

Dynamic Simulation and Forecasting of Spatial Expansion in Small and Medium-Sized Cities Using ANN-CA-Markov Models

Chengquan Gao

School of Architecture and Urban Planning, Henan University of Urban Construction, Ping'dingshan 467001, China

Abstract—This study utilizes the ANN-CA-Markov (Artificial Neural Network-Cellular Automata-Markov) model to address spatial planning and expansion challenges in China's small and medium-sized cities. With China's urbanization rate reaching 59.58% in 2018 and expected to hit 70% by 2030, the country is entering a mid-stage of urbanization, leading to rapid expansion of megacities and a gradual decline in smaller cities. The study aims to dynamically simulate urban spatiotemporal evolution and predict future land use changes, integrating land use data, DEM elevation, transportation, administrative centers, and ecological information. The model forecasts the ecological spatial layout of Wanzhou District by 2025, with results indicating a slight decrease in ecological space and an increase in construction land. This suggests a need to balance urban development with ecological sustainability amidst rapid urbanization. The study demonstrates the high accuracy of the ANN-CA-Markov model in predicting land use changes and provides valuable insights for urban planners in making informed land use decisions.

Keywords—ANN-CA; Markov; small and medium-sized cities; spatial; planning

I. INTRODUCTION

In 2018, China's urbanization rate reached 59.58%, and it is expected to hit 70% by 2030, marking the country's entry into the middle stage of urbanization [1-3]. Since 2011, China has been systematically revising its traditional urbanization approach, leading to the rapid expansion of megacities and the gradual decline of many small and medium-sized cities and towns. With the central government's strategic decisions to "establish a scientifically and reasonably structured urban pattern, where large, medium, and small cities and towns, as well as city clusters, are well-organized," new opportunities and broad prospects have emerged for the development of small and medium-sized towns. The healthy urbanization of these smaller cities requires more scientific planning and development of urban land resources by administrators [4, 5].

Scientifically defining urban development boundaries not only enables the intensive use of spatial resources and controls the disorderly sprawl of cities but also supports the sustainable socio-economic development of cities while protecting the natural ecological environment [6-9]. However, one of the major challenges in urban planning is accurately predicting the spatial and temporal evolution of land use, particularly in small and medium-sized cities where data may be less comprehensive, and urbanization patterns are more complex.

In recent years, numerous scholars have utilized the ANN-CA (artificial neural network-cellular automata) coupled model to dynamically simulate the temporal and spatial evolution of cities and predict future land use changes, assisting in urban land use simulation, urban development boundaries, and ecological redline protection in urban planning, achieving many meaningful results [10-12]. For instance, Xu et al. [13] integrated Artificial Neural Networks (ANN), Cellular Automata (CA), and Markov Chain (MC) to simulate urban expansion in rapidly urbanizing areas, revealing the nonlinear relationship between the expansion process and its drivers. Similarly, Zhao et al. [14] studied land-use changes in Yucheng District, Ya'an City, China using an ANN-CA model, emphasizing the improvement of simulation accuracy through appropriate thresholds and random variable parameters. Additionally, Asanza et al. [15] explored the integration of ANN and CA models for spatial-temporal land use forecasting, highlighting the enhanced forecasting accuracy through temporality and geospatial data analytics.

Despite these advances, challenges remain in effectively modeling and predicting land use changes in regions where data availability is limited or where the urbanization process involves complex interactions among multiple factors. The current study addresses these challenges by developing an enhanced ANN-CA-Markov coupled model that incorporates more comprehensive datasets and refined transition rules to achieve higher simulation precision. Specifically, this study leverages Wanzhou District's land use cover data, DEM elevation data, road traffic data, administrative center data, river data, ecological protection redline data, and natural conservation area data from 2000, 2006, 2012, and 2018. By integrating these data with the ANN-CA-Markov coupled model, we aim to predict the ecological spatial layout of Wanzhou District by 2025 more accurately.

The novel contribution of this work lies in overcoming the difficulties associated with limited data availability and complex urbanization processes by enhancing the precision of the ANN-CA-Markov model. This is achieved through the integration of additional data sources and the refinement of transition rules, offering a more comprehensive approach to modeling and predicting spatial changes. This study, therefore, represents a significant advancement in the field, particularly in the context of small and medium-sized cities where such challenges are most pronounced.

II. METHOD

A. ANN-CA Model

The ANN-CA model, short for Artificial Neural Network - Cellular Automata model, is a discrete time, space, and state grid dynamic model where spatial interactions and temporal causal relations are local. It enables the bottom-up simulation of the spatiotemporal evolution of complex systems. The state of each cell is determined by the states of its neighboring cells, and upon establishing transition rules, all cells can evolve autonomously following these rules, highlighting the core essence of transition rules. Artificial neural networks possess self-learning and associative capabilities, allowing for the rapid identification of optimized solutions [16-18]. By learning the rules of land use data changes through the neural network model and applying these extracted rules to the grid data of the starting year, simulation predictions can be completed within the cellular automata. The core principle involves training neurons with land use data from different periods, then determining the transition probabilities for each land use type based on the characteristics of influencing factors, culminating in the simulated prediction of land use planning. To ensure the model's accuracy, the input parameters were rigorously selected and optimized. The primary inputs include land use types, digital elevation model (DEM), neighborhood development density, and transition suitability. These parameters were chosen based on the terrain characteristics of the study area, the diversity of land use, and the complexity of its spatial distribution.

Extensive experiments were conducted to evaluate the impact of different parameter combinations on the model's predictive outcomes. Specifically, different neighborhood window sizes, DEM resolutions, and land use classification standards were tested. Sensitivity analysis revealed that the size of the neighborhood window significantly affects the model's spatial resolution and computational efficiency, while variations in DEM resolution notably influence the prediction accuracy. Ultimately, a 5×5 Moore neighborhood window and a 30-meter resolution DEM were selected as they offered the best balance between computational efficiency and prediction accuracy.

Additionally, the selection of transition suitability parameters was explored. These parameters primarily represent the likelihood of transitions between different land use types. During the parameter adjustment process, a stochastic disturbance factor was introduced to simulate unforeseen changes in real-world conditions. The final selection of these parameters was made based on a comparative analysis of multiple simulation results, tailored to the specific conditions of the study area. The mathematical expression is as follows:

$$P(k, t, l) = (1 + (-\ln \gamma)^\alpha) \times P_{ann}(k, t, l) \times \Omega_k^t \times \cos(S_k^t) \quad (1)$$

In this expression, it mainly expresses the transition probability P of a cell k at time t to the l -th type of land use as a function of random factors, artificial neural network calculated probabilities, neighborhood development density, and transition suitability. $(-\ln \gamma)^\alpha$ represents the random factor; $P_{ann}(k, t, l)$ is the transition probability of a certain land type calculated by the trained artificial neural network; Ω_k^t represents the urban land

density within the defined neighborhood window, i.e., the total number of urban land cells divided by the total number of grid cells in the neighborhood window; $\cos(S_k^t)$ represents the transition suitability between two land types, generally indicated by 0 or 1, mainly to signify whether a transition is possible.

B. Markov Model

The Markov model (Markov), based on a type of stochastic process is a mathematical method used to predict the prior probabilities and conditional probabilities of events [14-16]. Its changes over time are continuous, and when the process parameters take discrete time values, it is referred to as a Markov sequence. The primary feature of the Markov sequence is its Markov property, not time series property, meaning when the state of the process (system) at time t_0 is known, the state of the process (system) at time t ($t > t_0$) is independent of the state at time t_0 . The Markov model is an extremely important predictive model in geographic forecasting research. In constructing the Markov model, we meticulously adjusted the state transition probability matrix. The state transition probabilities primarily reflect the likelihood of transitions between different land use types within a specific time sequence. To ensure the accuracy of this probability matrix, we conducted several experiments to test the impact of different initial conditions and transition probabilities on the final predictive outcomes. Sensitivity analysis indicated that certain key parameters in the state transition probability matrix significantly influence the spatial distribution of the model's predictions. For instance, increasing the probability of converting arable land to construction land by 10% substantially increases the predicted area of construction land while decreasing the areas of forest land and water bodies. This highlights the need to adjust transition probabilities in accordance with real-world conditions. The Markov model mainly consists of states, state transition processes, state transition probabilities, and state transition matrices, detailed as follows:

1) *State*: Represents an outcome, indicating the result appearing at a specific point in time.

2) *State transition process*: The relationship between the state change of an event and time.

