Article

Not peer-reviewed version

A Spatiotemporal Encoder–Decoder and Competitive Patch-Growth Cellular Automaton for Land Use/Cover Change Simulation

Minghao Liu , Jian Xiong * , Kai Zhou

Posted Date: 23 March 2026

doi: 10.20944/preprints202603.1739.v1

Keywords: land use/cover change (LUCC); spatiotemporal aggregation residual network; competitive patch-growth cellular automaton; Chengdu–Chongqing twin-city economic circle

Article

A Spatiotemporal Encoder–Decoder and Competitive Patch-Growth Cellular Automaton for Land Use/Cover Change Simulation

Minghao Liu [1,2], Jian Xiong [2,3,*] and Kai Zhou [2,3]

[1] College of Artificial Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

[2] Chongqing Engineering Research Center of Spatial Big Data and Intelligent Technology, Chongqing 400065, China

[3] College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

* Correspondence: s240201022@stu.cqupt.edu.cn

Highlights

What are the main findings?

- This study proposes a STARNet-CPGCA framework for LUCC simulation by coupling a spatiotemporal encoder-decoder network with a competitive patch-growth cellular automaton.
- In the Chengdu–Chongqing Twin-City Economic Circle, STARNet-CPGCA improved OA, Kappa, and FoM by 0.0561, 0.0095, and 0.0427, respectively, over DST-CA, with better change-area reconstruction and boundary preservation.

What are the implications of the main findings?

- Integrating STARNet-based spatiotemporal potential generation with CPGCA-driven patch-scale competitive growth improves the realism of LUCC simulation by balancing overall landscape consistency with local change reconstruction.
- The proposed framework provides a reliable basis for territorial spatial planning, land-use optimization, and scenario-based LUCC simulation in rapidly urbanizing regions.

Abstract

Land use/cover change (LUCC) is shaped by the combined effects of natural conditions, socioeconomic activities, and policy interventions over the long term, and therefore exhibits pronounced spatiotemporal heterogeneity. Existing deep learning–CA coupled models still have limited ability to capture spatiotemporal patterns in long-term LUCC simulation. Moreover, conventional CA evolution mechanisms struggle to represent the transition from clustered patch emergence to outward edge expansion, resulting in limited morphological realism in simulated landscapes. To address these issues, this study proposes a STARNet-CPGCA framework for LUCC simulation in the Chengdu–Chongqing Twin-City Economic Circle by integrating the Spatiotemporal Aggregation Residual Network (STARNet) with the Competitive Patch-Growth Cellular Automaton (CPGCA). Within this framework, STARNet is developed as an embedded spatiotemporal encoder–decoder architecture to enhance spatiotemporal feature learning, while weighted feature fusion and refinement modules are incorporated to improve patch-boundary reconstruction. Based on the development potential generated by STARNet, CPGCA is further developed as a competitive patch-growth mechanism to simulate the synchronous competitive expansion of multiple land-use classes and better capture the dynamic process from clustered patch emergence to outward edge growth. The results show that STARNet-CPGCA outperforms ST-CA, KCLP-CA, and DST-CA. Compared with DST-CA, the proposed model improves overall accuracy (OA), the Kappa coefficient, and the figure of merit (FoM) by 0.0561, 0.0095, and 0.0427, respectively. These findings indicate that the proposed model can more effectively simulate change areas while maintaining consistency in the overall landscape pattern, thereby providing more reliable support for regional territorial spatial planning and land-use optimization.

Keywords: land use/cover change (LUCC); spatiotemporal aggregation residual network; competitive patch-growth cellular automaton; Chengdu–Chongqing twin-city economic circle

1. Introduction

Land use/cover change (LUCC) is a key manifestation of how human activities reshape Earth-surface processes and has long been a central topic in global environmental change research [1]. Numerous studies have shown that LUCC alters landscape patterns and the provision of ecosystem services, thereby influencing the carbon cycle, climate regulation, and the security of water and soil resources, with far-reaching implications for regional sustainable development [2,3]. Against the backdrop of rapid urbanization and accelerated socioeconomic development, accurately characterizing and predicting long-term LUCC dynamics to provide quantitative support for territorial spatial planning, ecological conservation, and China's carbon peaking and carbon neutrality goals has become an important research direction in geography and land science [4].

In LUCC simulation, cellular automata (CA) have been widely used for urban expansion and land-use change because they can generate complex spatial patterns from simple local rules [5]. CA describe the diffusion and competition among land-use classes on raster grids through local transition rules, making them suitable for multi-class and multi-scale spatial evolution problems [6]. In CA-based LUCC simulation, well-designed transition rules are critical for achieving high-accuracy results [7]. Land-unit transition rules are typically determined jointly by factors such as land-use type, development potential, and neighborhood effects [8]. These rules not only represent dependencies among spatial units but also specify how the state of each unit at the next time step is computed from its current state and neighborhood information [9]. Therefore, deriving reliable transition rules to enable more accurate LUCC simulation and prediction remains a key objective in this field [10].

Early studies mainly relied on researchers' subjective experience to calibrate transition rules, which was both time-consuming and limited in accuracy [11]. To enhance the representational capacity of transition rules, machine learning methods were gradually coupled with CA. Approaches such as random forests, support vector machines, and logistic regression have been widely used to estimate development potential and then integrated into CA frameworks, significantly improving simulation accuracy in many regions [12,13]. Building on this line of research, the PLUS model proposed the Land Expansion Analysis Strategy (LEAS) and a CA module based on Multiple Random Seeds (CARS). Through random seeding and a decreasing-threshold mechanism, PLUS enhances the emergence of natural patches and improves morphological realism [14]. However, traditional machine-learning approaches mainly extract features at the single-pixel level and tend to ignore nonlinear interactions among neighboring cells [15]. With the development of deep learning, deep learning–CA (DL–CA) models have been increasingly adopted for LUCC simulation. For example, the FLUS model employs an artificial neural network to learn land conversion probabilities from multi-source driving factors, effectively capturing nonlinear relationships among neighboring units, and incorporates adaptive competition and stochastic disturbance mechanisms in CA for multi-scenario land-use simulation [16]. Although models such as FLUS address nonlinear modeling to some extent, they typically extract features from a single temporal snapshot, neglecting the temporal nature of LUCC processes [17]. In essence, LUCC is a dynamic process with pronounced temporal characteristics [18].

Qian et al. employed a three-dimensional convolutional neural network (3DCNN) to capture local short-term spatiotemporal features in LUCC processes by sliding 3D kernels to extract short-term nonlinear interactions among neighboring cells [19]. Geng et al. further treated multi-period driving factors and land-use states as "spatiotemporal volumetric data" and achieved integrated spatiotemporal feature extraction in the ST-CA model using a 3D convolutional network, thereby improving long-term LUCC fitting capability [20]. These 3DCNN-based methods perform well in extracting local short-term spatiotemporal features [21], yet they remain insufficient for characterizing long-term spatiotemporal dependencies in LUCC. Recurrent neural networks (RNN) have therefore been introduced to extract temporal features from LUCC time series [22]. Subsequently, the long short-term memory network (LSTM), a variant of RNN, was used to build LSTM-CA models, demonstrating improved learning of long-term temporal dependencies in urban expansion simulation [23,24]. Yang et al. further combined 3DCNN and LSTM to extract both local short-term and long-term spatiotemporal features for LUCC [25].

To further learn spatiotemporal dependencies and improve the robustness of probability estimation, some studies have explored multi-model fusion strategies. Liu et al. combined CNN and LSTM to extract LUCC features and used a random forest (RF) to select high-quality seeds for LUCC simulation, thereby improving both accuracy and efficiency [26]. Huang et al. divided the study area into geographic subregions using k-means clustering, applied CNN-LSTM to learn within-region spatiotemporal transition rules, and coupled the resulting development potential with the LEAS–CARS mechanism in PLUS, effectively balancing regional heterogeneity, long-range dependence, and patch-level expansion processes [27]. Li et al. further developed a coupled CNN-LSTM–PLUS framework by weighted fusion of spatiotemporal probabilities learned by CNN-LSTM and probabilities computed by RF in PLUS, leveraging the complementary strengths of deep learning and traditional machine learning to improve simulation accuracy in cropland reserve resource areas [28]. Zhou et al. proposed the STUrban model for long-term urban growth simulation, in which RF captures spatial effects, LSTM captures temporal effects, and a 3D spatiotemporal fusion convolutional network integrates both to predict urban expansion [29]. To better capture complex spatial interactions, Xu et al. proposed TL-CA, which uses self-attention to quantify heterogeneous interactions among neighboring cells and thus alleviates the limited representation of spatial heterogeneity in conventional CA models [30]. Gui et al. argued that relying solely on small convolution kernels restricts the receptive field and limits contextual feature representation, weakening the ability to model complex urban dynamics; accordingly, they proposed an attention-enhanced U-Net coupled with CA for urban expansion simulation [31].

