Landscape and Urban Planning 205 (2021) 103960

Contents lists available at ScienceDirect

Landscape and Urban Planning

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

Research Paper

Mixed-cell cellular automata: A new approach for simulating the spatio-temporal dynamics of mixed land use structures

Xun Liang[a,b,c], Qingfeng Guan[a,b,*], Keith C. Clarke[c], Guangzhao Chen[d], Song Guo[a], Yao Yao[a,b]

[a] School of Geography and Information Engineering, China University of Geosciences, Wuhan, Hubei 430078, China
[b] National Engineering Research Center of GIS, China University of Geosciences, Wuhan 430078, Hubei Province, China
[c] Department of Geography, University of California Santa Barbara, Santa Barbara, CA 93106, USA
[d] Guangdong Key Laboratory for Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou, 510275, China

ARTICLE INFO

Keywords:
Mixed cell
Cellular automata
Land use structure
Mixed land use
Land use change

ABSTRACT

When used for land use change modeling, Cellular Automata (CA) usually assume that each cell has only one land use type at each time step, ignoring the mixed land use structures that are often found in land units. Mixed cells, composed of cover proportions of multiple land use types, provide a new perspective for modeling the spatio-temporal dynamics of mixed land use structures. Simulating land use change with mixed cells is challenging because mixed-cell CAs are fundamentally different from conventional CAs. This study develops a mixed-cell CA (MCCA). The structure of the CA is re-designed based on the cover proportion of land uses, including the representations of cell state, lattice, and neighborhood. The transition rules are automatically constructed by random-forest regression using historical data and a competition mechanism among multiple land use types at the sub-cell scale is proposed. In addition, a mixed-cell figure of merit (mcFoM) accuracy measure is proposed to validate the MCCA. The MCCA was applied to the Wuhan metropolitan area in China, and the results show that the MCCA was able to simulate the subtle changes of land use proportions within land units. The MCCA represents a new breed of geospatial CA models for spatio-temporal dynamics of mixed land use structures, which enables mixed land use research to leap from static analysis to dynamic simulation. The software for MCCA has been made available at https://github.com/HPSCIL/Mixed_Cell_Cellullar_Automata.

1. Introduction

Forecasts of land use and land cover (LULC) are needed to analyze the impacts of LULC change for a wide variety of socioeconomic and ecological processes, including population growth, economic development, carbon cycling, landscape dynamics, surface hydrology and climate change (Li et al., 2017; Pontius, Peethambaram, & Castella, 2011; Sohl, Loveland, Sleeter, Sayler, & Barnes, 2010). Land use modeling can therefore help understand the dynamics of the land use system and project future land use change in planning practices to achieve more sustainable development and help retain ecological security (Huang, 2014; Sohl et al., 2014; Verburg et al., 2002). Cellular automata (CA) have been widely used for simulating land use change at multiple scales (Basse, Omrani, Charif, Gerber, & Bódis, 2014; Dong, You, Cai, Li, & Lin, 2018; Liang et al., 2020), as they are simple and naturally spatio-temporally dynamic (Chaudhuri & Clarke, 2013; White, Engelen, & Uljee, 1997; He et al., 2020).

Traditionally, geospatial CA models assume each cell within the system to be of a uniform land use type and assign a discrete state label at each time step (Chen, Li, Wang, Liu, & Ai, 2013; Pontius et al., 2007; Yeh & Li, 2002; Zhai et al., 2020). In other words, the cell state of conventional CA models is pure and discrete (Clarke & Gaydos, 1998; Pijanowski, Alexandridis, & Müller, 2006). However, because of the complexity of land use patterns, especially in cities, a piece of land is usually a mixture of multiple land use types, serving multiple functions (Abdullahi, Pradhan, Mansor, & Shariff, 2015). Therefore, at the commonly used scales of CA models (e.g., 30 m × 30 m and coarser), the land space of a cell often contains various land use types with different cover proportions, which means the lattice (also termed cell space) of a CA model is not only composed of pure cells, but also a large number of mixed cells with multiple land use structures (Foody, 1996). For example, a single urban commercial cell may contain government offices and residences, and an agricultural cell may contain roads, houses and ponds. It is worth noting that although the term 'mixed cell' was first mentioned by Hu and Li (2004), their 'mixed cell' meant a mixture of points, polylines and polygons, different from the concept of 'mixed

* Corresponding author at: School of Geography and Information Engineering, China University of Geosciences, Wuhan, Hubei 430078, China.
E-mail address: guanqf@cug.edu.cn (Q. Guan).

https://doi.org/10.1016/j.landurbplan.2020.103960
Received 13 September 2019; Received in revised form 19 September 2020; Accepted 22 September 2020

cell' in this study that represents a mixture of multiple land use types within a cell.

The mixed structures inside land units are one of the key concerns of planning strategies (Song & Knaap, 2004), as they are closely related to human mobility and energy consumption (Abdullahi et al., 2015), and so affect the environmental sustainability and the functions of land units (Liu et al., 2018a; Yue et al., 2017). Previous studies concentrated on the identification, measurement and change analysis of mixed land uses (Shi & Yang, 2015), and only a few studies have focused on the simulation of the dynamics of mixed land use (Charif, Omrani, Abdallah, & Pijanowski, 2017; Omrani, Abdallah, Charif, & Longford, 2015). Understanding the dynamics can provide rich information for understanding the interactions between mixed land use and driving factors, and for making scientifically sound land use plans for a sustainable future. A new modeling approach is therefore needed to simulate the spatio-temporal dynamics of land use structures that cover proportions of land use categories within mixed land units. Given the success of CA in land use modeling, CA models with mixed cells appear to be a promising approach to achieve this purpose.

1.1. Mixed-cell CAs vs. pure-cell CAs

Simulating the change of land use structures based on mixed cells is a challenge, because mixed-cell CAs are fundamentally different from conventional pure-cell CAs. CA models are composed of five basic elements: cell, lattice (or cell space), neighborhood, a set of initial states and transition rules. All these basic elements must be re-designed for a CA with mixed cells. In addition, the evaluation methods must be re-designed since the commonly used methods were primarily designed for pure-cell CAs.

Each cell in a geospatial CA represents a land unit, and a state is associated with every cell representing the attribute/status of the land unit. In pure-cell CA models, a discrete state label from a finite set is assigned to a cell, representing the uniform land use type of the land unit (Li & Yeh, 2000). Unlike a pure cell, the state of a mixed cell is made up by an array of continuously measured components, each representing the cover proportion of a certain land use type (Fig. 1). By changing the state of each cell (i.e., changing the cover proportions of land use types within a cell) along time steps, a mixed-cell CA model is able to simulate the continuous structural change of land use mixture within each land unit, while pure-cell CA models can only simulate the discrete change of the whole cell (Li, Shi, He, & Liu, 2011; Liu & Phinn, 2003; Seto et al., 2012).

Usually, the lattice of a geospatial CA is composed of a group of cells arranged in a 2D space, representing the whole region of interest. Along time steps, the cells within the lattice change their states individually, hence to collectively simulate the spatio-temporal dynamics of a phenomenon (e.g., land use and/or land cover) in the region. With a discrete state label for each cell, a pure-cell CA model has only one layer of lattice for the target phenomenon, besides other layers of driving factors (Wu & Webster, 1998). As the state of a mixed cell is composed of an array of land use components (i.e., cover proportions of land use types), the lattice of a mixed-cell CA is a multi-layer structure, each layer representing the distribution of cover proportion of a certain land use type over the region.

Neighborhood effects are essential to CA models (Li & Yeh, 2002). With pure cells, CA models often use the numbers of cells of various land use types within the neighborhood (i.e., moving window) around a certain cell to represent the neighborhood condition (Chen, Li, Liu, & Ai, 2013; Shu et al., 2017; Wu, 2002). Therefore, the variability of land use states within a neighborhood is limited by its size. For example, when using a 3 × 3 window, there are no more than 8 (3 × 3–1) land use types in the neighborhood (Chen, Li, Liu, Ai, & Li, 2016). With the continuously measured cover proportions of multiple land use types as the states of mixed cells, the land use structure of the neighborhood can be represented in more detail (Fig. 1).

The transition rules of mixed-cell CA models are different from those of pure-CA models in two major ways. First, pure-cell CA models simulate land use change through competition among different land use types at the cell scale (Yang, Su, Chen, Xie, & Ge, 2016). However, mixed-cell CA models must estimate the proportion changes of land use types through competition among the land use components inside each cell. The transition rules of mixed-cell CA models must consider not only the effects at the cell scale (e.g., the influences of driving factors at the locations of cells), the neighborhood scale (e.g., neighborhood conditions) and the regional scale (e.g., land demands) as pure-cell CA models (Verburg & Overmars, 2009), but also sub-cell scale competition among multiple land use components. Second, compared with pure-cell CA models that simulate the qualitative change of land use for each cell, mixed-cell CA models simulate the quantitative changes among land use components inside each cell. This characteristic determines that the construction of transition rules of mixed-cell CA models should be based on the quantitative analysis of historical land use transitions. Thus, the prospect of mixed-cell CA models is of great significance for the move of CA models from qualitative simulation to quantitative simulation at the sub-cell scale when simulating the land use change of

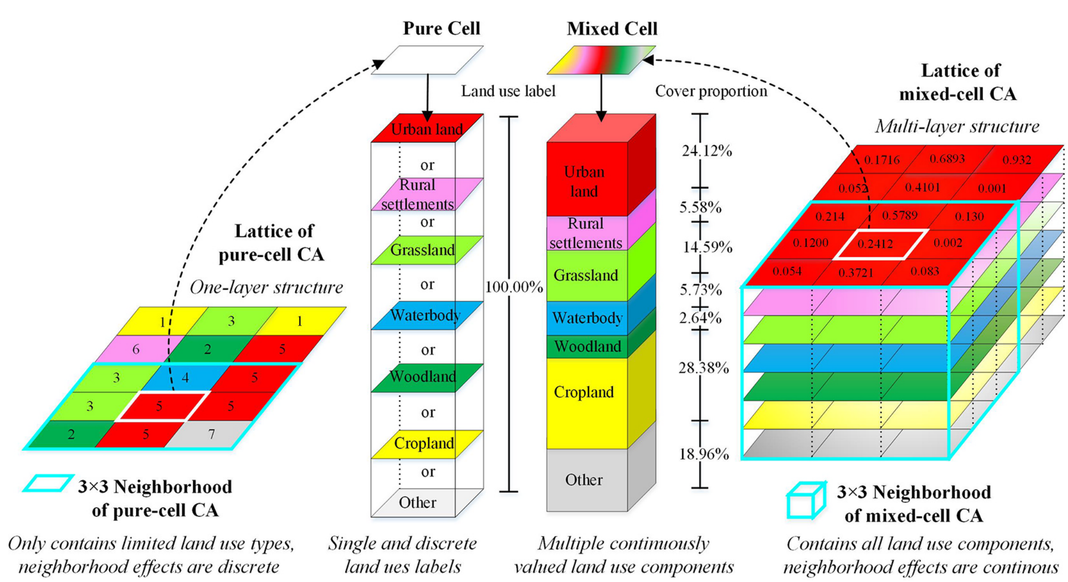

Fig. 1. Mixed-cell CA vs. pure-cell CA: representations of cell, neighborhood and lattice.

multiple land use types.

Finally, the simulation results of mixed-cell CA models are the distributions of cover proportions of multiple land use types (i.e., the multi-layer lattice in Fig. 1). The conventional evaluation methods, such as the 'confusion matrix' (Congalton, 1991) and 'figure of merit' (Pontius & Cheuk, 2006) are designed for discrete simulation results of pure-cell CA, and are unable to evaluate the continuous and multi-dimensional simulation results of mixed-cell CA models. Thus, evaluation of the accuracy of a mixed-cell CA model is an issue to be dealt with. The cell state of a mixed cell is a multi-dimensional array, representing the land use structure of the corresponding land unit. Therefore, the similarity of land use structure between simulation results and ground truth is an important part of the performance assessment of mixed-cell CA models. A thorough mixed-cell simulation framework needs reasonable and reliable evaluation methods, which can assess the accuracy of continuous and multidimensional distribution, the structural similarity of mixed land use between simulation results and ground truth, and even the change accuracy of mixed-cell simulation.

1.2. Relevant studies

Some scholars have been aware of the importance of simulating the dynamics of mixed urban land structures. For example, Li and Yeh (2000) proposed a grey-CA that can represent the percentages of urban within cells. Yeh and Li (2002) presented a CA model that incorporates density gradient in the simulation of urban development. Liu and Phinn (2003) developed a fuzzy-set CA to simulate the change of degree of membership of urban land in each cell. Sunde, He, Zhou, Hubbart, and Spicci (2014) proposed an I-CAT model that can provide quantitative information on impervious surfaces within each cell. Recently, Liu et al. (2018a) developed a gradient-CA model, which can express the temporal evolution characteristics of different urbanization stages. Mustafa et al. (2018) also developed a cellular automaton based on multinomial logistic regression and a genetic algorithm to simulate the densification change of urban land. However, these studies only focused on the growth of the urban fraction, and are not applicable to the simulation of the structural change of multiple land use types. Ching and S., Milne, G (2003) developed an Epidemic CA (ECA) model that includes population densities and mobility in each cell, and Tovar, Patel, Niebur, Sen, and Renaud (2006) also proposed a Hybrid CA model (HCA) for topology optimization in mechanical design. The cell states of ECA and HCA are composed of a set of variables, which are similar to the mixed cell in this study. However, because of the differences in research fields and modeling theories, the ECA and HCA models cannot be used in the field of geospatial studies to simulate the structural change of mixed land use.

