Advances And Prospects of Cellular Automata in Simulating Spatial Expansion of Metropolitan Areas

Shijie Chu *

Department of Surveying and Mapping Engineering, Heilongjiang University of Science and Technology, Harbin, China

* Corresponding Author Email: chushijie@usth.edu.cn

Abstract. This study systematically reviews research advances in Cellular Automata (CA) modeling for spatial expansion simulation. Bibliometric and model evolution analyses reveal an evolutionary trajectory of Cellular Automata models transitioning from basic morphological simulation to "multi-factor coupling, uncertainty quantification, and intelligent optimization". Methodological innovations in resolving the "high-density agglomeration–resource overload" paradox are examined, focusing on three directions: ecological constraint integration, deep learning optimization, and uncertainty quantification. Following the logical framework of "technological evolution–scale adaptation–paradigm innovation", future breakthroughs are proposed addressing bottlenecks in scale effects and social mechanism integration, thereby providing scientific support for territorial spatial planning. In addition, the review emphasizes the necessity of integrating cross-disciplinary methods to enhance the explanatory power of CA models. It also highlights the importance of policy-oriented applications in guiding sustainable urban expansion. Ultimately, the study aims to bridge theoretical modeling with practical governance, ensuring CA research contributes effectively to real-world planning challenges.

Keywords: Cellular Automata; Spatial Expansion; Simulation; Urban Expansion.

1. Introduction

Urbanization, characterized by quantitative increases in built-up areas, is fundamentally driven by demographic and economic factors. From a land-use perspective, it manifests as significant land cover transformation along urban-rural gradients—where non-urban lands (e.g., farmland, bare soil) are extensively converted into construction lands. This process constitutes the core driver depicting urbanization [1]. Rapid urbanization has led the spatial expansion of metropolitan areas to profoundly reshape regional eco-environments, resource carrying capacity, and sustainable development potential. In global megacities exceeding ecological carrying thresholds, structural dilemmas—including water scarcity, overloaded transportation networks, and excessive air pollutants—have become critical constraints on urban health [2]. Within methodological frameworks for urban expansion simulation, Cellular Automata (CA), as spatiotemporal discrete dynamic models, represent a core domain of complexity science. Their bottom-up modeling logic, complex system computational capability, and inherent parallel computing features provide a universal framework for studying highly heterogeneous geographical systems [3]. These attributes position CA models as pivotal tools for resolving the "high-density agglomeration–resource overload" paradox in metropolitan areas. Particularly in simulating nonlinear processes—such as industrial land retreat under Beijing’s de-growth policy and ecological redline management—they demonstrate superior adaptability to traditional statistical models.

2. Historical Background and Evolutionary Context

2.1. Cellular Automata

Cellular Automata (CA) are defined as a modeling framework characterized by spatiotemporal computational features. The modeling philosophy from local interactions to global patterns is widely applied to simulate complex systems with spatial and temporal discreteness. A standard CA comprises discrete cells with finite states distributed in regular grids, following identical transition rules and undergoing synchronous updates. Complex dynamic systems are formed through localized interactions among numerous cells [4].

2.2. Theoretical Foundations and Initial Applications

In the 1970s, geographic CA frameworks were systematically established by Michael Batty and Yichun Xie, advancing urban complex system simulations. During the 1980s, theoretical validation was provided by Helen Couclelis, demonstrating CA’s applicability to urban systems through its capability to generate self-organizing structures via simple rules, thereby overcoming limitations of traditional mathematical models.

The rise of GIS technology in the 1990s resolved fundamental spatial data limitations, enabling substantive breakthroughs in CA modeling. Represented by the SLEUTH model—an open-source CA implementation focusing on long-term urban land expansion at macro-to-meso scales—key advantages include its transparent architecture, computational efficiency, and portability. These attributes rendered it an effective tool for Western urban expansion studies with significant referential value for CA development. However, explicit limitations were identified: incapability to simulate urban decline, time-consuming parameter calibration, and computational inefficiency due to massive calculations. These characteristics collectively delineate the model’s primary application scope and potential boundaries [3].

Ma Aigong’s Lanzhou case study addressed topographic constraints in valley cities by developing a dual-constrained CA framework integrating rigid ecological constraints (e.g., terrain) and elastic constraints. Limitations of this phase were twofold: (1) transition rules primarily relied on expert-defined heuristics; (2) drivers such as economic dynamics and population shifts remained inadequately integrated, rendering them inadequate for planning requirements.

