Computers & Geosciences 137 (2020) 104430

Contents lists available at ScienceDirect

Computers and Geosciences

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

Research paper

A novel cellular automata model integrated with deep learning for dynamic spatio-temporal land use change simulation

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

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

ARTICLE INFO

Keywords:
Land use change
Spatio-temporal modeling
Deep learning
Model integration

ABSTRACT

Land use change (LUC) exhibits obvious spatio-temporal dependency. Previous cellular automata (CA)-based methods usually treated the LUC dynamics as Markov processes and proposed a series of CA-Markov models, which however, were intrinsically unable to capture the long-term temporal dependency. Meanwhile, such models used only numerical proportion of neighboring land use (LU) types to represent neighborhood effects of LUC, which inevitably neglected the complicated spatial heterogeneity and thus caused inaccurate simulation results. To address these problems, this paper presents a novel CA model integrated with deep learning (DL) techniques to model spatio-temporal LUC dynamics. Our DL-CA model firstly uses a convolutional neural network to capture latent spatial features for complete representation of neighborhood effects. A recurrent neural network then extracts historical information of LUC from time-series land use maps. A random forest is appended as binary change predictor to avoid the imbalanced sample problem during model training.

Land use data collected from 2000 to 2014 of the Dongguan City, China were used to verify our proposed DL-CA model. The input data from 2000 to 2009 were used for model training, the 2010 data for model validation, and the data collected from 2011 to 2014 were used for model evaluation. In addition, four traditional CA models of multilayer perceptron (MLP)-CA, support vector machine (SVM)-CA, logistic regression (LR)-CA and random forest (RF)-CA were also developed for accuracy comparisons. The simulation results demonstrate that the proposed DL-CA model accurately captures long-term spatio-temporal dependency for more accurate LUC prediction results. The DL-CA model raised prediction accuracy by 9.3%–11.67% in 2011–2014 in contrast to traditional CA models.

1. Introduction

Land use change (LUC) such as urbanization, deforestation, and desertification have increased dramatically in recent decades, transforming landscapes globally (Fu and Weng, 2016; Lyu et al., 2019). Rapid urban expansion and chaotic LUC has increased the stress on the environment, not only accelerating global warming but also causing pervasive and irreversible biological diversity losses (Lee, 2019; Shafizadeh-Moghadam et al., 2019). Therefore, it is of great necessity to precisely model the LUC process and deeply understand the spatial-temporal LUC patterns for proper land management and better urban planning (Wagner and Waske, 2016).

The last three decades have witnessed remarkable achievements in the application of cellular automata (CA) models to urban development and LUC simulation (Shafizadeh-Moghadam, 2019; Wang and Li, 2011; Yang et al., 2008). As a bottom-up dynamic modeling approach, CA was originally proposed by Ulam and Von Neumann in the 1940s and has been widely used to simulate the evolution of complex non-linear systems (Roodposhti et al., 2019; Von Neumann and Burks, 1966; Wagner and Fohrer, 2019). For a CA model, five essential components need to be defined, including cell space, cell states, neighborhood, time steps and transition rules (Feng et al., 2019a). Space can be represented as a grid of discrete cells. Each cell holds an initial state from a predefined finite set of states and evolves along discrete time steps. The neighbors of a given cell is defined as a collection of cells based on proximity. The transition rules are established to capture the spatial dependency of local interactions between cells. These rules will be applied on each cell at a regular discrete time steps, and calculate the new state for a given cell by taking into account its previous state and the surrounding cells in the defined neighborhood. For LUC-CA models, precise transition rules are a key determinant to the prediction capability of a model.

Researchers have proposed numerous CA-based models for LUC

* Corresponding author. Tel.: +86 27 68778969.
E-mail addresses: xingweiran@whu.edu.cn (W. Xing), guanxuefeng@whu.edu.cn (X. Guan).

https://doi.org/10.1016/j.cageo.2020.104430
Received 15 August 2019; Received in revised form 29 January 2020; Accepted 31 January 2020
Available online 8 February 2020

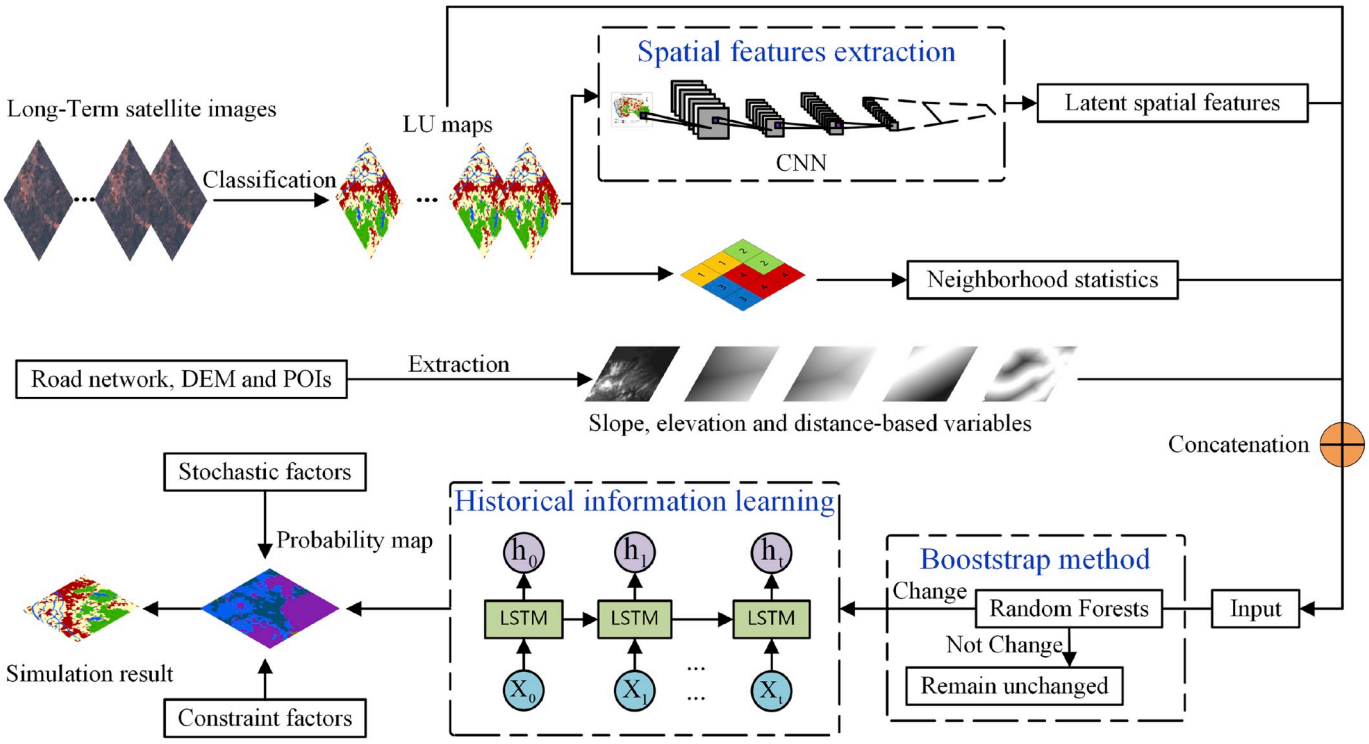

Fig. 1. The proposed deep-learning CA model framework.