In summary, studies represented by CNN-LSTM and its coupled variants (e.g., integrations with PLUS and RF) incorporate long-term temporal information but typically adopt stepwise multi-model coupling strategies. Such strategies increase integration complexity and often sacrifice shallow spatial details during deep feature extraction, making high-precision boundary reconstruction difficult. Consequently, simulated results may exhibit overly smooth boundaries and the disappearance of small patches, reflecting insufficient modeling of fine-grained spatial details and thus limiting simulation accuracy. Recent work represented by Xu et al. and Gui et al. alleviates the limited receptive field of conventional convolution and the simplicity of spatial interaction modeling[30,31]; however, these studies mainly focus on enhancing spatial contextual information, while effective mechanisms for handling temporal information in sequential inputs remain inadequate. Moreover, how to better attend to key historical years while maintaining the advantage of large-scale spatial feature extraction remains a critical challenge for improving LUCC simulation accuracy.

In addition, regarding CA iterative mechanisms, PLUS-CARS improves natural emergence to some extent through multiple random seeds and a decreasing threshold [14]. Liu et al. proposed a high-quality seed selection algorithm to reduce the computational cost caused by purely random seed selection [26]. Liu et al. also developed a dual-constraint RF-Patch-CA that uses driving-factor importance to estimate the maximum patch expansion area as a local constraint and historical total expansion as a global constraint, alleviating urban clustering and over-connectivity to some extent [32]. Nevertheless, the core update unit of these CA mechanisms remains primarily at the cell level: patch morphology and patch number are often emergent outcomes jointly determined by local rules and sampling, making it difficult to explicitly control patch generation and expansion at the patch scale. Meanwhile, probability-based random allocation can easily produce jagged boundaries and salt-and-pepper noise. Because patch structure and connectivity are directly related to the structural

doi:10.20944/preprints202603.1739.v1

realism of landscape patterns and the reliability of evaluations based on landscape metrics (e.g., fragmentation and aggregation), improving the capability to preserve patch-scale structures is of great importance.

To address the two core bottlenecks identified above—insufficient spatiotemporal feature modeling and the limitations of conventional CA iteration mechanisms—this study develops a novel coupled framework, termed STARNet-CPGCA, that integrates a spatiotemporal encoder–decoder architecture with a competitive patch-growth cellular automaton. Within this framework, the Spatiotemporal Aggregation Residual Network (STARNet) is designed to improve the learning of spatiotemporal dependencies and enhance the representation of complex land-use transition patterns, while the Competitive Patch-Growth Cellular Automaton (CPGCA) is introduced to overcome the deficiencies of traditional CA evolution in representing patch-level growth and multi-class synchronous expansion. Specifically, STARNet adopts a U-Net backbone with embedded residual blocks (ResBlocks) and introduces ConvLSTM units at the bottleneck layer to explicitly capture long-term evolutionary trends while preserving global spatial structures. This design enables the macroscopic evolution trends learned by ConvLSTM to be effectively integrated with the high-resolution boundaries retained by shallow features, thereby alleviating the common problems of boundary over-smoothing and small-patch disappearance in conventional temporal models and enabling more effective learning of spatial details. In addition, a temporal aggregation (TA) module is incorporated into the lateral skip connections, where an attention mechanism adaptively selects key-year spatiotemporal features to reduce temporal redundancy. At the CA iteration stage, CPGCA is developed to better characterize patch-level dynamic evolution during land-use transitions. Specifically, multi-class patches are initialized using high-suitability cells as seeds, and an iterative "bidding–auction–expansion" process is introduced: each patch bids for cells on its growth front based on integrated probabilities, and cells are allocated to the most competitive patches through auctioning. Under land-demand constraints, this mechanism enables synchronous competitive growth of multiple land-use classes at the patch scale, while providing explicit control over patch scale and boundary smoothness. In this way, patch-level growth and multi-class synchronous competition are jointly represented, resulting in a more realistic simulation of the overall spatial pattern.

2. Materials and Methods

2.1. Study Area and Data

The Chengdu–Chongqing Twin-City Economic Circle is strategically located at the intersection of the Belt and Road Initiative and the Yangtze River Economic Belt, encompassing extensive areas of Sichuan Province and Chongqing Municipality (Figure 1). It is regarded as an important region for promoting high-quality development in western China and for advancing the national dual-circulation strategy. By strengthening economic linkages, factor mobility, and regional coordination, this region plays an increasingly important role in unlocking development potential and enhancing the overall competitiveness of western China. As a core urban agglomeration in western China, the Chengdu–Chongqing Twin-City Economic Circle has undergone rapid urbanization in recent years, accompanied by intense land-use change and pronounced spatiotemporal heterogeneity. These characteristics make it a representative study area for long-term LUCC simulation. Therefore, investigating LUCC dynamics in this area not only helps reveal the evolution patterns and driving mechanisms of land-use change in rapidly urbanizing regions, but also yields findings that may be applicable to other urban agglomerations undergoing similar transitions.

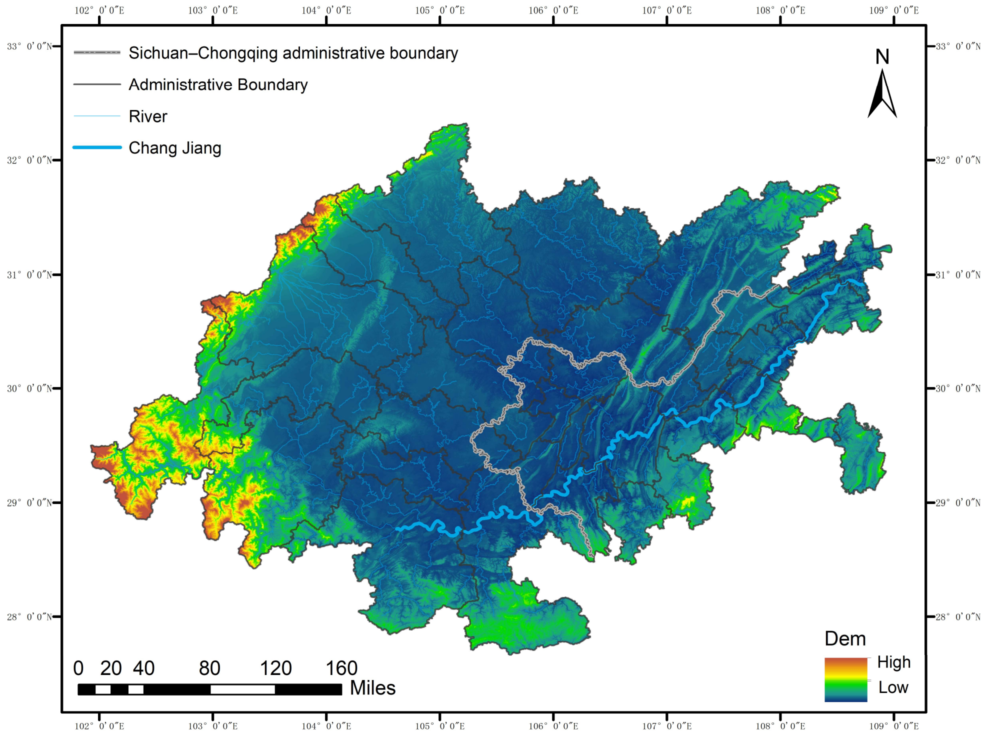

Figure 1. Chengdu–Chongqing Twin-City Economic Circle.

Land-use data were derived from the annual China Land Cover Dataset (CLCD) developed by the team of Huang et al. at Wuhan University [33]. The CLCD provides land-cover information at a spatial resolution of 30 m, with an overall classification accuracy exceeding 80%. In this study, five land-use maps at 5-year intervals from 2000 to 2020 were selected to characterize long-term land-use change. Considering the regional characteristics of the Chengdu–Chongqing Twin-City Economic Circle, which is dominated by mountainous and hilly terrain and has undergone rapid urbanization, as well as the driving mechanisms of long-term LUCC, this study selected a set of natural and socioeconomic driving factors by drawing on the methodological framework of Liu et al. for urban expansion simulation [34]. Specifically, elevation and slope were used as natural drivers to reflect topographic constraints on land-use transformation, while socioeconomic drivers included distance to urban centers, distance to rivers, distance to highways, distance to railways, population density, and nighttime lights, which together capture the effects of accessibility, location conditions, human activities, and urban development intensity on LUCC. Distance-related variables were derived using Euclidean distance analysis based on road, river, and urban-center data, whereas density-related variables were generated using kernel density analysis based on point data. The original data sources are summarized in Table 1, and Figure 2 presents the spatial patterns of the driving factors for 2015 as a representative year.

Table 1. Raw datasets.

Data category	Format	Source
Land-use data	Raster	CLCD（Annual China Land Cover Dataset）
Elevation (DEM)	Raster	https://search.earthdata.nasa.gov
Roads, railways, rivers, etc.	Vector	https://www.openstreetmap.org
Population	Raster	https://www.worldpop.org
Nighttime lights	Raster	http://www.geodata.cn

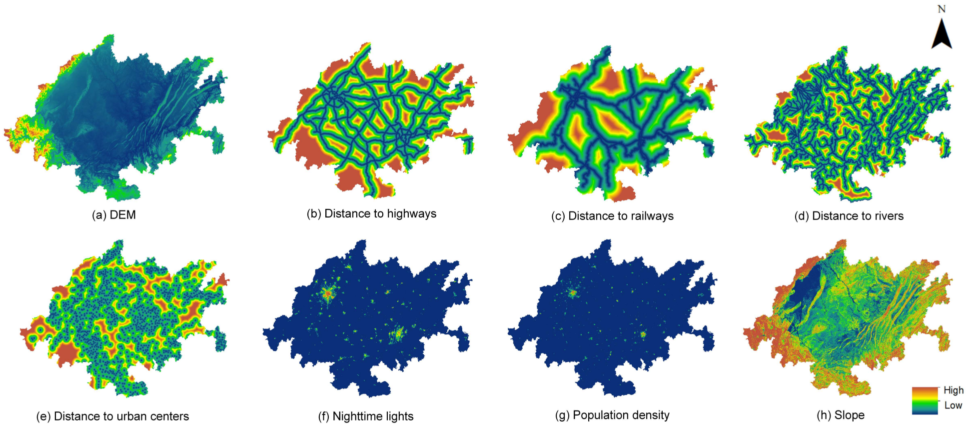

Figure 2. Driving factors used in the model: (a) DEM; (b) distance to highways; (c) distance to railways; (d) distance to rivers; (e) distance to urban centers; (f) nighttime lights; (g) population density; and (h) slope. Warmer colors indicate relatively higher values.