2.3. Technological Integration and Innovation Phase

The proliferation of remote sensing (RS) and Geographic Information Systems (GIS) in the 1990s, characterized by the deep integration of multi-source RS data and intelligent algorithms, provided high-precision spatiotemporal data support for CA, becoming a hallmark of this phase. Breakthroughs were achieved in two key areas of rule optimization: heuristic parameter optimization and machine learning-based rule mining. Li Xia's research team observed the issue of farmland loss due to disordered urban expansion in China (e.g., the Dongguan case study), which traditional CA models struggled to address by integrating policy and natural constraints. By leveraging GIS spatial analysis capabilities, constrained CA models were developed for the first time to quantitatively simulate sustainable urban forms, providing a scientific tool for "compact city" development [5]. This research is also regarded by subsequent scholars as one of the landmark achievements marking the transition of China's urban CA models from the "technology integration period" to the "diversified development period." The multi-model coupling framework of CA emerged as a new pathway for simulating complex systems. Examples include:

Beijing's "de-growth development" model was based on an economic constraint framework. This framework allocated the urbanized area, derived from an economic model, within the urban space, enabling prediction and simulation under dual policy and economic constraints.

$$Prob_{ij}^{t} = P_{dist}^{t-1} \times condition\left(S_{ij}\right) \times (1 + (-ln\gamma)^{\alpha}) \times N_{ij}^{t-1} \quad (1)$$

Where P_{dist}^{t-1} represents the influence of transportation accessibility and proximity to urban centers on the cell at time t-1;$condition(S_{ij})$denotes the impact of restrictive conditions on the i-th cell. $(1 + (-ln\gamma)^{\alpha})$ introduces a stochastic factor and N_{ij}^{t-1} reflects the influence of the state of neighboring cells on the central cell at time t−1 [6].

In the Beijing-Tianjin-Hebei regional simulation, a logistic regression CA quantified transition probabilities using spatial variables (distance to roads, population density). The challenge of simulating urban interactions in large-scale regions was addressed by incorporating varying factor weights, constructing a policy-driven CA framework.

In the Shenzhen scheme, a CA model incorporated the urban scale as a size constraint. Ecological protection redlines and habitat quality assessment results based on the InVEST model served as rigid and elastic ecological constraints, respectively, for delineating the Urban Growth Boundary (UGB) [7]. This approach established a dual-control system of "rigid ecological redlines + elastic habitat constraints," which reduced the mis-conversion rate of ecological land by 23%. It addressed the one-dimensional limitations of traditional UGB delineation methods, pioneered a "rigid elastic" dual ecological constraint framework, and achieved a dynamic balance between ecological conservation and urban expansion.

Despite these advancements, CA models in this phase exhibited limitations. Factor weights relied on regression of historical data, whereas actual urban expansion is influenced by dynamic factors like policy and economic shifts. Consequently, these models could not dynamically respond to abrupt policy changes, overlooked interaction effects between influencing factors (e.g., the nonlinear relationship between population density and traffic distance), and suffered from insufficient long-term sequence reliability.

2.4. Intelligent Simulation Deepening Phase

Deep learning and high-performance computing have propelled CA models into a new era of "intelligent simulation." The application of Convolutional Neural Networks (CNN) has overcome the limitations of traditional rule mining: The deep transfer learning CA model (MSCNN-CA) developed by Xie Zhiwen's team at Wuhan University utilized multi-scale convolutional kernels (2x2, 4x4, 6x6) to simultaneously extract regional features. Experiments in Wuhan and Pudong New Area showed that compared to logistic regression and neural networks, a single-structure CNN-CA model improved the Figure of Merit (FoM) index by 23.3% to 29.4%. Furthermore, the MSCNN-CA model improved upon the single-structure CNN-CA by an additional 0.8% to 4.8% [8]. In 2023, Guan Xuefeng's team proposed the HGAT-VCA model, which incorporated a graph attention mechanism. This model abstracted vector parcels into graph nodes and captured long-distance spatial interactions through a higher-order adjacency matrix, achieving a simulation accuracy of FoM = 0.3853 in the Moreton Bay region of Australia (a 40.7% improvement over traditional models). On the other hand, support from cloud computing platforms has enabled parallel computation for tens of thousands of cells, facilitating high-resolution simulations at a national scale and pushing multi-scenario simulation frameworks toward maturity (Table 1).