simulations, including the multilayer perceptron (MLP)-CA, support vector machine (SVM)-CA, logistic regression (LR)-CA and random forest (RF)-CA models (Karimi et al., 2019; Shafizadeh-Moghadam et al., 2017b). Nevertheless, 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 (Gounaridis et al., 2019; Grekousis, 2019; Guan et al., 2019; Karimi et al., 2019). However, since LUC is a long-term process, long-term temporal dependency is ignored and leads to inaccurate predictions. Until now, the long-term temporal dependency issue of LUC modeling has been rarely addressed in previous studies.

At the same time, many researchers have also adopted a collection of suitable driving factors to express the LUC process in the real world when building precise CA-based LUC models (Liang et al., 2018; Zhang et al., 2019). However, the inadequate representation problem of neighborhood effects has still not received much attention (Du et al., 2018; Rienow and Goetzke, 2015; Shafizadeh-Moghadam et al., 2017a, 2017c). Currently neighborhood effects are only expressed by numerical proportion of LU types, e.g. the cell number of built-up area in the neighbors of the target cell (Liu et al., 2018; Mustafa et al., 2018; Newland et al., 2018). However, even the same proportion of LU types may have different type distribution, so the numerical type proportion is insufficient to handle the complex spatial heterogeneity and will affect the accuracy of prediction results. Therefore, these latent spatial features are important and should be considered when representing neighborhood effects.

Most recently, deep learning methods have been developed and applied in the LUC modeling. Recurrent neural network (RNN) is characterized for its recurrent mechanism, which can transfer the historical information between time steps. Jia et al. (2017) and Wang et al. (2019) both use long short-term memory (LSTM) with long time series images to effectively learn long-term temporal dependency and make precise land cover classification. Convolutional neural network (CNN) is another class of deep neural network, which commonly uses convolutional filters to extract implicit features of local regions. He et al. (2018) considers spatial effects and utilizes CNN to extract spatial features of driving factors for LUC simulation, but fails to learn the temporal dependency existed in LUC. Although limitations exist, all these researches had shown the great power of deep learning methods in LU domain.

Furthermore, LUC is a long-term process and the LU types in most areas remain unchanged over short periods, which will cause the enormous amount of unchanged areas. In the domain of machine learning, imbalanced samples will cause the over-fitting problem and imprecise model implementation (Buda et al., 2018). Nevertheless, the problem of imbalanced samples has been rarely mentioned in the existing LUC studies, which should be handled before simulations.

Now two key questions can be raised here as: (1) how to accurately learn the long-term spatial temporal dependency in the CA models; (2) how to handle the imbalanced sampling problem during the model training in the machine-learning-based CA models. To tackle these two problems, this paper proposes a deep learning-based CA model (DL-CA) to simulate the LUC dynamics. In this DL-CA model, both RNN and CNN are integrated to derive the complex transition probability on which the LUC-CA model relies. Since LUC process exhibits typical spatial-temporal dependency, the LSTM model is capable of modeling time series with long time dependency, while the CNN model extracts local latent spatial features. Furthermore, Random Forest is employed as a bootstrap method to address the imbalanced sampling problem.

To evaluate the performance of our proposed DL-CA model, fifteen years of LU data from Dongguan City were collected from Landsat-7 satellite images. Four traditional MLP-CA, SVM-CA, LR-CA, RF-CA and other three RNN models of LSTM-CA, LSTM-RF-CA, LSTM–CNN–CA were developed for comparisons. From the experimental results, spatio-temporal long-term dependency is efficiently learned by our proposed DL-CA model and highest overall accuracy is obtained in contrast to the traditional CA models.

The remainder of this paper is organized as follows. Section 2 describes our DL-CA model framework, its components and evaluation metrics. Section 3 describes the study region and lists the data used in this study. Section 4 provides the experimental results, verifies the effect of learning the spatio-temporal dependency, compares the improvements of each component, analyzes the results of different step

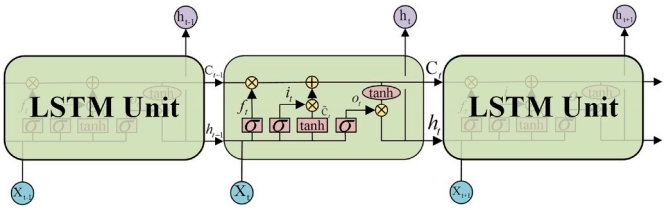

Fig. 2. The illustration of LSTM structure.

prediction. Finally, Section 5 gives the conclusions that summarize the research of this paper and provides the future direction.

2. Methodology

Fig. 1 presents our proposed DL-CA model framework. In our DL-CA framework, the initial LU maps are used to calculate the numerical proportion between different LU types and extract latent spatial features of neighborhood with CNN. After that, these initial LU maps, the latent spatial features, neighborhood proportion, slope, elevation and distance-based variables are concatenated as the input of RNN model, i.e. LSTM (long short-term memory). The RNN model assimilates historical land use information to make future predictions. A random forest (RF) is appended before LSTM as binary change prediction to filter changed/unchanged samples during model training. By stacking RF/CNN/RNN, this hybrid deep learning method can model the complex spatio-temporal dependency and derive the required transition probabilities for the underlying CA model and conduct the further LU simulations.

2.1. Deep learning models for transition probability calculation

2.1.1. The long short-term memory model for historical information assimilation

As a typical variant of RNN, LSTM consists of discrete units in the recurrent hidden layer, which contains units with self-connections to store the temporal state of the network. As shown in Fig. 2, each unit in the LSTM contains an input gate, an output gate and a forget gate (Gers et al., 1999). The input gate controls the flow of input activations into the unit, and the output gate controls the output flow of unit activations into the rest of the network. The forget gate is used to scale the internal state of the unit before adding it as input to the unit, therefore adaptively forgetting or resetting the unit's memory.

