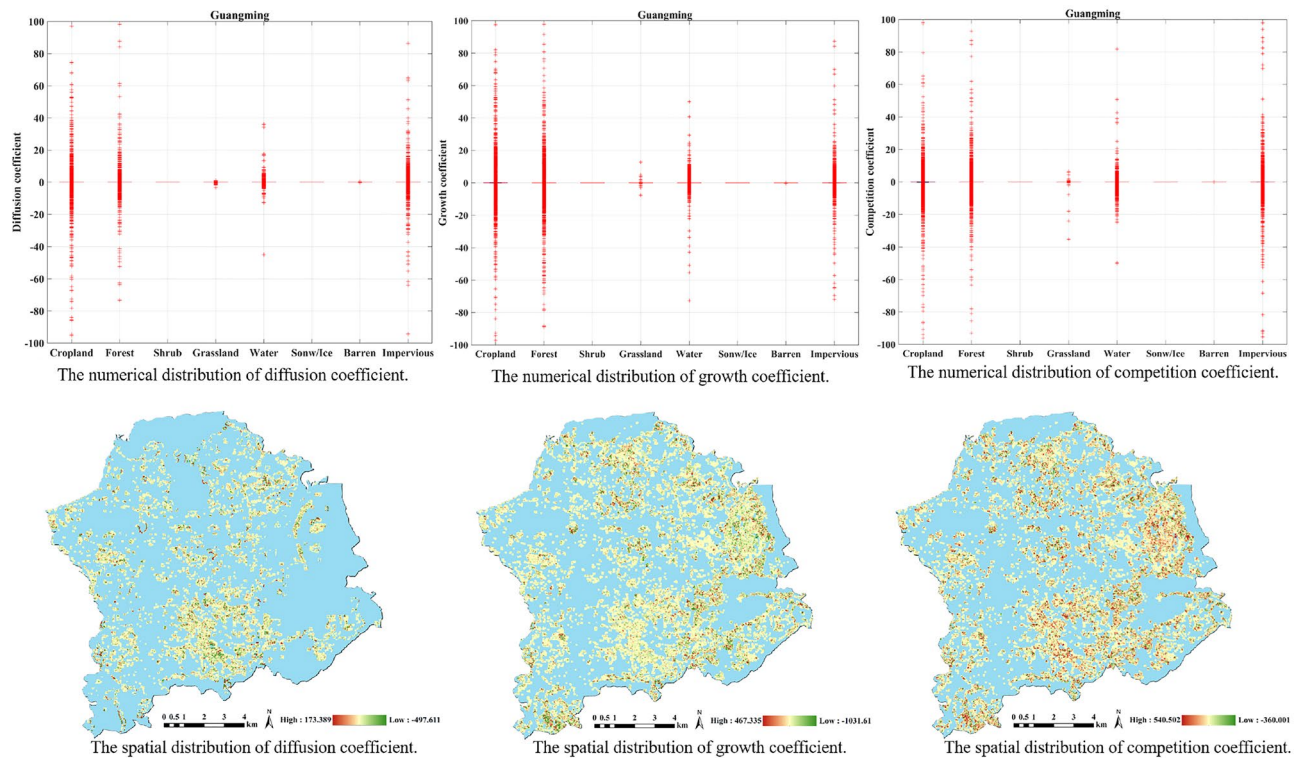


Figure 6. The distribution of different coefficients in Guangming. The figure was created by QGIS (QGIS Desktop 3.22.4, <https://www.qgis.org>).

The result of the competition coefficient is delineated in Fig. 6. From the perspective of numerical distribution, the distribution of the competition coefficient is similar to that of the growth coefficient, and it is mainly concentrated in the four types of cropland, forest, water and impervious. Similarly, from the perspective of spatial distribution, in areas where the growth coefficients of land cover types change drastically, the corresponding competition coefficients also change sharply.

The experiment in Yantian District. As an old city in Shenzhen, Yantian District is not as intense as the other three districts in terms of the diffusion coefficient, growth coefficient, or competition coefficient values. The results of the diffusion coefficients in Fig. 7 presents that the impervious is the most affected by neighbors, followed by cropland and water. Grassland and barren are less susceptible to neighborhood changes.