Table 1. Developmental Stages and Characteristic Comparison of Cellular Automata Models for Urban Expansion

Developmental Stage	Key Technological Breakthroughs	Representative Models	Precision Enhancement Mechanism
Foundational Model Formation Stage (1970s-1990s)	GIS spatial analysis techniques	SLEUTH, Constrained CA	Spatially explicit constraint rules
Technological Integration and Innovation Stage (2000s-2010s)	Nighttime light data calibration, Genetic algorithm optimization	ANN-CA. CA-Markov	Nonlinear rule learning
Intelligent Simulation Deepening Stage (From the 2020s Onward)	Multi-scale convolutional neural networks, Graph attention mechanisms	MSCNN-CA. HGAT-VCA	Adaptive regional feature extraction

3. Core Technological Breakthroughs: Analysis of Key Advances

3.1. Multidimensional Data Integration

The enhanced precision in metropolitan spatial simulation is attributed to the deep integration of remote sensing and social sensing data. Primary data sources currently employed are summarized in Table 2.

Table 2. Primary Data Sources for Urban Expansion Applications

Data Type	Application in Urban Expansion	Data Type	Application in Urban Expansion
Optical Remote Sensing Imagery (Landsat/Sentinel)	Multitemporal image analysis, land-use classification, and spatial metric quantification are utilized for urban expansion monitoring. High-resolution data enable precise identification of urban boundary dynamics, construction density, and expansion directionality, providing dynamic decision-making support for urban planning when integrated with GIS.	Optical Remote Sensing Imagery (Landsat/Sentinel)	Multitemporal image analysis, land-use classification, and spatial metric quantification are utilized for urban expansion monitoring. High-resolution data enable precise identification of urban boundary dynamics, construction density, and expansion directionality, providing dynamic decision-making support for urban planning when integrated with GIS.
Technological Integration and Innovation Stage (2000s-2010s)	Nighttime light data calibration, Genetic algorithm optimization	ANN-CA. CA-Markov	Nonlinear rule learning
Nighttime Light Data (NPP/VIIRS)	With a spatial resolution of ~500 m and daily global coverage, saturation issues inherent in traditional DMSP/OLS data are overcome. This compensates for the limitations of medium-to-low resolution imagery in identifying built-up areas, particularly suited for monitoring urban agglomerations at large regional scales.	Nighttime Light Data (NPP/VIIRS)	With a spatial resolution of ~500 m and daily global coverage, saturation issues inherent in traditional DMSP/OLS data are overcome. This compensates for the limitations of medium-to-low resolution imagery in identifying built-up areas, particularly suited for monitoring urban agglomerations at large regional scales.

The core challenge of data fusion lies in spatial resolution inconsistencies among remote sensing images and their misalignment with socio-spatial data. To address this developed the MTS-CA model incorporating Self-Organizing Maps (SOM) for geographical zoning, calibrating data scales within unified partitions.

Self-Organizing Maps (SOM) provide an unsupervised dimensionality reduction method that effectively organizes data points in low-dimensional spaces (e.g., grids) based on similarity. Their value derives from inherent self-organizing capability, environmental adaptability, and robust fault tolerance, facilitating exploration of urban expansion patterns across diverse spatial contexts [9].

In the MTS-CA case study by Li et al., a partition-based calibration strategy was implemented: local transition rules were extracted via Logistic regression with 1:1 balanced sampling of historical urban expansion patches within each subzone, quantifying driver contributions (e.g., higher distance weighting in core areas); landscape metrics (LEI, AWMEI, ENN_MN) were subsequently computed to assess pattern similarity between historical and target expansion phases, enabling dynamic-weighted fusion of multiperiod local rules (weight proportional to similarity); ultimately, total expansion volume was constrained by AWMEI to prevent imbalance, while neighborhood effects were introduced to fine-tune boundary cell states, thereby coordinating spatial heterogeneity with global consistency and ensuring natural transition of expansion trends.