A LSTM network computes a mapping from an input sequence $X = (X_1, ..., X_t)$ to an output sequence $h = (h_1, ..., h_t)$ by calculating the network unit activations using the following equations iteratively from Equations (1)–(7):

$$h_t = f(h_{t-1}, X_t) \tag{1}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{4}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{5}$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = O_t \cdot \tanh(C_t) \tag{7}$$

where W denotes the weight matrices, C_t is the cell state and b is the input bias vectors. The function f can be simply defined as an element-wise logistic sigmoid function.

In our proposed DL-CA model, the LSTM layers were used to extract temporal features from both historical LU data and driving factors. Historical LU data, and driving factors were stacked to construct a temporal input vector $X = (X_0, X_1, X_2, ..., X_{14})$ for the LSTM layers, and the temporal information learned from long time dependency was automatically extracted layer-by-layer. In this vector, which contains one cell state and 14 variables, X_0 represents the initial LU type; X_1~X_2 are used to represent the neighborhood effects that X_1 represent four LU types statistics of neighborhood and X_2 represent the latent spatial features extracted by CNN; X_3~X_4 represent the elevation and slope, X_5~X_{14} represent the Euclidean distance-based variables. Both X_1 and X_2 have 4 dimensions, each of X_3~X_{14} has one dimension, and the total dimension of X is 20.

2.1.2. Spatial features extraction using convolutional neural network

Illustrated by Fig. 3, convolutional neural network (CNN) is further developed to extract the spatial features from the LU map. The CNN model includes six layers, including two convolution layers, two max-pooling layers, one fully connected layer and one softmax layer. The activation function adopted in the convolution layers and the full-connection layer is the Rectified Linear Units (ReLU).

According to the distance-decay effect of CA models, the spatial features extraction by CNN can be restricted in the local extent with a given radius (Liao et al., 2014; Shafizadeh-Moghadam et al., 2017c). Thus, following empirical research results, the input size is defined as 25 × 25 (i.e. the radius is 1 km with an 80 m spatial resolution). As shown at the far left of Fig. 3, each pixel of the LU maps is chosen as a central pixel, and the center pixel and its neighbor pixels are clipped to form a 25 × 25 grid.

Our CNN adopts a 3 × 3 convolution kernel with a 2 × 2 pooling layer, and the first convolutional layer consists of 32 convolution kernels of 3 × 3, a feature image of 25 × 25 × 32 is exported. The second layer is composed of a pooling layer of 2 × 2 and output 13 × 13 × 32 feature images. As the third layer establishes 64 convolution kernels of 3 × 3, input data will be converted to 13 × 13 × 64. The fourth layer is also a pooling layer of 2 × 2, where a dimensional feature image of 7 × 7 × 64 is obtained after pooling the input data. Finally, the fully connected layer with 4 neurons using softmax regression was adopted to produce the final output, i.e. the captured spatial features of neighborhood.

2.1.3. A bootstrap method with random forest for imbalanced samples

Since the LUC is a long-term process, most of LU types will not

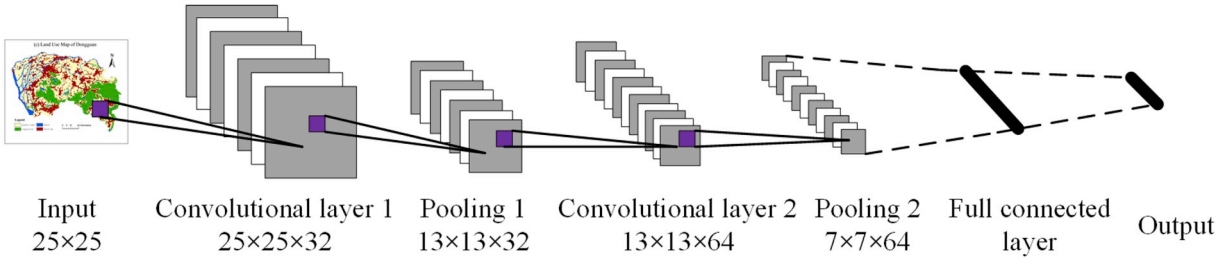

Fig. 3. The CNN structure for spatial feature extraction, including two convolutional layers, two pooling layers and one full connected layer.

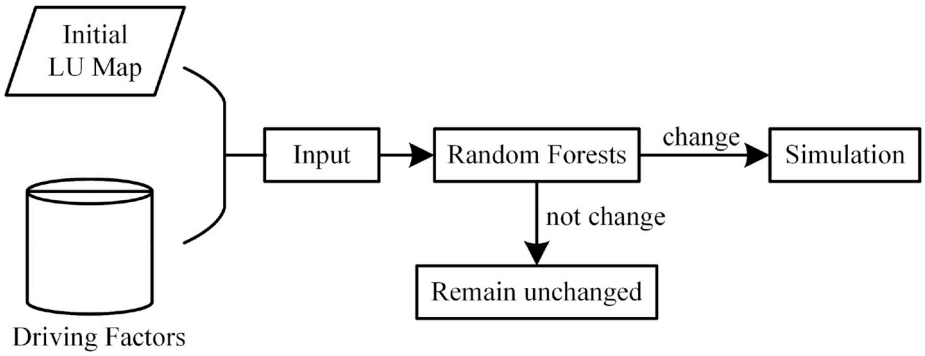

Fig. 4. A bootstrap method of random forest.

change over a short period, causing an imbalanced distribution of land use types in the data samples used to train the model and thus causing the over-fitting problem and further decreasing LUC prediction accuracy. To avoid this problem, random forest (RF) method was selected as a bootstrap method before simulation. The RF algorithm is a decision tree-based ensemble classifier that consists of multiple decision trees, and the final probability is obtained by combining the results of each decision tree. As a popular machine learning method, RF hardly causes over-fitting problem by increasing the number of decision trees. Therefore, we utilized RF as a bootstrap method to avoid imbalanced samples and the process of RF can be visualized in Fig. 4.

As shown in Fig. 4, initial land use maps and driving factors are processed as the input for random forest (RF). The RF will output a probability p for type change, which ranges from zero to one (i.e., $1-p$ for *not change*). With this derived probability p, a current cell remains unchanged if the probability of *change* is lower than that of *not change* (i.e., $p < 0.5$); if higher, the current cell will be treated as type change and the model continues to predict which type it will change to with the following LSTM models.