Land-use types were classified into six categories: cropland, forest, grassland, water, built-up land, and unused land. This classification was established with reference to the Current Land Use Classification standard (GB/T 21010-2017), so as to ensure both consistency with national land-use standards and applicability to simulation modeling. To balance simulation accuracy and computational efficiency while ensuring spatial consistency across multi-source data, all datasets were resampled to a spatial resolution of 200 m before model construction.

2.2. Overall Framework of STARNet-CPGCA

The overall workflow of the proposed approach is shown in Figure 3. First, multi-temporal land-use maps and multi-source driving factors are used to construct spatiotemporal input sequences. STARNet then encodes these inputs and performs spatiotemporal modeling to generate pixel-wise development potentials for each land-use class. Next, the potential maps serve as the spatial basis for CA transitions, while land demand in the target year is imposed as a global constraint. Within a unified effective simulation domain, CPGCA initializes patch seeds for each land-use class and conducts synchronous competitive patch growth. The land-use states are then iteratively updated under the demand constraints to generate the simulated land-use map for the target year. Finally, the simulated result is compared with the observed map to validate the effectiveness of the proposed framework.

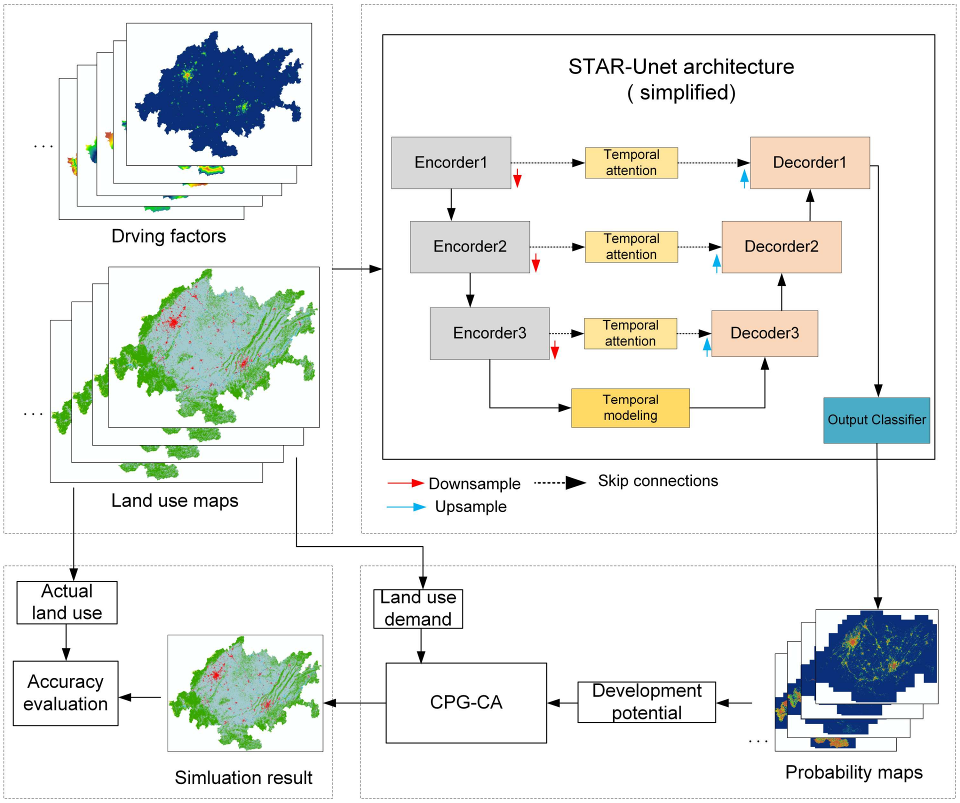

Figure 3. Framework of the STARNet-CPGCA model.

2.3. Development Potential Learning Model

This section introduces the development potential learning component of the proposed framework. Specifically, STARNet is designed to estimate the development potential of different land-use types from multi-temporal land-use maps and multi-source driving factors. Its overall architecture and the key modules, including the temporal aggregation (TA) module and feature refinement head (FRH), are described as follows.

2.3.1. Overall Architecture of STARNet

We propose a deep residual spatiotemporal encoder–decoder network for multi-year LUCC sequences. As shown in Figure 4, the framework uses a U-Net backbone with embedded residual blocks (ResBlocks), introduces recurrent temporal modeling at the bottleneck, and performs temporal aggregation along multi-scale skip connections. The input is a multi-temporal image sequence stacked by year $\mathrm{X} \in \mathbb{R}^{B\times T\times C\times H\times W}$, where B is the batch size, T is the number of years, C is the number of channels, and H and W denote the spatial size of each patch. Here, $t \in \{1,\dots,T\}$ denotes the temporal index. For each time step, the model shares the same ResNet-style encoder. Residual connections alleviate the vanishing-gradient problem and enable extraction of deeper multi-scale spatial feature sequences $\{x_t^{(s)}\}_{t=1}^{T}$, where $s \in \{1,2,3,4\}$ denotes the encoder scale corresponding to resolutions H, $H/2$, $H/4$, *and* $H/8$. Except for the first scale, which is kept at the original resolution, max pooling is used for spatial downsampling before entering the next encoder stage.

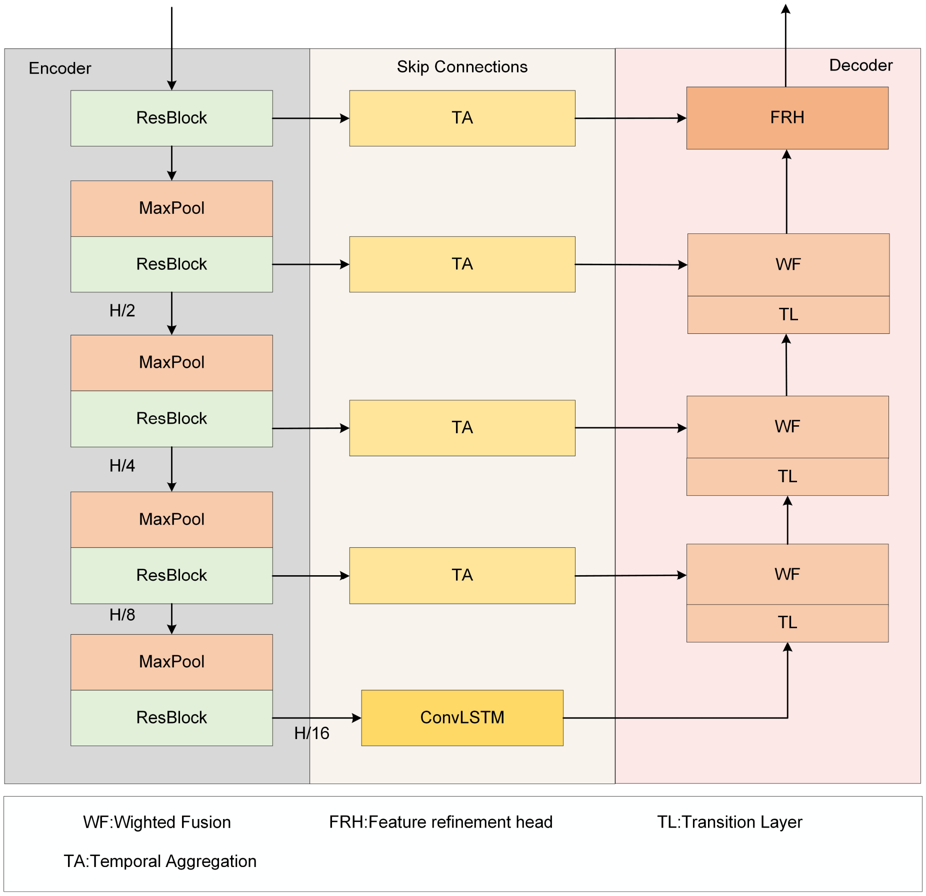

Figure 4. Architecture of STARNet.

At the deepest encoder stage, the bottleneck representation with the strongest semantics and lowest resolution is obtained as $x_t^{(5)} \in \mathbb{R}^{B\times C_5\times\left(\frac{H}{16}\right)\times\left(\frac{W}{16}\right)}$. To capture long-term dependencies, a ConvLSTM is applied at the bottleneck:

$$(h_t, c_t) = \mathrm{ConvLSTM}\left(x_t^{(5)}, (h_{t-1}, c_{t-1})\right) \tag{1}$$

The final hidden state h_T is used to initialize the decoder.