3) *State transition probability*: Refers to the likelihood of an event's state changing to another state, expressed mathematically as follows:

$$P_{ij} = P(E_i \rightarrow E_j) = P(E_j / E_i) \quad (2)$$

In the formula, P_{ij} and $P(E_i \rightarrow E_j)$ both represent the probability of the state transition of the event, $P(E_j / E_i)$ represents the conditional probability, and E_i and E_j respectively represent the state at moments i and j .

4) *State transition probability matrix*: If a specific event has n possible states, then the state transition probability from state E_i to state E_j is denoted as p_{ij} and expressed through a matrix as follows:

$$p = \begin{pmatrix} p_{11} & \dots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \dots & p_{nn} \end{pmatrix} \quad (0 \leq p_{ij} < 1) \quad (3)$$

In the above expression, p represents the state transition probability matrix.

C. Construction of the ANN-CA-Markov Model

To simulate and predict the ecological space pattern of Wanzhou District, this study employs the CA-Markov model,

which combines the transition prediction capability of the Markov model with the spatial distribution simulation of cellular automata [17-19]. Traditional CA-Markov models face challenges in handling nonlinear relationships, so an artificial neural network (ANN) model is introduced to learn and establish more accurate transition rules. The process is divided into two stages: training and simulation. The operational logic of the model is illustrated in Fig. 1.

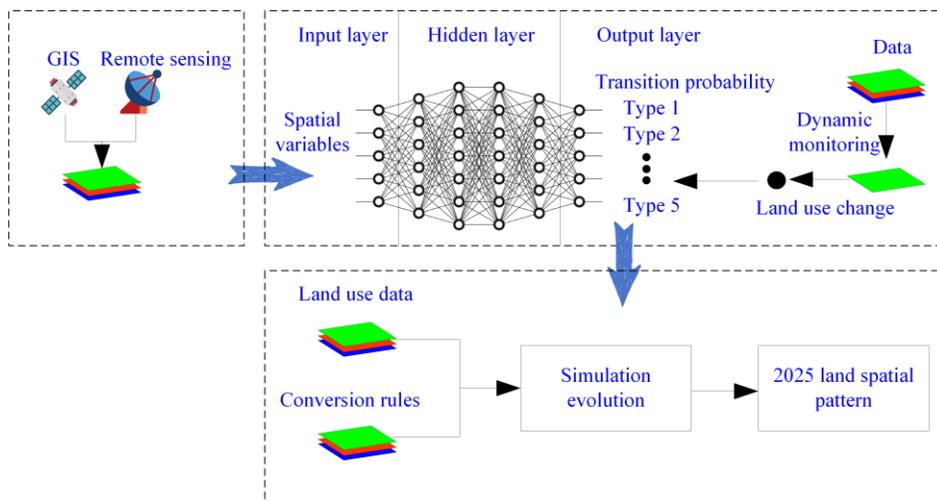


Fig. 1. Technical route map.

III. RESEARCH AREA AND DATA PROCESSING

A. Geographic Overview

Wanzhou District is located on the eastern edge of the Sichuan Basin, in the northeastern part of Chongqing. The Yangtze River flows into Wanzhou from the southwest, traverses northeastward, and then flows into Yunyang. Wanzhou District is situated between 107°55'22"E to 108°53'25"E and 30°24'00"N to 31°14'58"N. Lichuan and Shizhu are located to the south of Wanzhou District, Yunyang to the east, Liangping and Zhong County to the west, and Kaizhou District and Kaijiang to the north. The straight-line distance from Wanzhou District to Chongqing is 228 kilometers. Fig. 2 shows the location map of Wanzhou District.



Fig. 2. Location map of wanzhou district.

B. Data Sources

This study employed Landsat TM/ETM imagery for Wanzhou District from 2000, 2006, 2012, and 2018. The 2000, 2006, and 2012 data were sourced from Landsat-5 TM, while 2018 data came from Landsat-8, all acquired via the Geospatial Data Cloud (Table I). The imagery has a 30m×30m resolution, and the data were projected using the Albers projection. Additional data included 30m GDEM elevation data, OSM road, river, and administrative boundaries, and population data from the Chongqing Statistical Yearbook.

TABLE I. CLASSIFICATION STANDARDS FOR LAND USE TYPES IN WANZHOU DISTRICT

Land Category Code	Land Category Type	Detailed Types
1	Cultivated Land	Paddy fields, dry land
2	Forest Land	Forested land, shrub land, sparse forest, and other forest lands
3	Grassland	High, medium, and low coverage grasslands
4	Water Bodies	Rivers, lakes, reservoirs, ponds, tidal flats
5	Constructed Land	Urban areas, rural settlements, industrial and mining areas, transportation land
6	Unused Land	Bare land, sandy land, other types of unused land

C. Data Preprocessing

The study highlighted remote sensing image preprocessing importance using ERDAS 9.1 for band synthesis, geometric correction, and image cropping, enhancing data quality for

Wanzhou District. Band synthesis integrated different bands to improve image classification accuracy. Geometric corrections were applied using a third-order polynomial method to ensure pixel accuracy, while image cropping maintained research area consistency. The analysis involved visual and computer processing, adhering to “Current Land Use Classification” standards and incorporating field data and professional verification to achieve over 85% interpretation accuracy. These processed images provide a reliable basis for further analysis.

IV. ANALYSIS OF ECOLOGICAL SPACE CHANGE CHARACTERISTICS

A. Dynamic Degree and Transfer Rate

In studying the characteristics of ecological space change, the dynamic degree and transfer rate are mainly used to examine the changes in the quantity of ecological spaces within the study area. The formula for calculating the transfer rate of ecological space types is as follows:

$$K = \frac{u_b - u_a}{u_a} \times \frac{1}{T} \times 100\% \quad (4)$$

K represents the transfer rate of a certain ecological space type; u_a represents the quantity of a certain ecological space type at the beginning of the study; u_b represents the quantity of the same ecological space type at the end of the study; T represents the time span of the study period. The formula for calculating the overall dynamic degree of ecological space in the study area is:

$$Lc = \left(\frac{\sum_{i=1}^n \Delta Lu_{i-j}}{2 \sum_{i=1}^n Lu_i} \right) \times T^{-1} \times 100\% \quad (5)$$

Lu_i represents the quantity of the i th ecological space type at the beginning of the study; ΔLu_{i-j} represents the absolute value of the quantity of the i th ecological space type transformed into the non- i th ecological space type during the study period; T represents the time span of the study period, Lc represents the comprehensive dynamic degree of ecological space.

B. Transition Matrix

In the land-use transition matrix, the rows and columns represent the ecological space types at the beginning and end of a certain time sequence unit, respectively. Therefore, using the transition matrix to study ecological space types can reveal the changes in the area of each ecological space type from the beginning to the end of the period, as well as the transition situations of ecological space types. The formula for calculating the transition matrix is as follows:

$$S_{ij} = \begin{bmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{bmatrix} \quad (6)$$

In the formula, S represents the area of ecological space types within the time sequence unit, n is the total number of ecological space types within the time sequence unit, i represents the index of ecological space types at the beginning of the time

sequence unit; j represents the index of ecological space types at the end of the time sequence unit.

C. Landscape Pattern Indices

The spatial change characteristics of ecological spaces are mainly analyzed through the landscape pattern indices and classified landscape pattern indices of the study area. The analysis of these indices quantitatively describes the structure and distribution characteristics of the landscape itself. This paper selects the patch area, patch number, patch density, largest patch index, splitting index, and Shannon’s diversity index to analyze the landscape pattern indices and classified landscape pattern indices, thereby analyzing the spatial change characteristics of ecological spaces.

1) *Patch area (CA)*: The patch area (CA) represents the total area of a specific landscape type, and its calculation formula is as follows:

$$CA = \sum_{j=1}^n a_{ij} \times (1/10000) \quad (7)$$

In the formula, a_{ij} represents the area of patch i_j in a landscape type, n is the total number of all landscape types in the study area.

2) *Patch Number (NP)*: The patch number (NP) reflects the spatial pattern of the landscape. The larger the value of NP, the higher the degree of fragmentation in space, and vice versa. The spatial distribution characteristics of various landscape types can be determined to some extent by the patch number (NP), and its calculation formula is as follows:

$$NP = m(m \geq 1) \quad (8)$$

In the formula, m represents the number of patches of a certain ecological space type.

