



## Coupling cellular automata with area partitioning and spatiotemporal convolution for dynamic land use change simulation

Yuehui Qian, Weiran Xing, Xuefeng Guan\*, Tingting Yang, Huayi Wu

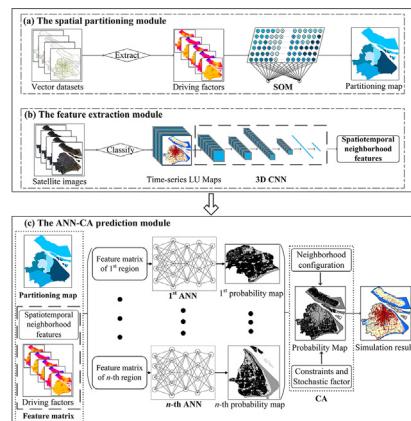
*State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China*



### HIGHLIGHTS

- A CA model is proposed by integration with deep learning methods for LUC modeling.
- SOM is employed to handle spatial heterogeneity by region partitioning.
- A 3D CNN model is adopted to capture spatiotemporal neighborhood features.
- Three patterns of "coverage rate-growth rate" are found in SOM partitioning result.
- The optimal time steps in CNN shows a positive correlation with urban growth rate.

### GRAPHICAL ABSTRACT



### ARTICLE INFO

#### Article history:

Received 27 December 2019

Received in revised form 16 February 2020

Accepted 3 March 2020

Available online 12 March 2020

Editor: Christian Herrera

#### Keywords:

Land use

Spatial heterogeneity

Neighborhood effects

Spatiotemporal features

Cellular automata

### ABSTRACT

Urbanization processes have accelerated over recent decades, prompting efforts to model land use change (LUC) patterns for decision support and urban planning. Cellular automata (CA) are extensively employed given their simplicity, flexibility, and intuitiveness when simulating dynamic LUC. Previous research, however, has ignored the spatial heterogeneity among sub-regions, instead applying the same transition rules across entire regions; moreover, most existing methods extract neighborhood effects with only one data time slice, which is inconsistent with the nature of neighborhood interactions as a long-term process exhibiting obvious spatiotemporal dependency.

Accordingly, we propose a hybrid cellular automata model coupling area partitioning and spatiotemporal neighborhood features learning, named PST-CA. We use a machine-learning-based partitioning strategy, self-organizing map (SOM), to divide entire regions into several homogeneous sub-regions, and further apply a spatiotemporal three-dimensional convolutional neural network (3D CNN) to extract the spatiotemporal neighborhood features. An artificial neural network (ANN) is then built to create a conversion probability map for each sub-region using both spatiotemporal neighborhood features and factors that drive the LUC. Finally, the dynamic simulation results of entire study area are generated by fusing these probability maps, constraints and stochastic factors.

Land use data collected from 2000 to 2015 in Shanghai were selected to verify our proposed method. Four traditional models were implemented for comparison, including logistic regression (LR)-CA, support vector machine (SVM)-CA, random forest (RF)-CA and conventional ANN-CA. Results illustrate that the proposed PST-CA

\* Corresponding author.

E-mail addresses: [yuehui.qian@whu.edu.cn](mailto:yuehui.qian@whu.edu.cn) (Y. Qian), [guanxuefeng@whu.edu.cn](mailto:guanxuefeng@whu.edu.cn) (X. Guan).

outperformed four traditional models, with overall accuracy increased by 4.66%~6.41%. Moreover, three distinctly different “coverage rate-growth rate” composite patterns of built-up areas are shown in the SOM partitioning results, which verifies SOM’s ability to address spatial heterogeneity; while the optimal time steps in 3D CNN generally maintained a positive correlation with the growth rate of built-up areas, which implies longer temporal dependency should be captured for rapidly developing areas.

© 2020 Elsevier B.V. All rights reserved.

## 1. Introduction

Land use and land cover change are among the most direct manifestations of the interaction between human activities and the natural environment (Fiener et al., 2011; Haase et al., 2012). Currently, rapid population growth and rampant urban sprawl in China have accelerated the tension among human-environment interactions (Aburas et al., 2019; Yang et al., 2018); this has resulted in a variety of land use problems, including farmland occupation, deforestation, soil pollution, etc. Thus, the research society has prompted efforts to explore patterns and disciplines in land use change (LUC) for decision support and planning (Cheshire and Hay, 2017; Liu et al., 2017; Sun et al., 2018).

Over the past few decades, a number of methods for LUC dynamic simulation have been proposed. These can be classified into several categories, including statistical and machine learning methods, economic-based models, cellular automata (CA) models, agent-based models, and hybrid approaches (Noszczyk, 2018; Ren et al., 2019). Statistical and machine learning models usually explore the LUC processes by establishing a mathematical relationship between conversion probability and various variables; but these methods are difficult to interpret the physical mechanism under the LUC (Takamatsu et al., 2014). The economic-based models can explain land use by employing distances and transport costs, assuming that landowners will use land in a way that maximizes their utility and expected profits; however, they usually ignore the spatial interactions between different land units (Azadi et al., 2017). CA models are discrete in terms of time and space, and have strong ability to represent the spatiotemporal processes of LUC (Roodposhti et al., 2019; White et al., 1997). Agent-based models aim to understand the LUC processes based on the prescribed rules and interactions between autonomous agents abstracted from a variety of decision makers and institutions; but huge computational and empirical resources are required (Groeneveld et al., 2017; Parker et al., 2003). Finally, hybrid approaches simulate the LUC processes according to practical demands, incorporating different conceptual frameworks, theories, and observations together; however, they suffer from complexity in model calibration/validation (Chang-Martínez et al., 2015).

Of these methods, CA has been extensively employed by researchers to simulate LUC because of its simplicity, flexibility, and intuitive ability to reflect spatiotemporal dependency in land use patterns. CA is a discrete dynamic process model that usually consists of five components: space, state, time step, neighborhood, and transitional rule (Islam et al., 2018; Wolfram, 1984). The space in a CA is usually tessellated into cells of equal size and shape. Cells can interact with each other according to defined transitional rules, and update their state at every time step during the model evolution. The definition of the transitional rules involves the states of neighbor cells and a series of external factors (Xia et al., 2017). Neighborhood interactions follow transition rules at discrete regular steps, which allows CA to accurately simulate the dynamic LUC processes in a bottom-up way. Therefore, the correct extraction of transition rules and the precise modeling of neighborhood effects are determinants of the model’s prediction capability (Cao et al., 2015; Feng and Tong, 2018).

Transition rules have been widely studied over the last few decades using a variety of methods, including logistic regression (LR), support vector machine (SVM), random forest (RF), artificial neural network (ANN) and so on (Gounaris et al., 2019; Islam et al., 2018; Mustafa et al., 2018; Rienow and Goetzke, 2015). Although the extraction

methods of transition rules have become more and more complex, while the accuracy of LUC dynamic simulations has gradually improved, most previous research has used the same transition rules to simulate and predict LUC across the entire study area, which ignores the spatial heterogeneity among sub-regions. Spatial heterogeneity is a non-negligible geographical phenomenon existing in LUC simulations, which can manifest as inconsistent development level and variable growth rate in different sub-regions. While growth rates can be very high in some areas where the economy or population is booming, they may progress at a relatively slow rate in well-developed or undeveloped areas (Xia et al., 2019). Therefore, neglecting spatial heterogeneity inevitably leads to limited simulation performance when the same transition rules are used across the entire research area.

In light of the above, some researchers have begun to pay attention to the existence of spatial heterogeneity in LUC simulations. Xia et al. (2019) applied partitioned transition thresholds in dynamic LUC simulations to address spatial heterogeneity, however, they encountered difficulty in selecting an appropriate threshold for different regions. Ke et al. (2015) and Yang et al. (2018) have partitioned their research area using land use change data by means of the *k*-means clustering algorithm, and consequently derived relatively homogeneous sub-regions; however, these approaches suffered from the random determination of initial cluster centers. Moreover, the topological properties of the entire study area were destroyed, since these studies used land use change data only, which obstructs normal cell interactions.

Neighborhood effects refer to a geographical phenomenon commonly observed in LUC processes, which exhibit spatial correlation and temporal dependency. Spatial correlation is the tendency of land use statuses that are closer together to be more similar than those that are further apart. Moreover, temporal dependency occurs when the land use statuses of adjacent units remain consistent over time; for example, a land use unit is unlikely to convert into a built-up land use category when most of the neighbors remain non-built areas for an extended period.

Some researchers have tried to qualify and analyze the neighborhood effects using various methods. Hughes Jr. and Turnbull Hughes and Turnbull (1996) analyzed neighborhood effects with a theoretical economic model by explicitly treating neighborhood externalities as uncertain processes, which is limited by both the extensive expertise required and the obviously subjective nature of the process. Zhou and Kockelman (2008) explored neighborhood impacts with parcel-level data by using a multinomial logit model from spatial perspective, whereas cell interactions within one parcel cannot be assessed and the temporal dependency of LUC is ignored. He et al. (2018) considered neighborhood effects with a convolutional neural network for united mining (UMCNN) implemented on spatial driving factors; this approach, however, explored neighborhood effects with only one data time slice, meaning that temporal dependency was neglected and historical information was lost.

To address these limitations, we propose a novel hybrid cellular automata model coupling area partitioning and spatiotemporal neighborhood features learning (PST-CA). It employs a partition zoning strategy with a self-organizing map (SOM) to generate multiple relatively homogeneous sub-regions. A spatiotemporal three-dimensional convolutional neural network (3D CNN) is adopted to capture both the spatial correlation and temporal dependency of neighborhood. Furthermore, taking the learned spatiotemporal neighborhood features

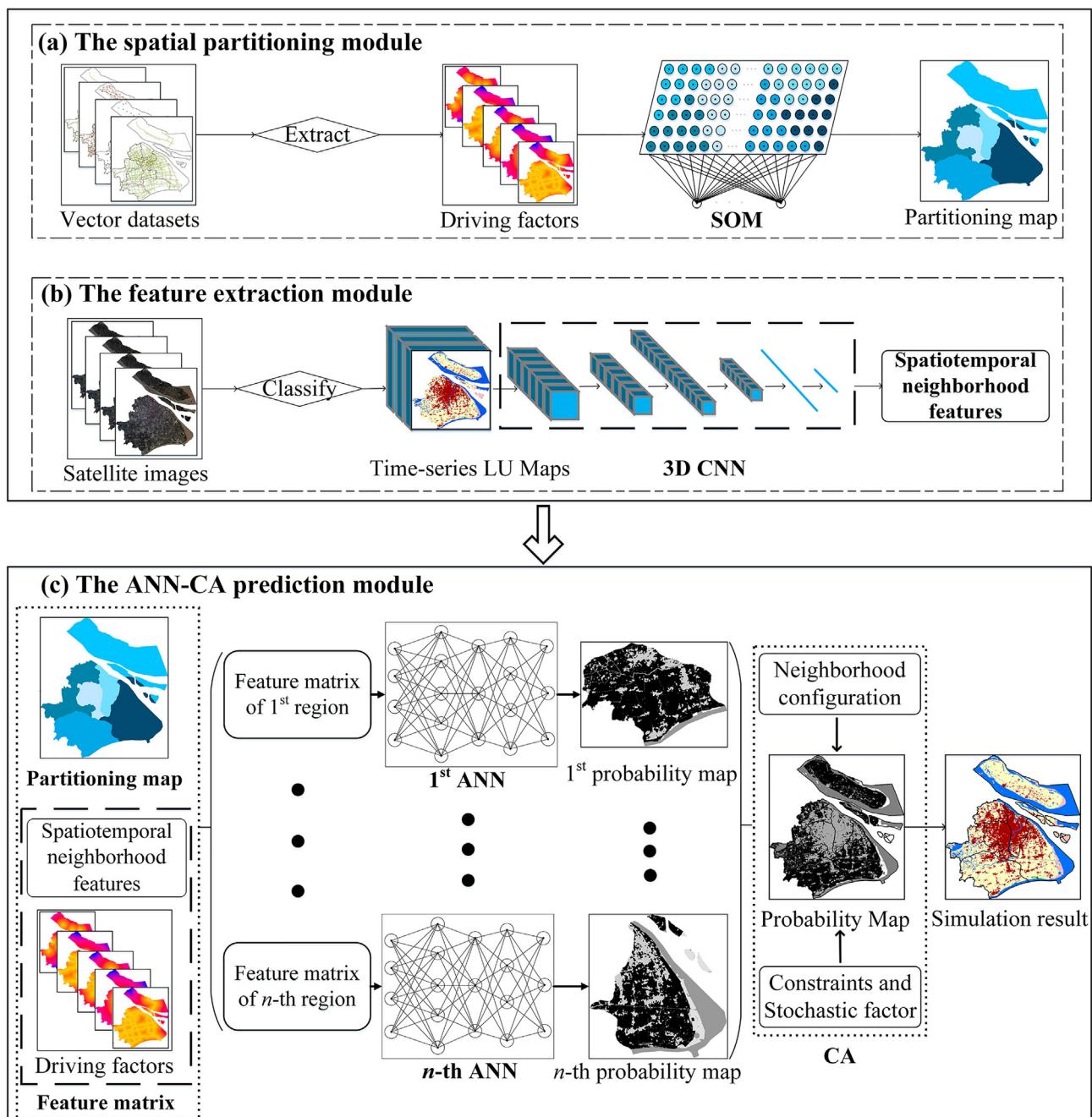
and driving factors as input, an artificial neural network (ANN)-CA model is built in each sub-region to simulate dynamic LUC of the entire research area. In addition, four conventional models, namely LR-CA, SVM-CA, RF-CA and ANN-CA, were applied on the same test data to facilitate model comparison. The city of Shanghai, in China is selected as a case study, with multivariate data from 2000 to 2015 used to validate the proposed models. The experimental results indicate that the proposed PST-CA model outperforms the conventional models, with an increase in overall accuracy, about 4.66%~6.41%.