Similarly, in the results for the growth coefficients, see Fig. 7, the distribution of growth coefficients of various types in the Yantian is on the low side. However, there are occasional outliers in the growth coefficients of cropland, forest and impervious, indicating that these three types are other main factors in the Yantian change. Figure 7 displays the regions where these major changes occurred.

In Fig. 7, the results for the competition coefficient indicates that the three types of cropland, forest and impervious still fluctuated greatly. While the competition coefficient is more spread out in value distribution than the growth coefficient, it is relatively uniform in spatial distribution.

Discussion

In this paper, we proposed a partial differential equation with cross-diffusion to model the evolutionary process of urban elements. The model explores the development and change process of urban elements and uses the diffusion coefficient to represent the impact on the surrounding neighborhood, and the cross-diffusion coefficient to represent the effects of other elements. The action process of the elements themselves is also reflected by the growth coefficient and the competition coefficient. Through the spatial distribution of these key coefficients, we observed the spatial heterogeneity of the development process of elements at the micro level. Combining the ecological dynamic meaning of these key elements, we were able to obtain a glimpse of the key driving elements or driving forces in the process of element evolution and gain a further understanding of urban elements. At the same time, our model still needs to be improved. The force that urban elements receive during the evolution process is far greater than that of the few items included in the model. Thus, this is still a field worth exploring to examine more action items according to the characteristics of element development.

In the model proposed in this paper, simulating changes in urban elements requires that the impact of the surrounding neighborhood be considered. Therefore, we first needed to define the range of the spatial neighborhood. In this paper, we used the classic Moore neighborhood format, which is also used in the CA model as a spatial neighborhood. In the simulation experiment for population and housing price elements, we regarded