3) *Patch Density (PD)*: Represents the number of patches per unit area of the entire ecological space or a certain ecological space type. The larger the value of PD, the greater the separation among individual patches; conversely, the closer the patches are to each other. The calculation method is as follows:

$$PD = N / A \quad (9)$$

In the formula, N represents the number of patches of the ecological space type, A is the area of the study region.

4) *Largest Patch Index (LPI)*: The largest patch index represents the ratio of the largest patch of an ecological space type to the total area of the study region. The larger the LPI value, the higher the connectivity among patches of ecological space types, and vice versa. The calculation method is as follows:

$$LPI = \frac{Maxa_{ij}}{A} \times 100 \quad (10)$$

In the formula, LPI represents the value of the largest patch index, a_{ij} is the area of patch i_j of the ecological space type, A is the total landscape area of the entire study region.

5) *Splitting Index (SPLIT)*: Used to indicate the degree of fragmentation of a landscape type. The larger the value of SPLIT, the higher the degree of fragmentation of the landscape type, and vice versa. The calculation method is as follows:

$$SPLIT = \frac{A^2}{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2} \quad (11)$$

In the formula, SPLIT represents the value of the splitting index, a_{ij} is the area of patch ij in a landscape type, A is the total landscape area of the study region.

6) *Shannon's Diversity Index (SHDI)*: Indicates the degree of evenness among different ecological space types within the study area. The larger the value of SHDI, the more even the distribution of ecological space type patches, and vice versa. The calculation method is as follows:

$$SHDI = -\sum_{i=1}^m p_i \quad (12)$$

In the formula, m represents the number of ecological space types, p_i is the proportion of ecological space type i in the area of the study region.

D. Analysis of Ecological Space Structure Type Change Characteristics

The ecological space structure types in Wanzhou District mainly include cultivated land, forest land, grassland, water bodies, and unused land. Using the land use type data of Wanzhou District from 2000, 2006, 2012, and 2018, the software ArcGIS 10.1 was used to analyze the changes in the area of ecological spaces in Wanzhou District.

Using ArcGIS 10.1's mapping features, land use type maps of Wanzhou District from 2000 to 2018 were created, as shown in Fig. 3. Additionally, ArcGIS's statistical tools were employed to compile data on ecological spaces and land use areas in Wanzhou District from 2000 to 2018, as presented in Table II.

Table II highlights that Wanzhou District's ecological spaces are predominantly woodland (over 50%) and arable land (over 44%), with water bodies and grassland as secondary types. Unused land is minimal. Transition matrices for 2000–2006, 2006–2012, 2012–2018, and 2000–2018, generated using ArcGIS and Excel, are shown in Tables III to VI.

As shown in Table III, during 2000–2006, the main ecological space type conversions were from arable land and woodland to water bodies and built-up land, primarily due to the water storage of the Three Gorges Reservoir Area and rapid urbanization.

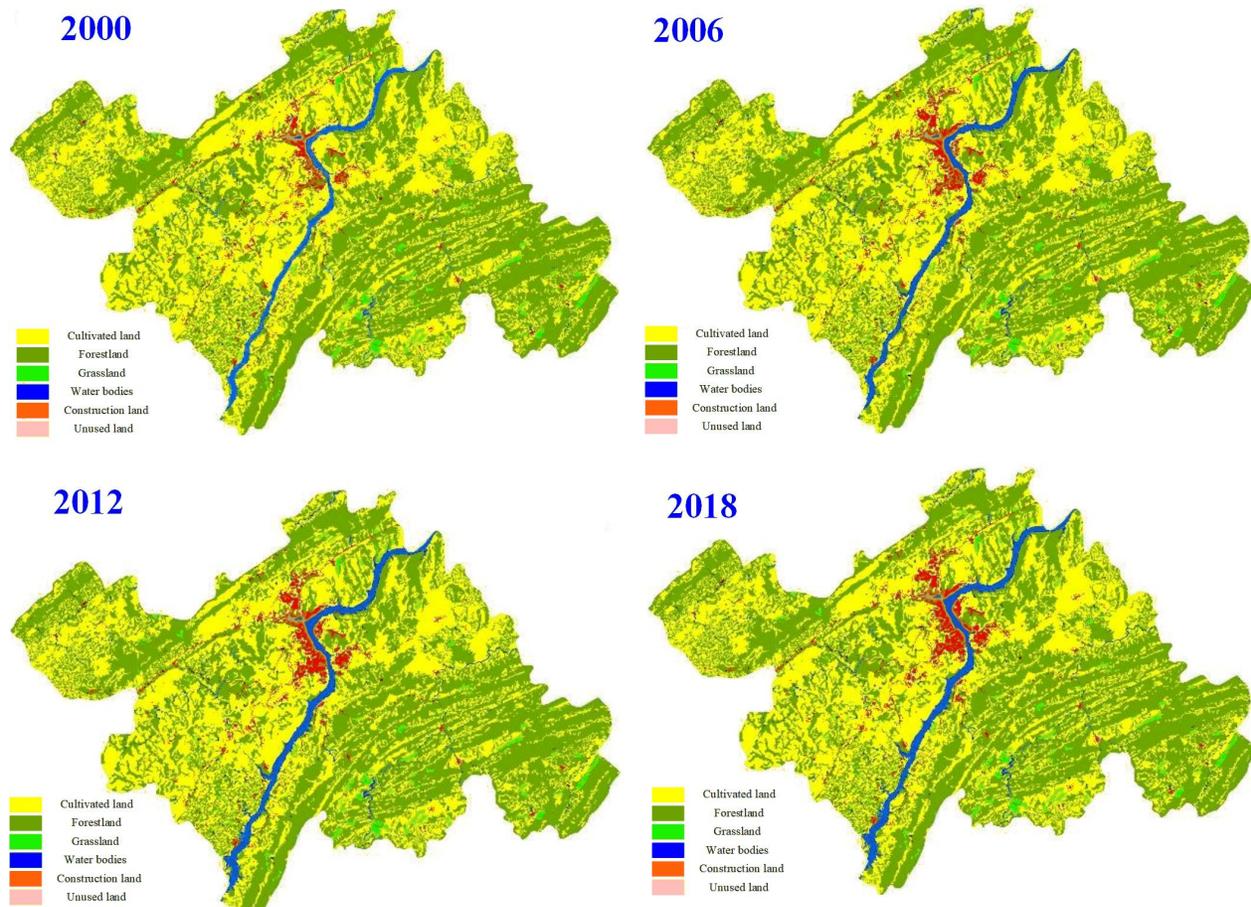


Fig. 3. Distribution map of ecological space structure types in wanzhou district from 2000 to 2018.

TABLE II. AREA STATISTICS OF ECOLOGICAL SPACE STRUCTURE TYPES IN WANZHOU DISTRICT FROM 2000 TO 2018

Type	2000		2006		2012		2018	
	Area (hm ²)	Percentage (%)						
Ecological Space	337826.13	98.33%	336026.89	97.80%	334956.05	97.49%	334279.17	97.29%
Arable Land	152779.02	44.47%	150913.92	43.92%	149056.68	43.38%	148385.62	43.19%
Woodland	169647.69	49.38%	169140.76	49.23%	168503.00	49.04%	168417.14	49.02%
Grassland	5906.24	1.72%	5826.97	1.70%	5754.00	1.67%	5848.84	1.70%
Water Bodies	9461.36	2.75%	10138.51	2.95%	11637.71	3.39%	11618.41	3.38%
Unused Land	31.82	0.01%	6.73	0.00%	4.66	0.00%	9.17	0.00%
Developed Land	5748.00	1.67%	7547.25	2.20%	8618.08	2.51%	9294.97	2.71%

Based on the 2006–2012 land use transition matrix (Table IV) for Wanzhou District, the conversion of arable land to built-up land significantly decreased compared to the period 2000–2006.

According to Table V, arable land no longer converts to forestland but solely transitions from arable land to forestland, with a conversion area of 19.57 hm². This indicates that during the 2012–2018 period, the policy of converting farmland back to forestland was strongly implemented in Wanzhou District. During this time, the primary land use transitions were from arable land and forestland to grassland and built-up land. The most significant conversion was from arable land to built-up land, with an area of 612.62 hm², followed by the conversion from forestland to built-up land, with an area of 64.27 hm².