During decoding, to effectively exploit multi-scale details extracted by the encoder, the skip features from the four scales are first passed through a temporal aggregation module to adaptively fuse the time dimension into a single detail representation. To address channel- and semantic-level mismatches between deep semantic features and shallow detail features, a transition layer (TL) is introduced at each decoding stage. Specifically, a 1×1convolution aligns channels of the feature from the previous decoder stage to match the channel width at the current scale, after which hierarchical reconstruction is conducted using a weighted fusion (WF) module. The WF module employs learnable parameters to adaptively balance the contributions from deep and shallow features. For the k-th decoding stage, the inputs include the upsampled deep feature z_k from the previous stage and the skip feature $\tilde{x}^{(5-k)}$ at the same scale. The WF module first upsamples z_k to the target resolution using bilinear interpolation and maps $\tilde{x}^{(5-k)}$ to an aligned channel dimension via convolution. It then introduces learnable weights w and generates fusion coefficients α_1, α_2 through ReLU activation and normalization:

$$\alpha_i = \frac{\mathrm{ReLU}(w_i)}{\sum_{j=1}^{2} \mathrm{ReLU}\left(w_j\right) + \epsilon}, i \in \{1,2\} \tag{2}$$

The fused feature z_{k-1} is computed as a weighted sum followed by a 3×3 convolution for smoothing:

$$z_{k-1} = \mathrm{Conv}_{3\times3}\left(\alpha_1 \cdot \mathrm{Conv}_{1\times1}\left(\tilde{x}^{(5-k)}\right) + \alpha_2 \cdot \mathrm{Up}(z_k)\right) \quad (3)$$

In this way, the WF module dynamically adjusts the contribution of multi-level features according to their importance, mitigating redundancy and improving reconstruction. After multi-stage WF-based upsampling restores spatial resolution, a feature refinement head (FRH) replaces the conventional final convolutional output. By leveraging both spatial and channel attention, FRH further sharpens boundaries and calibrates class responses, and the final pixel-wise prediction Y is produced by a 1×1 convolution.

2.3.2. Temporal Aggregation (TA)

We design a temporal aggregation module on multi-scale skip connections. The goal of TA is to fuse, along the temporal dimension, the cross-year feature sequence at the same scale into a single "time-aggregated" feature map for skip connections in the decoder. By explicitly performing temporal modeling on the skip branches, TA selects information from key years before spatial-detail recovery, thereby improving boundary and connectivity reconstruction under long time-series inputs. As illustrated in Figure 5, let the feature tensor at a given scale be $F \in \mathbb{R}^{B\times T\times C\times H\times W}$. TA adopts an attention-weighted aggregation strategy. For each time step t, global average pooling is first applied to $F_{:,t}$ to obtain a channel-wise descriptor g_t. Next, two 1×1 convolution layers with ReLU activation are used to generate a scalar score s_t. A softmax operation is then applied along the temporal dimension to obtain the attention weight α_t. Finally, the output $\tilde{x} \in \mathbb{R}^{B\times C\times H\times W}$ is computed as the weighted sum of frame-wise features over time, enabling the module to adaptively emphasize information from "key years."

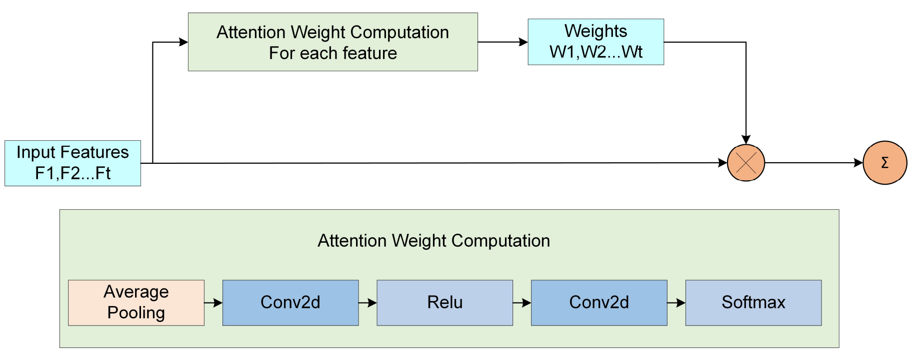

Figure 5. Temporal Aggregation.

2.3.3. Feature Refinement Head (FRH)

Although the multi-stage WF modules restore the feature maps to the original resolution, repeated downsampling in deep networks inevitably smooths spatial details, making small objects and complex boundaries less sharp. To mitigate this issue, we introduce a feature refinement head (FRH) at the end of the decoder. Inspired by the joint modeling of channel and spatial attention [35], FRH aims to recalibrate feature responses via dual attention, facilitating the final transition from coarse semantics to fine-grained classification. As shown in Figure 6, FRH adopts a parallel structure with two branches. In the channel path, a global average pooling operation is applied to the input feature to generate a channel-wise attention map $Mc \in \mathbb{R}^{1\times1\times C}$, where C denotes the number of feature channels. The resulting descriptor is then processed by a reduce–expand operation composed of two 1×1 convolutional layers, which first reduce the channel dimension by a factor of four and subsequently restore it to the original size. In parallel, the spatial path applies a depthwise convolution to produce a spatial attention map $Ms \in \mathbb{R}^{H\times W\times 1}$, where H and W denote the spatial

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1739.v1

resolution of the feature map. The attention features generated by the two paths are then combined through an element-wise summation operation. This design remains lightweight while explicitly improving boundary integrity and class separability, thereby enhancing the pixel-level accuracy of the final land-use map.

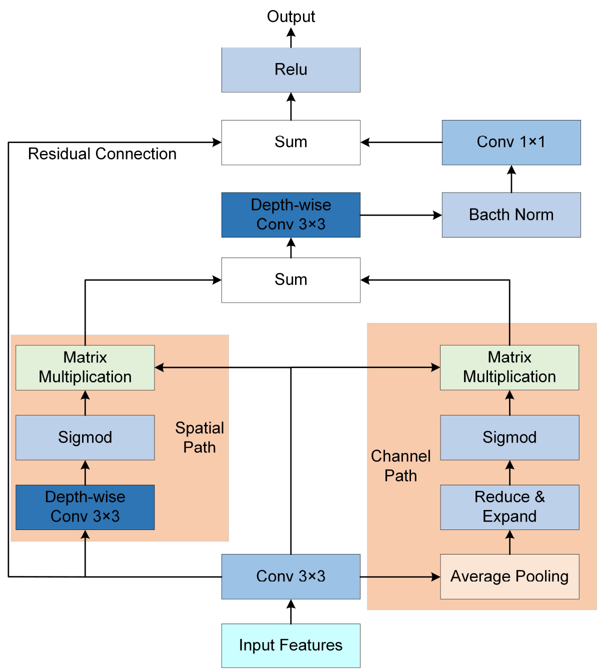

Figure 6. Feature refinement head.

2.4. *Competitive Patch-Growth Cellular Automaton*

We develop a land-use simulation framework based on a competitive patch-growth cellular automaton (CPGCA). As shown in Figure 7, the workflow consists of four stages: (1) unifying the effective simulation domain, (2) patch seeding initialization, (3) competitive patch expansion, and (4) stable convergence-based termination.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1739.v1

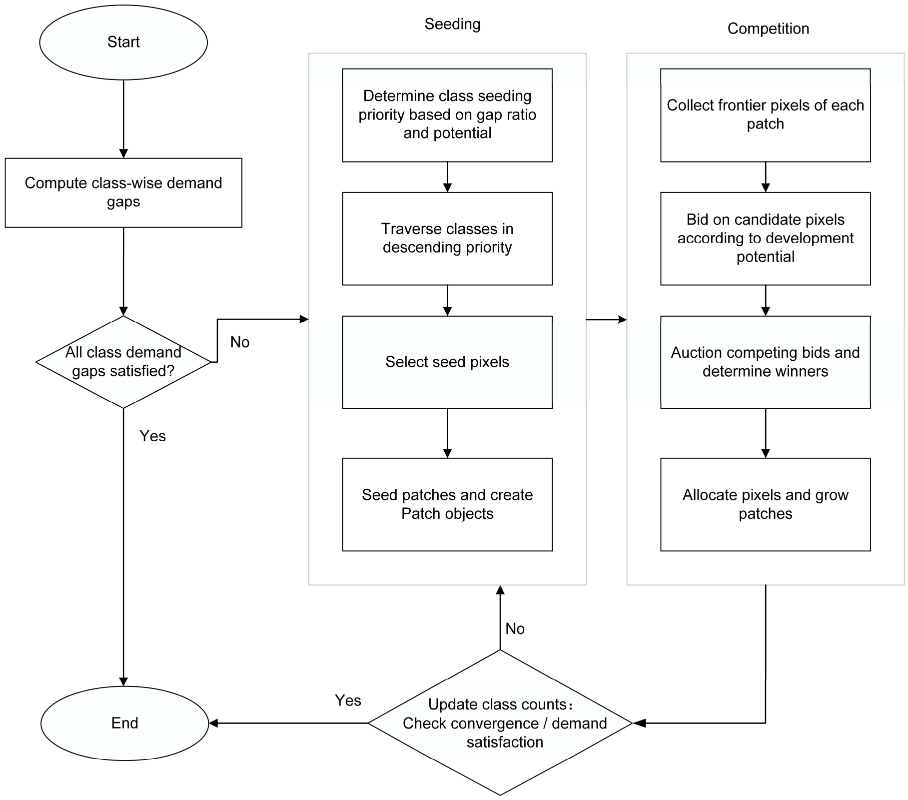

Figure 7. Workflow of the proposed CA method.