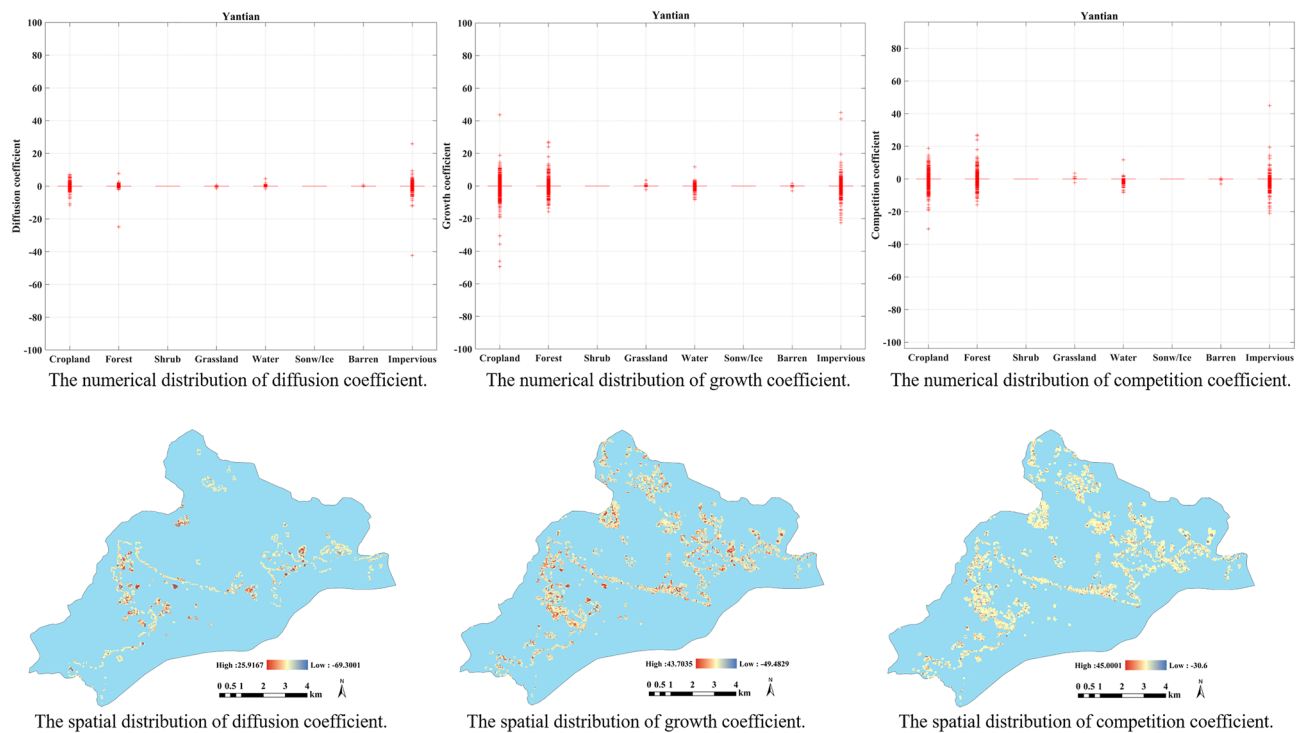


Figure 7. The distribution of different coefficients in Guangming. The figure was created by QGIS (QGIS Desktop 3.22.4, <https://www.qgis.org>).

housing price as the main factor and population density as the secondary factor, and we evaluated the impact of population density on the development process of housing prices in Shenzhen while considering the developmental characteristics of housing prices themselves. In the experiment with labeled data, specifically land-use data, we regarded the built-up area with the largest newly-added area as the main element, according to the characteristics of the data, and evaluated the impact of other categories on the expansion process of the built-up area while considering the developmental inertia of the built-up area itself.

Methods

In the process of urban evolution, different elements form specific spatial structures. Not only are elements constantly changing, but they are also constantly affecting other elements neighbored to them. Taking built-up as an example, the marginal diffusion and peripheral diffusion of urban built-up area are dominant in China⁴¹. In other words, the scope of urban built-up areas has been continuously developed and further expanded in the process of urban improvement^{42,43}. Second, the state of the neighborhood around the built-up area will also limit its outward diffusion process. For example, if there are many bodies of water around the built-up area, this will directly affect the direction and speed of the outward spread of the built-up areas. At the same time, the process of out-diffusion of settlement areas will also influence other factors in the city, such as the distribution of population density. After the construction of the built-up area is completed, the area will inevitably become susceptible to settlement, which will lead to changes in the distribution of urban population density.

Taking into account the features of urban development, we made the following assumptions about the process through:

- Constrained by resources and other environmental factors, it is difficult for urban elements to grow indefinitely. The closer a factor reaches the upper limit of capacity, the greater the resistance to sustaining growth. Referring to the competition dynamics model in ecology, we use the competition term to simulate this growth-limited phenomenon of urban elements.
- Under resource scarcity, the diffusion speed of different elements is affected by the density of elements in the surrounding area. When the spatial fraction or density of a given element is greater in a given spatial neighborhood, the subsequent continuous diffusion space becomes smaller, and its own evolution is affected, resulting in a slowdown in the rate of diffusion of elements.
- The growth of one factor is influenced by other factors, and this effect is complex. There may be some mutual attraction, such as in densely built-up areas, where the population density is often high, as are housing prices. Some factors are mutually exclusive; for example, in areas with dense water bodies, the population density will be lower.
- Finally, in a certain spatial neighborhood, this paper refers to the molar neighborhood, the urban elements move along the gradient direction of the concentration difference. Therefore, we assume that the ability of

the same element to diffuse along the concentration difference in the spatial neighborhood is the same, and that the diffusion ability of different elements is different.

Combined with the previous work⁴⁰, in this paper, the calculation method for the concentration of urban elements is as follows. For population density data and house price data, we take the value of population density and house price as the corresponding concentration. For land cover data, since each value only represents a category, we take the proportion of land use types within a certain neighborhood as the corresponding concentration value. With this concentration defined, we model the following reaction–diffusion model.

Construction of a reaction–diffusion model. It is assumed that the continuous change of urban elements can be divided into multiple microstages σt ($\sigma t \rightarrow 0$) within a certain time interval. Theoretically, the change of urban elements is accumulated by n microscopic processes. At the initial time point $t = 0$, the city element does not exist. For the time point $t = \sigma t, 2\sigma t, 3\sigma t, \dots$, the addition of urban elements in each micro-process occurs in the vacant growth area, which increases the space for the existing distribution of urban elements. If the urban area is regarded as a grid area, then in the expansion process of urban elements, the growth of urban elements in each grid of the urban area obeys the diffusion equation in every microscopic process. That is to say, during the microcosmic change process of urban elements in each urban grid, the change in urban elements obeys a diffusion equation involving time and space. Each item in the equation is constructed on the basis that the diffusion of elements is a spatiotemporal process, so the time rate of change in elements is related to spatial location. At the same time, the elements in each grid not only have an influence on the surrounding neighborhood, but also have a self-growth process and are also related to the distribution of other related elements. Therefore, the cross-diffusion system equation was used for fitting.