It is worth mentioning that Omrani et al. (2015) introduced the multi-label (ML) concept, where each spatial unit can belong to multiple classes simultaneously. Omrani, Tayyebi, and Pijanowski (2017) also simulated multi-label land use change with an ML-CA-LTM model, which was a great stride forward in simulating the dynamics of mixed land use. Charif et al. (2017) uses a multi-label learning method-a multi-label support vector machine, Rank-SVM-to define the transition rules of ML-CA that significantly improved the simulation accuracy. However, the multi-label land use data used in the ML-CA series model does not include the cover proportion of land use types in each cell. Therefore, a mixed-cell CA model, specifically designed to simulate the continuous and quantitative changes of cover proportions of multiple land use components within cells, is still missing.

The performance of CA largely depends on the transition rules (Yang, Liu, Li, Li, & Ge, 2018). In geospatial studies, and especially in land use change studies, the transition rules of CAs are often derived using one of two approaches: (1) transition rules are set by model designers, and the parameters/coefficients are then calibrated using historical data. Typical examples include the DUEM (Batty, Xie, & Sun, 1999), SLEUTH (Clarke & Gaydos, 1998) and multi-criteria evaluation (Yang et al., 2016) models; or (2) transition rules are automatically constructed by a data mining model/algorithm using historical data (Hagenauer, Omrani, & Helbich, 2019). In recent years, a large number of CA models have been developed using the second approach, as it makes fewer subjective assumptions and is more flexible. For example, the Artificial Neural Network (ANN) model (Liang, Liu, Li, Zhao, & Chen, 2018a; Yang, Guo, Li, Zhang, & Li, 2019; Yeh & Li, 2002), Random Forest (RF) model (Kamusoko & Gamba, 2015; Zhang et al., 2019), and cuckoo search algorithm (Cao, Tang, Shen, & Wang, 2015) have been used to derive the relationships between land use types/changes and their driving factors. Given the discrete state label of pure-cell CAs, previous studies usually regard the mining of transition rules as a classification problem. The transition rules output a discrete land use type (i.e., label) for a certain cell under the influences of driving factors. Such a classification approach can only obtain qualitative and ad hoc transition rules.

Different from conventional CA models, mixed-cell CA models are concerned with the continuous and quantitative changes of multiple land use components in each cell. Therefore, instead of classification, the construction of transition rules of mixed-cell CAs should be regarded as a regression problem, to discover the relationships between the quantitative changes of land use components and the driving factors. Regression methods have been used in CA models. For example, Liu et al. (2018a) employed Support Vector Regression (SVR) to mine the relationship between urban growth and its driving factors. However, this study only simulated the continuous change of one land use type (i.e., impervious surface), which cannot be used in the simulation of the more complex mutual transitions among multiple land use components (i.e., the structural change) inside mixed cells. Thus, previous studies lack a mining framework for quantitative transition rules of mixed-cell CA models.

In summary, mixed-cell CA models are fundamentally different from conventional CA models in many important ways, including the cell state, lattice, neighborhood, transition rules, and the evaluation methods. Previous methods are unable to simulate the structural changes of multiple land use components inside mixed cells. This study aims to develop a mixed-cell CA framework for land use structural change simulation, which includes a mining method for constructing quantitative transition rules based on a regression approach, a CA model for simulating mutual changes of land use components inside mixed cells, as well as ways to validate the simulation accuracy of mixed-cell CA models. The development and evaluation of mixed-cell CA models are important advances in land use models, which can provide an effective simulation method and important support for planners and researchers for regional policy making, as well as for exploring the causes and consequences of land use change.

2. Method

A mixed-cell CA framework for land use structural change simulation is proposed in this study. Such a framework is based on the conceptual representations of cell state, lattice and neighborhood for mixed cells, as mentioned in Section 1.1 (Fig. 1), and contains three main parts: (1) a mining method for discovering the quantitative relationships between the changes of land use components within mixed cells and various driving factors; (2) a CA model for simulating the structural changes of mixed cells; and (3) a set of evaluation methods for assessing the performance of the mixed-cell CA (Fig. 2).

2.1. Mining relationships between land use structural changes and driving factors

2.1.1. Random forest regression (RFR)

To enable a CA model to simulate the structural changes of mixed cells, the relationships between the changes of land use components

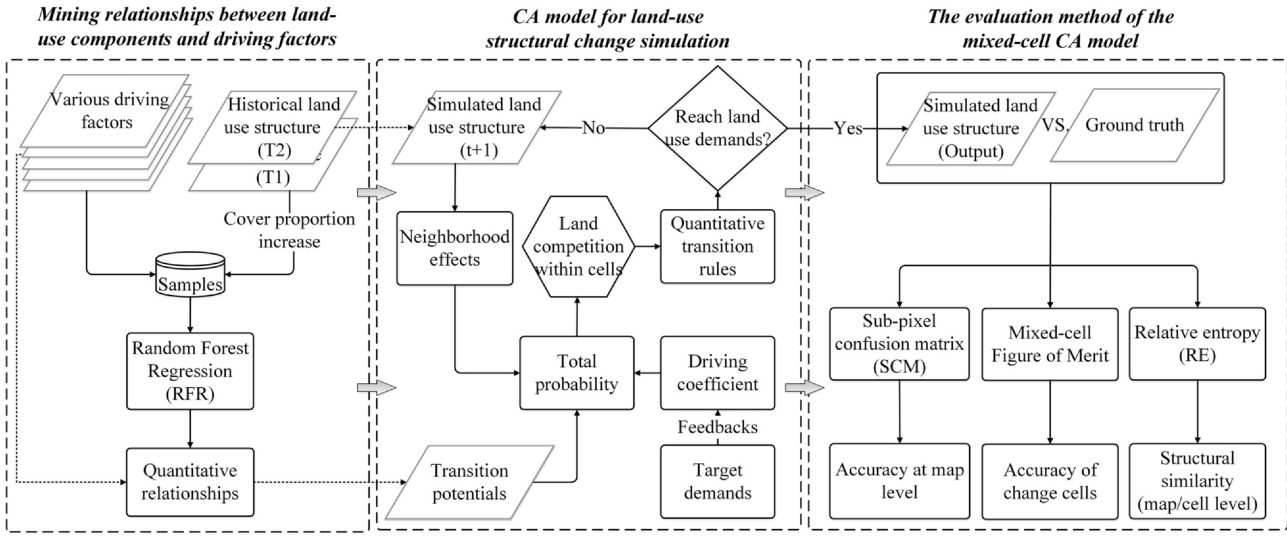

Fig. 2. Framework of mixed-cell CA model.

and driving factors must be derived. As mentioned above, the discovery of such relationships should be regarded as a regression problem instead of a classification problem, as the structural changes of cells are continuous rather than being discrete. Many regression methods can be used for this purpose, such as the Artificial Neural Network (ANN) model, the Support Vector Machine (SVM) or the Random Forest (RF). In this study, RF was used for its ability to overcome the multiple correlative problems among spatial variables, especially in higher-dimensional fitting situations (Palczewska, Palczewski, Marchese Robinson, & Neagu, 2014). RF is an aggregation of the decision-tree algorithm, in which an individual decision tree is constructed from each training sub-dataset. The generalization error of RF can be calculated by averaging the errors of these decision trees (Yao et al., 2017a). RF is commonly used in solving classification (RFC) and regression (RFR) problems, and has proved to be an effective method for mining the transition rules for land use change simulations (Gounaridis, Chorianopoulos, Symeonakis, & Koukoulas, 2019; Yao et al., 2017a, 2017b).

The principle of RFR is that for an arbitrary feature (e.g., a driving factor in this study) A, a point of demarcation s that divides the feature A into two datasets D_1 and D_2 can be determined, which makes both the mean square errors (MSEs) of D_1 and D_2 and the sum of the MSE of D_1 and D_2 minimum at the same time. The objective function of RFR is as follows:

$$\min_{A,s}\left[\min_{c_1}\sum_{x_i\in D_1(A,s)}(y_i-c_1)^2+\min_{c_2}\sum_{x_i\in D_2(A,s)}(y_i-c_2)^2\right] \tag{1}$$

where x_i is the features of the i-th training sample; y_i is the dependent variable of the ith training sample; c_1 is the sample mean of dataset D_1; and c_2 is the sample mean of dataset D_2. The regression value of the RFR algorithm is the mean values of the outputs of all regression trees (Fig. 3).

2.1.2. Mining the development potentials of land use components

The RFR is used to fit the relationship between the proportional change of each land use component and the driving factors with mixed cells. Once such a relationship is derived by the training process of RFR using historical samples, the change potential of the corresponding land use component of a certain mixed cell at a certain time step can be predicted, given a set of driving factor values at the corresponding location and time. Unlike classification problems that are commonly solved by training on multiple land use types, training each land use component in the regression problem is a more common way (Liu et al., 2018a). In addition, simultaneously fitting multiple arrays may affect the fitting precision of RFR. Therefore, we trained the development potential of each land use type separately in this study. This training method has been widely used in many other studies based on logistic regression (Verburg et al., 2002; Sohl, Wimberly, Radeloff, Theobald, & Sleeter, 2016). Although interactions between multiple land use components cannot be addressed in the training process, nevertheless, we still can address the interactions into the competition between different land use components in the simulation process, which is discussed in Section 2.2.

Before training the RFR, the proportional change of each land use component must be converted to the dependent variable of RFR:

$$Y_{i,k}=\begin{cases}PC_{i,k} & PC_{i,k}>0\\ 0 & PC_{i,k}\le 0\end{cases} \tag{2}$$

where $PC_{i,k}$ represents the proportion change of land use component k in mixed cell i between different periods; and $Y_{i,k}$ is the dependent variable of the RFR. Note that the cover proportion of a land use component within a mixed cell may increase ($PC_{i,k}>0$) or decrease ($PC_{i,k}\le 0$) between two periods. Since the land use components of a mixed cell are mutually constraining, an increase of a component means the decrease of others. Therefore, we only need to focus on the increases of land use components, and can regard the decreases as 'no increase', in order to avoid repeated calculations.

After such preprocessing, the relationship RF_k between the proportion change (i.e., increase) of land use component k and driving factors is derived through the training process of RFR, using randomly extracted samples from the historical dataset:

$$RF_k = RFR\_train(Y^s_{i,k}, DF^s_i) \tag{3}$$

where $Y^s_{i,k}$ denotes the sample dataset of $Y_{i,k}$; DF^s_l represents the sample dataset of various driving factors; and $RFR\_train(*)$ denotes the training process of the RFR.

Once RF_k is obtained, the development (i.e., increased) potential of land use component k can be predicted by the following equation:

$$DP_{i,k} = RFR\_predict(RF_k, DF_i) \tag{4}$$

Where DF_i represents the dataset of driving factors; $RFR\_predict(*)$ denotes the predicting process of RFR; and $DP_{i,k}$ denotes the development potential of land use component k (Gounaridis et al., 2019) at mixed cell i. For K land use components, K RFs must be trained. At each iteration of CA simulation, the development potential of each land use component of a certain cell is predicted and used as a key part of the transition rules.

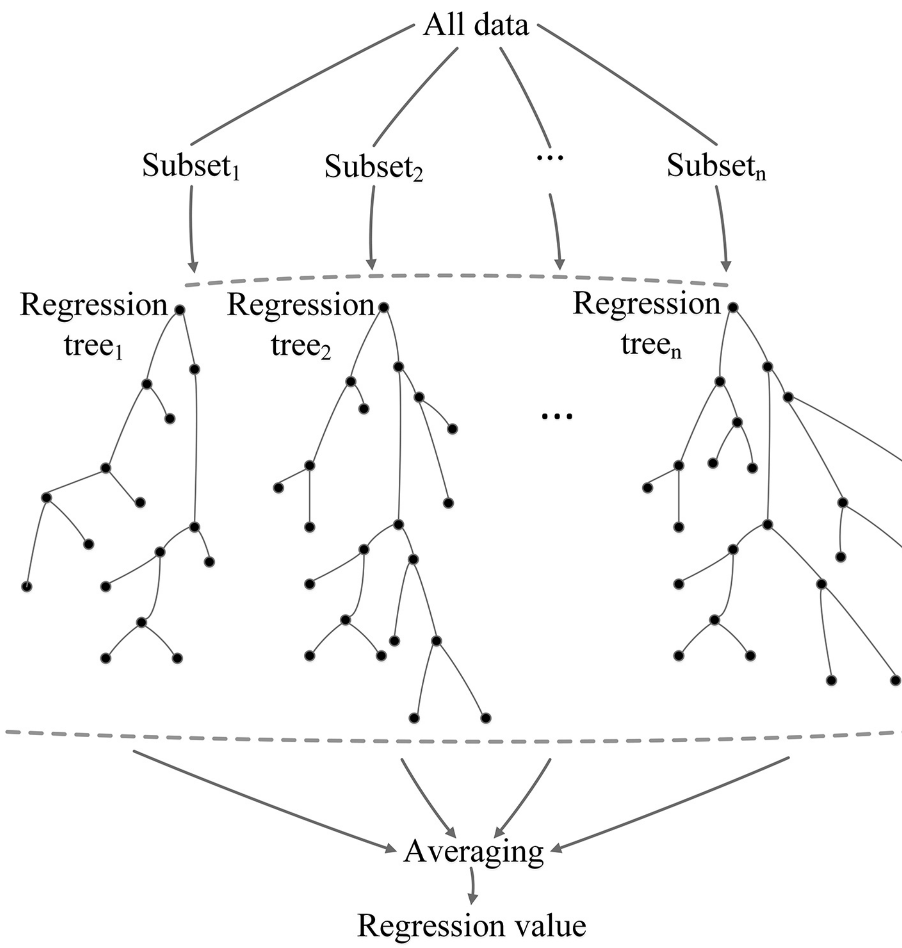

Fig. 3. The schematic diagram of Random Forest Regression (RFR).