3.2. Optimization and Innovation of Transition Rules

Transition rules, the core of cellular automata (CA) models, have experienced three major paradigm shifts. The first generation, characterized by experience-driven rules, relied heavily on linear assumptions, such as those embedded in logistic regression. These approaches were only capable of expressing simple relationships, for example, a higher probability of development being associated with proximity to roads. However, the reliance on fixed coefficients made it difficult to capture the complex, nonlinear effects that are intrinsic to urban systems [10]. To overcome these limitations, Zhang Yihan and colleagues proposed a maximum entropy-based CA framework. By leveraging the unbiased nature, adaptability, and generalization capacity of the maximum entropy principle, this framework effectively addressed problems of uncontrolled randomness and parameter sensitivity that constrained traditional urban expansion simulations. Its core innovation lies in integrating physical constraints with probabilistic optimization, thereby offering a more transferable and robust paradigm for modeling complex systems. Looking forward, future research could focus on combining the maximum entropy principle with deep learning techniques, such as convolutional recurrent networks, to further improve the representation and efficiency of high-dimensional spatial features [11].

Convolutional Neural Networks (CNN) and Graph Neural Networks (GNN) have revolutionized rule learning mechanisms: 1) The MSCNN-CA model employs a parallel convolutional architecture utilizing 3×3 kernels to capture local road network influences and 7×7 kernels to identify regional center radiation effects, with multi-scale features fused in the fully connected layer. 2) The HGAT-VCA model pioneers the integration of attention mechanisms into spatial interactions.

3.3. Multi-Scale Simulation and Scale Transformation

Metropolitan spatial expansion involves cross-scale interactions between macro regional structure evolution and micro land parcel morphological changes. Three methodological approaches are employed in CA modeling to address scale-related challenges: 1) Nested architectures: 1 km grids are utilized for simulating urban agglomeration expansion at regional scales, switching to 30 m grids to achieve refined simulation of parcel transformations at urban scales. 2) Vector cell innovations: Traditional raster cells are abandoned in the HGAT-VCA model, where vector parcels serve as fundamental units for constructing graph structures that preserve cadastral boundary authenticity. 3) Hierarchical rule transmission: Provincial policies act as macro-rules controlling total urban land use scale, while developer agent decisions function as micro-rules guiding local development sequences (Table 3).

Table 3. Comparison of Multi-scale Simulation Methods for Metropolitan Areas

Simulation Scale	Applicable Models	Spatial Unit	Advantages
Macro Regional Scale (Urban Agglomeration/Provincial)	CA-Markov System Dynamics-Coupled CA	1–5 km raster	Reveals regional development corridors
Technological Integration and Innovation Stage (2000s-2010s)	Nighttime light data calibration, Genetic algorithm optimization	ANN-CA、CA-Markov	Nonlinear rule learning
Meso Urban Scale (Individual Metropolitan Area)	ANN-CA MSCNN-CA	30–100 m raster	Balances accuracy and efficiency

4. Practical Applications: Scale-Specific Adaptation in Metropolitan Simulation

4.1. Macro Regional Scale: Structural Evolution of Urban Agglomerations

In urban agglomerations such as the Yangtze River Delta and Guangdong-Hong Kong-Macao Greater Bay Area, CA models are primarily employed to simulate formation mechanisms of polycentric network structures. The CA-SD model developed by The Chunyang's team couples system dynamics, translating macro variables (e.g., population growth, GDP growth rate) into land-use demand constraints, which successfully replicated the corridor-style expansion pattern between Shanghai and surrounding cities during 1990–2020. The core value of such simulations lies in identifying regional development corridors: by strategically configuring strategic nodes (e.g., high-speed rail stations, ports) to enhance neighborhood development probability, the model is guided to generate hub-and-spoke spatial structures consistent with planning visions. However, macro-scale simulations are often constrained by data granularity; for instance, nighttime light data can only identify built-up area contours while failing to differentiate functional differences between industrial zones and residential areas.