The main contributions of this paper can be summarized as follows:

- 1) We propose a PST-CA model to simulate the dynamic LUC processes that can address spatial heterogeneity and extract spatiotemporal neighborhood features.

- 2) We employ SOM to partition the whole area into several relatively homogeneous sub-regions in order to handle spatial heterogeneity.
- 3) We use a 3D CNN during the modeling process to more accurately characterize the spatial correlation and temporal dependency of neighborhoods.

The remainder of this paper is structured as follows: **Section 2** introduces the model basics and methodology used in this research; **Section 3** outlines the study area and data; **Section 4** introduces the simulation conditions and evaluation metrics of our proposed method; **Section 5** presents and analyzes the simulation results; finally, **Section 6** presents the discussion and **Section 7** ends with the conclusion.



**Fig. 1.** Framework of the PST-CA model.

## 2. Methodology

There are three modules in the proposed PST-CA model: (a) the spatial partitioning module; (b) the feature extraction module, and (c) the ANN-CA prediction module, as shown in Fig. 1.

As shown in Fig. 1, the entire study area is partitioned into several sub-regions by the SOM in module (a). The input of SOM is the driving factors extracted from various vector datasets of accessibility/terrain conditions, while the output is the partitioning results for the entire study area; this module will be described in more detail in Section 2.1. In module (b), a 3D CNN model is utilized to extract spatiotemporal neighborhood features for the entire study area, using time-series LU maps as input; this will be discussed in more detail in Section 2.2. In module (c), an individual ANN model with input feature matrix is employed to calculate the probability map for each sub-region. Finally, by integrating the probability map, neighborhood configuration, stochastic factors and constraints, the simulation result is acquired; integration details will be further presented in Section 2.3.

### 2.1. Partitioning strategy with self-organizing map

In the LUC process, unbalanced urbanization levels and uneven development speeds are common phenomena among different regions, and usually area partitioning are used to handle spatial heterogeneity. Due to various socioeconomic driving factors involved, the input for partitioning algorithms are very sparse in this established high-dimensional space, which will inevitably reduce the effectiveness of direct distance-based clustering methods, e.g. *k*-means (Napoleon and Pavalakodi, 2011). Here, we applied a dimension-reduction based clustering method, self-organizing map (SOM) to partition the study area into several relatively homogeneous sub-regions. SOM is an unsupervised artificial neural network first proposed by Kohonen (1982), and has been widely applied to the solving of clustering problems (Nguyen et al., 2015; Sacha et al., 2018). It can project high-dimensional, nonlinear data to a low dimensional map, using a neighborhood distance function while preserving the topological properties of input data (Chen et al., 2019).

A SOM consists of two main layers, namely the input layer and the output or competitive layer. The input neurons can transfer the input vectors to the clustering model, while the output neurons determine the distribution of the input vectors using a competitive and unsupervised method. Each neuron in the output layer is connected to the neurons in the input layer by means of a weight vector. The learning process of SOM involves finding the output neuron closest to each input vector and moving the “winning” output neuron closer to the input vector using three consecutive steps: competition between output neurons with Euclidian distance, selection of a winner output neuron, and updating of the weight vectors of each output neuron (Huang et al., 2017; Nguyen et al., 2015).

In this study, there are  $N$  neurons used as training vectors  $\{x_1, x_2, \dots, x_N\}$  in the input layer; each of these input vectors  $x_i$  has  $M$  features, all of which can be expressed as  $\{x_{i,1}, x_{i,2}, \dots, x_{i,M}\}$ .  $N$  denotes the number of geographical cells in the study area, while  $M$  denotes the number of features, which are extracted from the road network data, points of interest (POIs), and terrain data. The output layer contains  $P$  neurons, where  $P$  indicates the number of clusters  $\{Y_1, Y_2, \dots, Y_P\}$ . Each output neuron has a weight vector, all of which can be expressed as  $\{w_1, w_2, \dots, w_P\}$ . Moreover, each of these weight vectors  $w_j$  has the same dimensionality as the input vector  $x_i$ ; thus, each weight vector can be expressed as  $\{w_{j,1}, w_{j,2}, \dots, w_{j,M}\}$ . The SOM is trained iteratively, and the whole training procedure is as described below.

During the iterations, a sample input vector  $x_i$  is randomly selected. The Euclidian distance between  $x_i$  and each weight vector  $w_j$  is

calculated as in Eq. (1):

$$dis(x_i, w_j) = \sum_{k=1}^M (x_{ik} - w_{jk})^2 \quad (1)$$

As shown in Eq. (2), we select the output vector with minimum distance as the Best Matching Unit (BMU)  $c$ .  $x_i$  will be projected to BMU  $c$ :

$$c = \underset{1 \leq j \leq P}{\operatorname{argmin}} |dis(x_i, w_j)| \quad (2)$$

Then, the learning rates  $\psi(t)$ , neighborhoods  $\theta(t)$  of BMU  $c$ , and weight vectors  $w_j(t+1)$  of BMU  $c$  and neighboring cells will be updated by Eqs. (3)–(5):

$$\psi(t) = \psi_0 \exp\left(-\frac{t}{\tau}\right) \quad (3)$$

$$\theta(t) = \exp\left(-\frac{dis^2}{2\psi^2(t)}\right) \quad (4)$$

$$w_j(t+1) = w_j(t) + \theta(t) \cdot \psi(t) \cdot (x_i - w_j(t)) \quad (5)$$

Here,  $\psi_0$  is the initial value of the learning rate,  $t$  is the number of iterations,  $\tau$  is a time constant based on the iterations.

The BMU and its topological neighbors are moved closer to the input vector. As a result, those input neurons (geographical cells) with similar instances will be more likely to be projected to the same output neuron (cluster) in the map grid space.

### 2.2. Extracting neighborhood features with 3D convolutional neural network

Since LUC exhibit obvious spatial patterns and temporal dependency, it is of great necessity to accurately extract both spatial and temporal dependency contained in historical land use sequences to explore LUC patterns (Sidharthan and Bhat, 2012). Here, we employed a 3D convolutional neural network (3D CNN) to capture the spatiotemporal neighborhood features, because this approach can take both spatial patterns and temporal dependency into account simultaneously. 3D CNN is an extension of 2D CNN into three dimensions. A 2D convolution operation has been used to capture local spatial features in 2D images (He et al., 2018), as shown in Eq. (6):

$$y_{c,d}^l = \sigma \left( \sum_m \sum_{i=0}^H \sum_{j=0}^W w_{i,j}^{l,m} x_{c+i,d+j}^{l-1,m} + b^l \right) \quad (6)$$

Here,  $l$  denotes the current layer,  $y_{c,d}^l$  represents the output at location  $(c, d)$ ,  $m$  is the number of feature maps in the  $(l-1)^{th}$  layer,  $b^l$  is the bias,  $\sigma$  is the activation function,  $w_{i,j}^{l,m}$  is the value at position  $(i, j)$  of the kernel, and  $H$  and  $W$  are the height and width of the kernel respectively.

Directly applying 2D convolution operations onto 3D data is challenging, since 2D convolution operations only capture two-dimensional spatial information, and ignore the information contained in the temporal dimension (Zou et al., 2017). To extract spatiotemporal neighborhood features, we employ a 3D convolution operation to time-series LU maps as in Eq. (7):

$$y_{c,d,e}^l = \sigma \left( \sum_m \sum_{k=0}^T \sum_{i=0}^H \sum_{j=0}^W w_{c,d,e}^{l,m} x_{c+i,d+j,e+k}^{l-1,m} + b^l \right) \quad (7)$$

Here,  $T$  represents the size of the 3D kernel along the temporal dimension,  $w_{c,d,e}^{l,m}$  is a 3D tensor, and  $y_{c,d,e}^l$  denotes the output at location

( $c, d, e$ ), respectively. The difference between 2D and 3D convolution operations is illustrated in Fig. 2.

As Fig. 2(a) shows, two dimensional data were convolved with 2D kernels and the spatial features were extracted after 2D convolution operations. In Fig. 2(b), moreover, the three-dimensional data (time-series LU maps in this study) were convolved with 3D kernels, considering both spatial and temporal information.

More details about the 3D CNN model employed in this study are presented in Fig. 3. As the figure illustrates, we select each pixel of the LU maps as a central pixel; subsequently, the central pixel and its neighbor pixels for  $T$  consecutive years are clipped to form a  $T \times 35 \times 35$  grid. The input layer of this 3D CNN is a four-dimensional tensor  $M@T \times 35 \times 35$ . Here,  $M$  indicates independent channels, and is equal to 1 in our study because of the grayscale nature of the image, while  $T$  represents the temporal steps. To explore the neighborhood features more fully, 35 is selected as the width and height of each sample. The three convolutional layers in our model are used to capture implicit spatiotemporal neighborhood features. The 3D CNN model employs a  $3 \times 3 \times 3$  convolution kernel and a  $1 \times 2 \times 2$  pooling layer. Since additional pooling layers will reduce the information contained in the original data, only the first convolutional layers are followed by pooling layers, meaning that the spatiotemporal information is preserved as much as possible. All of these convolutional layers are applied with appropriate temporal padding. The output layer is a high-dimensional feature vector. The softmax regression layer is removed after the 3D CNN model is trained, and the high-dimensional features of the upper layer are exported.