2.2. CA model for land use structural change simulation

We propose a CA model combining top-down (i.e., macro land use demands) and bottom-up (i.e., local land use competition) effects to simulate the spatio-temporal dynamics of land use structures. The feedbacks between land use demands and land use structures are passed by a self-adaptive coefficient, driving the amount of land use to approach the target land use demands. In the simulation process, multiple land use components first compete with each other within each mixed cell through a roulette wheel approach, to determine whether or not the cover proportion of a land use component increases, and the amount of the increase. Then the amounts of other land use components converted to the increasing land use component are estimated through a set of quantitative transition rules. Details are given in Fig. 4.

2.2.1. Feedbacks between land use demands and land use structures

Based on the $DP_{i,k}$ calculated by RFR, this study proposes a demand-driven (or scenario-based) mixed-cell CA model for land use structural change simulation. First, the total change probability of land use component k can be represented as:

$$TP_{i,k}^{t} = DP_{i,k} \times \Omega_{i,k}^{t} \times Driv_{k}^{t} \tag{5}$$

where $TP_{i,k}^{t}$ is the total change probability of land use component k of mixed cell i at iteration t; $\Omega_{i,k}^{t}$ represents the neighborhood effects of mixed cell i, which are the cover proportions of land use components k within the neighborhood (cover proportions of all components represent the land use structure of the neighborhood); and $Driv_{k}^{t}$ is the feedback of future demand for land use type k, which is a self-adaptive coefficient that depends on the gap between the current amount at iteration t and target demand of land use k. The self-adaptive method of $Driv_{k}^{t}$ is as follows:

$$Driv'^{t}_{k} = \begin{cases} Driv_{k}^{t-1} & if|D_{k}^{t-1}| \le |D_{k}^{t-2}| \\ Driv_{k}^{t-1} \times \frac{|D_{k}^{t-2}|+1}{|D_{k}^{t-1}|+1} & if 0 > D_{k}^{t-2} > D_{k}^{t-1} \\ Driv_{k}^{t-1} \times \frac{|D_{k}^{t-1}|+1}{|D_{k}^{t-2}|+1} & if D_{k}^{t-1} > D_{k}^{t-2} > 0 \end{cases} \tag{6}$$

where $|D_{k}^{t-1}|$ and $|D_{k}^{t-2}|$ represent the absolute values of the differences between the cumulative amount and future demand of land use type k at the $t-1$th and $t-2$th iteration. We add 1 to them to prevent the numerator and denominator from becoming 0. However, the sum of all $Driv'^{t}_{k}$ may not be equal to 1. So a normalization is carried out through the following equation to ensure the range of $Driv_{k}^{t}$ is in the range 0–1:

$$Driv_{k}^{t} = \frac{Driv'^{t}_{k}}{\sum_{k=1}^{K} Driv'^{t}_{k}} \tag{7}$$

The $Driv_{k}^{t}$ drives the amount of land use k to approach the future demand during the iterative process, which is similar to the self-adaptive inertia coefficient proposed by Liu et al. (2017) and Liang et al. (2018b). It combines the "top-down" effects provided by the land use demands and the "bottom-up" influence that is rooted inside cells and neighborhoods (Liang et al., 2018b).

2.2.2. Land use competition within mixed cells

After the total development probability $TP_{k,l}^{t}$ is obtained, a roulette selection mechanism is used to address the land use competition at the sub cell scale during the simulation (Fig. 4). First, the total development

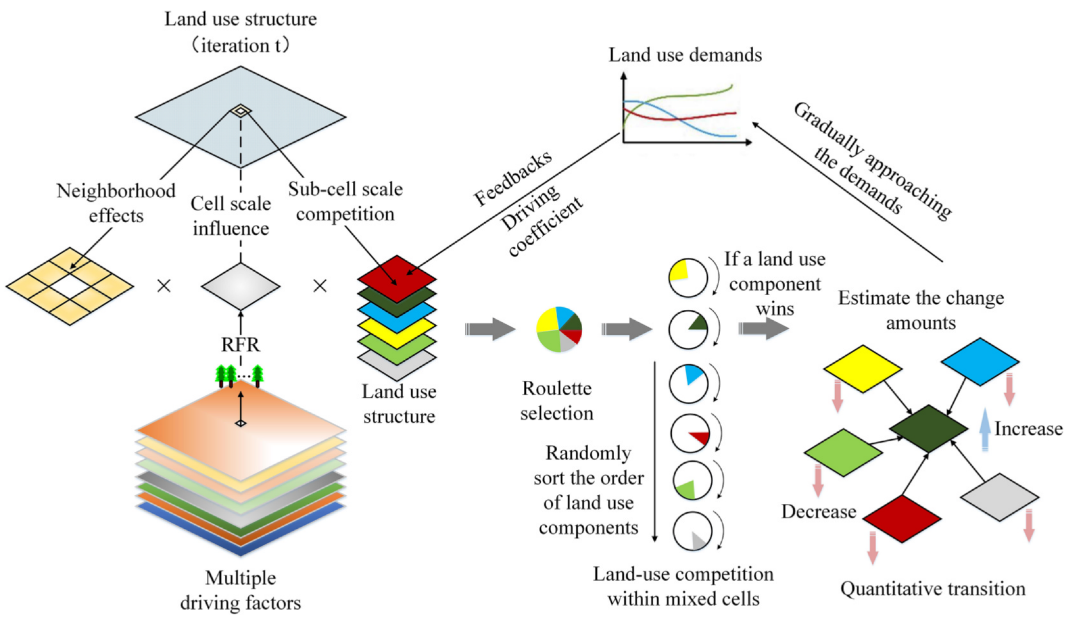

Fig. 4. The feedback, competition and quantitative transition mechanisms of the mixed-cell CA model.

probabilities of various land use components are normalized using the following equation:

$$TP^{t'}_{i,k} = \frac{TP^{t}_{i,k}}{\sum_{k=1}^{K} TP^{t}_{i,k}} \tag{8}$$

where $TP'^{t}_{n,k}$ is the normalized total change probability of land use component k of mixed cell i at iteration t (the sum of $TP'^{t}_{i,k}$ for all components is equal to 1). K represents the number of land use components in the study area. Then a roulette wheel R^t_i can be constructed according to the $TP'^{t}_{i,k}$, each sector of the roulette wheel represents a land use component. Then the order of land use components is randomly shuffled, each component competes with others through the roulette wheel R^t_i according to the shuffled order. That is, there are K rounds of competition for each cell, each round for a land use component k. If land use component k wins in its competition round, k is selected as the changed land use component. The changed amount of component k of cell i at iteration t can be calculated as:

$$IA^{t}_{i,k} = TP'^{t}_{i,k} \times Ra \times \epsilon_k \tag{9}$$

where $IA^{t}_{i,k}$ represents the increased proportion of land use component k, ranging within [0, 1]; Ra denotes a random number ranging from 0 to 1, which represents the stochastic perturbation of land use change; and ϵ_k is the step size of the amount change of component k at each iteration ranging from 0 to 1, which is an empirical parameter set by users.

2.2.3. The quantitative transition rules of land use components

When land use component k wins in its competition round, the amount of component k will increase in a mixed cell, which also means declines of other components. Similarly, the declining probability of a land use component can be estimated according to the growth probabilities of other components. Therefore, we define the declining probability of land use component o as the sum of the total development probabilities of other components:

$$SP^{t'}_{i,o} = \sum_{w=1}^{K} TP^{t'}_{i,w} \quad w \neq o \tag{10}$$

where $SP'^{t}_{i,o}$ represents the declining probability of component o of mixed cell i at iteration t. Thus when land use component k increases, the declining proportion of component o can be estimated using the following equation:

$$DA^{t}_{i,o} = \frac{SP^{t'}_{i,o}}{\sum_{v=1}^{K-1} SP^{t'}_{i,v}} \times IA^{t}_{i,k} \qquad v, \ o \neq k \tag{11}$$

$$\begin{cases} if con_{o\to k} = 1 then IA^{t}_{i,k} = IA^{t}_{i,k};\ DA^{t}_{i,o} = DA^{t}_{i,o} \\ if con_{o\to k} = 0 then IA^{t}_{i,k} = IA^{t}_{i,k} - DA^{t}_{i,o};\ DA^{t}_{i,o} = 0 \end{cases}$$

where $DA^{t}_{i,o}$ denotes the declining proportion of land use component o of mixed cell i at iteration t; and $con_{o\to k}$ represents a transition matrix that determines whether the original land use type o is allowed to convert to the target type k (1 denotes inevitable conversion and 0 denotes impossible conversion). If a conversion is impossible ($con_{o\to k} = 0$), the value of $IA^{t}_{i,k}$ is adjusted as: $IA^{t}_{i,k} = IA^{t}_{i,k} - DA^{t}_{i,o}$, then the value of $DA^{t}_{i,o}$ is set as 0. The schematic diagram of the competition and quantitative conversion mechanisms of the mixed cell CA model is shown in Fig. 4.

Driven by the future demands of various land use types, the above transition rules are applied to all cells, to determine the increasing and declining land use components and to evaluate the transition amount of each land use pair. When the simulated land use amounts are equal to the target land use demands, the mixed-cell CA model will output the simulated results with K layers. Each layer represents the simulated distribution of a land use component.

2.3. Evaluation of the mixed-cell CA model

Conventional evaluation methods, such as the ‘confusion matrix’ (Congalton, 1991) and ‘figure of merit’ (Pontius & Cheuk, 2006), are primarily designed for assessing the accuracy of discrete simulation results produced by pure-cell CAs, and cannot be used to evaluate the continuous and multidimensional simulation results of mixed-cell CA model. Therefore, we propose an evaluation scheme, which can assess the simulation accuracy of mixed-cell CA model from three aspects: (1) the total accuracy of the distributions of all land use components; (2) a new Figure of Merit indicator for mixed-cell simulation; and (3) the similarity of land use structure between the simulation and ground truth.

2.3.1. Sub-pixel confusion matrix

The Sub-pixel Confusion Matrix (SCM) is employed in this study to assess the total accuracy of the simulation results of the mixed-cell CA model. SCM, proposed by Pontius and Cheuk (2006), is an improved version of the conventional Confusion Matrix for evaluating the accuracy of soft classification, and is suitable for assessing the simulation result of mixed-cell CA models.

The first step is to randomly select a number of cells from both simulated and actual land use maps. The second step is to calculate the agreement of each land use component (u or v) for diagonal elements ($u = v$) of the sub-pixel confusion matrix, and the disagreement of each land use component for off-diagonal elements ($u \neq v$) of each sample cell i according to the following equation:

$$p_{iuv} = \begin{cases} MIN(s_{iu}, a_{iv}), \ u = v \\ (s_{iu} - p_{iuu})\left[\frac{(a_{iv} - p_{ivv})}{\sum_{v=1}^{N}(a_{iv} - p_{ivv})}\right], \ u \neq v \end{cases} \quad (12)$$

where p_{iuv} is the element of the SCM for the i-th sampled cell; s_{iu} is the simulated cover proportion of land use component u of cell i; a_{iv} is the actual cover proportion of land use component v of cell i; and $MIN(*)$ is the rule to select the minimum value among s_{iu} and a_{iv}. The final SCM is constructed by averaging the SCMs for all sample cells, and each element of the final SCM is represented as p_{vv}.

After the final SCM is built, several accuracy indices, including the Overall Accuracy (OA) at the map level and Producer's Accuracy (PA_v) and User's Accuracy (UA_v) at the category level, can be calculated through the following equations:

$$OA = \sum_{v=1}^{K} p_{vv} \quad (13)$$

$$PA_v = p_{vv}/a_v \quad (14)$$

$$UA_v = p_{vv}/s_v \quad (15)$$

We used the OA derived from SCM as the evaluation metric to assess the total precision of the mixed-cell CA model.