According to Table VI, during the period from 2000 to 2018, the primary transition for arable land was towards built-up land, with a conversion area of 2885.10 hm², followed by transition to water bodies, totaling 1455.37 hm². Similarly, forestland mainly converted to built-up land, with an area of 654.03 hm², and secondarily to water bodies, with 643.38 hm². Grassland predominantly transformed into built-up land (57.50 hm²) and forestland (42.93 hm²). Water bodies mainly converted to arable land (27.03 hm²) and built-up land. Built-up land primarily transitioned to water bodies, with an area of 63.22 hm². Unused land primarily transformed into arable land (11.42 hm²) and built-up land (7.86 hm²).

TABLE III. LAND USE TRANSITION MATRIX FOR WANZHOU DISTRICT (HM²) FOR 2000–2006

	2000-2006					
	Arable Land	Woodland	Grassland	Water Bodies	Developed Land	Unused Land
Arable Land	150852.15	40.01	0.00	418.10	1465.34	3.43
Woodland	3.00	169050.12	13.02	276.19	302.04	3.30
Grassland	18.52	42.93	5806.04	6.89	31.86	
Water Bodies	20.96		3.07	9437.32		
Developed Land	—	—	—	—	5748.00	—
Unused Land	19.28	7.70	4.84	—	—	—

TABLE IV. LAND USE TRANSITION MATRIX FOR WANZHOU DISTRICT (HM²) FOR 2006–2012

	2006-2006					
	Arable Land	Woodland	Grassland	Water Bodies	Developed Land	Unused Land
Arable Land	149023.48	33.02	—	1038.91	815.01	3.50
Woodland	10.68	168453.66	1.08	386.47	287.72	1.16
Grassland	11.39	13.02	5752.93	20.92	28.71	—
Water Bodies	7.71	—	—	10128.17	2.63	—
Developed Land	—	—	—	63.22	7484.02	—
Unused Land	3.43	3.3	—	—	—	—

TABLE V. LAND USE TRANSITION MATRIX FOR WANZHOU DISTRICT (HM²) FOR 2012–2018

	2010-2018					
	Arable Land	Woodland	Grassland	Water Bodies	Developed Land	Unused Land
Arable Land	148385.62	19.57	38.88	—	612.62	—
Woodland	—	168397.57	36.65	—	64.27	4.50
Grassland	—	—	5754.00	—	—	—
Water Bodies	—	—	19.29	11618.41	—	—
Developed Land	—	—	—	—	8618.08	—
Unused Land	—	—	—	—	—	4.66

TABLE VI. LAND USE TRANSITION MATRIX FOR WANZHOU DISTRICT (HM²) FOR 2000–2018

	2010-2018					
	Arable Land	Woodland	Grassland	Water Bodies	Developed Land	Unused Land
Arable Land	148303.57	92.59	38.88	1455.37	2885.10	3.50
Woodland	13.68	168273.92	57.02	643.38	654.03	5.66
Grassland	29.91	42.93	5748.09	27.81	57.50	—
Water Bodies	27.03	—	—	9428.63	5.69	—
Developed Land	—	—	—	63.22	5684.78	—
Unused Land	11.42	7.70	4.48	—	7.89	4.66

TABLE VII. TRANSFER RATE OF ECOLOGICAL SPACE IN WANZHOU DISTRICT FROM 2000 TO 2018 (%)

	2000—2006	2006—2012	2012-2018
Arable Land	0.09	0.05	0.03
Woodland	0.20	0.21	0.08
Grassland	0.05	0.06	0.01
Water Bodies	0.22	0.21	-0.27
Developed Land	-1.19	-2.46	0.03
Unused Land	13.14	5.12	-16.11
Comprehensive Dynamic Degree	0.08	0.10	0.05

E. Analysis of Ecological Space Dynamics

When studying the evolution characteristics of ecological space, it is necessary to examine the transfer rate and the activity level of the ecological space. Therefore, based on the area statistics table of Wanzhou District’s ecological space types from 2000 to 2018, the transfer rate and comprehensive dynamic degree of Wanzhou District’s ecological space from 2000 to 2018 are calculated, with the results shown in Table VII.

F. Influencing Factors

The changes in the ecological space layout of Wanzhou District are a complex process, influenced by a combination of various natural and human factors. Based on extensive literature review and adhering to the four principles of factor selection mentioned earlier, the preliminary selected influencing factors

include: elevation factor, slope factor, road traffic factor, administrative center factor, river factor, and policy control factor. The specific factors of elevation, slope, road traffic, administrative center, river, and policy control are illustrated in Fig. 4.

After the preliminary selection of influencing factors such as elevation, slope, road traffic, administrative center, river, and policy control factors, the Empirical Likelihood method is used to further select and analyze the elevation, slope, road traffic, administrative center, and river factors. The results are shown in Fig. 5. The reason for not applying the Empirical Likelihood method to further verify and select the policy control factor is that it serves as a constraint, significantly impacting the ecological space layout changes.

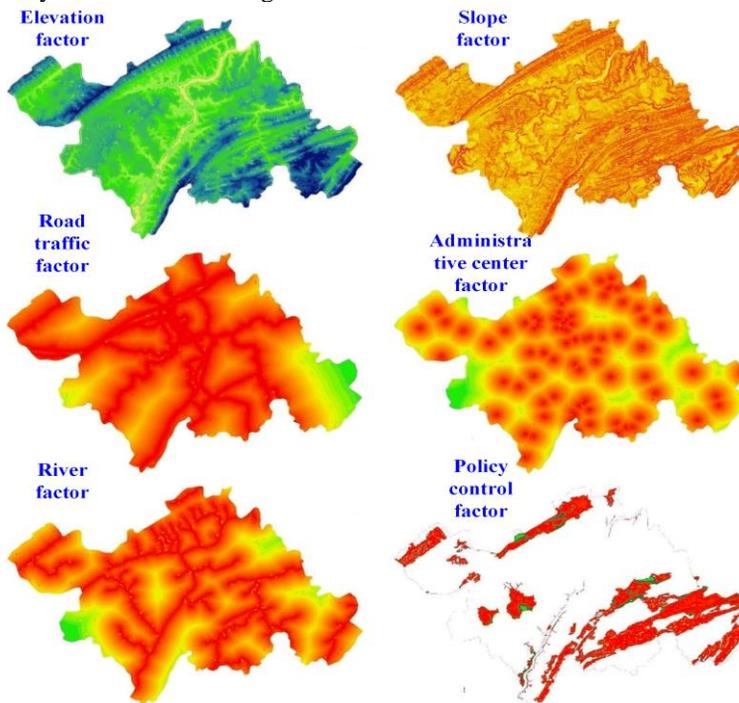


Fig. 4. Wanzhou district influencing factors map.

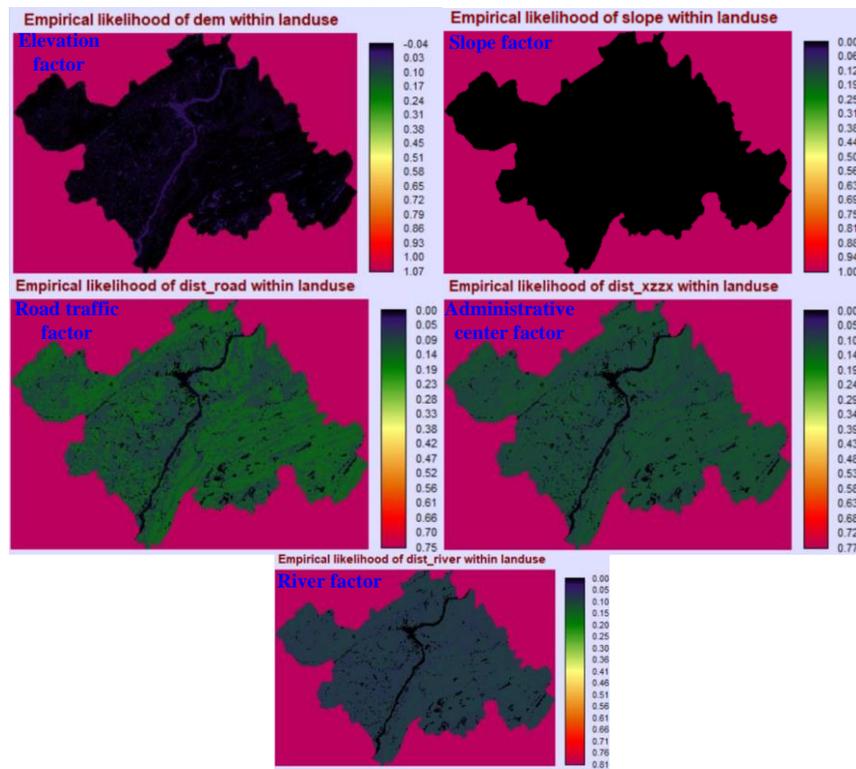


Fig. 5. Wanzhou district influencing factors impact map.