Diffusion is a material change process driven by a difference in the concentration of the material. Therefore, when diffusion theory is used to study the diffusion of urban elements, the concentration value of each element must first be determined for the diffusion of each urban element. For label data and non-label data in each spatial grid, the collection methods for concentration values were slightly different. In this study, we referred to a prior work⁴⁰ to obtain values for the solution concentrations for the different types of urban elements.

Next, we divided the spatial region of the entire city into a regular grid. In this case, we let $x \in \Omega$ represent the spatial grid, and Ω was the study area. $t > 0$ represents the moment when urban element data were sampled, and $\{i = 1, 2, \dots, n\}$ represents different urban element types. Then, $U_i(x, t)$ represents the concentration value of element i located on spatial grid x at time t .

Finally, we considered the influence of urban elements on the evolutionary process. The diffusion process of urban elements is first influenced by the change of the same element in the spatial neighborhood and also by other related elements. At the same time, the growth of urban elements is always constrained by various factors, making unlimited growth difficult. In summary, our diffusion model needed to reflect the impact of the above three functional relationships on the extent of urban element development.

In accordance with existing work^{44,45}, urban elements were divided into grids with equal spatial size according to the scope of time and space. The following reaction–diffusion equations were constructed to determine the evolution process of the elements in each of the grids. We used reaction–diffusion equations with cross-diffusion terms to describe the spatiotemporal evolution process of elements in the city. The equations had the following form:

$$\frac{\partial U_i(x, t)}{\partial t} = a_i(x)\Delta U_i + U_i(x, t)(r_i(x) - K_i(x)U_i(x, t)) + \sum_{j \neq i}^n b_j(x)\Delta U_j(x, t). \quad (1)$$

Among these variables, $U_i(x, t)$ and $U_j(x, t)$ show the state of the element in the grid of x at time t for the element $\{i, j = 1, 2, 3, \dots, n\}$. The $a_i(x)$ represents the influence of elements on the surrounding neighborhood during the evolution process and the promotion or restriction of the difference from the surrounding neighborhood to its own development. For example, when the built-up area is surrounded by bare land, this means that the area may be under development, and the scope of the built-up area will continue to expand.

$$\Delta = \frac{\partial^2 U_i(x, t)}{\partial x^2} \quad (2)$$

is Laplace operator, which represents the diffusion of urban elements in space.

The $r_i(x)$ refers to the growth process of the element itself, and the driving force of this process may come from the historical development trend or its own development (such as population growth). $K_i(x)$ represents the strength of the internal competition of elements. This means that when the concentration of elements is too high, the intensity of competition between neighbors will become larger.

The cross-diffusion term, $b_j(x)$, describes the interaction between different elements $U_j(x, t)$. When $b_j(x)$ was positive, it indicated that the current element was flowing in the direction of the low density of the influencing element. A negative cross-diffusion term coefficient indicates that the current element flowed from the low-density direction of the associated element to the high-density direction.

Since our model (1) aim to form the evolution of urban elements, the model has a non-negative initial value condition:

$$U_i(x, 0) \geq 0. \quad (3)$$

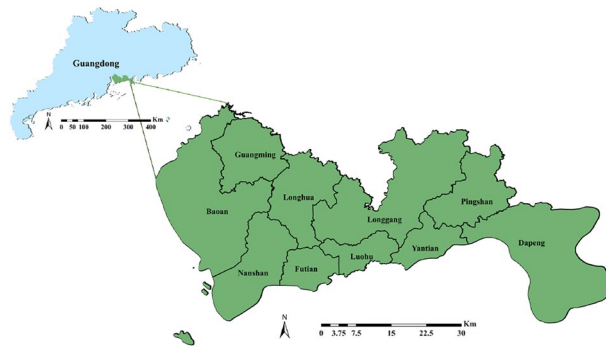


Figure 8. The geographical location and administrative division of Shenzhen. The figure was created by QGIS (QGIS Desktop 3.22.4, <https://www.qgis.org>).