### 2.3. Integrating the artificial neural network with cellular automata

Currently ANN has been widely used for classification and function approximation because of its efficiency and ability in modeling nonlinear relationships (Qiang et al., 2015). We employed an ANN model to calculate a probability map for each sub-region, then simulate the LUC results using CA. The CA model defines the state of a given cell at the next time step as a function of five parts: current cell state  $S_{i,t}$ , overall conversion probability  $P_o$ , neighborhood configuration  $\Omega_{i,m}$ , constraint coefficient  $CONS$  and stochastic factors  $RND$  (Feng et al., 2018; Liu and Feng, 2016). A conceptual formula for the transition rule can be written as in Eq. (8):

$$S_{i,t+1} = f(S_{i,t}, P_o, \Omega_{i,m}, CONS, RND) \quad (8)$$

Here,  $S_{i,t}$  and  $S_{i,t+1}$  denote the states of a cell at location  $i$  at time  $t$  and  $t + 1$ , respectively,  $f$  is the transition function,  $P_o$  represents the overall conversion probability, and  $\Omega_{i,m}$  denotes the configuration of the surrounding cells within the  $m \times m$  window, which can be calculated by Eq. (9) below. Moreover,  $CONS$  defines the constraint overlay and restricts which cells are able to change; a cell is assigned a value of 1 if this cell is available for development; otherwise assigned a value of 0.  $RND$  is a stochastic disturbance representing unknown perturbations, as shown in Eq. (10).

$$\Omega_{i,m} = \frac{\sum_{m \times m} (S_{i,t=k})}{m \times m - 1} \quad (9)$$

$$RND = 1 + (-\ln \gamma)^\alpha \quad (10)$$

Here,  $\gamma$  and  $\alpha$  are the controlling coefficients, used to adjust the effect of the stochastic factor in the LUC process,  $\gamma$  is a random real number ranging from 0 to 1, and  $\alpha$  is an integer in the range of [0,10].

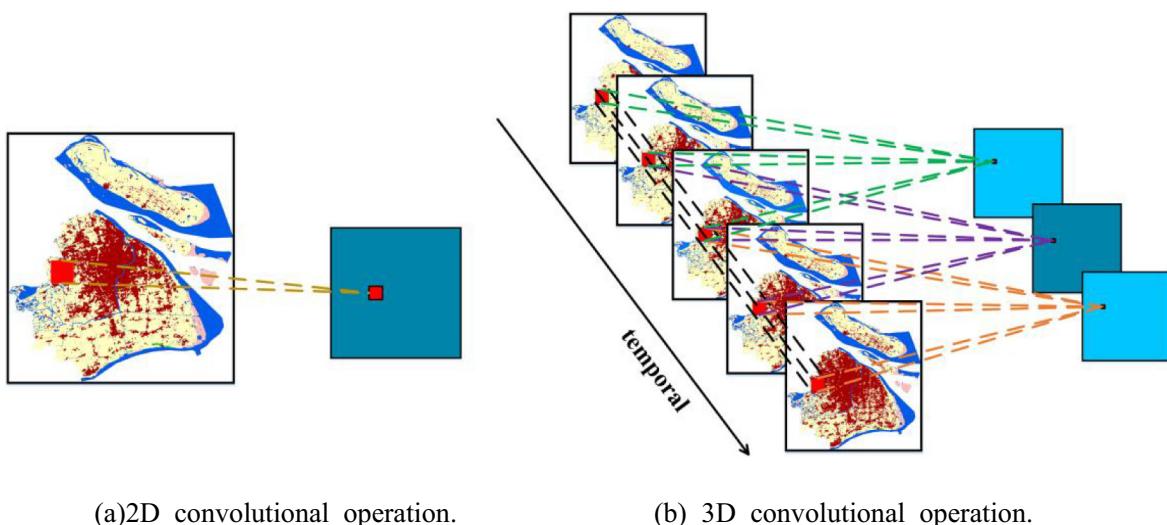
In this study, the overall conversion probability  $P_o$  at location  $i$  in each sub-region  $r$  is determined by an ANN model, expressed as  $P_{ann,i}^r$ .  $P_{ann,i}^r$  is affected by its current state  $X_{self}$ , a series of driving factors  $X_{dt}$  and neighborhood spatiotemporal information  $X_{st}$ , as shown in Eq. (11):

$$P_{ann,i}^r = g_{ann,i}^r(X_{self}, X_{dt}, X_{st}) \quad (11)$$

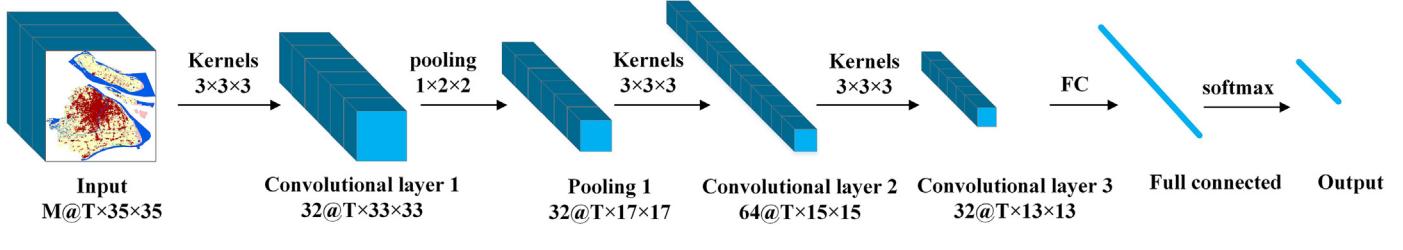
Here,  $X_{dt}$  is a vector constructed by a set of proximity factors calculated from the driving factors (i.e. road network data, public infrastructures data and terrain data),  $X_{st}$  is a vector depicting the neighborhood spatiotemporal features captured by 3D CNN, and  $g_{ann,i}^r$  is the ANN function in the sub-region  $r$ .

After calculating the conversion probability of each sub-region, a simulation result is generated using neighborhood configuration, constraint coefficient, and stochastic factors. Accordingly, Eq. (8) can be transformed into Eq. (12), as follows:

$$\begin{aligned} S_{i,t+1}^r &= f(S_{i,t}^r, P_{ann,i}^r, \Omega_{i,m}^r, CONS^r, RND^r) \\ &= f(S_{i,t}^r, g_{ann,i}^r(X_{self}, X_{dt}, X_{st}), \Omega_{i,m}^r, CONS^r, RND^r) \end{aligned} \quad (12)$$



**Fig. 2.** Illustration of (a) 2D convolutional operation and (b) 3D convolutional operation.



**Fig. 3.** Procedure of the 3D CNN.

### 3. Study area and data

#### 3.1. The Shanghai metropolitan region

Shanghai is known as China's economic, financial, trade, and technological innovation center (Fig. 4), covering an area of 6340.5 km<sup>2</sup>. Its GDP grew from 477.117 billion Yuan in 2000 to 2512.345 billion Yuan in 2015, while the population increased from 16.086 million to 24.153 million over this period (Bureau, 2016; Walcott and Pannell, 2006). The region covers latitude 30°40'–31°53' N and longitude 120°52'–122°12' E. There are sixteen administrative counties in Shanghai, namely Hongkou (HK), Jingan (JA), Huangpu (HP), Xuhui (XH), Chongming (CM), Yangpu (YP), Putuo (PT), Changning (CN), Jinshan (JS), Qingpu (QP), Fengxian (FX), Pudong New District (PND), Songjiang (SJ), Baoshan (BS), Minhang (MH), and Jiading (JD). Shanghai is adjacent to Jiangsu (JS) and Zhejiang (ZJ) provinces and located at the core of the Yangze River Delta urban agglomeration (You, 2016). The built-up area expanded by nearly 3.5 times from 2000 to 2015. Due to rapid urban development, land use has changed significantly in this region, resulting in a series of environmental, economic and urban development issues (He et al., 2018).

#### 3.2. Land use maps

Sixteen Landsat-7 Enhanced Thematic Mapper Plus (ETM+) images from 2000 to 2015 were used to derive the time-series land use maps for Shanghai City. Using the human settlement classification results of a study of Tsinghua University (Gong et al., 2019) as a reference, these

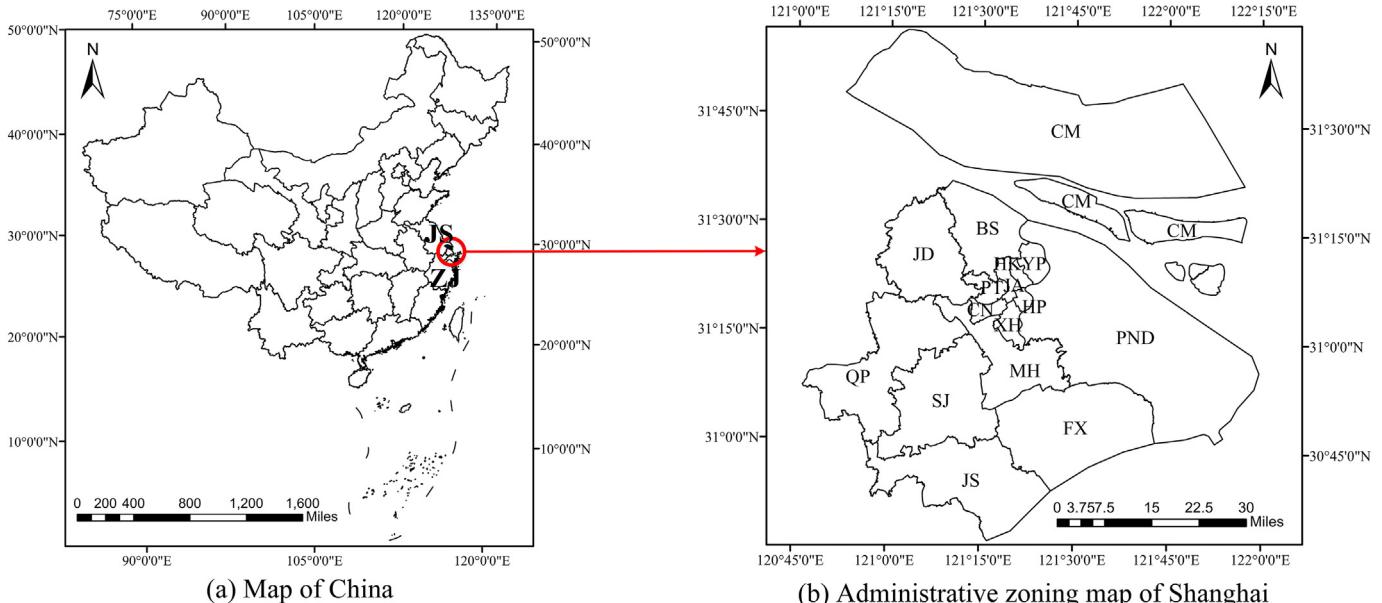
sixteen Landsat-7 images were classified by means of supervised classification. Data preprocessing included geometric corrections, radiometric corrections, cropping boundaries, supervised classification, and visual interpretation. All of these processing produces were conducted with ENVI and ArcGIS software. Finally, these time-series land use maps were resampled to 100 m to save calculation cost and accelerate simulation process.

The LU maps of these four years (2000, 2005, 2010, and 2015) are presented in Fig. 5. Six categories of land use were identified in the LU maps: namely, agricultural land, forest land, grass land, waterbodies land, urban land, and unused land. As shown in Fig. 5, the area of built-up is rapidly expanding, while the area of agricultural land is decreasing year by year.