Based on Fig. 5, the final selected influencing factors are elevation, road traffic, administrative center, and policy control factors.

V. SIMULATION OF ECOLOGICAL SPACE LAYOUT IN WANZHOU DISTRICT BASED ON THE ANN-CA-MARKOV MODEL

A. ANN-CA-Markov

The Ann-CA-Markov model, aimed at simulating Wanzhou District's ecological space pattern, is constructed as follows: (1) Cells are defined as 30m×30m grids based on TM image data, mirroring the district's ecological structure. (2) Cellular space consists of these grids. (3) Cell states represent ecological and non-ecological spaces, categorized into types like cultivated land, forest, grassland, water bodies, and construction land. (4) The neighborhood is defined using a 5×5 extended Moore setup, where 24 adjacent cells affect the central one. The choice of a 5×5 neighborhood window balances computational efficiency with spatial accuracy. This window size was selected based on sensitivity analysis, which demonstrated that smaller windows (e.g., 3×3) provided insufficient spatial context, while larger windows (e.g., 7×7) added unnecessary complexity without significantly improving model precision. (5) Transition rules are established using spatial and quantitative methods, with the former calculated via the MLP_ANN model and the latter through Markov model-derived transition probability matrices between space types. (6) Transition probabilities are calculated for 2000—2006, 2006—2012, and 2012—2018 using IDRISI software, aiding in predicting land use changes, shown in Tables VIII, IX, and X, respectively.

After selecting spatial variables and influencing factors, the MLP_ANN tool was utilized to construct the transition rules for

the cellular automaton model (Fig. 6). The MLP_ANN comprises a three-layer network structure, including an input layer, a hidden layer, and an output layer. In this study's MLP_ANN, the input layer consists of 15 neurons, corresponding to the 15 influencing factors identified earlier. The number of neurons in the hidden layer, representing the number of input samples, should be at least two-thirds of the number of input layer neurons, hence 11 neurons were set for the hidden layer. The output layer contains 6 neurons, each corresponding to the transition probabilities for Wanzhou District's five ecological space types and one non-ecological space type. However, due to the negligible area of unused land in the study area, the actual number of neurons in the output layer used is five, representing the transition probabilities for farmland, woodland, grassland, water bodies, and construction land.

Following the computation of transition probability maps for various land use types, including farmland, woodland, grassland, water bodies, and construction land, using the MLP_ANN model, the results were refined to reflect both the prevailing policies and the real-world conditions of the study area. To ensure that the study's outcomes were consistent with local land use regulations, specific constraints were applied. These constraints mandated that construction land and water bodies could not change, while ecological spaces and non-ecological areas within designated protected zones—such as ecological protection red lines, forest parks, nature reserves, geological parks, and scenic areas—were preserved from any transitions. As the ANN-CA-Markov model is built on the principles of cellular automata (CA), the refined transition rules, now represented as probability maps for each land use type,

were standardized to a 0–255 scale. The standardized transition probability maps for the respective land types are presented in Fig. 7. In the final step, these maps were integrated into a

comprehensive suitability mapset using the Collection Editor tool within the IDRISI software suite.

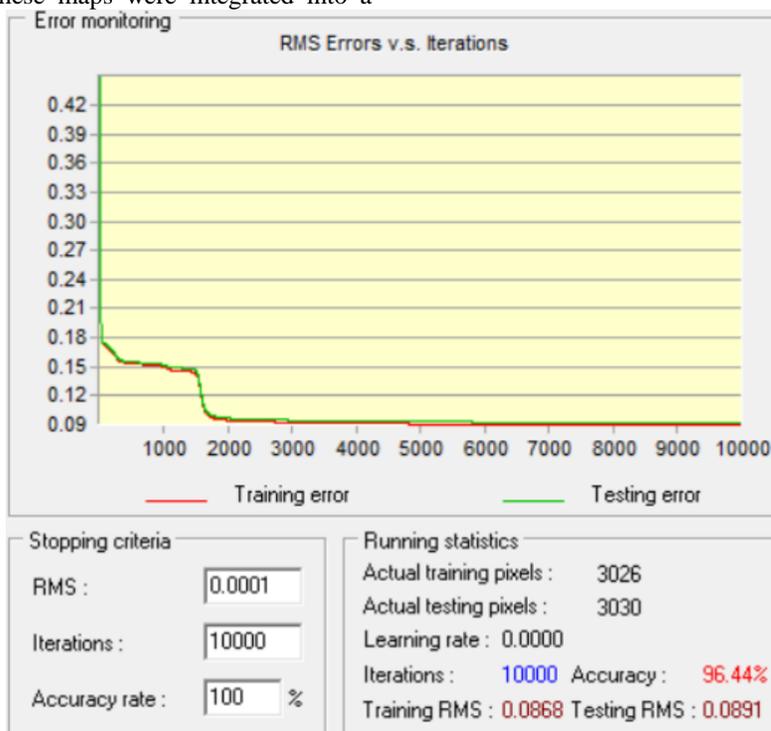


Fig. 6. MLP_ANN learning and calculation process diagram.

TABLE VIII. TRANSITION PROBABILITY MATRIX OF ECOLOGICAL SPACE TYPES IN WANZHOU DISTRICT FROM 2000—2006 (%)

	Arable Land	Woodland	Grassland	Water Bodies	Developed Land	Unused Land
Arable Land	0.8393	0.0033	0	0.035	0.1222	0.0003
Woodland	0.0007	0.847	0.0032	0.0708	0.0774	0.0008
Grassland	0.0307	0.0703	0.8356	0.0117	0.0517	0
Water Bodies	0.1345	0	0.0177	0.8478	0	0
Developed Land	0.03	0.03	0.03	0.03	0.85	0.03
Unused Land	0.6091	0.238	0.153	0	0	0

TABLE IX. TRANSITION PROBABILITY MATRIX OF ECOLOGICAL SPACE TYPES IN WANZHOU DISTRICT FROM 2006—2012 (%)

	Arable Land	Woodland	Grassland	Water Bodies	Developed Land	Unused Land
Arable Land	0.8393	0.0028	0	0.0882	0.0694	0.0003
Woodland	0.0023	0.8465	0.0003	0.0866	0.064	0.0003
Grassland	0.0252	0.0276	0.8393	0.0459	0.062	0
Water Bodies	0.1168	0	0	0.8491	0.0341	0
Developed Land	0	0	0	0.1571	0.8429	0
Unused Land	0.5068	0.4932	0	0	0	0

TABLE X. TRANSITION PROBABILITY MATRIX OF ECOLOGICAL SPACE TYPES IN WANZHOU DISTRICT FROM 20012—2018 (%)

	Arable Land	Woodland	Grassland	Water Bodies	Developed Land	Unused Land
Arable Land	0.8462	0.0045	0.0086	0	0.1407	0
Woodland	0	0.8495	0.053	0	0.0909	0.0066
Grassland	0.03	0.03	0.85	0.03	0.03	0.03
Water Bodies	0	0	0.1514	0.8486	0	0
Developed Land	0.03	0.03	0.03	0.03	0.85	0.03
Unused Land	0.03	0.03	0.03	0.03	0.03	0.85