2.3.2. Mixed-cell Figure of Merit (mcFoM)

The accuracy of mixed-cell simulation cannot be validated by the traditional Figure of Merit (FoM) (Pontius et al., 2008; Pontius & Millones, 2011) that is commonly used in pure-cell simulation. This study proposes a mixed-cell Figure of Merit (mcFoM) to validate the simulation accuracy of the mixed-cell CA model. First, we obtained the actual ($DA_{k,i}$) and simulated ($DS_{k,i}$) proportion change of cell i by subtracting the ground truth ($G_{k,i}$) and simulated components ($S_{k,i}$) with the initial components ($I_{k,i}$).

$$DA_{k,i} = G_{k,i} - I_{k,i} \quad (16)$$

$$DS_{k,i} = S_{k,i} - I_{k,i} \quad (17)$$

where k is the arbitrary land use type. Second, we compared each mixed cell of DA_k and DS_k and divided the change of all cells into four parts: A: area of error due to underestimating the change of components; B: the agreement of all classes (area of correct) according to the minimum rule (Eq. (16)), because the agreement cannot be more than the minimum value of the two proportion changes; C: area of error due to misestimating the change directions of components (e.g., observed increase but predicted decrease); and D: area of error due to overestimating the change of components.

$$A = \sum_i \sum_k (|DA_{k,i}| - |DS_{k,i}|) DA_{k,i} \times DS_{k,i} > 0, \ |DA_{k,i}| \geq |DS_{k,i}| \quad (18)$$

$$B = \sum_i \sum_k \min(|DA_{k,i}|, |DS_{k,i}|) \quad (19)$$

$$C = \sum_i \sum_k (|DA_{k,i} - DS_{k,i}|) DA_{k,i} \times DS_{k,i} < 0 \quad (20)$$

$$D = \sum_i \sum_k (|DS_{k,i}| - |DA_{k,i}|) DA_{k,i} \times DS_{k,i} > 0, \ |DA_{k,i}| < |DS_{k,i}| \quad (21)$$

The equations of the mixed cell Figure of Merit ($mcFoM$) can be expressed as:

$$mcFoM = \frac{B}{A + B + C + D} \quad (22)$$

$$PA = \frac{B}{A + B + C} \quad (23)$$

$$UA = \frac{B}{B + C + D} \quad (24)$$

Where PA and UA represent the producers accuracy and users accuracy respectively.

2.3.3. Relative entropy for land use structural similarity assessment

The similarity of land use structure is another important aspect for evaluating the simulation result of mixed-cell CA, which is a unique characteristic of multi-dimensional simulation results. The land use structure of a cell refers to the array of cover proportions of land use components of this cell, and the sum of all land use components is equal to 1. We computed the Relative Entropy (RE) as an indicator to evaluate the similarity of land use structure, which can represent the information decay of the simulation process (Song & Knaap, 2004). The RE of each cell is defined as:

$$RE_i = \sum_{k=1}^{K} P_i(k) \log\left(\frac{P_i(k)}{Q_i(k)}\right) \quad (25)$$

$$meanRE = \sum_{i=1}^{M} RE_i/M \quad (26)$$

where RE_i denotes the relative entropy of the actual and simulated land use structure of mixed cell i; P_i and Q_i represent the actual and simulated land use structure, respectively; while M is the total number of mixed cells. RE is able to measure the similarity between two vectors. In the calculation process, a very small constant value can be added ($\epsilon = 2.220 \times 10^{-16}$, DBL_EPSILON) to $Q_i(k)$ and $P_i(k)$ to avoid a zero denominator according to http://hanj.cs.illinois.edu/cs412/bk3/KL-divergence.pdf. When the actual and simulated land use structures of mixed cell i are identical, RE_i is 0. A larger RE_i value indicates a greater difference between the actual and simulated land use structure. After the REs for all cells are calculated, the mean RE of the whole region is calculated as the measurement of similarity of land use structure at the regional level.

3. Experiment and performance assessment

3.1. Study area

The proposed mixed-cell CA model was applied to a simulation of the Wuhan Metropolitan Area (WMA), which is located in central China and encompasses an area of 57,800 km^2 (Fig. 5). Wuhan is the central city of WMA, which is also the biggest city, transportation hub and education center in central China. The WMA contains eight large- and medium-sized cities around Wuhan, including Huangshi, Ezhou, Huanggang, Xiaogan, Xianning, Xiantao, Qianjiang and Tianmen. The WMA is one of the biggest urban agglomerations in central China, and is an important economic center of the Yangtze River Economic Zone of China. In 2017, the Gross Domestic Product (GDP) in WMA was about 305.93 billion U.S. dollars and the population was 38 million. According to the 'The Rise of Central China' strategy enacted by the Chinese government, the WMA is expected to develop into the fourth

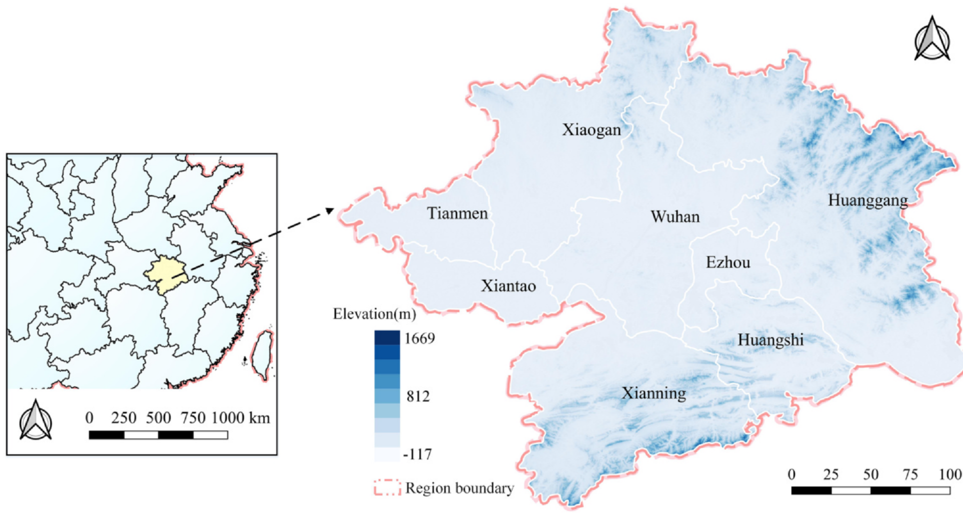

Fig. 5. Location and topography of the Wuhan Metropolitan Area.

economic growth pole in China (besides the Yangtze River Delta, Pearl River Delta and the Beijing-Tianjin-Hebei metropolitan area) with a resource-saving and environment friendly focus.

3.2. Data preparation

In practical applications, the mixed-cell land use data can be obtained using at least four approaches: (1) the remote sensing inversion method (Liu et al., 2018a); (2) decomposition of mixed pixels (Shi & Wang, 2014); (3) extracting the land use mix structure from multi-source big data (e.g., social network data, taxi trajectories, points of interest) (Liu et al., 2018b; Shi & Yang, 2015); or (4) aggregating fine-resolution land use data into coarse-resolution data (Omrani et al., 2017). The purpose of this experiment is to demonstrate and evaluate the proposed mixed-cell CA model, so we used the fourth approach – the most convenient method to obtain the mixed-cell land use data.

The China Land Use/Cover Dataset (CLUD), the most commonly used and highest-quality national land use/cover database for China (Kuang, Liu, Dong, Chi, & Zhang, 2016), was used in this study at its original fine resolution (30 m × 30 m). The CLUD includes six land use types: cropland, woodland, grassland, waterbody, built-up, and other. In this study, the built-up land use was further divided into urban land and rural settlement. Thus in total 7 land use types were considered for the experiment.

The study region was divided into 250 m × 250 m regular cells (in total 1,433,256 cells), and the cover proportions of 7 land use components for each cell were calculated from the original 30 m-resolution CLUD data. Through such a process, the mixed-cell land use data for 2000 and 2015 were generated. They were used to train the RFR model for the mining of relationships between the changes of land use components and driving factors, and to assess the simulation accuracy of the mixed-cell CA model.

In addition to the mixed land use data, we formulated a collection of spatial variables to represent the driving forces of land use structural change, including socioeconomic, climatic and environmental data (Table 1). These factors have been commonly used in previous studies (Li et al., 2017; Liu et al., 2017; Chen et al., 2020). All the spatial datasets were resampled to the resolution of 250 m × 250 m.

3.3. Model calibration and validation

3.3.1. Historical simulation

For the model calibration, the RFR for each land use type was trained using samples from the mixed land use data of the period of 2000–2015 in WMA. The sampling rate was set as 10%, and 100 regression trees were used to construct each RFR model. Other parameters of the mixed-cell CA model are listed in Table 2. The fitting precision of RFR was evaluated by the out-of-bag root-mean-square error (OOB RMSE). As shown in Table 2, the OOB RMSE of all land use components was lower than 0.04, indicating that the RFRs were well trained and capable of capturing the relationships between land use structural changes and driving factors. The step sizes represent the conversion rates of all land use types. The larger the step size for the MCCA model, the fewer the iterations for the simulation process. Considering that sufficient iteration is encouraged for the CA model (Yeh & Li, 2006), we assume that the conversion rates of all the land use types are the same and set the step sizes at 1 using trial-and-error (Feng & Tong, 2020), to balance the number of iterations and the runtime of the MCCA model.

After training, the mixed-cell CA model was used to simulate the land use structural change for the period of 2000–2015. Specifically, the trained RFRs were applied to calculate the development potentials of all land use components of each cell by importing all the driving factors, and the changes in proportion of land use components were calculated by the CA model for land use structural change simulation.

During the simulation, land use changes were constrained such that some conversions were not allowed, which is a typical step in land use modelling (Li et al., 2017; Schaldach et al., 2011; Verburg & Overmars, 2009). For instance, urban land cannot be converted to any other land use type, whereas cropland can be transformed to new rural settlements. Allowable conversions were specified in a conversion matrix for this experiment (Table 3). In addition, we assume that open water (a sub-category of waterbody) is not allowed to convert to other land use components. Therefore, a distribution map of the cover proportion of open water was used to provide the minimum quantity of waterbody of each cell, which means that the cover proportion of waterbody within each cell is not less than the cover proportion of open water. Meanwhile, the mutual conversions between other pairs of land use components are allowed in these cells.

The maps of simulated and actual land use structures in 2015 are shown in Fig. 6, with RGB images generated using different combinations of land use components to display details of the simulation results. The results showed that the simulated distribution of land use mixture matched well with the actual distribution at the regional scale.

3.3.2. Simulation accuracy at the regional scale

The Sub-pixel Confusion Matrix (SCM) and mcFoM were employed to quantitatively assess the consistency between the simulated result

Table 1
The spatial driving factors of land use change in this study.

Category	Data	Year[1]	Original Resolution	Data resource
Land use/cover data	Land use/cover data	2000–2015	30 *m*	CAS (http://www.resdc.cn)
Socioeconomic data	Population	2000–2015	1000 *m*	http://www.resdc.cn/Default.aspx
	GDP			
	Proximity to Wuhan	2014	–	World Urbanization Prospects: The 2014 Revision, CD-ROM Edition
	Proximity to city center			
	Proximity to Town center			
	Proximity to highway	2018	–	OpenStreetMap (https://www.openstreetmap.org/)
	Proximity to arterial road			
	Proximity to primary road			
	Proximity to secondary road			
	Proximity to tertiary road			
	Proximity to high-speed railway stations	2018		http://lbsyun.baidu.com/
Climatic and environmental data	Soil type	1995	1000 m	HWSD v 1.2 (http://westdc.westgis.ac.cn/data/844010ba-d359-4020-bf76-2b58806f9205)
	Annual Mean Temperature	1970–2000	30 arc-sec	WorldClim v2.0 (http://www.worldclim.org/)
	Annual Precipitation			
	DEM	2016	30 m	NASA SRTM1 v3.0
	Slope			

[1] Driving factors collected from different time periods is allowed (Long, Han, Lai, & Mao, 2013), and we have made the time periods of the driving factors as recent as possible.

Table 2
Parameters and accuracy indexes of mixed-cell CA model for the experiment.

	Parameters	Cropland	Woodland	Grassland	Waterbody	Urban land	Rural settlements	Other	
RFR	Parameter	Sampling rate	0.1						
		Number of regression trees	100						
	Accuracy index	OOB-RMSE	0.0191	0.0177	0.0070	0.0302	0.0372	0.0097	0.0076
CA for land use structural change	Parameter	Neighborhood	3 × 3						
		Step size	1	1	1	1	1	1	1
		Land use demand in 2015 (km^2)	22193.26	14717.40	1188.43	5186.35	2115.64	1706.14	148.83
	Accuracy index	OA	0.9303						
		mcFoM	0.2959 (PA = 0.3745, UA = 0.4011)						
		Mean RE	0.9768						

and the actual land use pattern. The overall accuracy (OA = 0.9303), the mixed cell Figure of Merit (mcFoM = 0.2959) at the map level, and their User's Accuracy and Producer's Accuracy were calculated (Table 4, Table 2). These accuracy indices indicate that the simulation result of the mixed-cell CA was acceptable. In addition, we compared the simulation results of the mixed-cell CA model with a widely used pure-cell CA model (Fig. 6), namely the FLUS model (Liu et al., 2017), with the same development potentials and land use demands. The simulation result of the FLUS model from 2000 to 2015 was validated with the traditional confusion matrix and Figure of Merit. The OA and FoM values of the pure-cell CA model are 0.8967 and 0.1530 respectively, which are lower than the result simulated by the mixed cell CA model. We also compared the change map simulated by both the mixed–cell CA model and the pure-cell CA model (Fig. 7). The patterns of simulated change of all land use components are similar to that of the actual change of land use components, compared to the discrete simulation result from the pure-cell CA model. The distribution of simulation results of the mixed-cell CA model is continuous in space and contains the quantitative information of the change of land use components within cells.