Define the boundary of the study area is $L \in \Omega$. We also assume that no external input is imposed from outside. Hence zero flux boundary conditions

$$\frac{\partial U}{\partial x} \Big|_L = 0 \quad (4)$$

is assumed.

The solution of the model. Since the evolution data of the actual urban elements are all discrete, in order to solve the problem, we need to discretize the above diffusion equation (1). Then we get:

$$\begin{aligned} \frac{U_i(x, t + T) - U_i(x, t)}{T} = & a_i(x) \left[\frac{U_i(x + X, t) - 2U_i(x, t) + U_i(x - X, t)}{X^2} \right] \\ & + r_i(x) U_i(x, t) \left[1 - \frac{U_i(x, t)}{K_i(x)} \right] \\ & + \sum_{j \neq i}^n b_j(x) \frac{U_j(x + X, t) - 2U_j(x, t) + U_j(x - X, t)}{X^2}, \end{aligned} \quad (5)$$

where T is the time interval, X is the space interval.

When X and T are given, after rearrangement, the following iterative equation is obtained:

$$\begin{aligned} U_i(x, t + T) = & \frac{a_i(x)T}{X^2} [U_i(x + X, t) - 2U_i(x, t) + U_i(x - X, t)] + U_i(x, t) \\ & + \frac{T}{X^2} \sum_{j \neq i}^n b_j(x) [U_j(x + X, t) - 2U_j(x, t) + U_j(x - X, t)] \\ & + Tr_i(x) U_i(x, t) \left[1 - \frac{U_i(x, t)}{K_i(x)} \right]. \end{aligned} \quad (6)$$

According to the iterative formula and the evolutionary data of the urban elements, we use a series of optimization algorithms to solve the parameters in the model. Due to the fast convergence speed of the particle swarm optimization (PSO) algorithm, the algorithm did not have many parameters that needed to be adjusted. For this paper, we used the PSO to identify the parameters in our model (6).

Materials

In this section, we describe the basic settings of the research area, and then we provide a detailed account of the source of data acquisition and the attributes of the data. In this article, we use the population density, housing price and land cover data of Shenzhen, China. We further demonstrate the preprocessing of the data for our experimental procedure. Finally, we illustrate some basic settings of the relevant parameters during our experiments.

Study area. Shenzhen is a sub-provincial city in Guangdong Province and a special economic zone in China. There are nine administrative districts and one new district under the jurisdiction of the city. As shown in Fig. 8, Shenzhen is located in the south of Guangdong Province, on the east bank of the Pearl River Estuary, between $113^\circ 43' E$ to $114^\circ 38' E$ and $22^\circ 24' N$ to $22^\circ 52' N$. Shenzhen is one of China's central economic cities. Its total economic output has long been ranked fourth among the cities in mainland China, marking it as one of mainland China's cities with the best economic benefits. Since 2012, Shenzhen's year-end resident population has increased from 11.9585 million to 17.6816 million (from the Shenzhen Municipal Bureau of Statistics). The

gross regional product also increased from 349.442 billion in the first quarter of 2015 to 573.403 billion in the first quarter of 2019.

Data and processing. In the study of urban elements, land-use change has attracted the most attention. Urban land change focuses on the simulation and prediction of the change trend, and even more so on the driving force of land change. Therefore, different models can be used, such as CA, the multi-agent model, and machine learning. In addition, more than half of the population lives in cities, and a shift in population distribution is one of the most significant changes in urbanization. House price, in certain areas of economics, has been evaluated with many statistical methods over the last 30 years, some of them focusing on spatial aspects.

Because of this, in this paper, we mainly used data on land cover, population density, and house price in Shenzhen. The land cover data for Shenzhen came from the annual China Land Cover Dataset (CLCD) dataset⁴⁶, which was published by the Wuhan University team. The CLCD dataset is an annual China map coverage dataset based on all available Landsat data on Google Earth Engine (GEE) combined with random forest classifiers to obtain classification results. It has a spatial resolution of 30 m and contains year-by-year land cover information for China from 1985 to 2021. Based on 5463 samples of visual interpretation, CLCD has an overall accuracy of 80%. Moreover, the CLCD dataset showed good consistency with global forest change, global surface water, and impermeable surface time-series datasets.