(1) Unified effective simulation domain:

The land-use raster in the initial year is used as the CA initial state. An initial mask is first constructed based on valid pixels in the baseline land-use map. The valid area of the development-potential maps predicted by the deep learning model is then intersected with the initial mask to obtain a unified effective simulation domain. The class-wise transition potential maps are stacked along the class dimension to form a multi-class potential field, which serves as the common driving force for both seeding and growth.

(2) Seeding stage:

Initializing patches from high-potential areas. At each CA iteration, we count the current number of pixels for each class, denoted by N_c, and compute the demand gap with respect to the target demand T_c (where c is the number of classes; $c = 6$ in this study):

$$gap_c = T_c - N_c \tag{4}$$

If all classes satisfy $gap_c \leq 0$, the iteration can terminate early. To determine the seeding order, we jointly consider the relative gap ratio and the potential strength of each class. Specifically, the gap ratio is computed as

$$r_c = \mathrm{clip}(gap_c / T_c) \tag{5}$$

where P_c denotes the mean value of the class potential map. Classes are then prioritized in descending order of $r_c \cdot P_c$.

Let $L(y,x)$ denote the land-use class label at pixel (y,x) in the current iteration, where $L(y,x) \in \{1,2,\dots,c\}$. Let $valid(y,x) \in \{0,1\}$ denote the unified domain mask, where $valid(y,x) = 1$ indicates that the pixel participates in CA evolution. For any target class c, the seeding potential is denoted by $P_c^{seed}(y,x)$. Seeding is only performed on pixels that are not currently class c, i.e., $L(y,x) \neq c$. The candidate set for class c is defined as:

 doi:10.20944/preprints202603.1739.v1

$$\Omega_c = \{(y,x) \mid P_c^{seed}(y,x) > 0, valid(y,x) = 1, L(y,x) \neq c\} \quad (6)$$

From Ω_c, the top 2% pixels ranked by $P_c^{seed}(y,x)$ (at least one pixel) are selected as seeds, and a patch is created for each seed. Each patch is associated with a unique land-use class. A patch state is described by its occupied pixel set and its growth front. The growth front is defined as the 4-neighborhood boundary pixels that (i) lie within the effective domain and (ii) have not been assigned in the current sub-iteration. During seeding, the gaps gap_c are updated accordingly. Once $gap_c \leq 0$, all patches of class $c$enter a dormant state and withdraw from subsequent competition.

(3) Competitive growth stage:

Bidding–auction–allocation on growth fronts. Each active patch attempts to expand on its growth front. For a frontier pixel (y,x), the bid of a patch of class c is defined as its growth potential P_c^{grow} (y,x). Bids from competing patches are compared at the pixel level, and the patch with the highest bid wins and is assigned the pixel. After allocation, the land-use map, the occupied sets, and the growth fronts are synchronously updated, and the corresponding gap gap_c is decreased. To control expansion step length and prevent excessive growth within a single sub-round, a maximum allocation limit (an empirical threshold) is imposed for each round. When $gap_c \leq 0$ for a class, all patches of that class become dormant and stop bidding, ensuring that competition focuses on classes with remaining demand gaps.

To prevent neighborhood effects from forcing occupation of pixels that are clearly more suitable for other classes, and to reflect cases where a class decreases overall but may still convert locally, we introduce a forced assignment rule to the best class. For each pixel, we first compute the maximum growth potential across classes:

$$P_{max}(y,x) = \max_c P_c^{growth}(y,x) \quad (7)$$

A patch of class c is allowed to bid at (y, x) only if

$$P_c^{growth}(y,x) \geq \theta \cdot P_{max}(y,x) \quad (8)$$

where $\theta \in (0,1]$ is an empirical threshold. If the condition is not met, bidding is prohibited and the pixel is directly assigned to the class with the maximum potential. Moreover, for non-gap classes, a stricter threshold is adopted, allowing only limited "compensatory" conversions at high-confidence pixels, thereby balancing overall quantity trends and local spatial processes. A schematic illustration of the competitive patch-growth mechanism is provided in Figure 8.

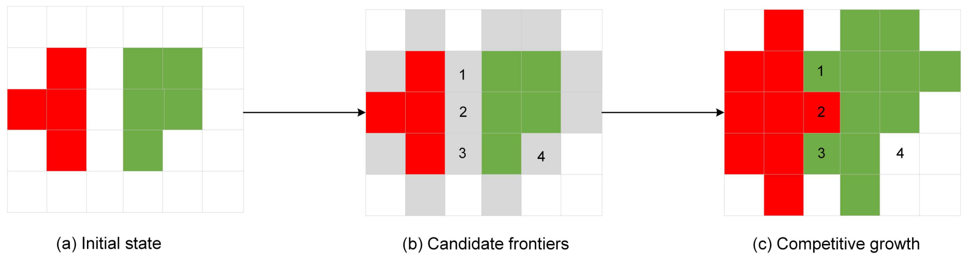

Figure 8. Schematic diagram of competitive patch growth: (a) initial state of two land-use patches (red and green); (b) extraction of candidate growth frontiers (gray pixels); and (c) competitive growth results. Pixel 2 is occupied by the red patch, pixels 1 and 3 are occupied by the green patch, and pixel 4 triggers the forced assignment rule and is directly assigned to the land category with the highest potential.

(4) Termination rule:

A competitive-growth sub-round stops when any of the following conditions hold: (i) the growth fronts of all active patches are empty; or (ii) the number of actually allocated pixels is zero for consecutive sub-rounds. The land-use map is then updated and class counts are recomputed before

 doi:10.20944/preprints202603.1739.v1

entering the next CA iteration. The CA process terminates and outputs the target-year simulation when class quantities meet the target demands, or when the class-count statistics remain unchanged over multiple consecutive rounds.

2.5. Experimental Design and Evaluation Metrics

2.5.1. Experimental Design

Land-use maps from 2000–2015 and the corresponding driving factors were used for model training. Specifically, land-use data from 2000–2010 were used as the input sequence, and the 2015 land-use map served as the supervisory label. To match the input requirements of the deep network and increase the number of training samples, the study-area rasters were first cropped into patches using a 256×256 window. To better preserve information from boundaries and transition zones while reducing information loss caused by sample redundancy, a sliding-window strategy with 50% overlap was adopted to generate sample patches. The samples were then split into training and validation sets at a ratio of 70%/30% for parameter learning and generalization assessment. During inference and simulation, multi-temporal land-use maps from 2005–2015 and the corresponding driving factors were fed into the trained model to produce development potential maps for 2020. These potential maps, together with land-demand constraints, were used as the spatial suitability basis for CA evolution, and the target-year land-use pattern was generated through CA iterations.

2.5.2. Evaluation Metrics

In land-use simulation, overall accuracy (OA), the Kappa coefficient, and the figure of merit (FoM) are commonly used to evaluate prediction performance [36]. The Kappa coefficient measures the agreement between simulated and observed maps. It ranges from -1 to 1: a value of 1indicates perfect agreement, 0 indicates no agreement beyond chance, and negative values indicate agreement worse than random guessing. Kappa is computed as follows:

$$P_0 = \sum_{i=1}^{6} \frac{w_i}{n} \tag{9}$$

$$P_c = \sum_{i=1}^{6} \frac{a_i \times b_i}{n^2} \tag{10}$$

$$kappa = \frac{P_0 - P_c}{1 - P_c} \tag{11}$$

where i denotes the land-use class; w_i is the number of pixels with observed class i that are also predicted as class i; a_i is the number of pixels of class i in the observed map; b_i is the number of pixels of class i in the predicted map; and n is the total number of pixels.

OA quantifies the overall proportion of correctly predicted pixels in the study area. It is calculated as the number of correctly predicted pixels divided by the total number of pixels; thus, a higher OA indicates better overall predictive accuracy. OA is computed as follows:

$$OA = \frac{G_p}{G_{all}} \times 100\% \tag{12}$$

where G_p is the number of correctly predicted pixels and G_{all} is the total number of pixels in the study area.

FoM is a quality metric that is particularly suitable when land-use classes are imbalanced. By jointly considering omission and commission components, FoM provides a more informative evaluation of performance under class imbalance and has been widely used in LUCC simulation studies [37]. FoM is computed as follows:

$$FoM = \frac{B}{A + B + C + D} \tag{13}$$

where A is the number of pixels that changed in reality but were predicted as unchanged; B is the number of pixels that changed in reality and were correctly predicted as changed; C is the number of pixels that changed in reality but were predicted as the wrong land-use class; and D is the number of pixels that did not change in reality but were predicted as changed.

3. Results

3.1. Ablation Study

To quantify the contribution of each component in STARNet-CPGCA, a series of ablation experiments was conducted. The simulation results of different module combinations are summarized in Table 2. For consistency, the baseline CA adopted a pixel-level sequential allocation strategy, in which land-use types were allocated in descending order of demand gap until the predefined land demand was satisfied or the iteration converged.