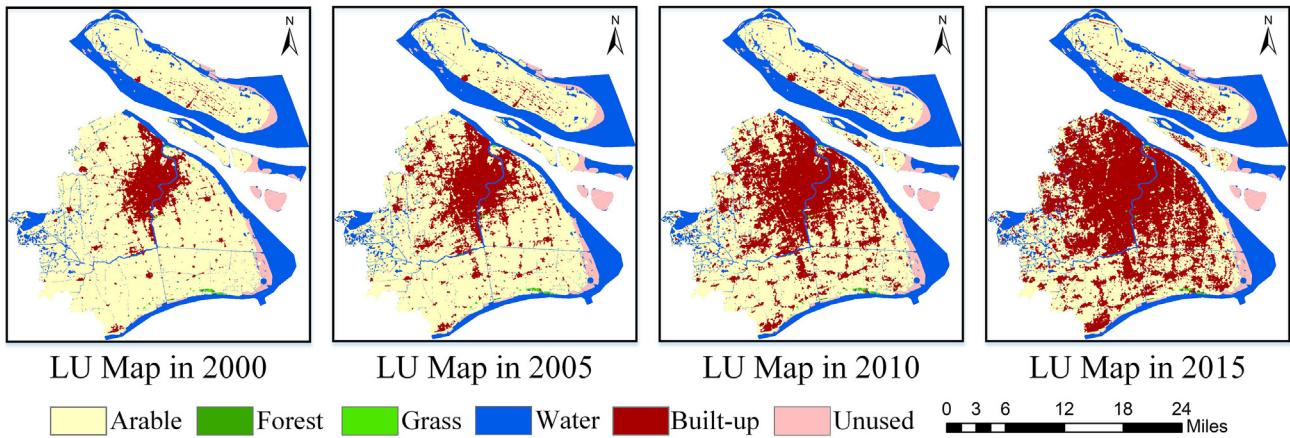
#### 3.3. Driving factors

Taking previous LUC studies into account (Chen et al., 2017; Gounaris et al., 2018), three major groups of driving factors were adopted to describe the LUC process in this research: accessibility to transport lines, accessibility to public infrastructures, and terrain conditions, which were in turn decomposed into 17 variables as listed in Table 1.

As Table 1 illustrates, six variables ( $X_1$ – $X_6$ ) indicating accessibility to transport lines are derived from road network and waterbodies data, nine variables ( $X_7$ – $X_{15}$ ) indicating accessibility to public infrastructures are derived from POIs, and two variables ( $X_{16}$ – $X_{17}$ ) indicating geographical conditions are derived from terrain data. Open Street Map, Baidu Map, and Geospatial Data Cloud are the main data sources in this



**Fig. 4.** Location of study area; (a) Map of China; (b) Administrative zoning map of Shanghai.



**Fig. 5.** Classification land use map in Shanghai city for the years 2000, 2005, 2010 and 2015.

research. All data are normalized to [0, 1] and resampled to 100 m to reduce calculation cost and accelerate the simulation process.

Transport lines have facilitated human travel, and many towns and counties have emerged along the transport lines. Thus, accessibility to transport lines has been frequently selected as a driving factor in previous research (Chen et al., 2018; Feng and Tong, 2018). Transport lines were identified and classified as one of six types in our research: motorways, trunk roads, primary roads, minor roads, railways, and rivers (i.e.  $X_1-X_6$  in Table 1). Accessibility is measured using the minimum Euclidean distance to the nearest transport lines (Fig. 6). “Low” denotes that the distance to these transport lines is short, while “High” denotes a distance far from these transport lines.

Public infrastructures afford us great convenience to our daily life and work. Locations closer to public infrastructures are thus more likely to convert into built-up areas. Thus, accessibility to nine types of common public infrastructures were selected as driving factors (Fig. 7), including the minimum Euclidean distance to city center, county center, town center, airport, train station, railway station, ferry terminal, coach station, and bus stop (i.e.  $X_7-X_{15}$  in Table 1). “Low” indicates that these type of public infrastructure are a short distance away, while “High” denotes a long distance from the public infrastructure.

Moreover, factors related to topography, such as elevation and slope (Fig. 8), also affect an area's suitability for built-up expansion

(Gounaris et al., 2019). Elevation, and slope are the two indicators used to depict geographical conditions in this study (i.e.  $X_{16}-X_{17}$  in Table 1). Elevations were obtained from the Shuttle Radar Topography Mission (SRTM). The original spatial resolution of the elevation data is 1 arc sec (often quoted as 30 m resolution), which was resampled to 100 m, while slope was calculated from the elevation. “Low” denotes a small elevation or slope value, while “High” indicates the elevation or slope value is large.

#### 4. Experiment implementations and evaluation metrics

##### 4.1. Experiment implementation

The proposed PST-CA model was coded in Python and implemented with Keras. As over-fitting problems usually occur while training deep learning algorithms (i.e. CNN, ANN), a dropout layer was added to avoid over-fitting, while Batch-Normalization was chosen to improve model performance. Furthermore, four traditional models, namely LRC-A, SVM-CA, RF-CA and ANN-CA, were employed to enable comparison of their results with the proposed PST-CA model. The four traditional models were built with Scikit-learn and standard Python libraries (Pedregosa et al., 2011; Rey and Anselin, 2010). In addition, to test the significance of area partitioning and spatiotemporal neighborhood

**Table 1**

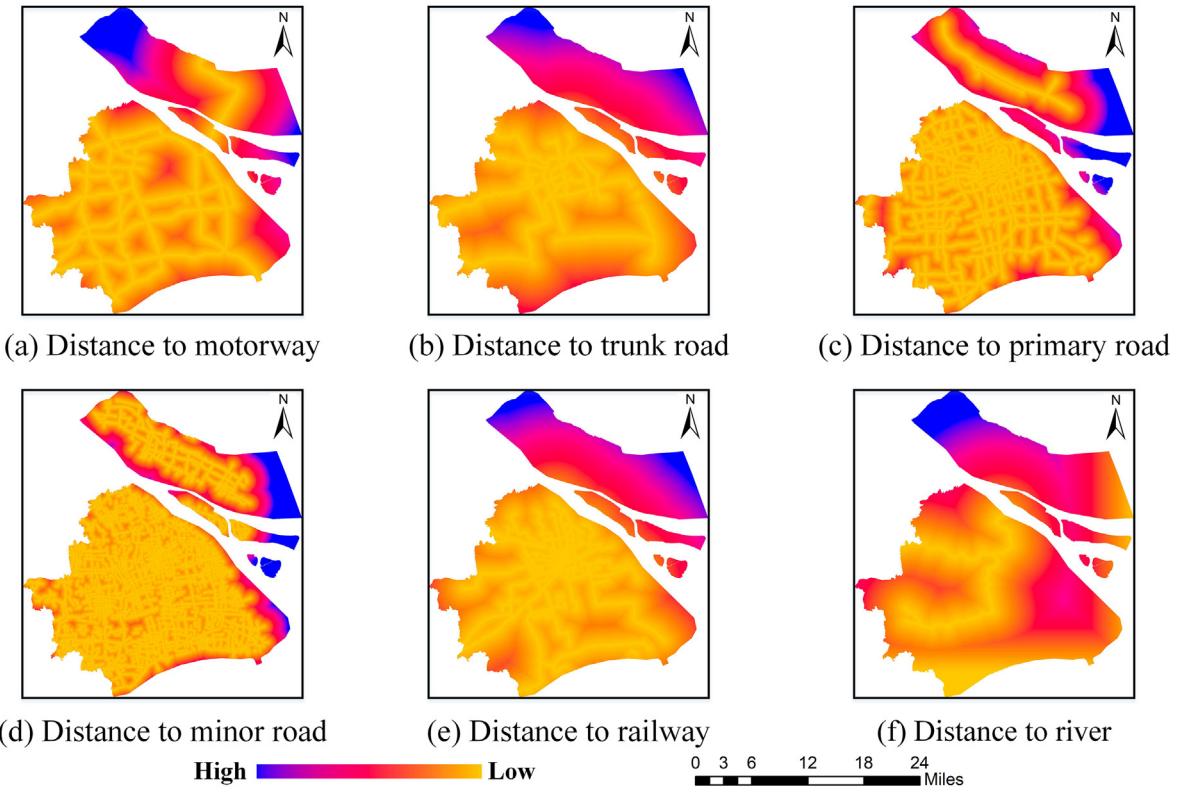
List of factors in this study.

Categories	Data	Variable	Data source	Spatial resolution
Accessibility to transport lines	Distance to motorway	$X_1$	Open Street Map <sup>a</sup>	100 m
	Distance to trunk	$X_2$	Open Street Map	100 m
	Distance to primary road	$X_3$	Open Street Map	100 m
	Distance to minor road	$X_4$	Open Street Map	100 m
	Distance to railway	$X_5$	Open Street Map	100 m
	Distance to river	$X_6$	Open Street Map	100 m
Accessibility to public infrastructures	Distance to city center	$X_7$	Baidu Map <sup>b</sup>	100 m
	Distance to county center	$X_8$	Baidu Map	100 m
	Distance to town center	$X_9$	Baidu Map	100 m
	Distance to airport	$X_{10}$	Baidu Map	100 m
	Distance to train station	$X_{11}$	Baidu Map	100 m
	Distance to railway station	$X_{12}$	Baidu Map	100 m
	Distance to coach station	$X_{13}$	Baidu Map	100 m
	Distance to bus station	$X_{14}$	Baidu Map	100 m
	Distance to ferry station	$X_{15}$	Baidu Map	100 m
Terrain conditions	Elevation	$X_{16}$	Geospatial Data Cloud <sup>c</sup>	100 m
	Slope	$X_{17}$	Geospatial Data Cloud	100 m

<sup>a</sup> Data of transport lines is provided by Open Street Map. <https://www.openstreetmap.org>

<sup>b</sup> Data of public infrastructures is provided by Baidu Map. <https://map.baidu.com>.

<sup>c</sup> Elevation is provided by Geospatial Data Cloud site, Computer Network Information Center, Chinese Academy of Sciences. <http://www.gscloud.cn>.



**Fig. 6.** Visualization of accessibility to transport lines in ArcGIS.

features, two variants of the proposed model (referred to as SOM-CA and CNN-CA) are also compared.

All these models are established through data from 2000 to 2015 for Shanghai. We split the experimental data into two sets: a training set from 2000 to 2010, and a test set composed of data from 2011 to 2015. A pixel-by-pixel comparison is used to derive a confusion matrix for later detailed evaluation.

#### 4.2. Evaluation metrics of simulation accuracy

Accuracy assessment is an indispensable step in the dynamic simulation of LUC. Confusion matrices are usually employed to derive measures of prediction accuracy, including overall accuracy (OA) and kappa coefficient, as shown in Eqs. (13)–(16) (Congalton et al., 1983).

$$OA = \frac{\sum_{i=1}^k n_{ii}}{\sum_{i=1}^k n_{ij}} \quad (13)$$

$$P_0 = \frac{\sum_{i=1}^k n_{ii}}{\sum_{i=1}^k n_{ij}} \quad (14)$$

$$P_e = \frac{\sum_{i=1}^k \left[ \left( \sum_{j=1}^k n_{ij} \right) \cdot \left( \sum_{j=1}^k n_{ji} \right) \right]}{\left( \sum_{i=1}^k \sum_{j=1}^k n_{ij} \right)^2} \quad (15)$$

$$Kappa = \frac{P_0 - P_e}{1 - P_e} \quad (16)$$

In Eqs. (13)–(16),  $k$  is the number of land use categories in this study, while  $n_{ii}$  denotes the number of cells that are simulated in category  $i$  and are actually in category  $i$  in the actual LU map; moreover,  $n_{ij}$  represents the number of cells that are simulated in category  $i$  but are actually in category  $j$  of the actual LU map, while  $n_{ji}$  denotes the number

of cells that are simulated in category  $j$  but are actually in category  $i$  in the actual LU map.  $P_0$  is the relative observed agreement among raters, and  $P_e$  is the hypothetical probability of chance agreement.