For population data, we used the top-down population density data provided on the WorldPop website. Existing studies have shown that in Guangdong Province, China, WorldPop has spatial distribution consistency with GPW v4 and two Chinese kilometer grid population distribution datasets. Compared with the sixth national census data as the true value, the overall accuracy of the WorldPop dataset is higher. The highest spatial resolution of the population density data provided on WorldPop is 100 m. To be consistent with the resolution of the land cover data, we resampled it to 30 m and cropped the data of Shenzhen. Finally, our house price data came from the data on Fangtianxia and Anjuke, which were converted into raster data with a spatial resolution of 30 m. Three kinds of data with the same coordinate system, spatial range, and temporal and spatial resolution were thus obtained. The time range of the three data types was from 2011 to 2019.

Experiment parameter setting. In the model proposed in this paper, simulating changes in urban elements requires that the impact of the surrounding neighborhood be considered. Therefore, we first needed to define the range of the spatial neighborhood. The currently used spatial neighborhood types include 4-neighborhood, 8-neighborhood, and 25-neighborhood formats. In this paper, we used the classic 8-neighborhood format, which is also used in the CA model as a spatial neighborhood. In the simulation experiment for population and housing price elements, we regarded housing price as the main factor and population density as the secondary factor, and we evaluated the impact of population density on the development process of housing prices in Shenzhen while considering the developmental characteristics of housing prices themselves. In the experiment with labeled data, specifically land-use data, we regarded the built-up area with the largest newly-added area as the main element, according to the characteristics of the data, and evaluated the impact of other categories on the expansion process of the built-up area while considering the developmental inertia of the built-up area itself.

Data availability

The CLCD dataset introduced in this article is freely available at <https://doi.org/10.5281/zenodo.4417810>. The population density data of China can be download at <https://www.worldpop.org>. The house price data can be found at https://github.com/minjincug/House_price_data. The administrative division data of Guangdong Province and Shenzhen City are downloaded from <https://guangdong.tianditu.gov.cn>.

Received: 24 March 2023; Accepted: 18 July 2023

Published online: 31 July 2023

References

- Henderson, J. V. *et al.* Urban development: Theory, fact, and illusion. *OUP Catal.* **20**, 20 (1991).
- Liu, X., Ming, Y., Liu, Y., Yue, W. & Han, G. Influences of landform and urban form factors on urban heat island: Comparative case study between Chengdu and Chongqing. *Sci. Total Environ.* **820**, 153395 (2022).
- Cohen, B. Urbanization, city growth, and the new united nations development agenda. *Cornerstone* **3**, 4–7 (2015).
- Buettner, T. Urban estimates and projections at the united nations: The strengths, weaknesses, and underpinnings of the world urbanization prospects. *Spat. Demography* **3**, 91–108. <https://doi.org/10.1007/s40980-015-0004-2> (2015).
- Portugali, J. What makes cities complex?. In *Complexity, Cognition, Urban Planning and Design: Post-proceedings of the 2nd Delft International Conference*, 3–19 (Springer, 2016).
- Batty, M. *Complexity in City Systems: Understanding, Evolution, and Design*. In *A Planner's Encounter with Complexity*. 99–122 (Routledge, 2016).
- Shi, W. *et al.* *Urban Informatics* (Springer, 2021).
- Li, J. & Sun, Z. Urban function orientation based on spatiotemporal differences and driving factors of urban construction land. *J. Urban Plan. Dev.* **146**, 05020011 (2020).
- Zhu, X., Zhang, P., Wei, Y., Li, Y. & Zhao, H. Measuring the efficiency and driving factors of urban land use based on the DEA method and the PLS-SEM model—a case study of 35 large and medium-sized cities in China. *Sustain. Cities Soc.* **50**, 101646 (2019).
- Wu, H., Lin, A., Xing, X., Song, D. & Li, Y. Identifying core driving factors of urban land use change from global land cover products and poi data using the random forest method. *Int. J. Appl. Earth Obs. Geoinf.* **103**, 102475 (2021).
- Verburg, P. H., De Koning, G., Kok, K., Veldkamp, A. & Bouma, J. A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. *Ecol. Model.* **116**, 45–61 (1999).
- Chen, W. Y., Hu, F. Z. Y., Li, X. & Hua, J. Strategic interaction in municipal governments' provision of public green spaces: A dynamic spatial panel data analysis in transitional china. *Cities* **71**, 1–10. <https://doi.org/10.1016/j.cities.2017.07.003> (2017).