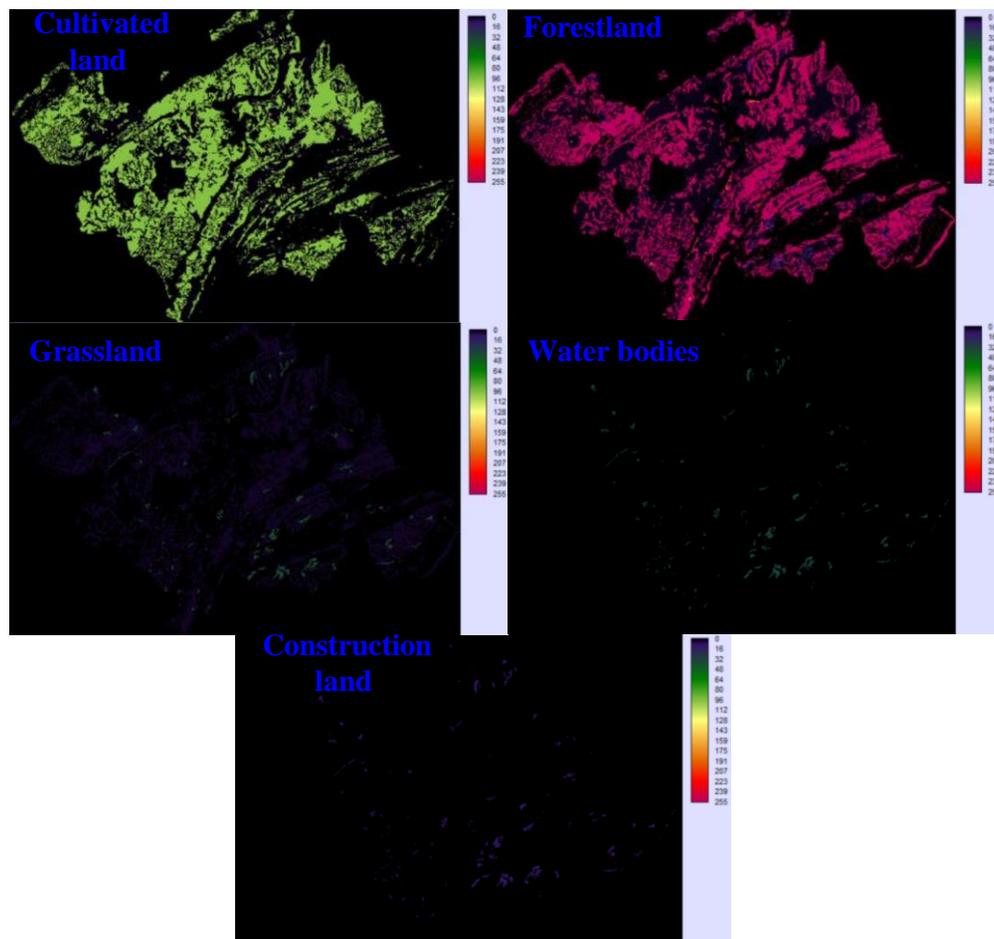


Fig. 7. Conversion probability maps for cultivated land, forest land, grassland, water area, and construction land.

B. Determining the Number of Iterations and the Forecast Year

After setting the iteration interval to 6 years, based on the base data and calculated transition probability matrices, the ecological space evolution in the study area for 2012 is simulated and predicted based on the 2000 and 2006 data. Then, the ecological space evolution in 2018 is simulated and predicted based on the 2006 and 2012 data. After verifying the accuracy of the simulated ecological space for 2012 and 2018 with the ANN-CA-Markov coupled model and meeting the accuracy requirements, the ecological space evolution in 2025 is simulated and predicted. Thus, in the ANN-CA-Markov coupled model, 2012, 2018, and 2025 are determined as the forecast years.

C. ANN-CA-Markov Coupled Model Accuracy Verification

After constructing the ANN-CA-Markov coupled model, the ecological space distribution in Wanzhou District for the years 2012 and 2018 is simulated based on the data from 2000 to 2006 and 2006 to 2012, respectively. The simulation results are shown in Fig. 8 and 9.

After completing the simulation maps of the ecological space structure types in Wanzhou District for 2012 and 2018, it is necessary to verify the simulation accuracy of the ANN-CA-Markov coupled model. This is done by comparing the simulation results for 2012 with the actual data for 2012, and the simulation results for 2018 with the actual data for 2018, to validate the simulation accuracy of the ANN-CA-Markov coupled model.

First, a quantitative verification is performed by comparing the simulated results for 2012 and 2018 with the actual data for those years. The specific quantitative verification results are shown in Tables XI and XII.

Based on Tables XI and XII, except for unused land, the simulation results for other ecological space types (cultivated land, forest land, grassland, water bodies) and non-ecological space type (construction land) show high accuracy rates. The lower accuracy for unused land is due to its minimal grid cell count, which does not reach 0.01% of the total grid cell count in the study area. Overall, the ANN-CA-Markov coupled model demonstrates high simulation accuracy in terms of quantity.

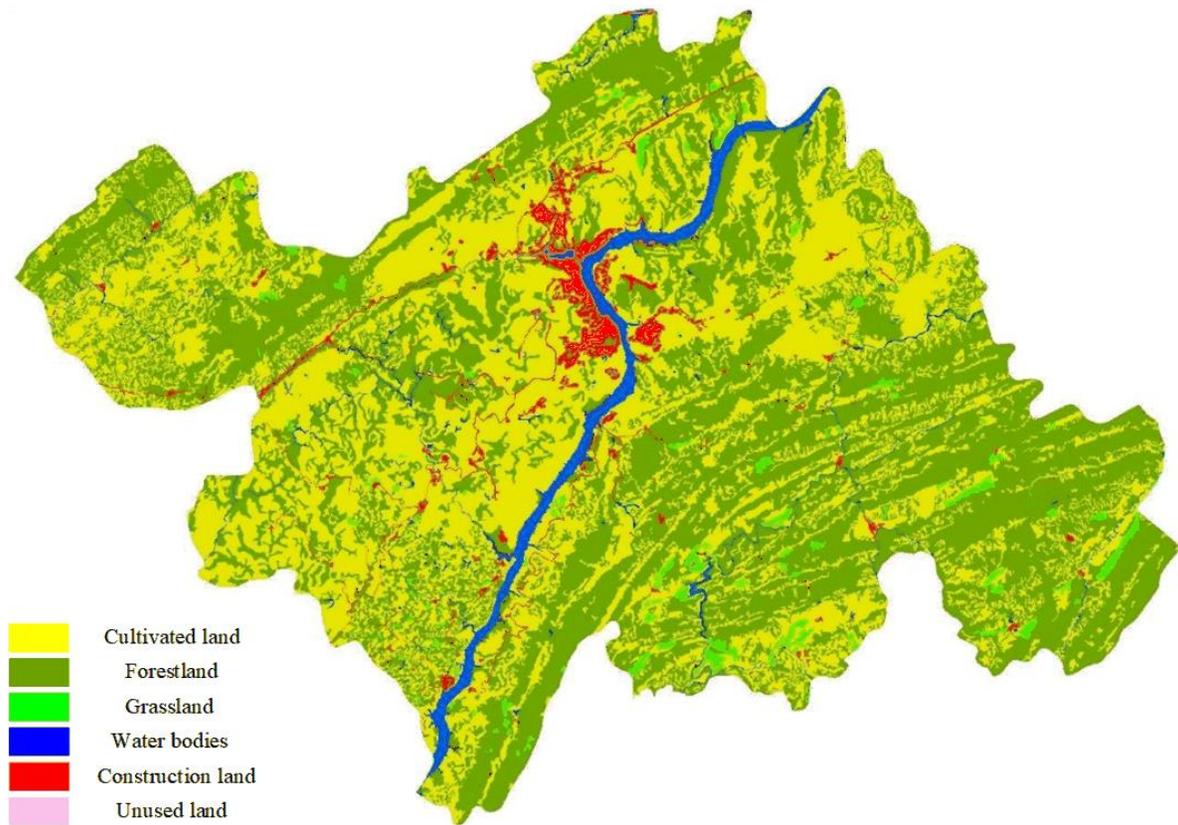


Fig. 8. Simulation map of the ecological space structure types in Wanzhou District for the year 2012.