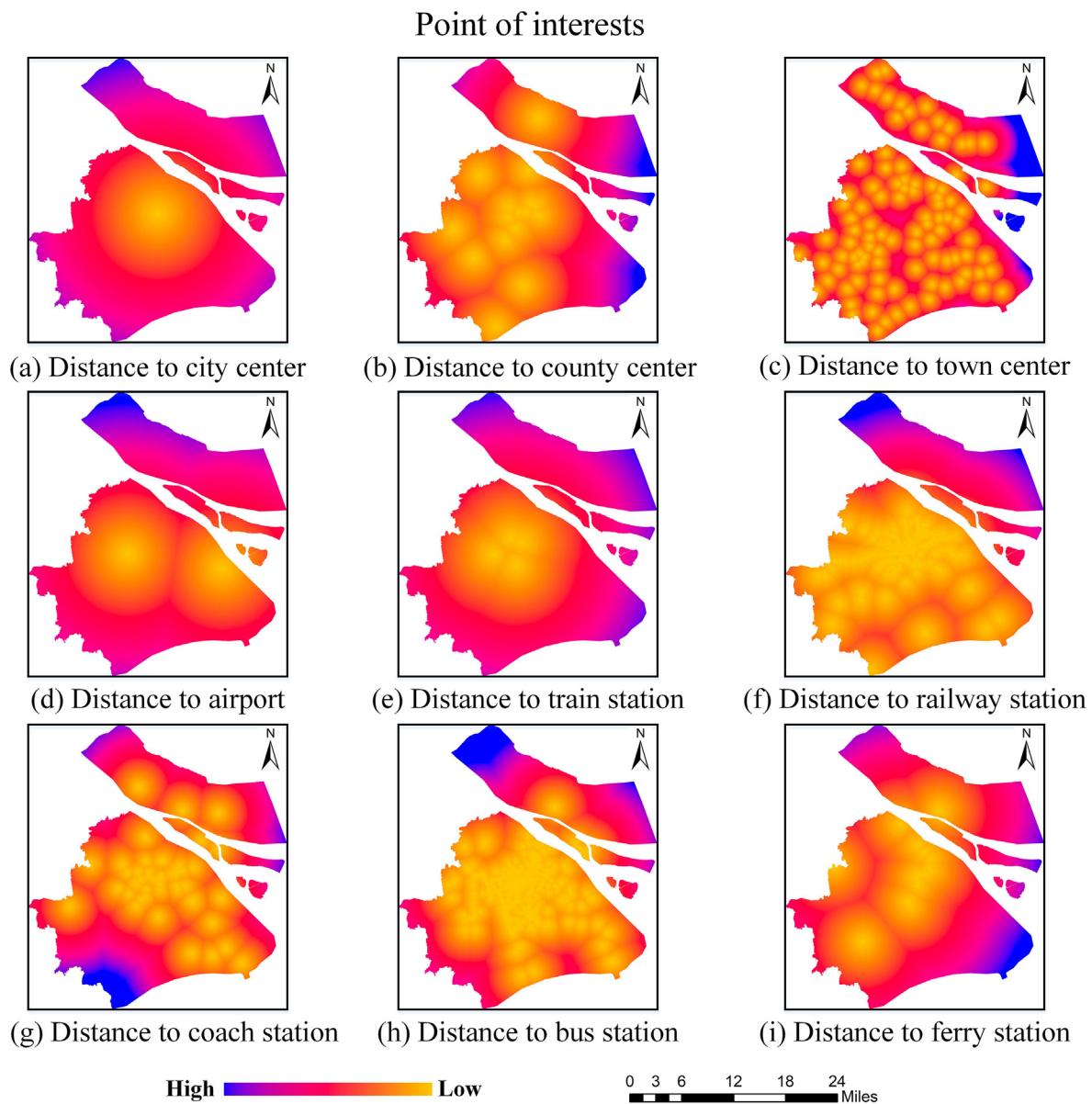
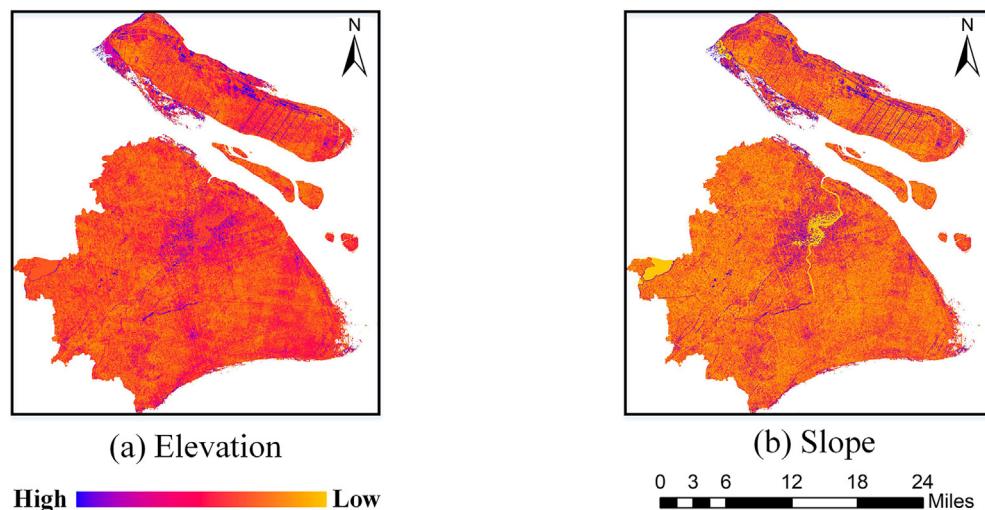
To facilitate a more comprehensive evaluation of the simulation results, a figure of merit ( $FoM$ ) was employed as shown in Eq. (17), which concentrates on the accuracy of changed areas rather than the whole research area (Pontius et al., 2008).

$$FoM = \frac{B}{A + B + C + D} \quad (17)$$

Here,  $A$  is the number of cells that were predicted to change in the simulation map, but where the cells remain unchanged in the actual LU map;  $B$  is the number of cells that changed in both the simulation map and the actual LU map and also changed correctly;  $C$  is the number of cells that changed in both simulation map and actual LU map, but changed incorrectly; finally,  $D$  is the number of cells that were predicted to remain unchanged in the simulation map, but that changed in the actual LU map.  $FoM$  ranges from 0 to 1, where a higher  $FoM$  demonstrates a higher simulation accuracy within changed areas.

#### 5. Experimental results and analysis

Three sets of comparative experiments were conducted in our study. In Section 5.1, we compare our methods with the baseline models from both quantitative and qualitative perspectives. Moreover, to validate the necessity of considering spatial heterogeneity, as well as to compare the performance of different partitioning strategies, an additional set of comparative experiments are conducted in Section 5.2. Finally, to elaborate on the importance of spatiotemporal neighborhood features and search for the optimal time steps in 3D CNN for each sub-region, a further set of comparative experiments are conducted in Section 5.3.

**Fig. 7.** Visualization of accessibility to public infrastructures in ArcGIS.**Fig. 8.** Visualization of terrain conditions in ArcGIS.

**Table 2**  
Comparison with conventional models.

Model	OA	Kappa	FoM
LR-CA	0.8537	0.7793	0.2092
SVM-CA	0.8604	0.7898	0.2219
RF-CA	0.8663	0.7999	0.2994
ANN-CA	0.8712	0.8072	0.3204
<b>SOM-CA</b>	<b>0.8916</b>	<b>0.8313</b>	<b>0.3271</b>
CNN-CA	0.9003	0.8507	0.3458
PST-CA*	0.9178	0.8774	0.4276

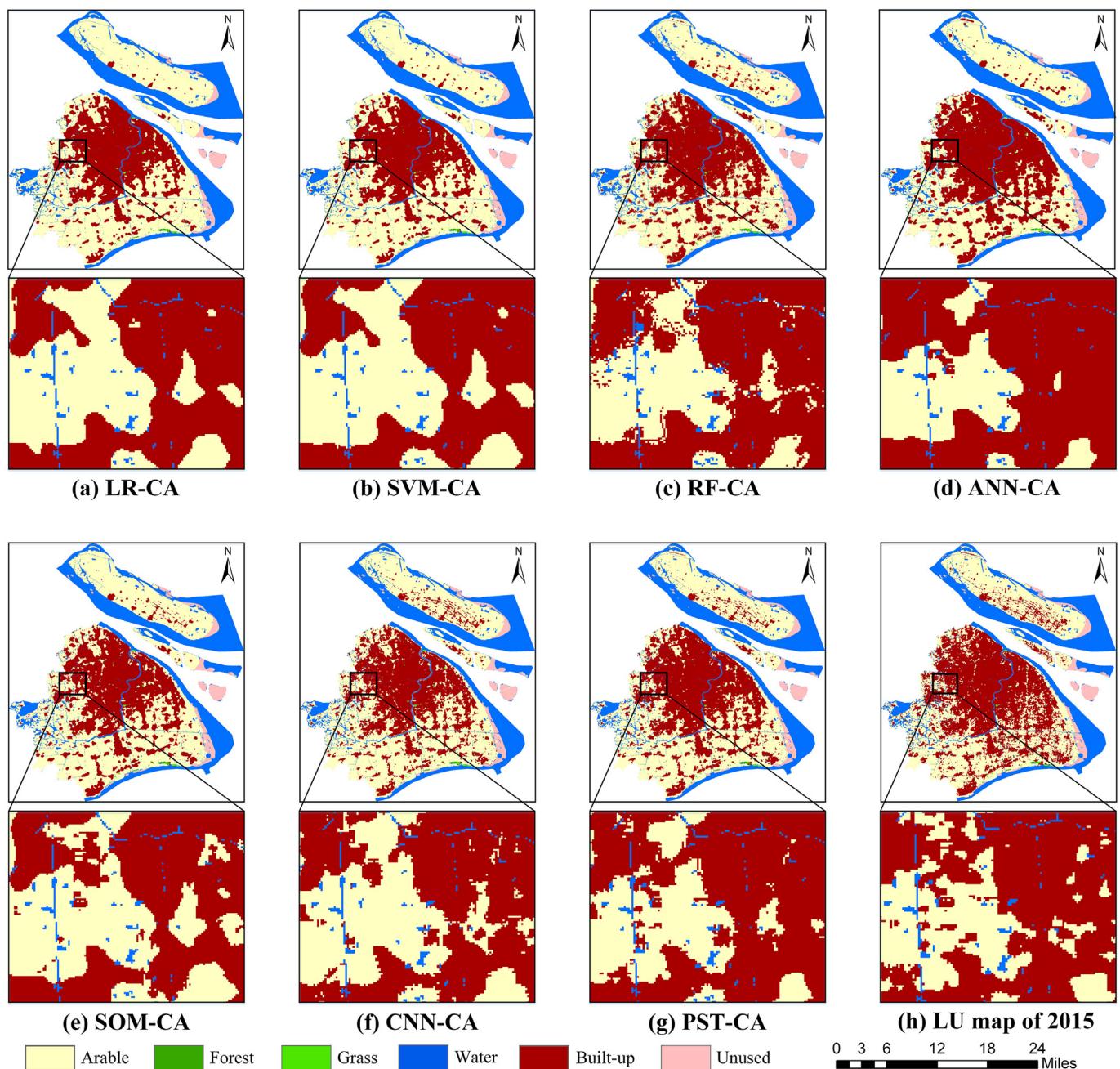
### 5.1. Comparison with baseline models

From the quantitative perspective, Table 2 illustrates the results of our proposed methods and the other baseline models, i.e. LR-CA, SVM-

CA, RF-CA, ANN-CA, SOM-CA, and CNN-CA. In the SOM-CA model, simulation was implemented along with ANN-CA using a SOM partitioning strategy. CNN-CA adopted a 3D CNN strategy along with ANN-CA, but without partitioning.

After analyzing the results in Table 2, we obtained the following findings:

- (1) LR-CA achieved the worst simulation OA at 85.37%, demonstrating its inferior capabilities in simulating a non-linear spatiotemporal LUC process relative to machine learning-based methods.
- (2) Compared to the four traditional CA models (LR-CA, SVM-CA, RF-CA, ANN-CA), SOM-CA achieved higher accuracy for all three metrics, indicating both that spatial heterogeneity across the entire study area should not be neglected and that accuracy can be improved by conducting area partitioning.