13. Lin, Y.-P., Chu, H.-J., Wu, C.-F. & Verburg, P. H. Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling—a case study. *Int. J. Geogr. Inf. Sci.* **25**, 65–87 (2011).
14. Thapa, R. B. & Murayama, Y. Drivers of urban growth in the Kathmandu valley, Nepal: Examining the efficacy of the analytic hierarchy process. *Appl. Geogr.* **30**, 70–83 (2010).
15. Du, X., Jin, X., Yang, X., Yang, X. & Zhou, Y. Spatial pattern of land use change and its driving force in Jiangsu province. *Int. J. Environ. Res. Public Health* **11**, 3215–3232 (2014).
16. Sang, L., Zhang, C., Yang, J., Zhu, D. & Yun, W. Simulation of land use spatial pattern of towns and villages based on ca-Markov model. *Math. Comput. Model.* **54**, 938–943 (2011).
17. Daba, M. H. & You, S. Quantitatively assessing the future land-use/land-cover changes and their driving factors in the upper stream of the awash river based on the ca-Markov model and their implications for water resources management. *Sustainability* **14**, 1538. <https://doi.org/10.3390/su14031538> (2022).
18. Weslati, O., Bouaziz, S. & Sarbeji, M. M. Modelling and assessing the spatiotemporal changes to future land use change scenarios using remote sensing and ca-Markov model in the mellegue catchment. *J. Indian Soc. Remote Sens.* **51**, 9–29 (2023).
19. Wang, H., Stephenson, S. R. & Qu, S. Modeling spatially non-stationary land use/cover change in the lower Connecticut river basin by combining geographically weighted logistic regression and the ca-Markov model. *Int. J. Geogr. Inf. Sci.* **33**, 1313–1334 (2019).
20. Liu, X., Li, X. & Ai, B. Multi-agent systems for simulating and planning land use development. *Acta Geograph. Sin. Chin. Ed.* **61**, 1112 (2006).
21. KulithalaiShiyamSundar, P. & Deka, P. C. Spatio-temporal classification and prediction of land use and land cover change for the Vembanad lake system, Kerala: A machine learning approach. *Environ. Sci. Pollut. Res.* **29**, 86220–86236 (2022).
22. Liu, X. *et al.* A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landsc. Urban Plan.* **168**, 94–116 (2017).
23. Verburg, P. H. *et al.* Modeling the spatial dynamics of regional land use: The clue-s model. *Environ. Manage.* **30**, 391–405 (2002).
24. Li, X. *et al.* The spatio-temporal pattern and transition mode of recessive cultivated land use morphology in the Huaibei region of the Jiangsu province. *Land* **11**, 1978 (2022).
25. Overmars, K. D., De Koning, G. & Veldkamp, A. Spatial autocorrelation in multi-scale land use models. *Ecol. Modell.* **164**, 257–270 (2003).
26. Chomitz, K. M. & Thomas, T. S. Determinants of land use in Amazonia: A fine-scale spatial analysis. *Am. J. Agr. Econ.* **85**, 1016–1028. <https://doi.org/10.1111/1467-8276.00504> (2003).
27. Chen, W., Shen, Y., Wang, Y. & Wu, Q. The effect of industrial relocation on industrial land use efficiency in China: A spatial econometrics approach. *J. Clean. Prod.* **205**, 525–535. <https://doi.org/10.1016/j.jclepro.2018.09.106> (2018).
28. Yu, D. Exploring spatiotemporally varying regressed relationships: The geographically weighted panel regression analysis. *Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* **38**, 134–139 (2010).
29. Jin, Y., Zhang, H., Yan, Y. & Cong, P. A semi-parametric geographically weighted regression approach to exploring driving factors of fractional vegetation cover: A case study of guangdong. *Sustainability* **12**, 7512 (2020).
30. Li, Z., Fotheringham, A. S., Li, W. & Oshan, T. Fast geographically weighted regression (FASTGWR): A scalable algorithm to investigate spatial process heterogeneity in millions of observations. *Int. J. Geogr. Inf. Sci.* **33**, 155–175 (2019).