2.2. Deep learning model integration with cellular automata

For each cell at each time step, the LSTM model continuously yields probability predictions of LU categories (natural-water, arable land, vegetation and built-up area), which are produced by a softmax function layer. The obtained transition probabilities of each pixel will be imported to CA for LUC simulation and prediction (Batty et al., 1999; Clarke and Gaydos, 1998; Li and Yeh, 2002).

The total transition probability of one CA model can be determined by Equation (8):

$$P = S \times C \times \Omega \times \alpha \tag{8}$$

where P is the development probability, S denotes the suitability of considered factors, C is the constraint of specific land use, Ω represents the impact of neighborhood configuration, and α is the stochastic perturbation reflecting the effect of random factors.

The transition probability CP exported by the proposed DL model combines S with Ω, can represent both the suitability of driving factors and neighborhood effects. Equation (8) can then be transformed as:

$$P = CP \times C \times \alpha \tag{9}$$

In Equation (9), the CP will be obtained by the LSTM component of the DL-CA model, the C will be determined by the natural characteristics of LU and policies and the α is set randomly. After calculating the value P, there are no other explicit transition rules required in the DL-CA model.

2.3. Model evaluation

In order to estimate the LUC prediction, we use overall accuracy, macro F1-score and figure of merit (FoM) to evaluate the prediction capability of the proposed CA models (Pontius and Millones, 2011). These metrics and FoM are expressed as these following equations:

$$overall\ \ accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$F1 - score = 2 \times \frac{\frac{TP}{TP+FP} \times \frac{TP}{TP+FN}}{\frac{TP}{TP+FP} + \frac{TP}{TP+FN}} \tag{11}$$

where TP denotes true positive, FP denotes false positive, TN denotes

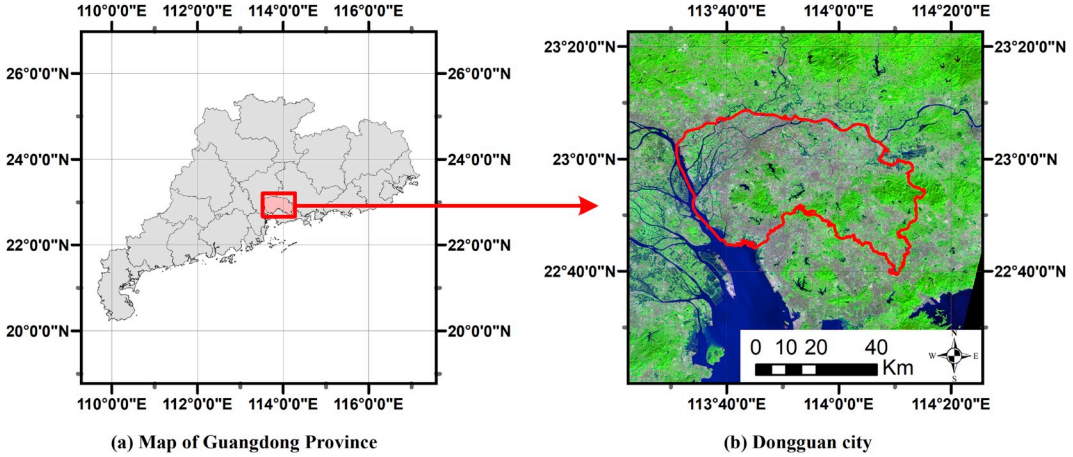

Fig. 5. Location of study area; (a) Administrative division map of Guangdong; (b) Classification land use map in Dongguan city.

Table 1
Datasets used in this study.

Data Type	Data Source	Time Span	Spatial Resolution	Purpose
Points	Administration	2000–2014	1:1000	Factors driving LUC as distance measurements between cells
Polylines	Road networks	2000–2014	1:1000	
Imagery	Landsat-7 (SLC-on)	2000–2003	30 m	Land use map
Imagery	Landsat-7 (SLC-off)	2003–2014	30 m	
DEM	Administration	2000–2014	10 m	Land surface slope

true negative, *FN* denotes false negative. Compared with overall accuracy, macro F1-score measures the classification performance of positive label, which enables specific evaluation of the model capability to predict the transitions from one LU type to another.

$$FoM = \frac{B}{A + B + C + D} \tag{12}$$

Furthermore, figure of merit (FoM) was also selected for model evaluation methods due to its focus on the change between observed and predicted. Shown in Equation (12), *A* is the number of error cells that the observed change is predicted as persistence, *B* is the number of correct cells that the observed change is predicted as change, *C* is the number of error cells that the observed change is predicted as wrong gaining category, and *D* is the number of error cells that the observed persistence is predicted as change. Higher value of FoM indicates a higher cell-level agreement (Tong and Feng, 2019).

3. Study area and datasets

The study area is chosen in the Dongguan city, which is located in the central Guangdong Province of South China (Fig. 5). As an important industrial city in the Pearl River Delta, Dongguan borders the provincial capital Guangzhou to the north, Huizhou to the northeast, Shenzhen to the south, and the Pearl River to the west. The region covers latitudes 22°39′-23°09′ N and longitudes 113°31′-114°15′ E, with an area of 2465 km^2.

Rapid urban expansion occurred in Dongguan since 2000 due to fast economic development. Many new urban lands emerged near the town centers due to transportation improvements and government policies. Communication between towns became more convenient and frequent with the development of express highways and railroads. Many rural areas along highways and major roads have been transformed into new urban lands.

All datasets used in this study are listed in Table 1. In addition to Landsat imagery, factors that influence LUC are also collected and analyzed, including POIs, road network, DEM, etc. (Wu et al., 2019). The DEM data were used to calculate land slope and elevation for later land development and suitability analyses. The road networks and POIs were used to obtain the Euclidean distance-based variables. For consistence, all the input data were resampled to 80 m spatial resolution to construct the target CA models.