4.2. Meso Urban Scale: Polycentric Expansion Patterns

For individual metropolitan areas (e.g., Shanghai, Wuhan), multi-temporal historical data comparison forms the foundation for uncovering expansion patterns. Analysis of Shanghai's 2000–2020 land-use data by Li Qiyuan et al. revealed that urban expansion does not follow a simple Markov process: the 2000–2005 period was dominated by peripheral expansion in Pudong New Area (LEI=12.3), while the 2015–2020 phase shifted to infill development in Baoshan District (LEI=68.9). The MTS-CA model proposed based on these findings quantifies similarity indices between historical and target expansion characteristics (Equation 3), enabling dynamic weighting of historical rules.

$$IES,P = 10 - |IREF,P - IREF,Q| \tag{2}$$

Where $IREF$ represents the relative expansion feature value and Q denotes the target period. This approach achieved a Kappa coefficient of 0.81 for simulating Shanghai's 2015–2020 expansion, demonstrating the effectiveness of multi-temporal pattern fusion [8].

4.3. Micro Parcel Scale: Refined Simulation of Cadastral Boundaries

In urban renewal projects, traditional raster-based Cellular Automata (CA) models are often challenged in precisely simulating complex, irregular urban parcel morphologies due to their inherent jagged boundary limitations. The HGAT-VCA model overcomes these constraints through three key innovations: First, a vector-based cell construction approach is adopted, directly utilizing irregular parcel polygons from cadastral databases as fundamental simulation units, thereby authentically reflecting urban spatial patterns. Second, a higher-order neighborhood definition mechanism is introduced, extending the traditional model's scope from 1st-order to a more flexible Kth-order, effectively capturing the complex spatial influences of significant facilities (e.g., schools, large commercial districts) on surrounding multiple parcels. Finally, a Graph Attention Network (GAT) mechanism is innovatively integrated, enabling the adaptive learning and generation of directional influence weights between parcels (e.g., precisely quantifying the promotive effect of a "commercial parcel" on "residential development potential"), significantly enhancing the accuracy and adaptability of urban dynamic simulations.

5. Challenges and Prospects: Bottlenecks and Frontier Directions

5.1. Existent Technical Bottlenecks

Current CA models systematically constrain the precise implementation of urban renewal practices, based on five fundamental conflicts: data fusion bottlenecks (e.g., spatiotemporal resolution misalignment in multi-source remote sensing data), human-computer interaction black boxes (difficulty in interpreting deep learning rules), inadequate policy quantification (lack of standardized parameters for Territorial Spatial Planning constraints), absence of dynamic feedback (ignoring real-time responses such as Not-In-My-Backyard (NIMBY) effects), and computational efficiency bottlenecks (requiring supercomputing support for nationwide high-resolution simulations).

5.2. Future Breakthrough Directions

5.2.1 Intelligent Algorithm Deepening

Interpretable AI: Techniques such as SHAP (SHapley Additive exPlanations) values and LIME (Local Interpretable Model-agnostic Explanations) are applied to decipher the decision-making logic of CNN-based rules, generating transparent reports with statements such as "a parcel is identified as a priority development zone due to its proximity (800 meters) to a subway station and a neighborhood development density exceeding 60%". For instance, Guan Xuefeng et al. integrated an attention weight visualization module into the HGAT-VCA model, quantifying the directional influence mechanism of commercial parcels on residential development potential and providing causal chains for planning decisions [11, 12].

Deep Reinforcement Learning: Government entities, developers, and residents are designed as competitive agents that learn optimal land development strategies through reward functions.

5.2.2 Human-Land Coupling Mechanisms

The incorporation of mobile phone signalling data enables researchers to conduct more refined quantitative analyses of urban work-residence balance based on large-scale population movement trajectories. In this process, the study constructs a commuting pressure feedback loop by characterising commuter flow intensity and direction, thereby dynamically calibrating transfer rules within urban expansion models. Compared to traditional methods reliant on static statistical data, mobile phone signalling data not only reflects spatial mobility patterns across different time periods but also reveals spatio-temporal matching relationships between workplaces and residences. This provides more authentic and timely evidence for assessing the rationality of urban spatial expansion. Such a big data-driven feedback mechanism helps mitigate work-residence imbalances during urban expansion, enhancing the scientific validity and policy relevance of simulation outcomes.

5.2.3 Technological Integration and Innovation

Quantum Cellular Automata: The superposition states of qubits are explored to simulate uncertainty in land-use types, addressing the stochasticity limitations of traditional Monte Carlo methods. Edge Computing Deployment: Lightweight CA models are deployed at municipal planning departments, enabling real-time scenario testing at the parcel scale.