3.3.3. Similarity between simulated and actual land use structures

In addition, the Relative Entropy (RE) among the simulated and actual land use structures of all cells was calculated according to Eqs. (23) and (24). The spatial distribution of RE is presented in Fig. 8, the mean RE (*mean RE*) of the study region was 0.9768 (Table 2), which represents the information loss of the simulation process. The *meanRE* can thus be regarded as an indicator to describe the similarity between the actual and simulated land use structures of the whole study region. For each cell or the whole region, the lower the RE, the more similar are the simulated and actual land use structures. Fig. 8 shows that the mixed cells with significant structural differences mainly distribute at some of the regions along the new roads and at the edges of the urban area, but most of the simulated land use structures are similar to the

Table 3
Land use conversion matrix (1 = conversion possible; 0 = not possible).

Change to →	Cropland	Woodland	Grassland	Waterbody	Urban land	Rural settlements	Other
Cropland	1	1	1	1	1	1	0
Woodland	1	1	1	0	1	1	0
Grassland	1	1	1	0	1	1	0
Waterbody	1	0	0	1	1	1	0
Urban land	0	0	0	0	1	0	0
Rural settlements	0	0	0	0	1	1	0
Other	1	1	1	1	1	1	1

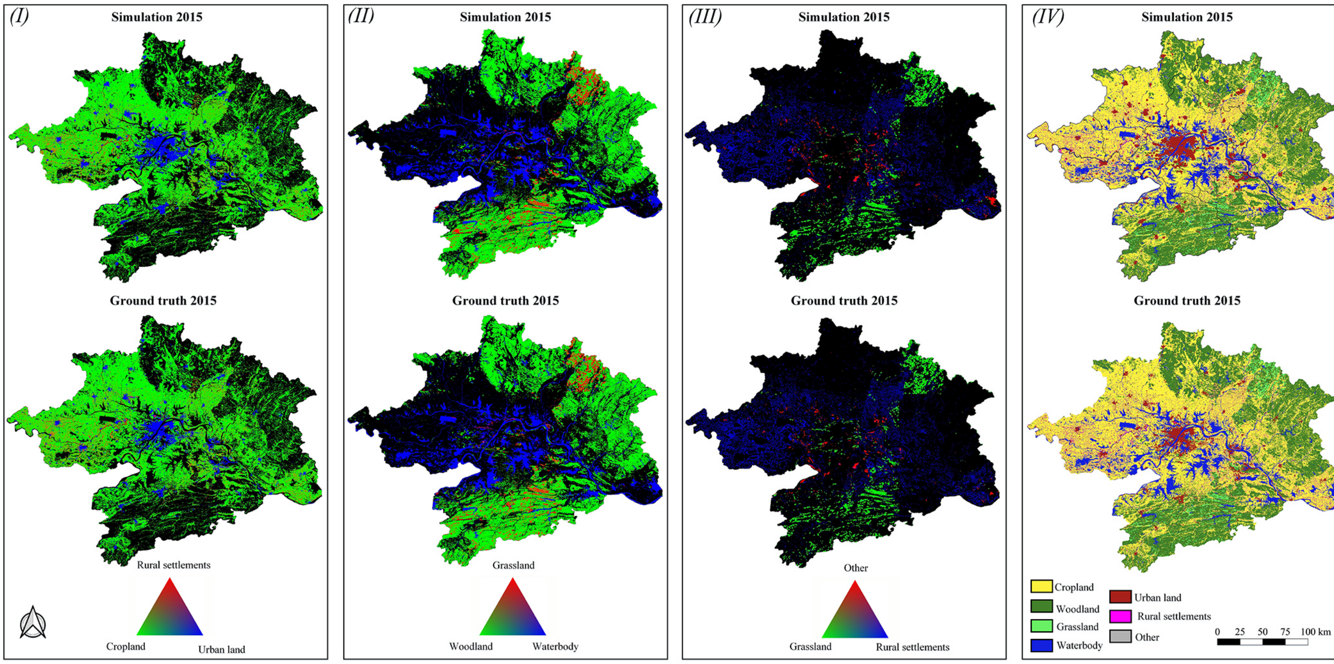

Fig. 6. RGB images to exhibit the mixing distribution of three land use components. Panel (I) depicts the mixing distribution of Rural settlements (R), Crop land (G) and urban land (B); Panel (II) shows the mixing distribution of Glass land (R), Woodland (G) and Waterbody (B); Panel (III) represents the mixing distribution of Other (R), Rural settlements (G) and Glassland (B). In the dark areas, the proportions of the three land use components are all very low. Panel (IV) shows the simulation result of a pure-Cell CA model and the pure-Cell ground truth.

Table 4
The Sub-pixel Confusion Matrix for the simulation from 2000 to 2015 (PA represents the producer's accuracy, UA is the user's accuracy).

Category	Cropland	Woodland	Grassland	Waterbody	Urban land	Rural settlements	Other	UA
Cropland	0.444164	0.004671	0.00039	0.008108	0.010073	0.002566	0.000241	0.944597
Woodland	0.004995	0.302976	0.000568	0.000907	0.002811	0.000287	0.000036	0.969271
Grassland	0.000291	0.002269	0.023485	0.000312	0.000244	0.000026	0.000017	0.881398
Waterbody	0.0105	0.000483	0.0002	0.095759	0.001168	0.000166	0.000781	0.878053
Urban land	0.008055	0.001521	0.00024	0.00252	0.029881	0.00146	0.000078	0.682902
Rural settlements	0.001708	0.000205	0.000037	0.000204	0.000428	0.032024	0.000018	0.924857
Other	0.000233	0.000015	0.000014	0.00103	0.000103	1.24E-05	0.001914	0.521871
PA	0.945237	0.970575	0.942138	0.881714	0.668537	0.876336	0.626748	
OA = 0.9303								

actual structures in the study region. Although most of the land use change happens in the regions along new roads and edges of urban areas, these places are the most difficult parts for simulating land use change. It is also very normal that most errors come from these regions. A lack of planning data may lead to this shortage because Chinese cities are largely impacted by planning policies. In our future work, we will try to use the planning data to improve the simulation of the mixed-cell CA model.

3.3.4. Effect of spatial aggregation on the model's performance

To examine the spatial aggregation effect of different aggregated grids on the MCCA's simulation accuracy. We also aggregated the simulation data to 500 m, 750 m, 1000 m, 1250 m, and 1500 m grids and tested the different accuracy indicators of the mixed-cell CA model respectively (Fig. 9). The variation of model performance with the growth of the aggregated grid size is shown in Fig. 9. We found that the mcFoM indicator shows a significant decreasing trend with the growth of the aggregated grid size, and the OA slightly decreases with the growth of the aggregated grid size. The mixed-cell CA model obtained the highest OA and mcFoM when the aggregated grid size was 250 m. Different from the previous two indicators, the changing trend of the meanRE is descending first and ascending last. The mixed-cell CA model obtained the highest structural similarity when the aggregated grid size is 750 m (with the lowest mean RE value).

3.4. Simulation of future land use structure

3.4.1. Projecting future land use demands

The mixed-cell CA is a scenario-driven model. The first step of prediction is to determine the total areas of land use types for the future period. The quantities of changes are then spatially allocated to individual cells during the prediction process of the mixed-cell CA model.

Future land use demands can be determined by many methods, such as using expert knowledge (Sohl, Sayler, Drummond, & Loveland, 2007), linear regression (Pontius & Malanson, 2005), Markov chains (Yang, Zheng, & Chen, 2014), the system dynamics model (Huang, 2014), or an integrated assessment model (Dong et al., 2018; Sohl et al., 2014; Verburg & Overmars, 2009). Considering that the purpose of this experiment is to demonstrate the capability of the proposed mixed-cell CA model, we employed linear regression, one of the simplest forecasting methods, to project the future land use demands based on the historical data of 2000, 2005, 2010 and 2015. Fig. 10 depicts the fitting equations and the projected trajectories of seven land use types from 2015 to 2035.

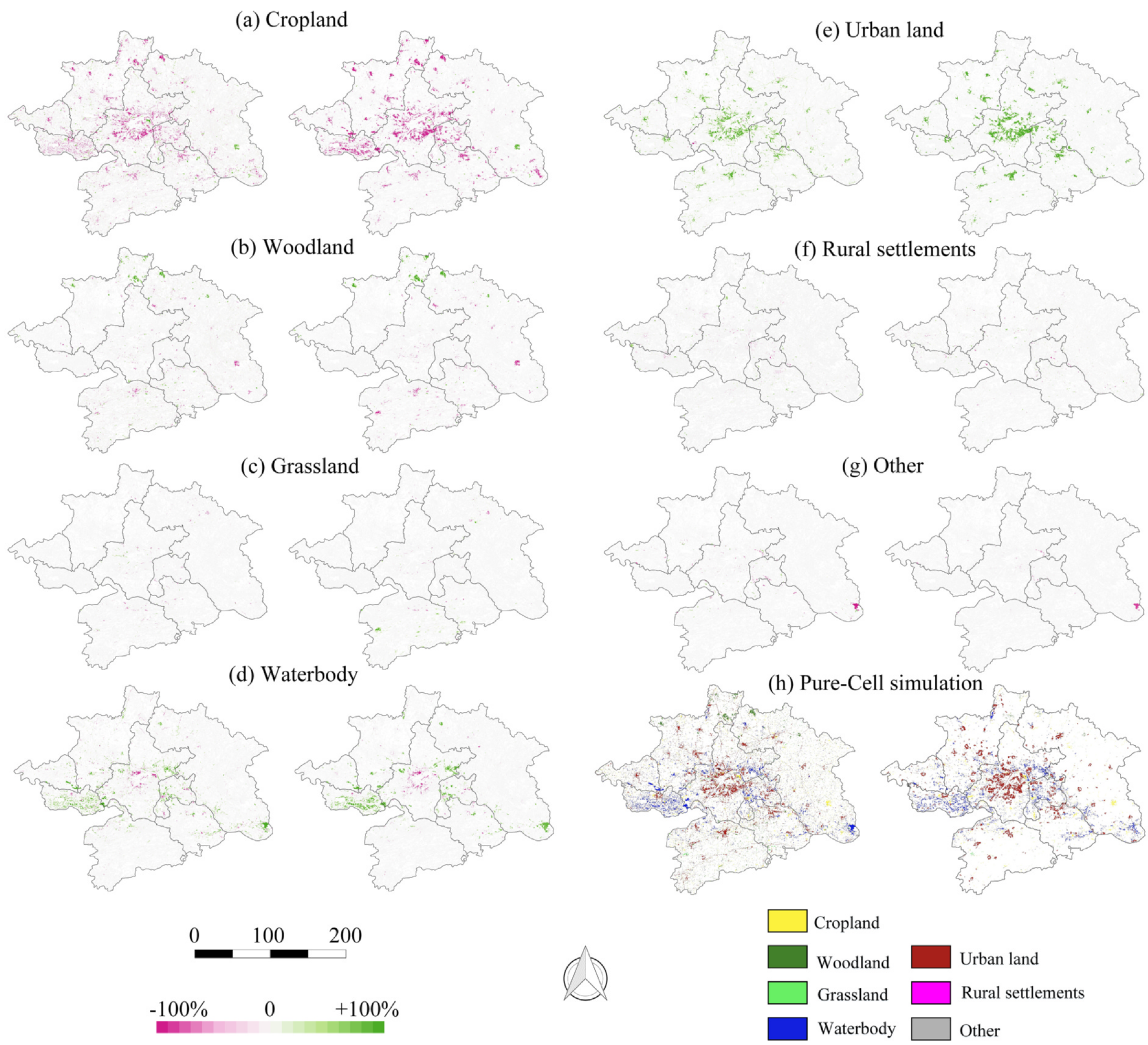

Fig. 7. Panel (a–g): actual (left) and simulated (right) change of each land use component from 2000 to 2015 in the mixed cell simulation; Panel (h): actual (left) and simulated (right) change cells of the pure-Cell simulation.