3.1. Time-series of land use maps

Fifteen Landsat-7 Enhanced Thematic Mapper Plus (ETM+) images from 2000 to 2014 were used to derive time-series land use map for Dongguan City. To compensate for the failure of the Landsat-7 ETM + scanline corrector (SLC), the data gaps of the Landsat-7 ETM + images were repaired using the gap-filling algorithm (Scaramuzza and Barsi,

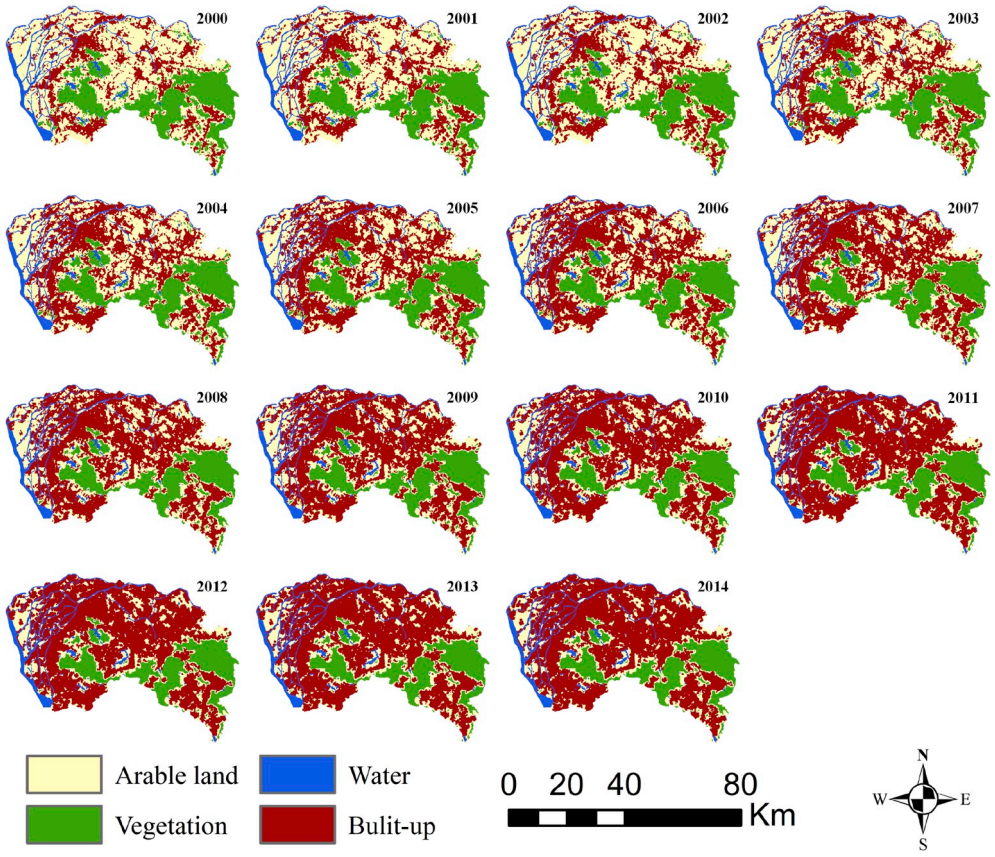

Fig. 6. Classified land use maps of Dongguan from 2000 to 2014.

Table 2
Descriptions of driving factors.

Name	Variable	Descriptions
LU proportion	X_1	Numerical proportion between different LU types of neighborhood cells collected by a 25 × 25 grid
Implicit spatial features	X_2	Implicit spatial features of neighborhood cells and extracted from LU maps
Elevation	X_3	Elevation from terrain data
Slope	X_4	slope from terrain data
Distance to river	X_5	Euclidean distance to rivers
Distance to railway	X_6	Euclidean distance to railways
Distance to highway	X_7	Euclidean distance to highways
Distance to first class road	X_8	Euclidean distance to first class roads
Distance to minor road	X_9	Euclidean distance to minor roads
Distance to city road	X_{10}	Euclidean distance to city roads
Distance to railway station	X_{11}	Euclidean distance to railway station
Distance to bus station	X_{12}	Euclidean distance to bus station
Distance to main POIs	X_{13}	Euclidean distance to main POIs including administration, hospital and university
Distance to central city	X_{14}	Euclidean distance to central city

2005). Supervised classification was carried out to derive time-series LU map from the ETM + images using the ERDAS Imagine software. The land use data was classified into four categories as arable land, vegetation, water and built-up. Fig. 6 shows the land use maps derived for the years from 2000 to 2014.

3.2. Driving factors

Physical factors, accessibility and neighborhood characteristics are commonly selected as driving factors in the previous studies (Chen et al., 2019; Feng and Tong, 2018; Feng et al., 2019b). Therefore, illustrated in Table 2, fourteen driving factors were used to build the LUC model in our study. The derived elevation and slope denote the land surface condition, while the other factors can be formulated as distance-based variables, i.e. using proximity effect (Karimi et al., 2019; Rienow and Goetzke, 2015). To represent accessibility, we used road network and main POIs to calculate the proximity effect and obtained ten distance-based variables for our established CA model. Elevation (X_3), slope (X_4) and ten distance-based variables (X_5~X_{14}) were visualized with ArcGIS, as shown in Fig. 7. In addition, the LU proportion (X_1) and latent spatial features (X_2) were extracted from the neighborhood cells in land use maps to express the neighborhood characteristics.

4. Experimental results and discussions

In the implementation of DL-CA model, a dropout layer was added for model regularization to prevent the LSTM component from overfitting; and cross-entropy was used as the loss function. The RMSprop algorithm was selected for LSTM optimization; and the Batch-Normalization was chosen to improve the model performance. All these models were written in Python with Tensorflow as the backend. Other related python libraries are also used in the models, including Pysal, Numpy, Pandas, Scipy, Arcpy (ESRI Inc.), and Scikit-learn (Abadi et al., 2016; Pedregosa et al., 2011; Rey and Anselin, 2010).