31. Wang, J. & Xu, C. Geodetector: Principle and prospective. *Acta Geogr. Sin.* **72**, 116–134 (2017).
32. Bruno, G. & Improta, G. Using gravity models for the evaluation of new university site locations: A case study. *Comput. Oper. Res.* **35**, 436–444. <https://doi.org/10.1016/j.cor.2006.03.008> (2008).
33. Zincenko, A., Petrovskii, S., Volpert, V. & Banerjee, M. Turing instability in an economic–demographic dynamical system may lead to pattern formation on a geographical scale. *J. R. Soc. Interface* **18**, 20210034 (2021).
34. Goh, S., Choi, M. Y., Lee, K. & Kim, K.-M. How complexity emerges in urban systems: Theory of urban morphology. *Phys. Rev. E* **93**, 052309. <https://doi.org/10.1103/PhysRevE.93.052309> (2016).
35. Barthélemy, M. Spatial networks. *Phys. Rep.* **499**, 1–101. <https://doi.org/10.1016/j.physrep.2010.11.002> (2011).
36. Friesen, J., Tessmann, R. & Pelz, P. F. Reaction–diffusion model describing the morphogenesis of urban systems in the US. In *GISTAM*, 88–96 (2019).
37. Raimbault, J. *et al.* Worldwide estimation of parameters for a simple reaction–diffusion model of urban growth. Tech. Rep. (2019).
38. Raimbault, J. A comparison of simple models for urban morphogenesis. [arXiv:2008.13277](https://arxiv.org/abs/2008.13277) (arXiv preprint) (2020).
39. Ribeiro, F. L. & Rybski, D. Mathematical models to explain the origin of urban scaling laws. *Phys. Rep.* **1012**, 1–39. <https://doi.org/10.1016/j.physrep.2023.02.002> (2023).
40. Jin, M., Wang, L., Ge, F. & Xie, B. Understanding the dynamic mechanism of urban land use and population distribution evolution from a microscopic perspective. *ISPRS Int. J. Geo-Inf.* <https://doi.org/10.3390/ijgi11110536> (2022).
41. He, Q., Song, Y., Liu, Y. & Yin, C. Diffusion or coalescence? Urban growth pattern and change in 363 Chinese cities from 1995 to 2015. *Sustain. Cities Soc.* **35**, 729–739. <https://doi.org/10.1016/j.scs.2017.08.033> (2017).
42. Sudhira, H., Ramachandra, T. & Jagadish, K. Urban sprawl: Metrics, dynamics and modelling using GIS. *Int. J. Appl. Earth Obs. Geoinf.* **5**, 29–39. <https://doi.org/10.1016/j.jag.2003.08.002> (2004).
43. Chakraborty, S., Maity, I., Dadashpoor, H., Novotny, J. & Banerji, S. Building in or out? Examining urban expansion patterns and land use efficiency across the global sample of 466 cities with million+ inhabitants. *Habitat Int.* **120**, 102503. <https://doi.org/10.1016/j.habitatint.2021.102503> (2022).
44. Dubey, B., Das, B. & Hussain, J. A predator–prey interaction model with self and cross-diffusion. *Ecol. Model.* **141**, 67–76. [https://doi.org/10.1016/S0304-3800\(01\)00255-1](https://doi.org/10.1016/S0304-3800(01)00255-1) (2001).
45. DeAngelis, D. L., Ni, W.-M. & Zhang, B. Dispersal and spatial heterogeneity: Single species. *J. Math. Biol.* **72**, 239–254. <https://doi.org/10.1007/s00285-015-0879-y> (2016).
46. Yang, J. & Huang, X. The 30 m annual land cover dataset and its dynamics in china from 1990 to 2019. *Earth Syst. Sci. Data* **13**, 3907–3925. <https://doi.org/10.5194/essd-13-3907-2021> (2021).

Acknowledgements

This work is supported by the International Research Center of Big Data for Sustainable Development Goals (CBAS2022GSP05), the National Natural Science Foundation of China under Grant 41925007 and the Open Research Program of the International Research Center of Big Data for Sustainable Development Goals under Grant CBAS2023ORP03.

Author contributions

M.J. conceived and conducted the experiments, L.W. and J.Y. contributed to the conception, F.G. helped to design of the model. All authors reviewed the manuscript.

Conflict of interest

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023