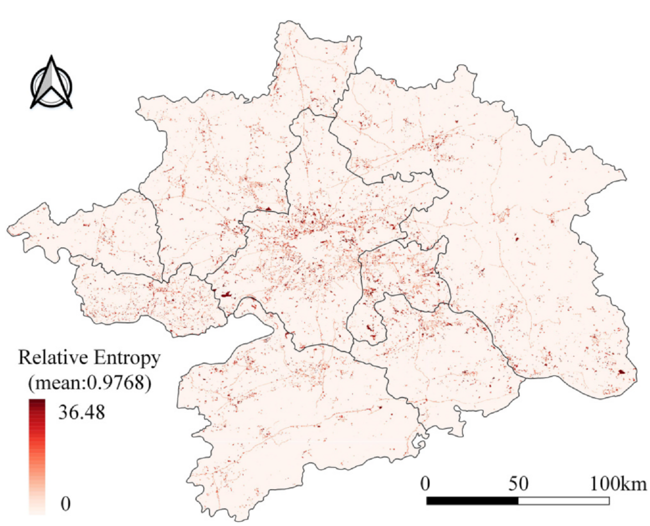

Fig. 8. The distribution of the relative entropy of the simulation result from 2010 to 2015.

3.4.2. Future prediction from 2015 to 2035

The mixed cell model was used to produce future land use structures for the period of 2015–2035, using the transition rules obtained from the period of 2010–2015 through RFR. Other simulation parameters are also listed in Table 2, which have been successfully calibrated. Fig. 11 depicts five sub-regions in WMA that respectively show the simulated changes of 5 land use components. In Fig. 11, panels (a1) to (a4) show a large increase of urban land in the central area of the WMA, mainly in Wuhan and its surrounding areas (e.g., Ezhou and Xiaogan). The urban area in Wuhan will expand in multiple directions simultaneously, among which the eastward expansion to Ezhou and the southwestward expansion along the Yangtze River are the most obvious trends. Stimulated by the fast urban expansion in Wuhan, the urban land in north Ezhou will develop rapidly and is eventually connected with the urban area in Wuhan in 2025, which matches the development strategy of "the Integration of Wuhan and Ezhou" formulated by the provincial government. By 2035, the urban layout of Ezhou will be composed of several big urban patches, including two new urban patches along the main road. The urban growth in Wuhan will first expand with relatively low density (Fig. 11(a1) and (a2)), then the cover proportion of urban land in mixed cells will increase from the inner city to the edge of the new urban areas (Fig. 11(a2) and (a3)). The mixed cells with relatively

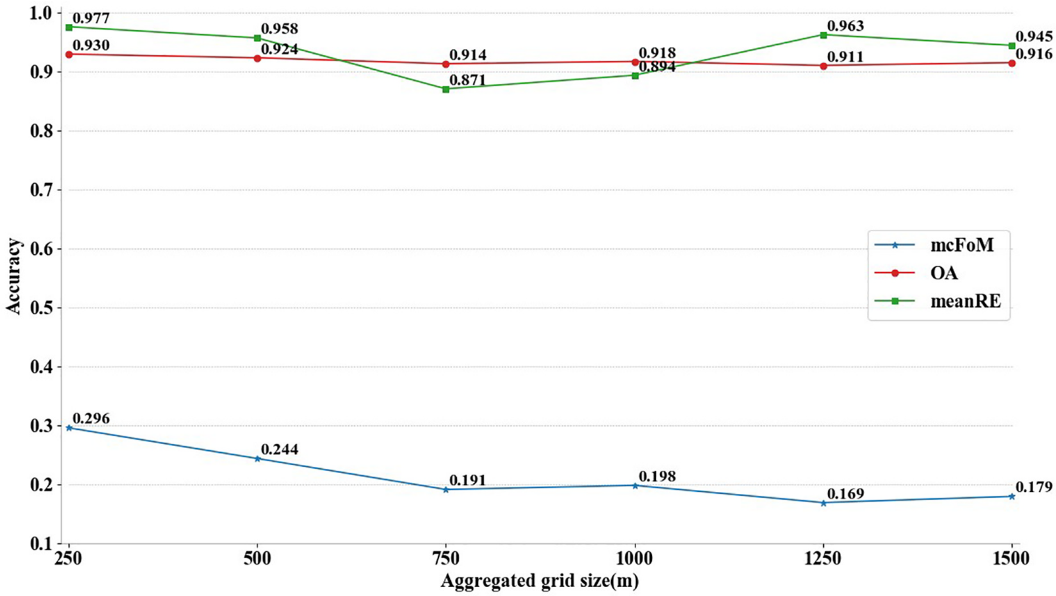

Fig. 9. Simulation accuracies versus aggregated grid size.

low urban density mainly distribute along the rim of the urban areas, with only a few scattered inside the high-density urban areas. These characteristics are consistent with the development regulation of the cities and reflect the spatial heterogeneity of land use distribution. Such characteristics cannot be expressed by traditional pure-cell CA models.

The fast development of the urban land is accompanied by a rapid loss of cropland in WMA. Fig. 11 (b1)–(b4) depicts the process of cropland loss from 2015 to 2035 in a sub-region in the western WMA, where the cropland is the most intensive and contiguous. Therefore, the protection and sustainable use of cropland in this region is of great significance for WMA. The urban expansion in this place will certainly have a negative impact on the integrity and connectivity of cropland in this region. By generating spatially continuous simulation results, the mixed-cell CA model improves our capability of assessing the influence of urban growth on the landscape connectivity.

In addition, rural settlements are an important source of urban land. Although the cover proportion of rural settlements is small, rural settlements scatter across China and the sprawl of rural settlements also causes a large amount of cropland loss (Tian, Qiao, & Gao, 2014). However, the rural settlements have been given less attention in previous studies (Kuang et al., 2016). According to the predicted land use demands for WMA (Fig. 10), the net change of rural settlements is very small. In this case, the result generated by traditional pure-cell CAs would only show very limited change. However, the simulation results of the mixed-cell CA model can explicitly reflect the density change of rural settlements at a fine scale under such a slow increase scenario (Fig. 11(c1, c2, c3)). These hot-spots of rural settlement change cannot be expressed by traditional pure-cell CA models, because they are unable to simulate the change in density of land use.

The fast development of urban land will also lead to the decline of woodland in various areas. In Fig. 11, panels (d1)–(d4) depict the change of woodland in Xianning, southern WMA. The woodland around high-density urban land has a higher risk of occupation by the rapid urbanization. In the southern and northern mountainous areas of WMA, many woodlands are most likely to degrade to grassland and result in a density decrease in the following decades. While cells with increasing woodland proportions will also scatter across the region and most of them occur in the eastern WMA. The amounts of increase and decrease in woodland proportion are approximately balanced, and the total amount of woodland in the WMA will remain almost unchanged in the following decades. In Fig. 11 panels (e1)–(e4) show the simulated expansion of grassland in the northern WMA. The process of the density increase of grassland in this region can also be projected by the mixed-cell CA model.

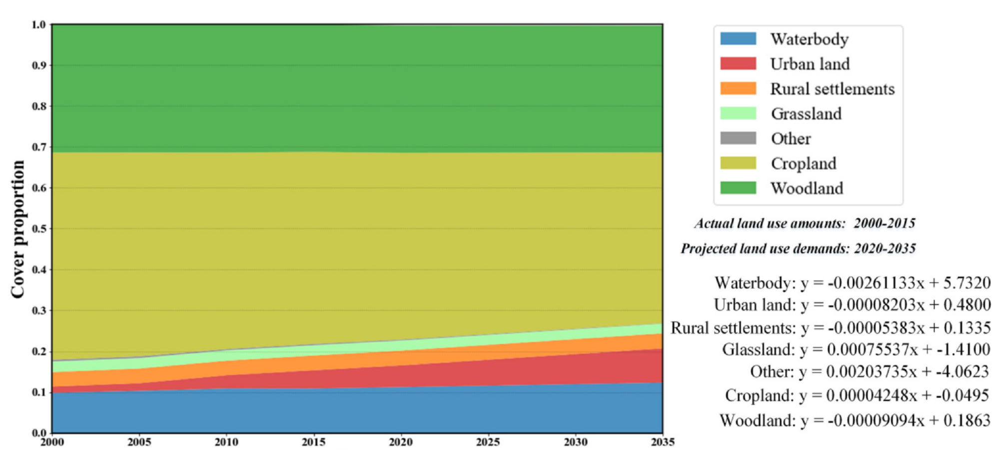

Fig. 10. Trajectories of the actual (2000–2015) and predicted (2015–2035) land use proportions in Wuhan Metropolitan Area.

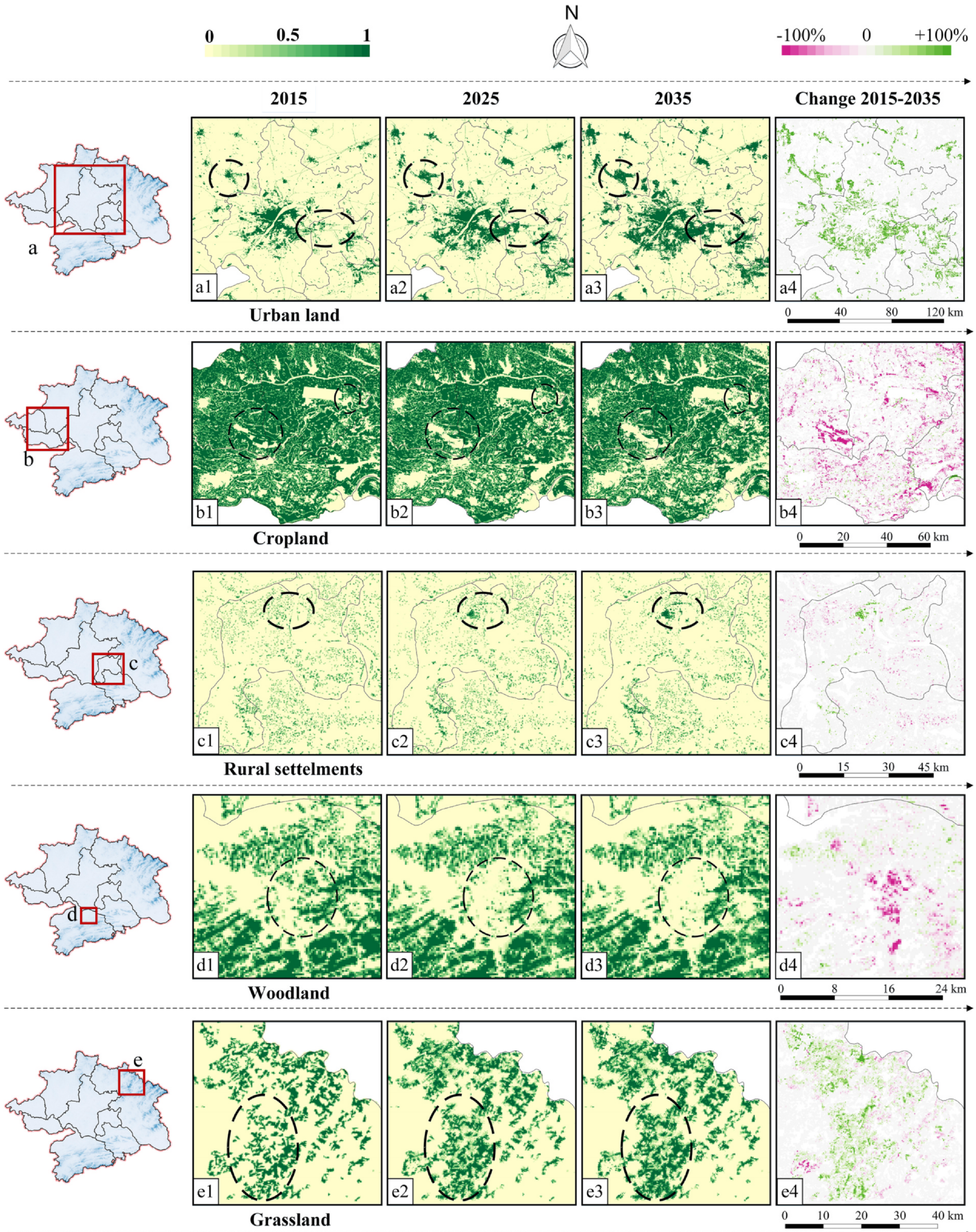

Fig. 11. The predicted cover proportions of four important land use components in WMA.