In addition to our proposed DL-CA model, four traditional models of MLP-CA, SVM-CA, LR-CA, RF-CA and other three different RNN models (LSTM-CA, RF-LSTM-CA, CNN-LSTM-CA) were developed for comparative purposes. Four traditional models with LSTM-CA were compared to evaluate the effect of long-term temporal dependence in LUC dynamic simulations. To assess the improvement of balanced samples preprocessing and spatial features extraction using CNN, we conduct further ablation studies to compare four different RNN models separately, LSTM-CA, LSTM-RF-CA, LSTM–CNN–CA and DL-CA. Compared with LSTM-CA, the LSTM-RF-CA model adds random forest to solved the issue of imbalanced samples during model training, the LSTM–CNN–CA model employs CNN to extract latent spatial features, and the DL-CA

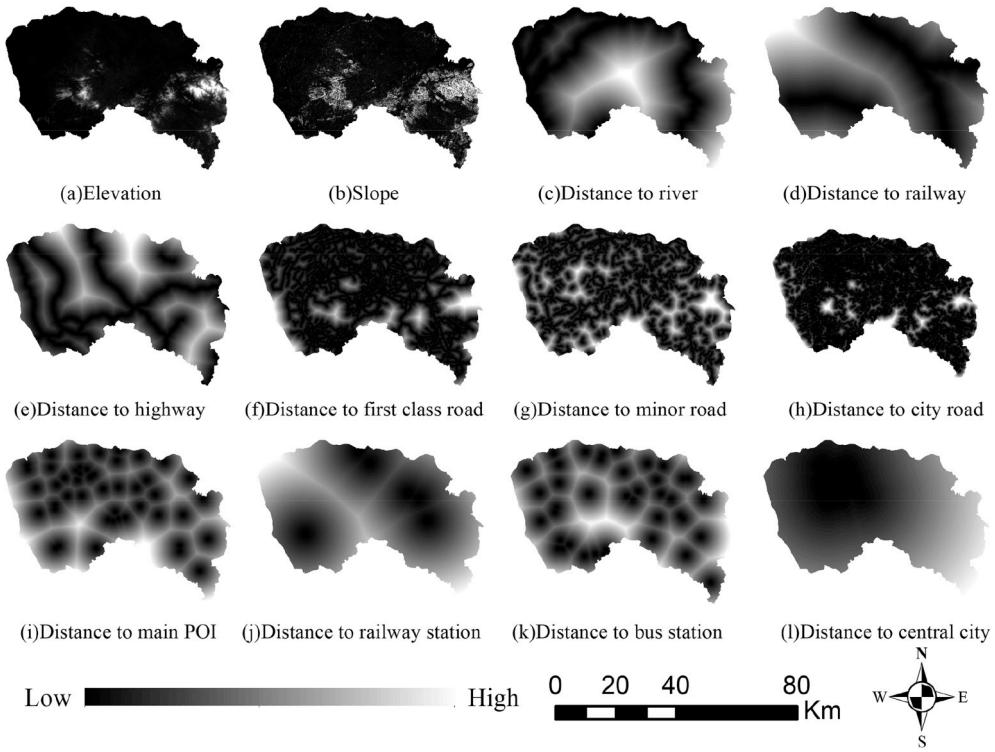

Fig. 7. Visualization of elevation, slope, and distance-based variables.

Table 3
Prediction results of different models.

Model	Overall accuracy	Macro F1-score	FoM
RF-CA	0.8584	0.8685	0.1763
SVM-CA	0.8703	0.8762	0.1727
LR-CA	0.8722	0.8745	0.1944
MLP-CA	0.8770	0.8823	0.2401
LSTM-CA	**0.9394**	**0.9491**	**0.2701**
DL-CA*	**0.9586**	**0.9626**	**0.4075**

model integrates both random forest and CNN.

For experimental data, we split the spatio-temporal data into three sets: a training set from 2000 to 2009; a validation set in 2010; and a test set including data collected from 2011 to 2014. To avoid the influence of spatial autocorrelation, only 50 percent of the origin dataset (years 2000–2009) are randomly sampled to form the training data.

4.1. Model comparisons

Table 3 lists the average prediction results of different traditional CA models (RF-CA, SVM-CA, LR-CA, MLP-CA) and two RNN models (LSTM-CA, DL-CA) on test dataset (2011–2014).

Table 3 demonstrate that LSTM-CA and DL-CA models significantly improves the overall accuracy of prediction compared with these four traditional models, i.e. about 0.06–0.1 increase in overall accuracy. These two RNN models outperform four traditional models by a great margin, which verifies the importance of long-term temporal dependency. From Table 3, we can also figure out that the DL-CA model achieves the highest overall accuracy of prediction than LSTM-CA, which shows the excellent power of our proposed DL-CA model in LUC simulation, and validates the contribution of both CNN extraction and RF bootstrap.

As shown in Fig. 8, we randomly choose an area to illustrate the difference in the simulation results among all the implemented CA models. From Fig. 8, the simulation results generated by our proposed DL-CA model can better match the actual LU map in contrast to the others. The DL-CA model, shows its great superiority in long-term spatio-temporal dependency capturing and imbalanced sampling problem handling, using RNN/CNN and RF components.

4.2. Ablation studies

In order to verify the effectivity of each component of the DL-CA model, we conduct a group of ablation studies with four models (LSTM-CA, LSTM–CNN–CA, LSTM-RF-CA DL-CA), and the results are shown in Table 4.

In these studies, we select LSTM-CA as the base model, and separately add two auxiliary components (RF/CNN) to evaluate the model improvement. First of all, LSTM-RF-CA model uses RF as a bootstrap method to solve the imbalanced data sampling and improves the overall accuracy from 0.9394 to 0.9447. Meanwhile, we add CNN component to handle the spatial heterogeneities and extract latent spatial features from cell neighbors, and we find that the LSTM–CNN–CA model improves the overall accuracy to 0.9466. Furthermore, we add both RF and CNN to construct our DL-CA model and the results demonstrate that the incorporation of RF and CNN gradually improve the prediction overall accuracy of our proposed DL-CA model from 0.9394 to 0.9586. This means both CNN and RF components are beneficial for LUC model building. By comparing LSTM–CNN–CA and LSTM-RF-CA models, we find that the improvement of CNN component is higher than RF, and the reason can be explained that the insufficient spatial representation of neighborhood is more serious than the problem of imbalanced samples.

Table 4
Model comparisons composed by different components.