**Fig. 9.** Simulation results of baseline models and PST-CA in Shanghai city for 2015.

- (3) CNN-CA also performed better than the four traditional CA models on the three selected metrics. Moreover, CNN-CA achieved a bigger increase in accuracy than SOM-CA, which demonstrates that neighborhood effects are pivotal to LUC and that simulation performance can be improved by spatiotemporal neighborhood feature learning.
- (4) Taking both spatial heterogeneity and neighborhood spatiotemporal features into account, our proposed PST-CA outperformed all other baseline models, which verifies the effectiveness and superiority of our method.
- (5) Furthermore, the increase in *FoM* was greater than that of the other two metrics in our proposed PST-CA, which demonstrates that our method is capable of simulation in changed areas.

From the qualitative perspective, the simulation results of these models and the actual LU map of Shanghai city at 2015 are presented in Fig. 9. In general, all these models achieved acceptable simulation performance, and were visually in good agreement with the actual map. However, there are slight differences between the models; some examples are presented below, using typical built-up areas depicted in rectangular boxes.

Simulation results of: (a) LR-CA; (b) SVM-CA; (c) RF-CA; (d) ANN-CA; (e) SOM-CA; (f) CNN-CA.

As Fig. 9 demonstrates, these four traditional models (LR-CA, SVM-CA, RF-CA and ANN-CA) present a relatively compact pattern compared to the actual LU map. Moreover, the proportion of built-up areas according to ANN-CA is higher than that of the actual LU map, which may suggest the occurrence of over-fitting. The results of SOM-CA tend to be more dispersed than that of ANN-CA, but more compact than the actual LU map; furthermore, CNN-CA is visually consistent with the actual LU map in terms of spatial distribution, which validates that spatiotemporal neighborhood features can be extracted effectively via 3D CNN. In addition, PST-CA outperforms other baseline models, and the obtained results are the best intuitive match with the actual LU map.

## 5.2. Simulation with different partitioning strategies

K-means clustering and SOM are employed to evaluate the efficiency of spatial partitioning in this study. Fig. 10 illustrates the changes in the

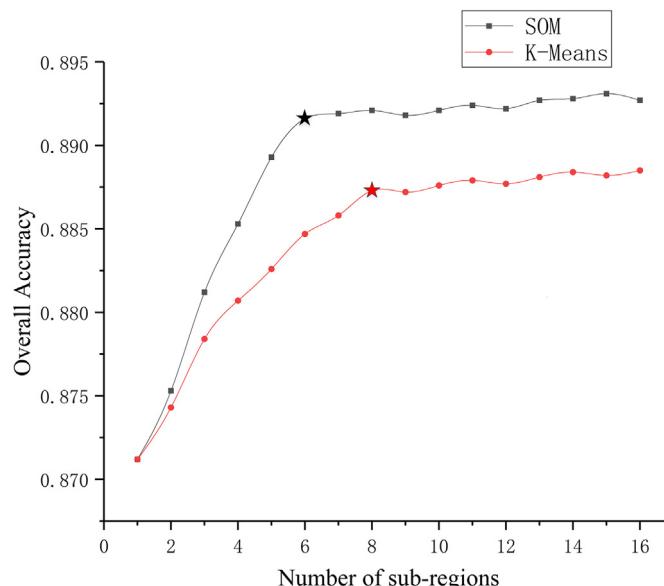


Fig. 10. Accuracy of *k*-means and SOM with different numbers of sub-regions.

trend of simulation accuracy with different numbers of partitions in *k*-means-CA and SOM-CA.

After analyzing the results in Fig. 10, we can conclude that:

- (1) As the number of partitions increased, both strategies achieved better performance, indicating that the accuracy of LUC dynamic simulation can be improved by partitioning the study area into sub-regions.
- (2) When the number of partitions reached a certain threshold value, the accuracy became stable, suggesting that the spatial heterogeneity may have been adequately addressed and the accuracy cannot be further improved to a significant degree.
- (3) Overall, SOM-CA performed better than *k*-means-CA at the same number of partitions; this may indicate that the sub-regions partitioned by SOM are more homogeneous than those partitioned using *k*-means.
- (4) The optimal partitioning number of SOM-CA is six, while this figure is eight for *k*-means-CA. At the same time, the accuracy of SOM-CA with six sub-regions is higher than *k*-means-CA with eight sub-regions, which validates the assumption that SOM is more suitable for LUC partitioning.

To further compare the performance of these two partitioning strategies, Fig. 11 depicts the optimal partitioning results (i.e. eight for *k*-means and six for SOM) and the administrative division map of Shanghai.

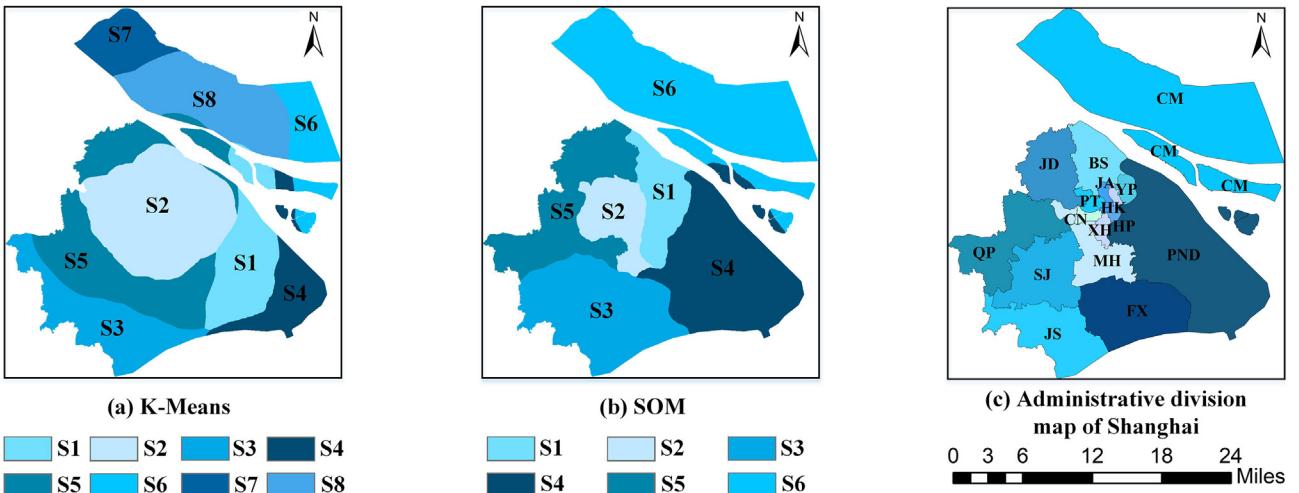
As shown in Fig. 11, the leftmost image is the partitioning result using *k*-means with eight sub-regions, which exhibits a circular pattern. The middle image is the partitioning result using SOM with six sub-regions, which are more explainable and visually similar to the administrative divisions.

To explain why the results of partitioning using SOM are more visually similar, we calculated the built-up proportion of each county in the year 2000 and the built-up growth rate of various counties between 2000 and 2015. Results are presented in Fig. 12.

From Fig. 12, it can be seen that the built-up area proportion in 2000 and the growth rate between 2000 and 2015 (shown in Fig. 12(a) and (b) respectively) differ substantially from county to county, which reflects that each county differs in terms of its level of economic development. Moreover, developed counties (i.e. HK, JA, etc.) that contained a large proportion of built-up areas in 2000 expanded slowly between 2000 and 2015, whereas counties with a small proportion of built-up areas (i.e. SJ, JD, etc.) grew much more rapidly in terms of built-up areas during the same period. We next analyzed the development status of each sub-region, including the counties it contains, built-up proportion rank in 2000 and growth rate rank between 2000 and 2015; the results are presented in Table 3. The built-up proportion in 2000 represented the level of urban development to some extent, and the growth rate between 2000 and 2015 was used to reflect the speed of urban development.

As Table 3 illustrates, HK, JA, HP, XH and YP have the slowest growth rate, with a highest proportion of built-up areas, and have been divided into S1 by SOM. At the same time, JA, XH, PT and CN are mostly divided into S2 and exhibit a growth rate that is still slow, but higher than that of S1. Both S1 and S2 belong to the central area of Shanghai, which can be characterized with high population density and developed economy. Thus, these two sub-regions (S1 & S2) could be categorized as areas with high coverage of existing urban lands and slow growth rate of urban lands, i.e. "High-Slow" composite patterns.

Moreover, JS and SJ are allocated to S3, which features a higher growth rate, and contains a built-up proportion of 2.24% and 2.32% respectively in 2000. FX district and PDN are divided into S4, with growth rate of 31.17% and 34.56% respectively. Furthermore, JD, MH, BS and QP are mostly segregated into S5, which exhibits the highest growth rate. From a geographical perspective, S5 is adjacent to Jiangsu province



**Fig. 11.** Partitioning results; (a) Partitioning result of *k*-means with eight sub-regions; (b) Partitioning result of SOM with six sub-regions; (c) Administrative division map of Shanghai.

and the central area of Shanghai, which are the cores of the Yangze River Delta urban agglomeration and have experienced a dramatic urban expansion. These three sub-regions (S3–S5) hold low coverage of existing urban lands and rapid growth of urban lands, i.e. “Low-Rapid” composite patterns.

Furthermore, CM is divided into S6 with 1.33% built-up areas and 5.64% growth rate. Limited by isolated geographical locations and inconvenient transport, the links between CM and Shanghai centers are weaker, which has resulted in a slow urbanization process. With low coverage of existing urban lands and slow growth of urban lands, CM shows obvious “Low-Slow” composite patterns.

From all the information shown by Figs. 11, 12 and Table 3, we can find that the partitioning results of SOM generally coincide with the actual development of Shanghai to some extent. There are three distinctly different composite patterns of the “coverage rate-growth rate” in the partitioning results, which verifies the ability to address spatial heterogeneity of SOM.

### 5.3. Convolution with different time steps

We explored the effectiveness of spatiotemporal neighborhood features using input data with different time steps. Table 4 illustrates the

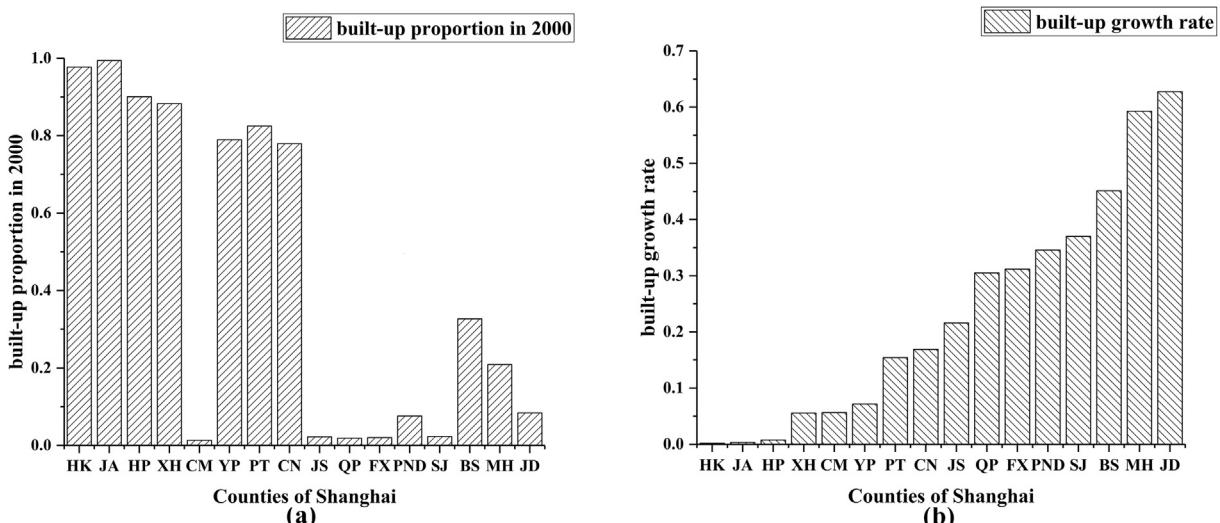
simulation performance of CNN-CA over the entire study area with different time steps, which indicate the length of historical LU maps used in convolutional operations. The model with time steps of 0 is ANN-CA, in which spatiotemporal neighborhood features were not captured. Moreover, the model with time steps of 1 means that spatiotemporal neighborhood features were captured with only one year of LU map, indicating convolutional operations equivalent to 2D CNN.