3.4.3. Change of land use mix

The mixture of land use components in land units is closely related to the socioeconomic activities, environmental functions and landscape amenities, and is very important for regional sustainable development (Abdullahi et al., 2015; Manaugh & Kreider, 2013; Musakwa & Niekerk, 2013). Different from pure-cell CA models, mixed cell CA models have the advantage of simulating the changes of land use structure within individual cells. The cell-level mixture can be directly measured by their entropy according to the cover proportions of land use components:

$$H_i = \frac{-\sum_{k=1}^{K} p_{i,k}\ln(p_{i,k})}{\ln(K)} \tag{27}$$

where H_i represents the entropy of mixed land use within cell i,

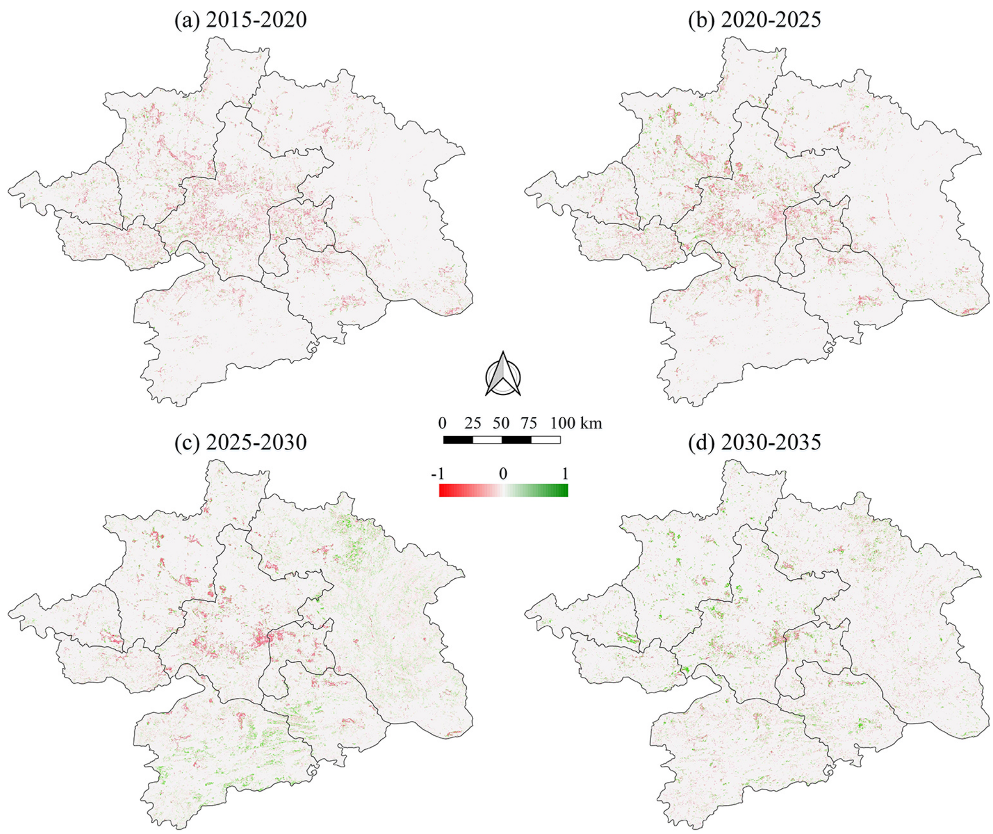

Fig. 12. The changes of land use mixture in WMA during 2015–2035.

ranging from 0 to 1; $p_{i,k}$ is the cover proportion of land use component k in cell i. A higher entropy means a higher degree of mixture or diversity of land use in an individual cell.

Fig. 12 demonstrates the changes of land use mixture during the prediction from 2015 to 2035. In the period of 2015–2020, the cells with decreasing mixture are obviously more than the cells with increasing mixture. In the cells whose mixture is increasing, the cover proportions of one or more land use components start increasing, leading to an increase in land use diversity. The later period (2020–2025) exhibits the opposite trend: the number of cells with increasing mixture tends to increase and they mainly distribute in the surroundings of the western WMA. The decrease of mixture in these areas is mainly due to the increase of the urban density. The periods of 2025–2030 and 2030–2035 continue the trend, but the cells with changes of land use mixture decentralize over time. In general, the land use mixture in WMA will decrease first in the west, and then increase in the east, caused by the continuous growth of urban land and grassland, and the evolution of woodland.

Table 5
A summary of the differences between mixed-cell CAs and pure-cell CAs.

Basic elements	Attributes	Pure-cell CA model	Mixed-cell CA model
Cell state	*State*	Discrete	Continuous
	Dimension	One-dimension	Multi-dimension
Neighborhood	*State*	Discrete	Continuous
	Classes	Limited classes sometimes	All classes
Lattice	–	One layer	Multiple layers
Transition rules	–	Qualitative	Quantitative
	Competition scale	Cell scale	Sub-cell scale
	Mining method	Classification	Regression
Functions	*Type*	Yes	Yes
	Structure	No	Yes
	Mixed land use	No	Yes
	Cell level mixture	No	Yes
Evaluation methods	*Map level accuracy*	Confusion Matrix	Sub-pixel confusion Matrix
	Changed cell accuracy	Figure of merit (FoM)	Mixed-cell FoM
	Structural similarity	–	Relative entropy

4. Discussion

This study introduces a new breed of CA models with mixed cells. The cell state, lattice, and neighborhood are re-designed to represent the mixed land use structures within land units. Consequently, the transition rules and evaluation methods are also re-designed to accommodate the unique characteristics of mixed cells. The differences between traditional pure-cell CA models and the mixed-cell CA model are summarized in Table 5. One of the biggest advantages of mixed-cell CA models is the capability of simulating the quantitative and continuous changes of multiple land use components inside cells, while pure-cell CA models can only simulate the qualitative and discrete change of land uses at the cell level. Therefore, mixed-cell CA models are able to simulate subtle changes in land use structures caused by minor variations of socio-economic, eco-environmental and political driving factors, providing a detailed perspective for understanding the land use change process.

Also, the quantitative and continuous simulations generated by mixed-cell CA models that contain the information for land use structure in each cell have the potential to help researchers to more precisely evaluate the impacts of land use change on many environment variables, such as air quality, the urban heat island effect, landscape connectivity, net primary production (NPP), ecological service value, energy consumption and more. Mixed-cell CA models may better support space-time continuous analysis and the quantitative calculation of environment variables. In addition, mixed-cell CA models provide an enabling approach to the simulation of structural changes of mixed land use, as most previous studies focused on the measurement and static analysis of mixed land use structures and ignored their dynamic evolution. The mixed-cell CA model can simulate gradual changes in land use structures and help researchers understand how the multiple driving factors interact to generate the future distribution of mixed land uses. It is worth mentioning that the simulation results of the mixed-cell CA model can be easily converted to traditional discrete land use data for specific uses by extracting the dominant land use type of a cell. Therefore, the simulation results of mixed-cell CA models are capable of covering all the functions of pure-cell simulation results.

Despite many obvious advantages of mixed-cell CA models, they are more complex to implement than pure-cell CA models. Therefore, this study also developed a software package for the mixed-cell CA (freely available at https://github.com/HPSCIL/Mixed_Cell_Cellullar_Automata). This software is written in the C++ programming language and contains data pre-processing, simulation and evaluation modules. The authors are responsible for the long-term maintenance and updates of the mixed-cell CA software. In our future work, the MCCA model will be coupled with mixed pixel decomposition algorithms to simulate the quantitative change of mixed land use structure. The mixed pixel decomposition algorithms can be used to provide multi-period mixed pixel land use data with higher resolution (e.g., Landsat imagery with a 30 m resolution), and the mixed-cell CA model will simulate the change trend of mixed land use structure of each pixel. The mixed pixel decomposition algorithms directly provide the data source for mixed-cell simulation, and the mixed-cell Cellular Automata can act as a new application of the output of mixed pixel decomposition studies.

5. Conclusion

This paper presents a new approach for the simulation of land use structural change – mixed-cell CA, which is fundamentally different from conventional pure-cell CA models. Specifically, the cell state of mixed-cell CA is composed of an array of continuously valued land use components, each representing the cover proportion of a certain land use type within a mixed land unit. Consequently, the lattice and neighborhood of CA are re-designed to accommodate the unique characteristics of mixed cells. Mixed-cell CA models are able to simulate the continuous change of multiple land use components, hence the structural change of mixed land units, which cannot be achieved by pure-cell CA models with discrete land use labels for cells.

The multi-dimensional representation of cell state also leads to a fundamental re-design of transition rules. To enable the simulation of continuous change of land use components within mixed cells, the discovery of the relationships between land use structural change and land use change driving factors must be regarded as a regression problem, instead of a classification problem as in pure-cell CA models. Also, the competition and the mutual conversions among multiple land use components within mixed cells (at the sub-cell scale) must be considered when constructing transition rules, in addition to the effects of cell-scale conditions, neighborhood conditions, and regional demands. Also, the evaluation methods for assessing the performance of CA models must be re-designed, because the commonly used methods are primarily designed for pure-cell CA and are not applicable for multi-dimensional and continuously valued land use structures.

Therefore, this study proposes a CA modeling framework to accommodate all these unique characteristics of mixed-cell CA. The relationships between land use structural change and driving factors are mined from historical data using random forest regression (RFR). Based on the development potentials of land use components derived from trained RFRs, transition rules determine the inter-conversion of land use components within cells by considering the feedbacks among land use demands, neighborhood effects, and sub-cell scale land use competition. Finally, the sub-pixel confusion matrix and relative entropy are used to evaluate the simulation accuracy. A new validation indicator, the mixed cell Figure of Merit (mcFoM), is proposed to assess the accuracy of the simulation of the changed cells of the mixed-cell CA model.

The proposed mixed cell CA model was applied to the land use structural change simulations in WMA. The mixed-cell CA model was calibrated using the land use structure data from 2000 to 2015. The simulation result achieved a high simulation accuracy (OA = 0.9303, mcFoM = 0.2959, mean relative entropy = 0.9768). The mixed-cell CA model was then used to project future land use structure change from 2015 to 2035, driven by the future land use demands that were forecasted through linear regression according to the historical trends. The results showed that the mixed-cell CA model can not only discover the hot-spots of land use density, but also can reflect the continuous change and spatial heterogeneity of the land use distribution. Finally, the change trends of land use mixture were analyzed.

In summary, mixed-cell CAs can effectively simulate the continuous and quantitative change of multiple land use components within mixed land units, providing key information for assessing the causes and consequences of land use change. The effects of land use structural change on socioeconomic and ecological environments can be evaluated by combining the mixed-cell CA model with other models. Mixed-cell CA models have a high potential in studying the function and structural change of land units, which can provide critical support for land management and urban planning. The mixed-cell CA model is a key link to the next stage of mixed land use study, moving from static analysis to dynamic simulation at the sub-cell scale, and can be regarded as an important addition to the theory and applications of Cellular Automata. In the future, using the mixed pixel decomposition as inputs, the mixed-cell CA model can also act as a bridge between mixed pixel decomposition and mixed cell simulation.

CRediT authorship contribution statement

Xun Liang: Conceptualization, Methodology, Software, Data curation, Validation, Formal analysis, Resources, Visualization, Writing - original draft, Funding acquisition. **Qingfeng Guan:** Conceptualization, Methodology, Resources, Writing - original draft, Funding acquisition, Supervision, Project administration. **Keith C. Clarke:** Conceptualization, Writing - original draft, Resources, Supervision, Project administration. **Guangzhao Chen:** Conceptualization, Methodology. **Song Guo:** Validation, Visualization. **Yao Yao:** Funding acquisition, Visualization.

Acknowlegments

This research was funded by the National Natural Science Foundation of China (Grant No. 41901332, U1711267, 41671408, 41801306); and the China Postdoctoral Science Foundation (Grant No. 2020T130616, 2019M652729). Guangzhao Chen and Xun Liang proposed the algorithm of the mcFoM indicator.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2020.103960.

References

Abdullahi, S., Pradhan, B., Mansor, S., & Shariff, A. R. M. (2015). GIS-based modeling for the spatial measurement and evaluation of mixed land use development for a compact city. *GIScience & Remote Sensing, 52*(1), 18–39.