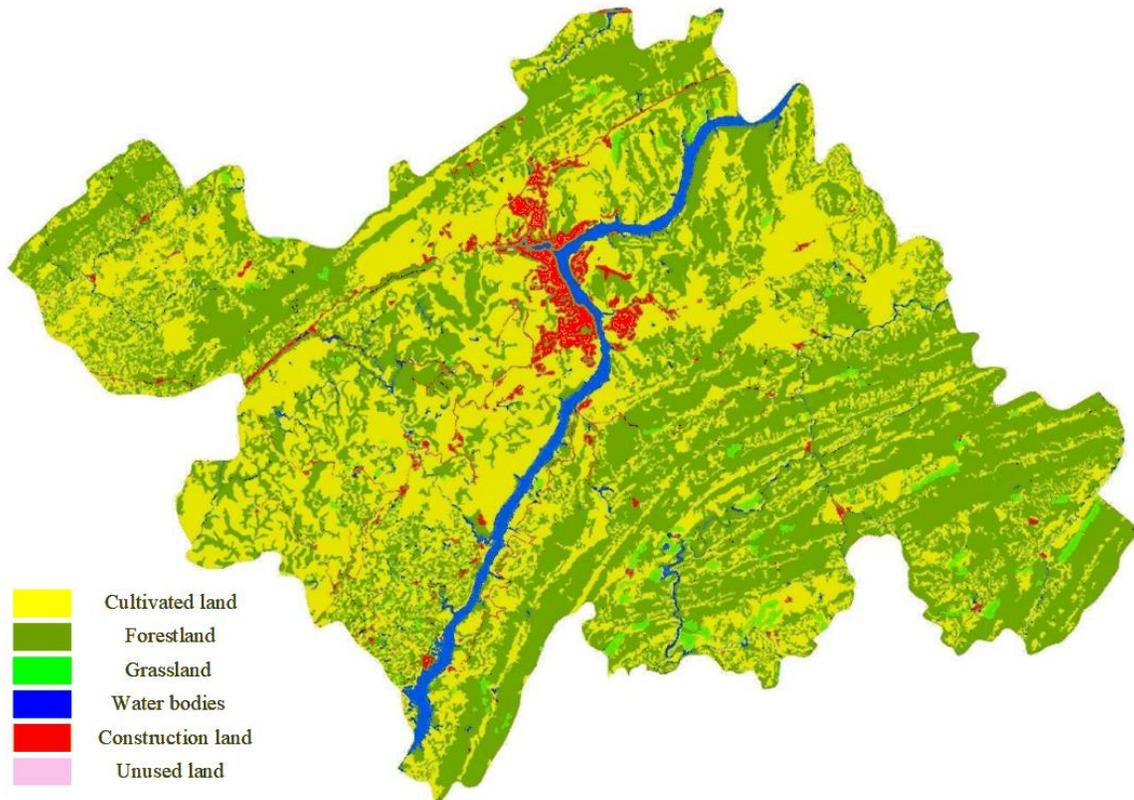


Fig. 9. Simulation map of the ecological space structure types in Wanzhou district for the year 2018.

TABLE XI. VERIFICATION OF THE SIMULATED AND ACTUAL NUMBER OF GRIDS FOR 2012

Land Type	Actual Number of Grids (2012)	Simulated Number of Grids (2012)	Accuracy (%)
Cultivated Land	1,655,590	1,672,030	99.01
Forest Land	1,870,820	1,873,846	99.84
Grassland	63,945	65,175	98.08
Water Area	129,273	113,517	87.81
Unused Land	55	73	67.27
Construction Land	95,783	90,825	94.82

TABLE XII. VERIFICATION OF THE SIMULATED AND ACTUAL NUMBER OF GRIDS FOR 2018

Land Type	Actual Number of Grids (2018)	Simulated Number of Grids (2018)	Accuracy (%)
Cultivated Land	1,648,162	1,661,779	99.17
Forest Land	1,869,859	1,870,178	99.98
Grassland	64,994	55,553	85.47
Water Area	129,054	128,992	99.95
Unused Land	107	55	51.40
Construction Land	103,290	98,909	95.76

D. Simulation and Result Analysis of Ecological Space Layout in the Study Area for 2025

After validating the Ann-CA-Markov model’s high accuracy, it was applied to project Wanzhou District’s 2025 ecological space. For the 2025 simulation, the model used 2018 data for neighboring variables and current space types, adjusted conversion probability maps, and derived the transition matrix from 2012 and 2018 data. These modifications enabled the projection of the district’s 2025 ecological layout, depicted in Fig. 10. A corresponding map illustrating the predicted ecological space distribution for 2025 is shown in Fig. 11.

Based on the simulation results for Wanzhou District in 2025 (Fig. 10 and 11) and the actual data from 2018, the areas of

ecological spaces in Wanzhou District were calculated and summarized, as shown in Table XIII.

In 2025, ecological space in Wanzhou District is projected to cover 334,022.95 hectares (97.28%), dominated by cultivated land (147,743.20 hectares) and forest land (168,944.40 hectares). Grassland (5,698.71 hectares) and water bodies (11,627.26 hectares) contribute smaller portions, while unused land remains negligible. Constructed land expands to 9,349.42 hectares (2.72%). From 2018 to 2025, the ecological space decreased slightly by 81.02 hectares, as cultivated land and grassland declined, while forest land and water bodies saw modest increases. The transfer matrix in Table XIV highlights the shifts between these land types during the period.

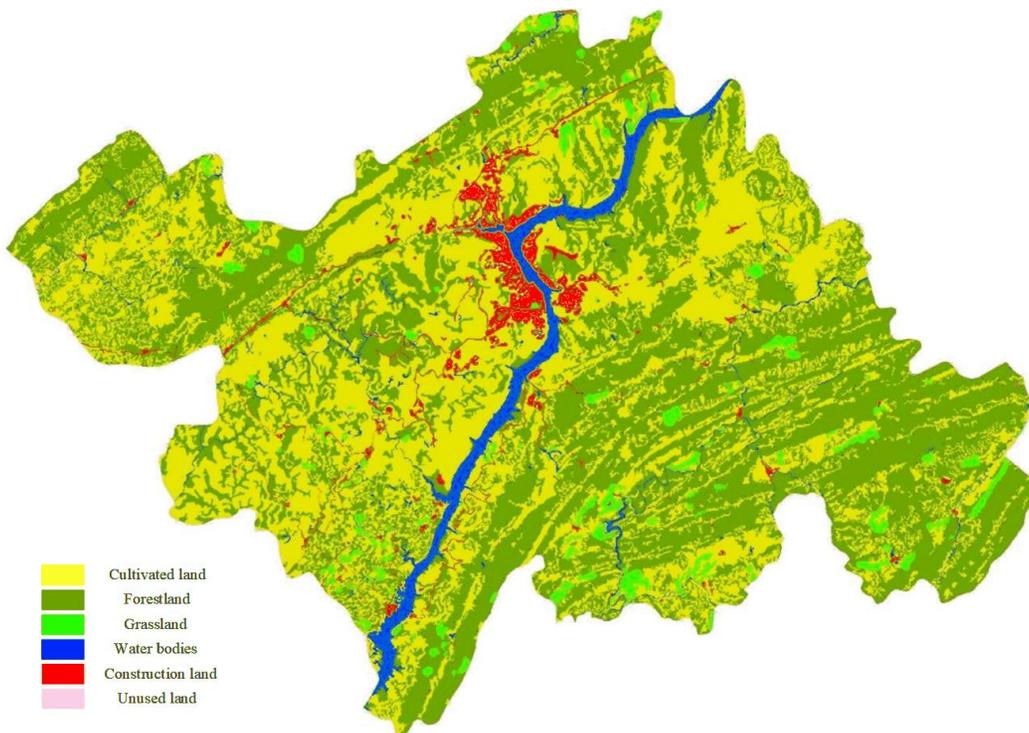


Fig. 10. Simulated distribution map of ecological space structure types in wanzhou district for 2025.

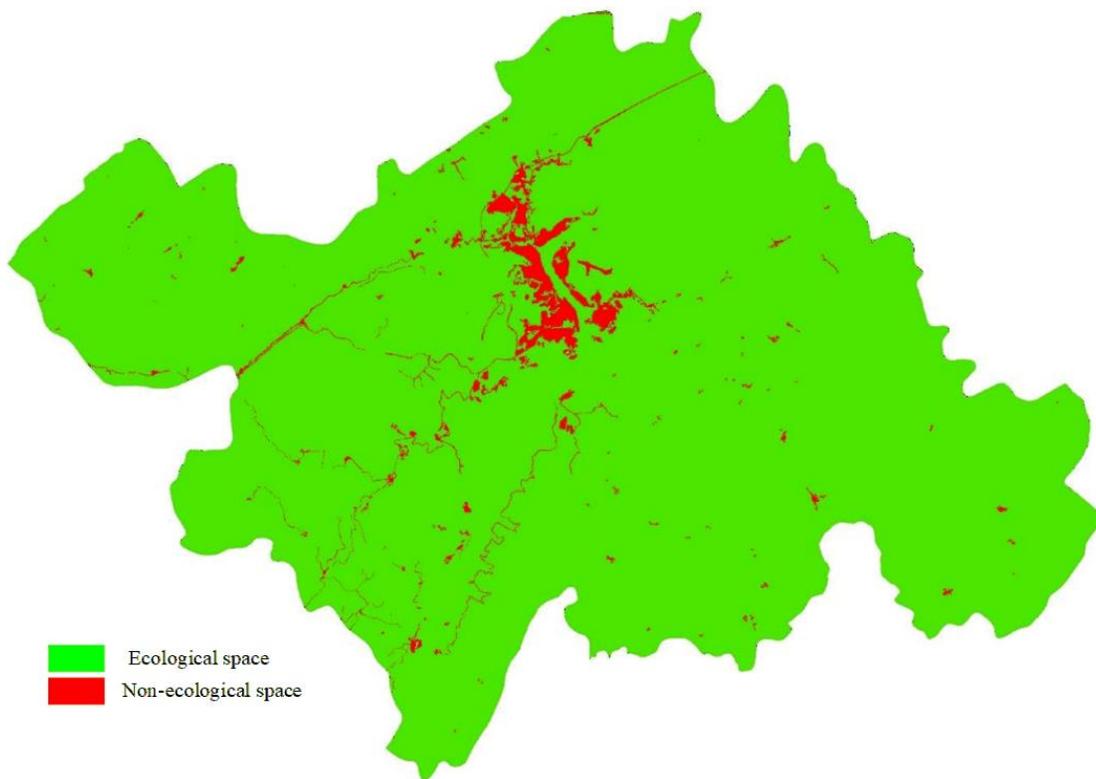


Fig. 11. Simulated ecological space distribution map in wanzhou district for 2025.