Table 2. Simulation accuracy of the ablation experiments.

Model	OA	Kappa	FOM
CNNLSTM-CA	0.8842	0.7717	0.1806
CNNLSTM-TA-CA	0.8864	0.7760	0.2161
UnetConvLSTM-CA	0.8824	0.7716	0.2521
UnetConvLSTM-FRH-CA	0.8844	0.7716	0.2566
STARNet-CA	0.8845	0.7739	0.2652
STARNet-CARS	0.8543	0.7140	0.2681
STARNet-CPGCA	0.9423	0.7852	0.2709

Using CNNLSTM-CA as the baseline, the model achieved OA, Kappa, and FoM values of 0.8842, 0.7717, and 0.1806, respectively. After introducing the temporal aggregation (TA) module, the model performance improved, and the FoM increased to 0.2161, indicating that adaptive aggregation along the temporal dimension is beneficial for simulating changed areas under multi-temporal inputs. By further embedding ConvLSTM into the bottleneck layer of U-Net, the resulting UnetConvLSTM-CA increased the FoM to 0.2521, showing that the encoder-decoder structure is more effective in reconstructing changed areas. After additionally incorporating the FRH module, UnetConvLSTM-FRH-CA achieved an OA of 0.8844 and a FoM of 0.2566, suggesting that feature refinement at the end of the decoder helps improve local prediction quality. The complete deep-learning component, STARNet-CA, further improved the performance, with the FoM reaching 0.2652.

On this basis, to further examine the effect of different CA mechanisms, the basic pixel-level sequential CA in STARNet-CA was replaced by CARS and the proposed CPGCA, respectively. The results show that STARNet-CARS slightly improved the FoM to 0.2681, but the OA and Kappa decreased to 0.8543 and 0.7140, respectively, indicating that although this mechanism slightly enhanced change detection, it weakened the overall spatial consistency. In contrast, STARNet-CPGCA achieved the best performance in all three metrics, with OA, Kappa, and FoM values of 0.9423, 0.7852, and 0.2709, respectively.

In summary, the ablation experiments demonstrate that all components in the proposed framework positively contributed to the final simulation performance. Specifically, STARNet improved the accuracy of development potential learning, while CPGCA further enhanced the quality of spatial allocation, enabling the model to achieve better overall pattern consistency and stronger capability in identifying changed areas.

3.2. Comparative Experiments

 doi:10.20944/preprints202603.1739.v1

To evaluate the effectiveness of the proposed STARNet-CPGCA framework in LUCC simulation, several representative conventional and recent models were selected for comparison. To ensure a fair comparison, all competing methods used the same input data and land demand constraints. In addition, except for STARNet-CPGCA, the other methods were implemented within the same pixel-level sequential allocation CA framework. Therefore, the differences among the competing baseline models mainly reflect their ability to learn development potentials, whereas the performance of STARNet-CPGCA reflects the combined effects of improved potential learning and the proposed patch-level allocation mechanism. Among them, RF-CA represents a conventional paradigm in which transition rules are learned by traditional machine learning and then coupled with CA for spatial allocation.

The quantitative results for the 2020 land-use simulation are reported in Table 3. RF-CA achieved an OA of 0.8834 and a Kappa of 0.7701, indicating that it can reconstruct the overall land-use pattern with reasonable stability. However, its FoM was only 0.1357, suggesting limited ability to identify actual change areas. After introducing 3D convolution, ST-CA improved the FoM to 0.2190, indicating that joint modeling of local spatiotemporal features can enhance the identification of changed areas, although its OA and Kappa were slightly lower than those of RF-CA. KCLP-CA further combined zoning strategies with CNN-LSTM and the PLUS model, achieving OA, Kappa, and FoM values of 0.8864, 0.7731, and 0.2084, respectively. Although its overall performance was improved compared with RF-CA and ST-CA, the enhancement in change-area simulation remained limited. DST-CA further improved change-area simulation by combining 3DCNN and LSTM to extract short-term and long-term temporal features in stages, resulting in a FoM of 0.2282, with OA and Kappa values of 0.8862 and 0.7757, respectively. In contrast, STARNet-CPGCA achieved the best performance in all three metrics, with OA, Kappa, and FoM values of 0.9423, 0.7852, and 0.2709, respectively, demonstrating higher overall consistency and stronger capability in identifying changed areas.

Table 3. Comparison of simulation accuracy among different models.

Model	OA	Kappa	FOM
RF-CA	0.8834	0.7701	0.1357
ST-CA[20]	0.8817	0.7668	0.2190
KCLP-CA[27]	0.8864	0.7731	0.2084
DST-CA[25]	0.8862	0.7757	0.2282
STARNet-CPGCA	0.9423	0.7852	0.2709

Figure 9 presents the simulated results of different models in a representative local area. As shown in the figure, in the mixed land-use region containing cropland, forest, grassland, built-up land, water, and unused land, the observed map includes a small built-up patch with a clear boundary in the upper-left corner. STARNet-CPGCA successfully reproduced both the location and shape of this patch, whereas RF-CA and DST-CA failed to generate a corresponding built-up patch in the same area. In addition, several scattered cropland and unused-land patches can be observed in the enlarged local region, forming a distinct mosaic pattern. Compared with the other models, STARNet-CPGCA better preserved these scattered patches and their boundary details, whereas the simulated patterns produced by the other methods were relatively smoother and showed different degrees of simplification in local category structures.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1739.v1

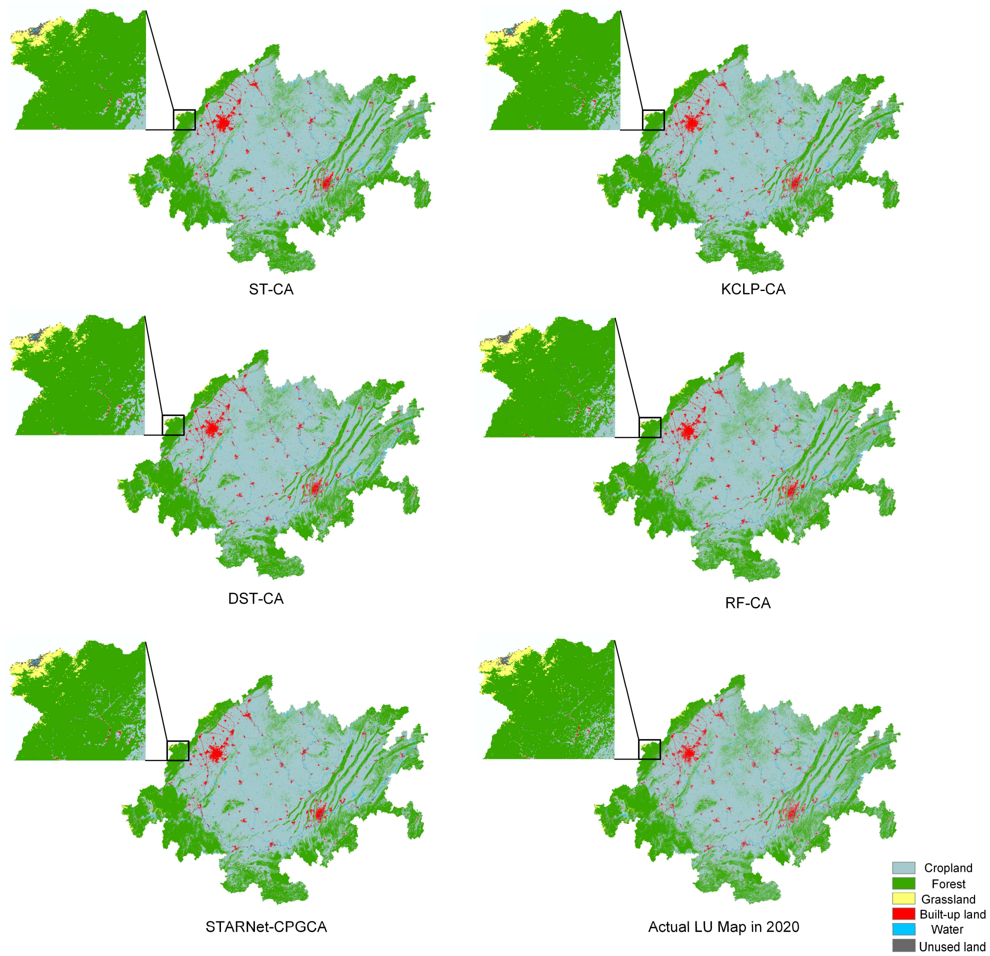

Figure 9. Comparison between the simulated and observed land-use maps in 2020.

Overall, the proposed model not only outperformed the competing methods in quantitative metrics, but also produced more realistic local spatial patterns. These results indicate that STARNet-CPGCA can improve overall simulation accuracy while more effectively identifying change areas and preserving land-use patch structures.

4. Discussion

4.1. Advantages of the Proposed Model

Compared with representative LUCC simulation models, the proposed STARNet-CPGCA framework shows several notable advantages in both spatiotemporal representation and spatial allocation.