Basse, R. M., Omrani, H., Charif, O., Gerber, P., & Bódis, K. (2014). Land use changes

modelling using advanced methods: Cellular automata and artificial neural networks. The spatial and explicit representation of land cover dynamics at the cross-border region scale. *Applied Geography, 53*, 160–171.
Batty, M., Xie, Y., & Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems, 23*(3), 205–233.
Cao, M., Tang, G., Shen, Q., & Wang, Y. (2015). A new discovery of transition rules for cellular automata by using cuckoo search algorithm. *International Journal of Geographical Information Science, 29*(5), 806–824.
Charif, O., Omrani, H., Abdallah, F., & Pijanowski, B. (2017). A multi-label cellular automata model for land change simulation. *Transactions in GIS, 21*(6), 1298–1320.
Chaudhuri, G., & Clarke, K. C. (2013). The SLEUTH land use change model: A review. *The International Journal of Environmental Resources Research, 1*(1), 88–104.
Chen, G., Li, X., Liu, X., Chen, Y., Liang, X., Leng, J., ... Huang, K. (2020). Global projections of future urban land expansion under shared socioeconomic pathways. *Nature Communications, 11*(1).
Chen, Y., Li, X., Liu, X., & Ai, B. (2013). Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automaton with a patch-based simulation strategy. *International Journal of Geographical Information Science, 28*(2), 234–255.
Chen, Y., Li, X., Liu, X., Ai, B., & Li, S. (2016). Capturing the varying effects of driving forces over time for the simulation of urban growth by using survival analysis and cellular automata. *Landscape and Urban Planning, 152*, 59–71.
Chen, Y., Li, X., Wang, S., Liu, X., & Ai, B. (2013). Simulating urban form and energy consumption in the pearl river delta under different development strategies. *Annals of the Association of American Geographers, 103*(6), 1567–1585.
Ching, F.u., & S., Milne, G (2003). *Epidemic Modelling Using Cellular Automata.* University of Western Australia.
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. *International Journal of Geographical Information Science, 12*(7), 699–714.
Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*(2), 270–279.
Dong, N., You, L., Cai, W., Li, G., & Lin, H. (2018). Land use projections in China under global socioeconomic and emission scenarios: Utilizing a scenario-based land-use change assessment framework. *Global Environmental Change, 50*, 164–177.
Feng, Y., & Tong, X. (2020). A new cellular automata framework of urban growth modeling by incorporating statistical and heuristic methods. *International Journal of Geographical Information Science, 34*(1), 74–97.
Foody, G. M. (1996). Relating the land-cover composition of mixed pixels to artificial neural network classification. *Photo.eng.rem.sens, 62(5):págs*, 491–500.
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. *Science of The Total Environment, 646*, 320–335.
Hagenauer, J., Omrani, H., Helbich, M., 2019, Assessing the performance of 38 machine learning models: the case of land consumption rates in Bavaria, Germany, Taylor & Francis, pp. 1399-1419.
He, Z., Deng, M., Cai, J., Xie, Z., Guan, Q., & Yang, C. (2020). Mining spatiotemporal association patterns from complex geographic phenomena. *International Journal of Geographical Information Science: Big Spatiotemporal Data Analytics, 34*(6), 1162–1187.
Hu, S., & Li, D. (2004). Vector cellular automata based geographical entity., in. *International conference on geoinformatics – geospatial information research: bridging the Pacific and Atlantic* (pp. 249–256). Sweden: University of Gävle.
Huang, Q. H. C. L., 2014, Modeling the impacts of drying trend scenarios on land systems in northern China using an integrated SD and CA model, 中国科学:地球科学英文版 57(4):839-854.
Kamusoko, C., & Gamba, J. (2015). Simulating Urban Growth Using a Random Forest-Cellular Automata (RF-CA) Model. *ISPRS International Journal of Geo-Information, 4*(2), 447–470.
Kuang, W., Liu, J., Dong, J., Chi, W., & Zhang, C. (2016). The rapid and massive urban and industrial land expansions in China between 1990 and 2010: A CLUD-based analysis of their trajectories, patterns, and drivers. *Landscape and Urban Planning, 145*, 21–33.
Yeh, A. G., & Li, X. (2006). Errors and uncertainties in urban cellular automata. *Computers, Environment and Urban Systems, 30*(1), 10–28.
Li, X., Chen, G., Liu, X., Liang, X., Wang, S., Chen, Y., ... Xu, X. (2017). A New Global Land-Use and Land-Cover Change Product at a 1-km Resolution for 2010 to 2100 Based on Human-Environment Interactions. *Annals of the American Association of Geographers, 107*(5), 1040–1059.
Li, X., Shi, X., He, J., & Liu, X. (2011). Coupling simulation and optimization to solve planning problems in a fast-developing area. *Annals of the Association of American Geographers, 101*(5), 1032–1048.
Li, X., & Yeh, A. G. (2000). Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *International Journal of Geographical Information Science, 14*(2), 131–152.
Li, X., & Yeh, A. G. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science, 16*(4), 323–343.
Liang, X., Liu, X., Chen, G., Leng, J., Wen, Y., & Chen, G. (2020). Coupling fuzzy clustering and cellular automata based on local maxima of development potential to model urban emergence and expansion in economic development zones. *International Journal of Geographical Information Science, 34*(10), 1930–1952.
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. *International Journal of Geographical Information Science, 32*(11), 2294–2316.
Liang, X., Liu, X., Li, X., Chen, Y., Tian, H., & Yao, Y. (2018). Delineating multi-scenario urban growth boundaries with a CA-based FLUS model and morphological method. *Landscape and Urban Planning, 177*, 47–63.
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. *International Journal of Geographical Information Science, 32*(1), 73–101.
Liu, X., Liang, X., Li, X., Xu, X., Ou, J., Chen, Y., ... Pei, F. (2017). A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landscape and Urban Planning, 168*, 94–116.
Liu, X., Niu, N., Liu, X., Jin, H., Ou, J., Jiao, L., & Liu, Y. (2018). Characterizing mixed-use buildings based on multi-source big data. *International Journal of Geographical Information Science, 32*(4), 738.
Liu, Y., & Phinn, S. R. (2003). Modelling urban development with cellular automata incorporating fuzzy-set approaches. *Computers, Environment and Urban Systems, 27*(6), 637–658.
Long, Y., Han, H., Lai, S., & Mao, Q. (2013). Urban growth boundaries of the Beijing Metropolitan Area: Comparison of simulation and artwork. *Cities, 31*, 337–348.
Manaugh, K., & Kreider, T. (2013). What is mixed use? *Presenting an interaction method for measuring land use mix, Journal of Transport & Land Use, 6*(1), 63–72.
Musakwa, W., & Niekerk, A. V. (2013). Implications of land use change for the sustainability of urban areas: A case study of Stellenbosch. *South Africa, Cities, 32*(32), 143–156.
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. *Computers, Environment and Urban Systems, 67*, 147–156.
Omrani, H., Abdallah, F., Charif, O., & Longford, N. T. (2015). Multi-label class assignment in land-use modelling. *Taylor & Francis,* 1023–1041.
Omrani, H., Tayyebi, A., & Pijanowski, B. (2017). Integrating the multi-label land-use concept and cellular automata with the artificial neural network-based Land Transformation Model: An integrated ML-CA-LTM modeling framework. *Taylor & Francis,* 283–304.
Palczewska, A., Palczewski, J., Marchese Robinson, R., & Neagu, D. (2014). Interpreting Random Forest Classification Models Using a Feature Contribution Method, in: *Integration of Reusable Systems*. In T. S. H. Rubin (Ed.). *Bouabana-Tebibel* (pp. 193–218). Cham: Springer International Publishing.
Pijanowski, B. C., Alexandridis, K. T., & Müller, D. (2006). Modelling urbanization patterns in two diverse regions of the world. *Journal of Land Use Science, 1*(2–4), 83–108.
Pontius, G. R., & Malanson, J. (2005). Comparison of the structure and accuracy of two land change models. *International Journal of Geographical Information Science, 19*(2), 243–265.
Pontius, R. G., Boersma, W., Castella, J., Clarke, K., de Nijs, T., Dietzel, C., ... Verburg, P. H. (2008). Comparing the input, output, and validation maps for several models of land change. *The Annals of Regional Science, 42*(1), 11–37.
Pontius, R. G., & Cheuk, M. L. (2006). A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions. *International Journal of Geographical Information Science, 20*(1), 1–30.
Pontius, R. G., & Millones, M. (2011). Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing, 32*(15), 4407–4429.
Pontius, R. G., Peethambaram, S., & Castella, J. (2011). Comparison of three maps at multiple resolutions: A case study of land change simulation in cho don district Vietnam. *Annals of the Association of American Geographers, 101*(1), 45–62.
Pontius, R. G., Walker, R., Yao-Kumah, R., Arima, E., Aldrich, S., Caldas, M., & Vergara, D. (2007). Accuracy assessment for a simulation model of amazonian deforestation. *Annals of the Association of American Geographers, 97*(4), 677–695.
Schaldach, R., Alcamo, J., Koch, J., Kölking, C., Lapola, D. M., Schüngel, J., & Priess, J. A. (2011). An integrated approach to modelling land-use change on continental and global scales. *Environmental Modelling & Software, 26*(8), 1041–1051.
Seto, K. C., Reenberg, A., Boone, C. G., Fragkias, M., Haase, D., Langanke, T., ... Simon, D. (2012). Urban land teleconnections and sustainability. *Proceedings of the National Academy of Sciences, 109*(20), 7687–7692.
Shi, B., & Yang, J. (2015). Scale, distribution, and pattern of mixed land use in central districts: A case study of Nanjing. *China, Habitat International, 46*, 166–177.
Shi, C., & Wang, L. (2014). Incorporating spatial information in spectral unmixing: A review. *Remote Sensing of Environment, 149*, 70–87.
Shu, B., Bakker, M. M., Zhang, H., Li, Y., Wei, Q., & Carsjens, G. J. (2017). Modeling urban expansion by using variable weights logistic cellular automata: A case study of Nanjing China. *International Journal of Geographical Information Systems, 31*(7), 1314–1333.
Sohl, T. L., Loveland, T. R., Sleeter, B. M., Sayler, K. L., & Barnes, C. A. (2010). Addressing foundational elements of regional land-use change forecasting. *Landscape Ecology, 25*(2), 233–247.
Sohl, T. L., Sayler, K. L., Bouchard, M. A., Reker, R. R., Friesz, A. M., Bennett, S. L., ... Van Hofwegen, T. (2014). Spatially explicit modeling of 1992–2100 land cover and forest stand age for the conterminous United States. *Ecological Applications, 24*(5), 1015–1036.
Sohl, T. L., Sayler, K. L., Drummond, M. A., & Loveland, T. R. (2007). The FORE-SCE model: A practical approach for projecting land cover change using scenario-based modeling. *Journal of Land Use Science, 2*(2), 103–126.
Sohl, T. L., Wimberly, M. C., Radeloff, V. C., Theobald, D. M., & Sleeter, B. M. (2016). Divergent projections of future land use in the United States arising from different models and scenarios. *Ecological Modelling, 337*, 281–297.
Song, Y., & Knaap, G. (2004). Measuring the effects of mixed land uses on housing values. *Regional Science and Urban Economics, 34*(6), 663–680.
Sunde, M. G., He, H. S., Zhou, B., Hubbart, J. A., & Spicci, A. (2014). Imperviousness

Change Analysis Tool (I-CAT) for simulating pixel-level urban growth. *Landscape and Urban Planning, 124*, 104–108.
Tian, G., Qiao, Z., & Gao, X. (2014). Rural settlement land dynamic modes and policy implications in Beijing metropolitan region China. *Habitat International, 44*, 237–246.
Tovar, A., Patel, N. M., Niebur, G. L., Sen, M., & Renaud, J. E. (2006). Topology optimization using a hybrid cellular automaton method with local control rules. *Journal of Mechanical Design, 128*(6), 1205–1216.
Verburg, P. H., & Overmars, K. P. (2009). Combining top-down and bottom-up dynamics in land use modeling: Exploring the future of abandoned farmlands in Europe with the Dyna-CLUE model. *Landscape Ecology, 24*(9), 1167–1181.
Verburg, P. H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., & Mastura, S. S. A. (2002). Modeling the spatial dynamics of regional land use: The CLUE-S Model. *Environmental Management, 30*(3), 391–405.
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: Planning and Design, 24*(3), 323–343.
Wu, F. (2002). Calibration of stochastic cellular automata: The application to rural-urban land conversions. *International Journal of Geographical Information Science, 16*(8), 795–818.
Wu, F., & Webster, C. J. (1998). Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B: Planning and Design, 25*(1), 103–126.
Yang, J., Guo, A., Li, Y., Zhang, Y., & Li, X. (2019). Simulation of landscape spatial layout evolution in rural-urban fringe areas: A case study of Ganjingzi District. *GIScience & Remote Sensing, 56*(3), 388–405.
Yang, J., Liu, W., Li, Y., Li, X., & Ge, Q. (2018). Simulating Intraurban Land Use Dynamics under Multiple Scenarios Based on Fuzzy Cellular Automata: A Case Study of Jinzhou District. *Dalian, Complexity, 2018*, 1–17.
Yang, J., Su, J., Chen, F., Xie, P., & Ge, Q. (2016). A local land use competition cellular automata model and its application. *ISPRS International Journal of Geo-Information, 5*(7), 106.
Yang, X., Zheng, X., & Chen, R. (2014). A land use change model: Integrating landscape pattern indexes and Markov-CA. *Ecological Modelling, 283*, 1–7.
Yao, Y., Liu, X., Li, X., Liu, P., Hong, Y., Zhang, Y., & Mai, K. (2017). Simulating urban land-use changes at a large scale by integrating dynamic land parcel subdivision and vector-based cellular automata. *International Journal of Geographical Information Science, 31*(12), 2452–2479.
Yao, Y., Liu, X., Li, X., Zhang, J., Liang, Z., Mai, K., & Zhang, Y. (2017). Mapping fine-scale population distributions at the building level by integrating multisource geospatial big data. *International Journal of Geographical Information Science, 31*(6), 1220–1244.
Yeh, A. G., & Li, X. (2002). A cellular automata model to simulate development density for urban planning. *Environment and Planning B: Planning and Design, 29*(3), 431–450.
Yue, Y., Zhuang, Y., Yeh, A. G. O., Xie, J. Y., Ma, C. L., & Li, Q. Q. (2017). Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. *International Journal of Geographical Information Systems, 31*(4), 658–675.
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. *GIScience & Remote Sensing, 56*(2), 282–308.
Zhai, Y., Yao, Y., Guan, Q., Liang, X., Li, X., Pan, Y., ... Zhou, J. (2020). Simulating urban land use change by integrating a convolutional neural network with vector-based cellular automata. *International journal of geographical information science : IJGIS, 34*(7), 1475–1499.