As Table 4 demonstrates, these three metrics of CNN-CA improve gradually as the time steps increases, which confirms the necessity of taking spatiotemporal neighborhood features into account. Moreover, CNN-CA performs best with time steps of 3. As the time steps reach a certain value, the simulation accuracy drops slightly, indicating that too long time steps may interfere with the LUC dynamic simulation.

For six sub-regions in our proposed PST-CA, Fig. 13 illustrates the relationship between overall accuracy and time steps in different sub-regions.

We observed the following findings in Fig. 13:

- (1) As time steps increased, the overall simulation accuracy continued to improve, indicating that long-term spatiotemporal neighborhood features were well captured and are also beneficial to LUC dynamic simulation.



**Fig. 12.** (a) Built-up proportion in each county at 2000; (b) Built-up growth rate in each county between 2000 and 2015.

**Table 3**  
The status quo of sub-region development.

Sub-regions	Contained counties	Built-up proportion rank in 2000	Growth rate rank between 2000 and 2015
S1	HK	2	16
	JA	1	15
	HP	3	14
	XH	4	13
	YP	6	11
S2	JA	1	15
	XH	4	13
	PT	5	10
	CN	7	9
S3	MH	9	2
	JS	13	8
S4	SJ	12	4
	FX	14	6
S5	PND	11	5
	JD	10	1
	MH	9	2
S6	BS	8	3
	QP	15	7
S6	CM	16	12

- (2) The optimal time steps differed between each sub-region, confirming that spatial heterogeneity indeed exists in LUC simulation.
- (3) When time steps reached a certain threshold value in each sub-region, the simulation accuracy became stable, suggesting that the spatiotemporal neighborhood features may have been well captured and the accuracy cannot be further improved with longer time steps.

Furthermore, we explored the relationship between optimal time steps and the growth rate of the built-up areas. Fig. 14 depicts the optimal time steps and built-up growth rate between 2000 and 2015 in each sub-region.

As Fig. 14 demonstrates, the optimal time steps for S2 and S5 are larger than those for other sub-regions. Moreover, the built-up growth rates in these two sub-regions are higher than the others. This demonstrates that the growth trends in these two sub-regions are consistent, which also indicates that more historical data may be required in order to fully capture the spatiotemporal neighborhood features in areas with rapid urban growth rates.

## 6. Discussion

### 6.1. Comparison with literature

Spatial heterogeneity is a significant existence in LUC process simulation. Different parts of a given research area may have various driving forces and variable growth rates during the LUC. Treating the research area as a whole and establishing a unified framework for transition rule extraction will incur either low simulation accuracy or high

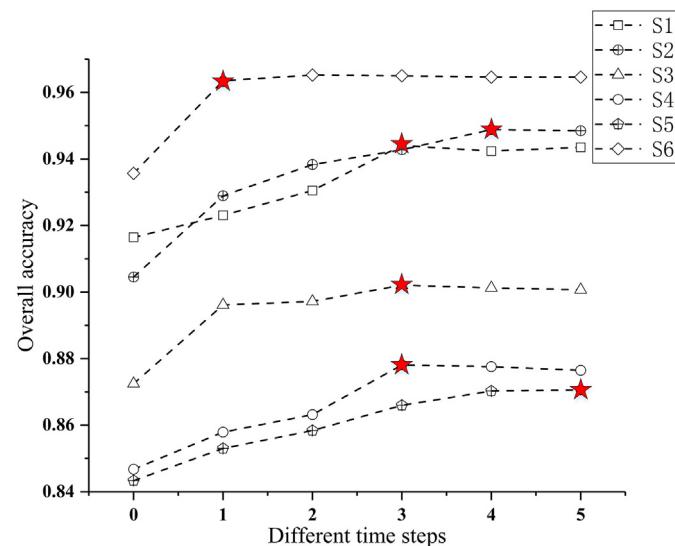


Fig. 13. Relationship between overall accuracy and time steps in different sub-regions.

model complexity. For current researches to deal with spatial heterogeneity problem in LUC, partitioned transition thresholds (Xia et al., 2019) suffer from the blind selection of appropriate thresholds and contained significant uncertainties in the conversion probability map. By contrast, SOM can divide the entire area into relatively homogeneous sub-regions through unsupervised learning, and no explicit partitioning thresholds are needed. The *k*-means clustering algorithm (Ke et al., 2015; Yang et al., 2018) treats each LUC pixel as discrete unit and neglect spatial proximity between them. This will destroy the topological properties of entire study area. It also suffers from the random determination of initial cluster centers, which come from the nature of *k*-means algorithm. Nevertheless, the topological properties of entire study area are well preserved with SOM, in which a neighborhood distance constraint is imposed on the high-to-low dimension projection.

Neighborhood effects are also an important factor for transition rules extraction and conversion probability calculation in LUC simulation (Liao et al., 2016). Recent researchers have utilized numerous statistics learning-based methods for transition rules extraction, including logistic regression, support vector machine, the multilayer perceptron, and random forest, etc. Deep learning-based methods, 2D CNN (He et al.,

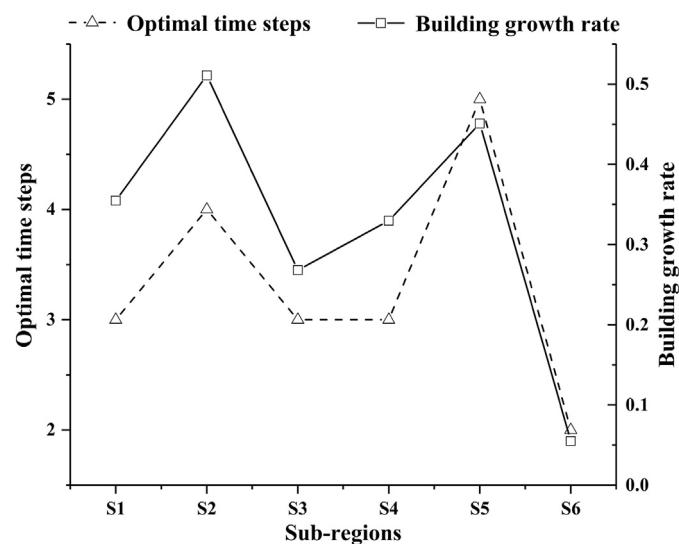


Fig. 14. Relationship between optimal time steps and the growth rate of the built-up areas.

**Table 4**  
Performance of CNN-CA with different time steps.

Time steps	Accuracy	Kappa	FoM
0	0.8712	0.8072	0.3204
1	0.8825	0.8238	0.3248
2	0.8916	0.8382	0.3256
<b>3</b>	<b>0.9003</b>	<b>0.8507</b>	<b>0.3271</b>
4	0.8992	0.8489	0.3265
5	0.8962	0.8443	0.3261

2018), are also employed to capture spatial neighbor interaction. All of these traditional CA-based LUC models are essentially based on the Markov hypothesis, and assume that the state of a given cell in next time step is only relevant in terms of the state of last time step. As a dynamic geographical process, LUC exhibits obvious long-term temporal dependency (Aburas et al., 2017), neglecting of the long-term temporal dependency will significantly reduce the simulation accuracy. However, LUC was regarded as a geographical phenomenon with historical dependent information in our 3D CNN; both spatial and temporal neighborhood features were well captured through spatiotemporal convolutions on time-series LU maps. The simulation results also demonstrate that CA model with 3D CNN can significantly increase the overall accuracy and outperforms all these traditional CA models, i.e. LR-CA, SVM-CA, RF-CA, ANN-CA, etc. This also further verified the importance of modeling the neighborhood effects in LUC.

## 6.2. Limitations and prospective future work

While the PST-CA model proposed in this study has achieved good performance in the simulation of LUC in Shanghai, there are some shortcomings of the present research due to the limitations of the data acquisition and research methods employed. Firstly, data expressed at a coarse resolution may contain information and patterns that are undetectable at finer resolutions (Gounaris et al., 2019; van Delden et al., 2011). We extracted the spatiotemporal neighborhood features using input data with the spatial resolution of 100 m, which cannot be applied at LUC simulations with other different resolutions. Secondly, driving factors are assumed to be temporally stationary in CA based LUC simulations (Feng et al., 2019), and the partitioning and simulation are implemented based on these temporally stationary driving factors. Therefore, the proposed model cannot support the projection of LUC scenarios over longer periods of time, since the driving factors may change substantially during these longer periods, which limited the scope of the method to some extent.

To make our proposed method more practical, there are still several aspects that should be improved:

- 1) Adopting more driving factors such as social-economic variables. Land use change is closely related to social and economic development; thus, more driving factors, including Gross Domestic Product (GDP) and population, should be adopted to improve simulation performance.
- 2) Analyzing the importance of different driving factors. LUC may be determined by some key factors or by combinations of different drivers. Analyzing the importance of different driving factors could thus enable simulation performance to be more accurate.
- 3) Exploring the LUC processes under distinct economic performance scenarios. LUC trends may differ under distinct socioeconomic environments. Thus, simulations that consider different scenarios will make the proposed method more convincing and robust.

## 7. Conclusion

The PST-CA model proposed in this study incorporates SOM, 3D CNN, and ANN to simulate the LUC processes. Area partitioning are acquired using the SOM model, while a 3D CNN model is employed to capture spatiotemporal neighborhood features. By integrating spatiotemporal neighborhood features, driving factors, stochastic factors and constraints, an ANN model is established to generate the probability map of each sub-region, after which the dynamic LUC at Shanghai city from 2000 to 2015 can be simulated and validated.

Compared to these four traditional models (LR-CA, SVM-CA, RF-CA and ANN-CA), the proposed model achieves increased simulation accuracy by 4.66%~6.41%, which demonstrates the effectiveness of taking spatial heterogeneity and spatiotemporal neighborhood features into account. Moreover, we find that: (1) the partitioning results of SOM exhibit three distinctly different “coverage rate-growth rate” patterns

(“High-Slow”, “Low-Rapid” and “Low-Slow”), which verifies SOM’s capability to address spatial heterogeneity; (2) there is a positive correlation between the optimal time steps in 3D CNN and the built-up growth rate of each sub-region to some extent, which indicates longer temporal dependency should be captured for rapidly developing areas. Our study can provide guides according to different sub-regions from spatiotemporal perspective and eventually help decision making and city planning.

## CRediT authorship contribution statement

**Yuehui Qian:** Conceptualization, Data curation, Methodology, Software, Formal analysis, Writing - original draft. **Weiran Xing:** Methodology, Formal analysis, Software. **Xuefeng Guan:** Conceptualization, Methodology, Formal analysis, Writing - review & editing, Funding acquisition. **Tingting Yang:** Methodology, Formal analysis. **Huayi Wu:** Funding acquisition, Project administration, Resources.

## Acknowledgments

The authors are grateful to the editor and reviewers for their careful and valuable suggestions. This work is supported by the National Natural Science Foundation of China program (Grant No.: 41971348) and the National Key Research and Development Program of China (Grant No.: 2017YFB0503802).

## Funding

The National Natural Science Foundation of China program (Grant No.: 41971348). The National Key Research and Development Program of China (Grant No.: 2017YFB0503802).