First, the results suggest that the proposed framework is more capable of preserving long-term temporal dependencies while maintaining multi-scale spatial information. The comparative experiments indicate that methods relying mainly on conventional machine learning or short-term local spatiotemporal modeling still have difficulty simultaneously capturing long-term land-use evolution and fine-grained spatial heterogeneity. Even when both short-term and long-term temporal information are considered, module-wise serial feature extraction may still weaken the interaction between deep temporal representations and high-resolution spatial details. In contrast, the performance of STARNet-CPGCA suggests that combining ConvLSTM-based long-sequence modeling with temporal aggregation in skip connections helps retain key historical information during multi-scale reconstruction. This may help explain why the proposed framework performs

 doi:10.20944/preprints202603.1739.v1

better in changed-area identification while also preserving boundary details and local spatial heterogeneity.

Second, the proposed framework places greater emphasis on patch-level competitive growth during spatial allocation, which appears to be important for generating more realistic land-use patterns. In many conventional CA-based LUCC simulation approaches, land-use change is still dominated by pixel-level iterative updating, and patch structures emerge implicitly from local neighborhood transitions. Although such strategies can produce reasonable overall patterns, they are more likely to generate isolated conversions, fragmented patches, and overly smoothed boundaries in areas with strong class competition. By contrast, the results indicate that the CPGCA mechanism can represent patch generation, inter-class competition, and outward growth more explicitly through the seed initialization–competitive bidding–patch growth process. This patch-oriented allocation strategy may reduce noisy pixel-level transitions and improve boundary continuity, thereby producing simulated landscapes that are more consistent with observed patch morphology and local spatial organization.

Third, the proposed framework appears to achieve a better balance between change-area identification and overall landscape consistency. A persistent challenge in LUCC simulation is that some models may maintain relatively high overall agreement while failing to capture actual changed areas, whereas others may improve the detection of changed pixels at the cost of introducing excessive false allocations and damaging the global spatial pattern. The results of this study indicate that STARNet-CPGCA can improve the representation of changed areas without sacrificing the consistency of the overall simulated landscape. This balanced performance is particularly important for practical LUCC applications, such as territorial spatial planning and land-use optimization, where both reliable global patterns and realistic local changes are required.

Overall, the main advantage of the proposed framework lies not only in the improvement of quantitative metrics, but also in its coordinated enhancement of spatiotemporal feature learning and patch-scale spatial evolution. The results indicate that STARNet strengthens temporal representation and multi-scale spatial reconstruction, while CPGCA improves the realism of patch-level allocation. Together, these two aspects may help explain the improved temporal rationality, spatial realism, and practical value of the simulation results produced by STARNet-CPGCA.

4.2. Limitations and Future Work

Although the proposed approach achieves strong overall accuracy and improved change detection, learning and simulating rare land-use classes remains challenging, particularly for unused land. This class accounts for a small proportion of both area and change within the six-class system (approximately 1% of the changed area in the study region), resulting in limited effective transition samples during training. Its spatially scattered distribution further increases the difficulty of learning stable representations.

To mitigate class imbalance, we applied class-weighted losses based on class proportions during training to reduce dominance of major classes and enhance gradient contributions from minority classes. However, as indicated by the confusion matrix derived from the comparison between simulated and observed maps for 2020 (Figure 10), unused land is still more likely to be confused with adjacent or more prevalent classes. This suggests that conventional supervised learning with simple class reweighting is insufficient to fully capture the transition behavior of rare classes.

 doi:10.20944/preprints202603.1739.v1

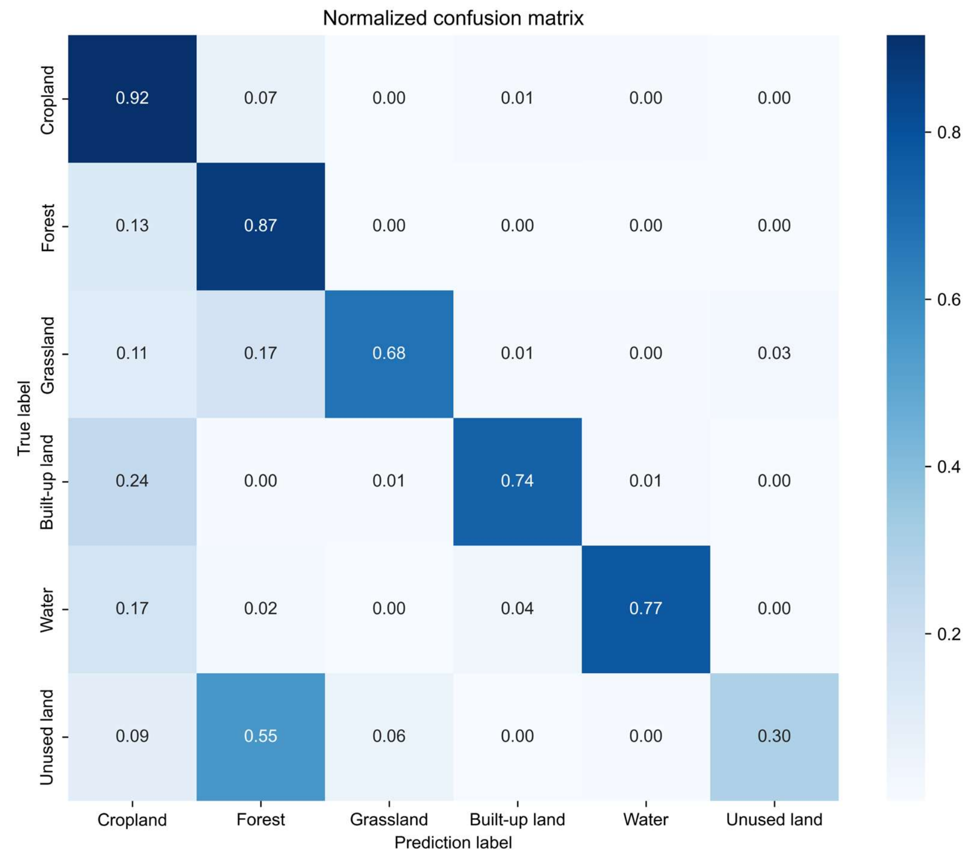

Figure 10. Confusion matrix of the predicted results.

Future work can be advanced along two main directions. First, more process-relevant constraints and driving information (e.g., soil/lithology and indicators of human activity intensity) can be incorporated to improve the separability of rare classes. Second, training strategies and evolutionary constraints tailored to rare classes should be explored to enhance the stability and generalizability of rare-class simulation.

5. Conclusions

Focusing on multi-temporal land-use data and multi-source driving factors in the Chengdu–Chongqing Twin-City Economic Circle, this study proposes a LUCC simulation framework, STARNet-CPGCA. By integrating deep-learning-based potential generation with competitive patch-growth cellular automaton, the framework simulates the target-year land-use pattern under land-demand constraints. The main contributions are summarized as follows:

(1) STARNet for spatiotemporal potential modeling. STARNet learns land-use transition potentials from multi-temporal land-use sequences and driving factors, improving the identification of change areas.

(2) CPGCA for patch-centric competitive growth. In the CA stage, a patch-based “competition–growth” process is introduced to enable synchronous expansion of multiple land-use classes and landscape evolution. This mechanism enhances the representation of patch morphology and improves the preservation of boundary details.

(3) Comprehensive validation in the study area. Comparative experiments in the Chengdu–Chongqing Twin-City Economic Circle demonstrate that STARNet-CPGCA achieves higher simulation accuracy than RF-CA, ST-CA, KCLP-CA, and DST-CA, with stronger characterization of change areas and better preservation of patch-boundary details.

Despite these advances, rare classes remain a major challenge in LUCC simulation. Limited effective change samples make their transition patterns difficult to learn; even with class-weighted training, the confusion matrix still indicates notable misclassification of unused land. Future work will focus on incorporating more process-relevant factors for rare classes, using higher-resolution land-use data to better represent small patches and boundary details, and improving training strategies and evolutionary constraints to enhance stable recognition of rare classes and the model's generalization in complex regions.