Model	Overall accuracy	Macro F1-score	FoM
LSTM-CA (base)	0.9394	0.9491	0.2701
LSTM-RF-CA	0.9447	0.9511	0.3163
LSTM–CNN–CA	0.9466	0.9528	0.3211
DL-CA*	**0.9586**	**0.9626**	**0.4075**

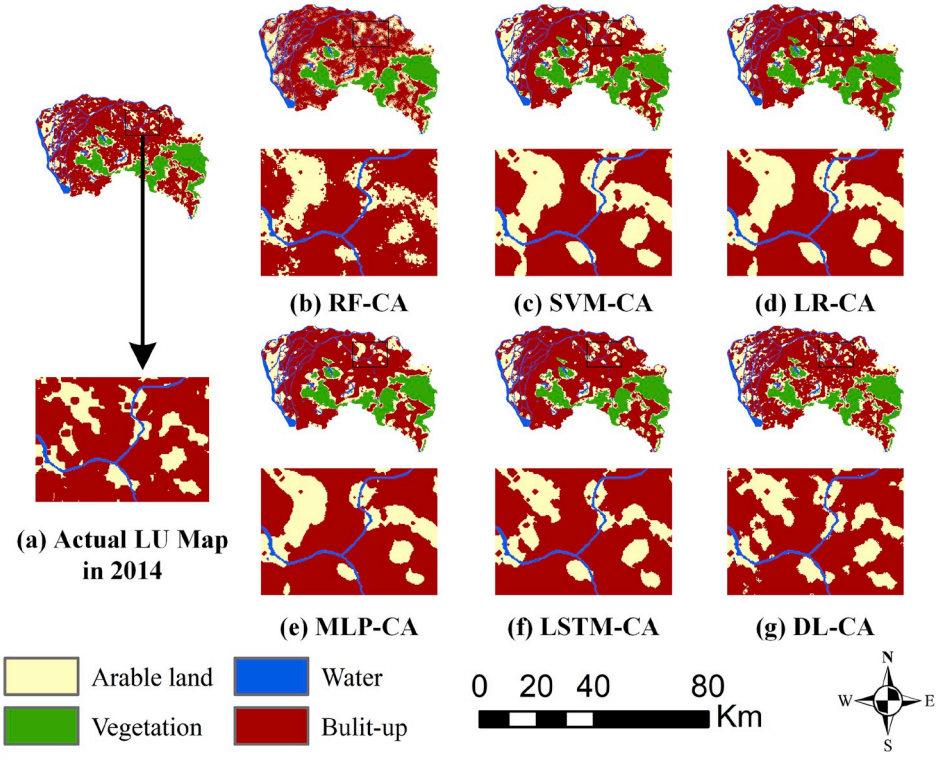

Fig. 8. Comparison of simulated results for 2014 generated by several models in study area. (a) actual LU map in 2014, (b)–(g) simulation of different models.

Table 5
Evaluation metrics of multi-step prediction.

Model	Years	Overall accuracy	Macro F1-score	FoM
DL-CA	**2011**	**0.9677**	**0.9730**	**0.4420**
	2012	0.9636	0.9688	0.4198
	2013	0.9551	0.9577	0.3929
	2014	0.9479	0.9508	0.3755

4.3. Explore the performance of multi-step prediction

To explore the performance of multi-step prediction, we use our proposed DL-CA model to simulate the LUC in Dongguan city for the years from 2011 to 2014, and the results are shown in Table 5.

Table 5 illustrates that one-step prediction is most accurate and the performance of our proposed DL-CA model decreases gradually when conduct multi-step prediction (i.e., 2012, 2013 and 2014 respectively). This can be explained that multi-step prediction cannot work well without the nearest temporal information, i.e., the temporal information of 2011 and 2012 were not provided when make the prediction of 2013.

5. Conclusions

This paper proposes a novel cellular automata model integrated with deep learning methods to effectively learn the spatio-temporal dependency and precisely simulate the LUC dynamics. The main contributions of this paper can be summarized as follows:

1) A novel cellular automata model, namely deep learning (DL)-CA model is proposed here. In this DL-CA model, long-term temporal dependency is learned by LSTM; to achieve efficient representation of neighborhood effects, CNN is built to capture the latent spatial features of neighborhood. By integration of RNN/CNN, our DL-CA model can effectively learn the spatio-temporal dependency.
2) In order to avoid the imbalanced sampling problem during model training, we employ random forest (RF) as a bootstrap method to conduct binary change prediction before the model prediction. The imbalanced sampling problem is successfully handled by RF from the experimental results.
3) Dongguan city, China was selected as a case study to verify the performance of our proposed DL-CA model. Experimental results demonstrate that DL-CA model achieved the highest LUC prediction accuracy in contrast to traditional models, with an overall accuracy of 0.9586, a macro F1-score of 0.9626 and FoM index of 0.4075.

Limitations also exist in the DL-CA model as the latent emergent structures and information about structural land use evolution is influenced by the resolution of datasets, and different driving factors have different effects on LUC. Future work should use high-resolution datasets, take more driving factors such as land-use policy, population density, and spatial, geophysical and socioeconomic factors into account and analyze the effects of different factors.

Declaration of competing interest

No conflict of interest exits 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.

Acknowledgements

This work is supported by the National Key Research and Development Program of China (Grant No.: 2017YFB0503802).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cageo.2020.104430.