## Declaration of competing interest

No conflict of interest exists in the submission of this manuscript. The manuscript is approved by all authors for publication. I would like to declare on behalf of my co-authors that the work described is original research that has not been published previously, and not under consideration for publication elsewhere, in whole or in part. All the authors listed have approved the manuscript that is enclosed.

## References

- Hughes Jr, W.T., Turnbull, G.K., 1996. Uncertain neighborhood effects and restrictive covenants. *J. Urban Econ.* 39 (2), 160–172.
- Aburas, M.M., Ho, Y.M., Ramli, M.F., Ash'aari, Z.H., 2017. Improving the capability of an integrated CA-Markov model to simulate spatio-temporal urban growth trends using an analytical hierarchy process and frequency ratio. *International Journal of Applied Earth Observation Geoinformation* 59, 65–78.
- Aburas, M.M., Ahamad, M.S.S., Omar, N.Q., 2019. Spatio-temporal simulation and prediction of land-use change using conventional and machine learning models: a review. *Environ. Monit. Assess.* 191 (4), 205.
- Azadi, H., Barati, A.A., Rafiaani, P., Taheri, F., Gebrehiwot, K., Witlox, F., et al., 2017. Evolution of land use-change modeling: routes of different schools of knowledge. *Landscape Ecol.* Eng. 13 (2), 319–332.
- Bureau SS, 2016. *Shanghai Statistical Yearbook 2015*. China Statistics Press, Beijing, pp. 5–59.
- Cao, M., Tang, Ga, Shen, Q., Wang, Y., 2015. A new discovery of transition rules for cellular automata by using cuckoo search algorithm. *Int. J. Geogr. Inf. Sci.* 29 (5), 806–824.
- Chang-Martínez, L.A., Mas, J.-F., Valle, N.T., Torres, P.S.U., Folan, W.J., 2015. Modeling historical land cover and land use: a review from contemporary modeling. *ISPRS Int. J. Geo Inf.* 4 (4), 1791–1812.
- Chen, Y., Liu, X., Li, X., Liu, X., Yao, Y., Hu, G., et al., 2017. Delineating urban functional areas with building-level social media data: a dynamic time warping (DTW) distance based k-medoids method. *Landscape Urban Plan.* 160, 48–60.
- Chen, Y., Li, X., Liu, X., Huang, H., Ma, S., 2018. Simulating urban growth boundaries using a patch-based cellular automaton with economic and ecological constraints. *Int. J. Geogr. Inf. Sci.* 33 (1), 55–80.
- Chen, N., Chen, L., Ma, Y., Chen, A., 2019. Regional disaster risk assessment of China based on self-organizing map: clustering, visualization and ranking. *International Journal of Disaster Risk Reduction* 33, 196–206.
- Cheshire, P.C., Hay, D.G., 2017. Problems of urban decline and growth: a review. *Urban Problems in Western Europe: An Economic Analysis.* 3, pp. 28–41.

- Congalton, R.G., Oderwald, R.G., Mead, R.A., 1983. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogramm. Eng. Remote Sens.* 49 (12), 1671–1678.
- van Delden, H., van Vliet, J., Rutledge, D.T., Kirkby, M.J., 2011. Comparison of scale and scaling issues in integrated land-use models for policy support. *Agric. Ecosyst. Environ.* 142 (1–2), 18–28.
- Feng, Y., Tong, X., 2018. Dynamic land use change simulation using cellular automata with spatially nonstationary transition rules. *GIScience & Remote Sensing* 55 (5), 678–698.
- Feng, Y., Liu, Y., Tong, X., 2018. Comparison of metaheuristic cellular automata models: a case study of dynamic land use simulation in the Yangtze River Delta. *Comput. Environ. Urban. Syst.* 70, 138–150.
- Feng, Y., Wang, R., Tong, X., Shafizadeh-Moghadam, H., 2019. How much can temporally stationary factors explain cellular automata-based simulations of past and future urban growth? *Comput. Environ. Urban. Syst.* 76, 150–162.
- Fiener, P., Auerswald, K., Van Oost, K., 2011. Spatio-temporal patterns in land use and management affecting surface runoff response of agricultural catchments—a review. *Earth Sci. Rev.* 106 (1–2), 92–104.
- Gong, P., Li, X., Zhang, W., 2019. 40-year (1978–2017) human settlement changes in China reflected by impervious surfaces from satellite remote sensing. *Science Bulletin* 64 (11), 756–763.
- Gounaris, D., Chorianopoulos, I., Koukoulas, S., 2018. Exploring prospective urban growth trends under different economic outlooks and land-use planning scenarios: the case of Athens. *Appl. Geogr.* 90, 134–144.
- Gounaris, D., Chorianopoulos, I., Symeonakis, E., Koukoulas, S., 2019. A random forest-cellular automata modelling approach to explore future land use/cover change in Attica (Greece), under different socio-economic realities and scales. *Sci. Total Environ.* 646, 320–335.
- Groeneveld, J., Müller, B., Buchmann, C.M., Dressler, G., Guo, C., Hase, N., et al., 2017. Theoretical foundations of human decision-making in agent-based land use models—a review. *Environ. Model Softw.* 87, 39–48.
- Haase, D., Haase, A., Kabisch, N., Kabisch, S., Rink, D., 2012. Actors and factors in land-use simulation: the challenge of urban shrinkage. *Environ. Model Softw.* 35, 92–103.
- He, J., Li, X., Yao, Y., Hong, Y., Jinbao, Z., 2018. Mining transition rules of cellular automata for simulating urban expansion by using the deep learning techniques. *Int. J. Geogr. Inf. Sci.* 32 (10), 2076–2097.
- Huang, F., Yin, K., Huang, J., Gui, L., Wang, P., 2017. Landslide susceptibility mapping based on self-organizing-map network and extreme learning machine. *Eng. Geol.* 223, 11–22.
- Islam, K., Rahman, M.F., Jashimuddin, M., 2018. Modeling land use change using cellular automata and artificial neural network: the case of Chunati Wildlife Sanctuary, Bangladesh. *Ecol. Indic.* 88, 439–453.
- Ke, X., Qi, L., Zeng, C., 2015. A partitioned and asynchronous cellular automata model for urban growth simulation. *Int. J. Geogr. Inf. Sci.* 30 (4), 637–659.
- Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. *Biol. Cybern.* 43 (1), 59–69.
- Liao, J., Tang, L., Shao, G., Su, X., Chen, D., Xu, T., 2016. Incorporation of extended neighborhood mechanisms and its impact on urban land-use cellular automata simulations. *Environ. Model Softw.* 75, 163–175.
- Liu, Y., Feng, Y., 2016. Simulating the impact of economic and environmental strategies on future urban growth scenarios in Ningbo, China. *Sustainability* 8 (10), 1045–1061.
- Liu, X., Liang, X., Li, X., Xu, X., Ou, J., Chen, Y., et al., 2017. A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landscape Urban Plan.* 168, 94–116.
- Mustafa, A., Heppenstall, A., Omrani, H., Saadi, I., Cools, M., Teller, J., 2018. Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic algorithm. *Comput. Environ. Urban. Syst.* 67, 147–156.
- Napoleon, D., Pavalkodi, S., 2011. A new method for dimensionality reduction using k-means clustering algorithm for high dimensional data set. *International Journal of Computer Applications* 13 (7), 41–46.
- Nguyen, T.T., Kawamura, A., Tong, T.N., Nakagawa, N., Amaguchi, H., Gilbuena, R., 2015. Clustering spatio–seasonal hydrogeochemical data using self-organizing maps for groundwater quality assessment in the Red River Delta, Vietnam. *J. Hydrol.* 522, 661–673.
- Noszczyk, T., 2018. A review of approaches to land use changes modeling. *Human and Ecological Risk Assessment: An International Journal* 25 (6), 1377–1405.
- Parker, D.C., Manson, S.M., Janssen, M.A., Hoffmann, M.J., Deadman, P., 2003. Multi-agent systems for the simulation of land-use and land-cover change: a review. *Ann. Assoc. Am. Geogr.* 93 (2), 314–337.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al., 2011. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12 (Oct), 2825–2830.
- Pontius, R.G., Boersma, W., Castella, J.-C., Clarke, K., de Nijs, T., Dietzel, C., et al., 2008. Comparing the input, output, and validation maps for several models of land change. *Ann. Reg. Sci.* 42 (1), 11–37.
- Qiang, Y., Lam, N.S., Assessment, 2015. Modeling land use and land cover changes in a vulnerable coastal region using artificial neural networks and cellular automata. *Environmental monitoring* 187 (3), 57.
- Ren, Y., Lü, Y., Comber, A., Fu, B., Harris, P., Wu, L., 2019. Spatially explicit simulation of land use/land cover changes: current coverage and future prospects. *Earth Sci. Rev.* 190, 398–415.
- Rey, S.J., Anselin, L., 2010. PySAL: a Python library of spatial analytical methods. *Handbook of Applied Spatial Analysis*. Springer, Berlin, Heidelberg, pp. 175–193.
- Rienow, A., Goetze, R., 2015. Supporting SLEUTH-enhancing a cellular automaton with support vector machines for urban growth modeling. *Comput. Environ. Urban. Syst.* 49, 66–81.
- Roodposhti, M.S., Aryal, J., Bryan, B.A., 2019. A novel algorithm for calculating transition potential in cellular automata models of land-use/cover change. *Environ. Model. Softw.* 112, 70–81.
- Sacha, D., Kraus, M., Bernard, J., Behrisch, M., Schreck, T., Asano, Y., et al., 2018. SOMFlow: guided exploratory cluster analysis with self-organizing maps and analytic provenance. *IEEE Trans. Vis. Comput. Graph.* 24 (1), 120–130.
- Sidharthan, R., Bhat, C.R., 2012. Incorporating spatial dynamics and temporal dependency in land use change models. *Geogr. Anal.* 44 (4), 321–349.
- Sun, Z., Meng, Q., Zhai, W., 2018. An improved boosting learning saliency method for built-up areas extraction in Sentinel-2 images. *Remote Sens.* 10 (12), 1863–1888.
- Takamatsu, M., Kawasaki, A., Rogers, P.P., Malakie, J.L., 2014. Development of a land-use forecast tool for future water resources assessment: case study for the Mekong River 3S sub-basins. *Sustain. Sci.* 9 (2), 157–172.
- Walcott, S.M., Pannell, C.W., 2006. Metropolitan spatial dynamics: Shanghai. *Habitat International* 30 (2), 199–211.
- White, R., Engelen, G., Uljee, I., 1997. The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. *Environment and Planning B: Urban Analytics and City Science* 24 (3), 323–343.
- Wolfram, S., 1984. Cellular automata as models of complexity. *Nature* 311 (5985), 419–424.
- Xia, C., Wang, H., Zhang, A., Zhang, W., 2017. A high-performance cellular automata model for urban simulation based on vectorization and parallel computing technology. *Int. J. Geogr. Inf. Sci.* 32 (2), 399–424.
- Xia, C., Zhang, A., Wang, H., Zhang, B., 2019. Modeling urban growth in a metropolitan area based on bidirectional flows, an improved gravitational field model, and partitioned cellular automata. *Int. J. Geogr. Inf. Sci.* 33 (5), 877–899.
- Yang, J., Guo, A., Li, Y., Zhang, Y., Li, X., 2018. Simulation of landscape spatial layout evolution in rural-urban fringe areas: a case study of Ganjingzi District. *GIScience & Remote Sensing* 56 (3), 388–405.
- You, H., 2016. Quantifying megacity growth in response to economic transition: a case of Shanghai, China. *Habitat International* 53, 115–122.
- Zhou, B., Kockelman, K.M., 2008. Neighborhood impacts on land use change: a multinomial logit model of spatial relationships. *Ann. Reg. Sci.* 42 (2), 321–340.
- Zou, L., Zheng, J., Miao, C., McKeown, M.J., Wang, Z.J., 2017. 3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI. *IEEE Access* 5, 23626–23636.