Author Contributions: Conceptualization, J.X.; methodology, J.X.; resources, K.Z.; data curation, K.Z.; writing—original draft preparation, J.X.; writing—review and editing, M.L.; supervision, M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used in this study were obtained from publicly available third-party sources cited in the main text, including CLCD, NASA Earthdata, OpenStreetMap, WorldPop, and the National Earth System Science Data Center. These datasets are available from the original providers subject to their respective access procedures and licensing terms. The source code of the proposed model is publicly available at GitHub: https://github.com/XJ2233/LUCC.git.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. He, C.; Zhang, J.; Liu, Z.; Huang, Q. Characteristics and progress of land use/cover change research during 1990–2018. *Journal of Geographical Sciences* **2022**, *32*, 537-559, doi:10.1007/s11442-022-1960-2.
2. Dadashpoor, H.; Azizi, P.; Moghadasi, M. Land use change, urbanization, and change in landscape pattern in a metropolitan area. *Science of the Total Environment* **2019**, *655*, 707-719, doi:10.1016/j.scitotenv.2018.11.267.
3. Foley, J.A.; DeFries, R.; Asner, G.P.; Barford, C.; Bonan, G.; Carpenter, S.R.; Chapin, F.S.; Coe, M.T.; Daily, G.C.; Gibbs, H.K.; et al. Global consequences of land use. *Science* **2005**, *309*, 570-574, doi:10.1126/science.1111772.
4. He, C.; Shi, P.; Chen, J.; Li, X.; Pan, Y.; Li, J.; Li, Y.; Li, J. Developing land use scenario dynamics model by the integration of system dynamics model and cellular automata model. *Science in China Series D: Earth Sciences* **2005**, *48*, 1979-1989, doi:10.1360/04yd0248.
5. Halmy, M.W.A.; Gessler, P.E.; Hicke, J.A.; Salem, B.B. Land use/land cover change detection and prediction in the north-western coastal desert of Egypt using Markov-CA. *Applied Geography* **2015**, *63*, 101-112, doi:10.1016/j.apgeog.2015.06.015.
6. Feng, Y.; Qi, Y. Modeling patterns of land use in Chinese cities using an integrated cellular automata model. *ISPRS International Journal of Geo-Information* **2018**, *7*, 403, doi:10.3390/ijgi7100403.
7. Cao, M.; Tang, G.; Shen, Q.; Wang, Y. A new discovery of transition rules for cellular automata by using cuckoo search algorithm. *International Journal of Geographical Information Science* **2015**, *29*, 806-824, doi:10.1080/13658816.2014.999245.
8. Wu, F. Calibration of stochastic cellular automata: the application to rural-urban land conversions. *International Journal of Geographical Information Science* **2002**, *16*, 795-818, doi:10.1080/13658810210157723.
9. Wolfram, S. Statistical mechanics of cellular automata. *Reviews of Modern Physics* **1983**, *55*, 601-644, doi:10.1103/RevModPhys.55.601.
10. Zhang, J.; Wu, D.; Zhu, A.X.; Zhu, Y. Modelling urban expansion with cellular automata supported by urban growth intensity over time. *Annals of GIS* **2023**, *29*, 337-353, doi:10.1080/19475683.2022.2079574.
11. Clarke, K.C.; Hoppen, S.; Gaydos, L. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay Area. *Environment and Planning B: Planning and Design* **1997**, *24*, 247-261, doi:10.1068/b240247.
12. Kamusoko, C.; Gamba, J. Simulating urban growth using a random forest–cellular automata (RF-CA) model. *ISPRS International Journal of Geo-Information* **2015**, *4*, 447-470, doi:10.3390/ijgi4020447.

 doi:10.20944/preprints202603.1739.v1

13. Yang, Q.; Li, X.; Shi, X. Cellular automata for simulating land use changes based on support vector machines. *Computers & Geosciences* **2008**, *34*, 592-602, doi:10.1016/j.cageo.2007.08.003.
14. Liang, X.; Guan, Q.; Clarke, K.C.; Liu, S.; Wang, B.; Yao, Y. Understanding the drivers of sustainable land expansion using a patch-generating land use simulation (PLUS) model: A case study in Wuhan, China. *Computers, Environment and Urban Systems* **2021**, *85*, 101569, doi:10.1016/j.compenvurbsys.2020.101569.
15. He, J.; Li, X.; Yao, Y.; Hong, Y.; Jinbao, Z. Mining transition rules of cellular automata for simulating urban expansion by using the deep learning techniques. *International Journal of Geographical Information Science* **2018**, *32*, 2076-2097, doi:10.1080/13658816.2018.1480783.
16. Liu, X.; Liang, X.; Li, X.; Xu, X.; Ou, J.; Chen, Y.; Li, S.; Wang, S.; Pei, F. A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landscape and Urban Planning* **2017**, *168*, 94-116, doi:10.1016/j.landurbplan.2017.09.019.
17. Li, H.; Liu, Z.; Lin, X.; Qin, M.; Ye, S.; Gao, P. A novel spatiotemporal urban land change simulation model: Coupling transformer encoder, convolutional neural network, and cellular automata. *Journal of Geographical Sciences* **2024**, *34*, 2263-2287, doi:10.1007/s11442-024-2292-1.
18. Zhou, Z.; Chen, Y.; Wang, Z.; Lu, F. Integrating cellular automata with long short-term memory neural network to simulate urban expansion using time-series data. *International Journal of Applied Earth Observation and Geoinformation* **2024**, *127*, 103676, doi:10.1016/j.jag.2024.103676.
19. Qian, Y.; Xing, W.; Guan, X.; Yang, T.; Wu, H. Coupling cellular automata with area partitioning and spatiotemporal convolution for dynamic land use change simulation. *Science of the Total Environment* **2020**, *722*, 137738, doi:10.1016/j.scitotenv.2020.137738.
20. Geng, J.; Shen, S.; Cheng, C.; Dai, K. A hybrid spatiotemporal convolution-based cellular automata model (ST-CA) for land-use/cover change simulation. *International Journal of Applied Earth Observation and Geoinformation* **2022**, *110*, 102789, doi:10.1016/j.jag.2022.102789.
21. Alameen, S.A.; Alhothali, A.M. A lightweight driver drowsiness detection system using 3DCNN with LSTM. *Computer Systems Science & Engineering* **2023**, *44*, 895-912, doi:10.32604/csse.2023.024643.
22. Cao, C.; Dragićević, S.; Li, S. Short-term forecasting of land use change using recurrent neural network models. *Sustainability* **2019**, *11*, 5376, doi:10.3390/su11195376.
23. Liu, C.L.; Xia, J.X. Dynamic simulation of land use based on the LSTM–CA model. *Remote Sensing for Natural Resources* **2022**, *34*, 122-128, doi:10.6046/zrzyyg.2021375.
24. Xing, W.; Qian, Y.; Guan, X.; Yang, T.; Wu, H. A novel cellular automata model integrated with deep learning for dynamic spatio-temporal land use change simulation. *Computers & Geosciences* **2020**, *137*, 104430, doi:10.1016/j.cageo.2020.104430.
25. Yang, W.; Zhang, Y.; Hou, K.; Wang, X. A hybrid cellular automaton model integrated with 3DCNN and LSTM for simulating land use/cover change. *International Journal of Digital Earth* **2025**, *18*, 2447337, doi:10.1080/17538947.2024.2447337.
26. Liu, M.; Chen, H.; Qi, L.; Chen, C. LUCC simulation based on RF-CNN-LSTM-CA model with high-quality seed selection iterative algorithm. *Applied Sciences* **2023**, *13*, 3407, doi:10.3390/app13063407.
27. Huang, C.; Zhou, Y.; Wu, T.; Zhang, M.; Qiu, Y. A cellular automata model coupled with partitioning CNN-LSTM and PLUS models for urban land change simulation. *Journal of Environmental Management* **2024**, *351*, 119828, doi:10.1016/j.jenvman.2023.119828.
28. Li, S.; Zhang, C.; Chen, C.; Yang, C.; Zhao, L.; Bai, X. Optimization simulation and comprehensive evaluation coupled with CNN-LSTM and PLUS for multi-scenario land use in cultivated land reserve resource area. *Remote Sensing* **2025**, *17*, 1619, doi:10.3390/rs17091619.
29. Zhou, Z.; Chen, Y. STUrban: A novel spatial-temporal deep learning model to simulate long-term urban growth. *Information Geography* **2025**, *1*, 100004, doi:10.1016/j.infgeo.2025.100004.
30. Xu, Q.; Guan, X.; Chen, X.; Yang, C.; Yang, X.; Wu, H. A local Moran's I guided transformer cellular automata for simulating heterogeneous urban growth. *International Journal of Geographical Information Science* **2025**, *39*, 2151-2176, doi:10.1080/13658816.2025.2478471.
31. Gui, B.; Bhardwaj, A.; Sam, L. A novel multi-scale deep learning framework for adaptive urban expansion simulation. *Sustainable Cities and Society* **2025**, *130*, 106594, doi:10.1016/j.scs.2025.106594.

 doi:10.20944/preprints202603.1739.v1

32. Liu, T.L.; Liu, M.H.; Jing, L.; Li, T. A dual-constraint RF-Patch-CA urban expansion simulation method considering the importance of driving factors (in Chinese). *Geography and Geo-Information Science* **2021**, *37*, 63-70, doi:10.3969/j.issn.1672-0504.2021.02.009.
33. Yang, J.; Huang, X. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. *Earth System Science Data* **2021**, *13*, 3907-3925, doi:10.5194/essd-13-3907-2021.
34. Liu, X.; Ou, J.; Chen, Y.; Wang, S.; Li, X.; Jiao, L.; Liu, Y. Scenario simulation of urban energy-related CO2 emissions by coupling the socioeconomic factors and spatial structures. *Applied Energy* **2019**, *238*, 1163-1178, doi:10.1016/j.apenergy.2019.01.173.
35. Wang, L.; Li, R.; Zhang, C.; Fang, S.; Duan, C.; Meng, X.; Atkinson, P.M. UNetFormer: A UNet-like transformer for efficient semantic segmentation of remote sensing urban scene imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* **2022**, *190*, 196-214, doi:10.1016/j.isprsjprs.2022.06.008.
36. Feng, Y.; Tong, X. Incorporation of spatial heterogeneity-weighted neighborhood into cellular automata for dynamic urban growth simulation. *GIScience & Remote Sensing* **2019**, *56*, 1024-1045, doi:10.1080/15481603.2019.1603187.
37. Zhang, Y.; Li, C.; Zhang, L.; Liu, J.; Li, R. Spatial simulation of land-use development of Feixi county, China, based on optimized productive–living–ecological functions. *Sustainability* **2022**, *14*, 6195, doi:10.3390/su14106195.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.