Authorship statement

Weiran Xing host the study, carried out the experiments and wrote the draft. Yuehui Qian developed the main Python code. Tingting Yang processed the raw data. Xuefeng Guan and Huayi Wu helped to design some comparisons and discussed and revised the manuscripts.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., 2016. Tensorflow: a System for Large-Scale Machine Learning, Operating Systems Design and Implementation (Savannah, GA, US).
Batty, M., Xie, Y., Sun, Z., 1999. Modeling urban dynamics through GIS-based cellular automata. Comput. Environ. Urban Syst. 23 (3), 205–233.
Buda, M., Maki, A., Mazurowski, M.A., 2018. A systematic study of the class imbalance problem in convolutional neural networks. Neural Network. 106, 249–259.
Chen, Y., Li, X., Liu, X., Huang, H., Ma, S., 2019. Simulating urban growth boundaries using a patch-based cellular automaton with economic and ecological constraints. Int. J. Geogr. Inf. Sci. 33 (1), 55–80.
Clarke, K.C., Gaydos, L.J., 1998. Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. Int. J. Geogr. Inf. Sci. 12 (7), 699–714.
Du, G., Shin, K.J., Yuan, L., Managi, S., 2018. A comparative approach to modelling multiple urban land use changes using tree-based methods and cellular automata: the case of Greater Tokyo Area. Int. J. Geogr. Inf. Sci. 32 (4), 757–782.
Feng, Y., Tong, X., 2018. Dynamic land use change simulation using cellular automata with spatially nonstationary transition rules. GIsci. Remote Sens. 55 (5), 678–698.
Feng, Y., Wang, J., Tong, X., Shafizadeh-Moghadam, H., Cai, Z., Chen, S., Lei, Z., Gao, C., 2019a. Urban expansion simulation and scenario prediction using cellular automata: comparison between individual and multiple influencing factors. Environ. Monit. Assess. 191 (5).
Feng, Y., Wang, R., Tong, X., Shafizadeh-Moghadam, H., 2019b. How much can temporally stationary factors explain cellular automata-based simulations of past and future urban growth? Comput. Environ. Urban Syst. 76, 150–162.
Fu, P., Weng, Q., 2016. A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with Landsat imagery. Remote Sens. Environ. 175, 205–214.
Gers, F.A., Schmidhuber, J., Cummins, F., 1999. Learning to forget: continual prediction with LSTM. In: 9th International Conference on Artificial Neural Networks, vol. 99. ICANN, Edinburgh, UK.
Gounaridis, 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.
Grekousis, G., 2019. Artificial neural networks and deep learning in urban geography: a systematic review and meta-analysis. Comput. Environ. Urban Syst. 74, 244–256.
Guan, D., Zhao, Z., Tan, J., 2019. Dynamic simulation of land use change based on logistic-CA-Markov and WLC-CA-Markov models: a case study in three gorges reservoir area of Chongqing, China. Environ. Sci. Pollut. Res. Int. 26 (20), 20669–20688.
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.
Jia, X., Khandelwal, A., Nayak, G., Gerber, J., Carlson, K., West, P., Kumar, V., 2017. Incremental dual-memory LSTM in land cover prediction. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, pp. 867–876.
Karimi, F., Sultana, S., Babakan, A., Suthaharan, S., 2019. An enhanced support vector machine model for urban expansion prediction. Comput. Environ. Urban Syst. 75, 61–75.
Lee, C., 2019. Impacts of urban form on air quality in metropolitan areas in the United States. Comput. Environ. Urban. Syst. 77, 101362.
Li, X., Yeh, A.G.-O., 2002. Neural-network-based cellular automata for simulating multiple land use changes using GIS. Int. J. Geogr. Inf. Sci. 16 (4), 323–343.
Liang, X., Liu, X., Li, D., Zhao, H., Chen, G., 2018. Urban growth simulation by incorporating planning policies into a CA-based future land-use simulation model. Int. J. Geogr. Inf. Sci. 32 (11), 2294–2316.
Liao, J., Tang, L., Shao, G., Qiu, Q., Wang, C., Zheng, S., Su, X., 2014. A neighbor decay cellular automata approach for simulating urban expansion based on particle swarm intelligence. Int. J. Geogr. Inf. Sci. 28 (4), 720–738.
Liu, X., Hu, G., Ai, B., Li, X., Tian, G., Chen, Y., Li, S., 2018. Simulating urban dynamics in China using a gradient cellular automata model based on S-shaped curve evolution characteristics. Int. J. Geogr. Inf. Sci. 32 (1), 73–101.
Lyu, R., Clarke, K.C., Zhang, J., Jia, X., Feng, J., Li, J., 2019. The impact of urbanization and climate change on ecosystem services: a case study of the city belt along the Yellow River in Ningxia, China. Comput. Environ. Urban Syst. 77, 101351.

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.
Newland, C.P., Zecchin, A.C., Maiera, H.R., Newman, J.P., Delden, H.v., 2018. Empirically derived method and software for semi-automatic calibration of Cellular Automata land-use models. Environ. Model. Software 108, 208–239.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12 (Oct), 2825–2830.
Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int. J. Rem. Sens. 32 (15), 4407–4429.
Rey, S.J., Anselin, L., 2010. PySAL: a Python library of spatial analytical methods. In: Fischer, M.M., Getis, A. (Eds.), Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 175–193.
Rienow, A., Goetzke, 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. Software 112, 70–81.
Scaramuzza, P., Barsi, J., 2005. Landsat 7 scan line corrector-off gap-filled product development. In: Proceeding of Pecora, Sioux Falls, South Dakota, US.
Shafizadeh-Moghadam, H., 2019. Improving spatial accuracy of urban growth simulation models using ensemble forecasting approaches. Comput. Environ. Urban 76, 91–100.
Shafizadeh-Moghadam, H., Asghari, A., Taleai, M., Helbich, M., Tayyebi, A., 2017a. Sensitivity analysis and accuracy assessment of the land transformation model using cellular automata. GIsci. Remote Sens. 54 (5), 639–656.
Shafizadeh-Moghadam, H., Asghari, A., Tayyebi, A., Taleai, M., 2017b. Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth. Comput. Environ. Urban Syst. 64, 297–308.
Shafizadeh-Moghadam, H., Minaei, M., Feng, Y., P Jr., R.G., 2019. GlobeLand30 maps show four times larger gross than net land change from 2000 to 2010 in Asia. Int. J. Appl. Earth Obs. Geoinf. 78, 240–248.
Shafizadeh-Moghadam, H., Tayyebi, A., Helbich, M., 2017c. Transition index maps for urban growth simulation: application of artificial neural networks, weight of evidence and fuzzy multi-criteria evaluation. Environ. Monit. Assess. 189 (6), 300.
Tong, X., Feng, Y., 2019. A review of assessment methods for cellular automata models of land-use change and urban growth. Int. J. Geogr. Inf. Sci. https://doi.org/10.1080/13658816.2019.1684499.
Von Neumann, J., Burks, A.W., 1966. Theory of self-reproducing automata. IEEE Trans. Neural Network. 5 (1), 3–14.
Wagner, P.D., Fohrer, N., 2019. Gaining prediction accuracy in land use modeling by integrating modeled hydrologic variables. Environ. Model. Software 115, 155–163.
Wagner, P.D., Waske, B., 2016. A review of current calibration and validation practices in land-change modeling. Environ. Model. Software 83, 245–254.
Wang, H., Zhao, X., Zhang, X., Wu, D., Du, X., 2019. Long time series land cover classification in China from 1982 to 2015 based on Bi-LSTM deep learning. Rem. Sens. 11 (14), 1639.
Wang, Y., Li, S., 2011. Simulating multiple class urban land-use/cover changes by RBFN-based CA model. Comput. Geosci. 37 (2), 111–121.
Wu, H., Li, Z., Clarke, K.C., Shi, W., Fang, L., Lin, A., Zhou, J., 2019. Examining the sensitivity of spatial scale in cellular automata Markov chain simulation of land use change. Int. J. Geogr. Inf. Sci. 33 (5), 1040–1061.
Yang, Q., Li, X., Shi, X., 2008. Cellular automata for simulating land use changes based on support vector machines. Comput. Geosci. 34 (6), 592–602.
Zhang, D., Liu, X., Wu, X., Yao, Y., Wu, X., Chen, Y., 2019. Multiple intra-urban land use simulations and driving factors analysis: a case study in Huicheng, China. GIsci. Remote Sens 56 (2), 282–308.