6. Conclusion

The development of Cellular Automata (CA) in simulating spatial expansion within metropolitan areas is undergoing a transition from static empirical models toward dynamic intelligent systems. This transformation is embodied not only through technological iterations (e.g., the evolution of simulation rules from traditional methods such as Logistic regression to advanced techniques like Graph Neural Networks), but also in a multidimensional shift in research paradigms: in terms of data foundation, a transformation has been achieved from reliance on single-source remote sensing images to the integration of multi-source social sensing big data; in rule generation, an evolution has been completed from expert experience-driven approaches to adaptive learning utilizing deep learning; in application scenarios, an expansion has been realized from purely predicting land-use changes to functioning as an integrated platform incorporating "spatial simulation, impact assessment, and planning decision-making".

Future research must further break down disciplinary barriers, deeply integrating the strengths of geographical spatial analysis with computer science intelligent algorithms and sociological behavioral theories. Particularly within the context of territorial spatial planning reform, CA models should enhance their connection with the "Dual Evaluations" (assessments of resource environmental carrying capacity and territorial spatial development suitability), constructing an intelligent decision-support system capable of supporting the entire planning lifecycle. As Michael Batty stated: "The goal of urban simulation is not to predict a definite future, but to understand the possibilities of change." This philosophy will guide CA models to continue innovating on a path that balances accuracy and interpretability, ultimately serving the urbanization process towards harmonious coexistence between humans and nature.

References

[1] Li Xianping. Study on the dynamic expansion of Beijing and its impact on the ecological environment [D]. Harbin Normal University, 2016.

[2] Dong Suyu, Wu Mengda, Wang Junlu, et al. What has Beijing done right to solve the “big city disease”? [N]. Xinhua Daily Telegraph, 2025-07-02(001). DOI: 10.28870/n.cnki.nxhmr.2025.005785.

[3] Ma Aigong. Research on the valley-type urban expansion based on cellular automata [D]. Lanzhou University, 2009.

[4] Shen Tiyan, Zhang Heng, Zhang Jinjie. Urban simulation of Beijing-Tianjin-Hebei region based on MODIS remote sensing images and cellular automata [C]// Asia Pacific Environmental Science Research Center, Hong Kong; Huazhong Normal University, China; Institute of Geodesy and Geophysics. Proceedings of 2010 International Conference on Remote Sensing (ICRS 2010), Volume 3. Beijing University School of Government; Tianjin Urban Planning and Design Institute, 2010: 250-254.

[5] Zhao Hongrui, Liu Xintong, Yao Yi, et al. Urban development boundary delimitation of Shenzhen based on urban expansion simulation and ecological-scale constraints [J]. Geography and Geo-Information Science, 2022, 38(05): 1-8.

[6] Xie Zhiwen. Construction of urban expansion cellular automata model using deep learning [D]. Wuhan University, 2020. DOI: 10.27379/d.cnki.gwhdu.2020.000974.

[7] Feng Huihui, Liu Huiping. Variable coefficient urban growth model based on error feedback control [J]. Scientia Geographica Sinica, 2014, 34(5): 594-600. https://doi.org/10.13249/j.cnki.sgs.2014.05.594.

[8] Li Qiyuan, Wang Haijun, Zeng Haoran, et al. Cellular automata simulation of urban expansion coupled with multi-period rules and spatial heterogeneity [J]. Geography and Geo-Information Science, 2024, 40(01): 26-36.

[9] Zhang Yihan, Liu Xiaoping, Chen Guangliang, et al. Cellular automata model based on maximum entropy and its simulation of urban expansion [J]. Science China: Earth Sciences, 2020, 50(03): 339-352.

[10] Zhao Li, Yang Jun, Li Chuang, Ge Yuting, Han Zenglin. Research progress of geographical cellular automata models [J]. Scientia Geographica Sinica, 2016, 36(8): 1190-1196. https://doi.org/10.13249/j.cnki.sgs.2016.08.009.

[11] Guan X F, Xing W R, Li J B, et al. HGAT-VCA: Integrating high-order graph attention network with vector cellular automata for urban growth simulation [J]. Computers, Environment and Urban Systems, 2023, 99: 101900.

[12] Li Xia, Yeh Anthony Gar-On. Constrained cellular automata model and the simulation of sustainable urban development patterns [J]. Acta Geographica Sinica, 1999, (04): 3-12.