TABLE XIII. SIMULATED RESULTS OF ECOLOGICAL SPACE IN WANZHOU DISTRICT FOR 2025

Type	2018		2025	
	Area (hm ²)	Proportion(%)	Area (hm ²)	Proportion(%)
Ecological Space	334103.97	97.30	334022.95	97.28
Cultivated Land	148328.32	43.20	147743.20	43.03
Forest Land	168317.33	49.02	168944.40	49.20
Grassland	5843.57	1.70	5698.71	1.66
Water Body	11605.37	3.38	11627.26	3.39
Unused Land	9.37	0.00	9.37	0.00
Constructed Land	9268.41	2.70	9349.42	2.72

TABLE XIV. TRANSFER MATRIX OF ECOLOGICAL SPACE STRUCTURE TYPES IN WANZHOU DISTRICT FOR 2018-2025 (HM²)

2018	2025					
	Arable Land	Woodland	Grassland	Water Bodies	Developed Land	Unused Land
Arable Land	147528.64	742.57	9.07	4.20	43.85	0.00
Woodland	88.62	168128.55	23.34	11.90	64.93	0.00
Grassland	48.58	60.61	5653.00	34.73	46.64	0.00
Water Bodies	3.25	8.61	8.30	11572.30	12.90	0.00
Developed Land	74.11	4.06	5.00	4.13	9181.11	0.00
Unused Land	0.00	0.00	0.00	0.00	0.00	9.37

According to Table XIV, during the period from 2018 to 2025, the area of cultivated land mainly transferred to forest land and constructed land, with transfer areas of 742.57 hm² and 43.85 hm², respectively. The primary transfers from forest land were to cultivated land and constructed land, with transfer areas of 88.62 hm² and 64.93 hm², respectively. Grassland transfers were relatively balanced among various types, with transfer

areas to cultivated land, forest land, water bodies, and constructed land being 48.58 hm², 60.61 hm², 34.73 hm², and 46.64 hm², respectively. Water bodies had smaller transfer areas to other types, with the largest being the transfer to constructed land, at 12.90 hm². Constructed land primarily transferred to cultivated land, with a transfer area of 74.11 hm². Unused land did not undergo any transfers.

VI. CONCLUSION

The deployment of the ANN-CA-Markov model in this study provides a detailed and forward-looking analysis of the expected land use changes in Wanzhou District by the year 2025. The results of the analysis suggest that ecological spaces will remain predominant, accounting for 97.28% of the district's total land area. However, a minor reduction in ecological spaces is forecasted, accompanied by a corresponding increase in construction land, indicating the growing impact of urbanization on the region's ecological zones. Specifically, the study anticipates a decline in arable land and grassland areas, while forested regions and water bodies are projected to expand. This shift may reflect the influence of regional policies and planning initiatives aimed at conserving forests and water resources. Although the increase in construction land is relatively small, it nonetheless reflects the broader trend of intensified land use and development driven by urbanization pressures. Furthermore, an examination of the transition dynamics between various land use categories reveals a pattern where arable land is increasingly converted to forested areas and construction sites, while transitions involving forest land typically lead to its conversion into arable land or construction areas. These land use transformations are likely driven by a combination of policy enforcement, economic development, and resource management practices.

REFERENCES

- [1] Wang L, Wang Y, Cai Y. An ANN-CA Modeling Method for Land Cover Change in the Karst Area of China: A Case Study of Maotiao River Basin. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2012, 48(1):116-122. DOI:10.1007/s11783-011-0280-z.
- [2] Zeshan M T, Mustafa M R U , Baig M F .Monitoring Land Use Changes and Their Future Prospects Using GIS and ANN-CA for Perak River Basin, Malaysia. *Multidisciplinary Digital Publishing Institute*, 2021(16). DOI:10.3390/W13162286.
- [3] Zhang Y , Qiao J , Wu B ,et al.Simulation of oil spill using ANN and CA models[C]//International Conference on Geoinformatics.IEEE, 2016. DOI:10.1109/GEOINFORMATICS.2015.7378560.
- [4] Tingting D , Lijun Z , Zengxiang Z .A Study on Spacetime Evolution of Soil Erosion Based on ANN-CA Model. *Journal of Geo-Information Science*, 2009, 11(1):132-138. DOI:10.3969/j.issn.1560-8999.2009.01.020.
- [5] Xi-Bao X U, Gui-Shan Y , Jian-Ming Z .Scenario Modeling of Urban Land Use Changes in Lanzhou with ANN-CA. *Geography and Geo-Information Science*, 2008, 24(6):80-84. DOI:10.3724/SP.J.1047.2008.00114.
- [6] João Lita da Silva, Caeiro F ,Isabel Natário,et al.Advances in regression, survival analysis, extreme values, Markov processes and other statistical applications. Selected papers based on the presentations of the 17th congress of the Portuguese Statistical Society, Sesimbra, Portugal, September 30–October. 012.
- [7] Cryan M, Goldberg L A , Goldberg P W .Evolutionary trees can be learned in polynomial time in the two-state general Markov model. *SIAM Journal on Computing*, 2002. DOI:10.1137/s0097539798342496.
- [8] Deruiter S L, Langrock R , Skirbutas T ,et al.A multivariate mixed hidden Markov model for blue whale behaviour and responses to sound exposure. *Annals of Applied Statistics*, 2017, 11(1):362-392. DOI:10.1214/16-AOAS1008.
- [9] Zhan Q, Tian J , Tian S .Prediction Model of Land Use and Land Cover Changes in Beijing Based on Ann and Markov_CA Model[C]//IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium.IEEE, 2019. DOI:10.1109/IGARSS.2019.8898388.
- [10] Ynoguti C A, Morais E D S , Violaro F .A comparison between HMM and hybrid ANN-HMM-based systems for continuous speech recognition[C]//Telecommunications Symposium, 1998. ITS '98 Proceedings. SBT/IEEE International.IEEE, 1998. DOI:10.1109/ITS.1998.713105.
- [11] Mondal M S. Modeling of Land Use Land Cover Change: CA Markov Modeling Approaches[M]. 2013.
- [12] Liu J, Cheng F , Zhu Y ,et al.Urban Land-Use Type Influences Summertime Water Quality in Small- and Medium-Sized Urban Rivers: A Case Study in Shanghai, China. *Land*, 2022, 11. DOI:10.3390/land11040511.
- [13] Xu T, Zhou D, Li Y. Integrating ANNs and cellular automata–Markov chain to simulate urban expansion with annual land use data[J]. *Land*, 2022, 11(7): 1074.
- [14] Zhao J, Zhu X, Zhou Y, et al. Examining land-use change trends in yucheng district, Ya'an city, China, using ANN-CA modeling[J]. *Journal of Urban Planning and Development*, 2023, 149(1): 05022042.
- [15] Zambrano-Asanza S, Morales R E, Montalvan J A, et al. Integrating artificial neural networks and cellular automata model for spatial-temporal load forecasting[J]. *International Journal of Electrical Power & Energy Systems*, 2023, 148: 108906.
- [16] Xie S, Hu Z, Wang J. Two-stage robust optimization for expansion planning of active distribution systems coupled with urban transportation networks. *Applied Energy*, 2020, 261: 114412.
- [17] Oyalowo B. Implications of urban expansion: land, planning and housing in Lagos. *Buildings & Cities*, 2022, 3(1).
- [18] Woldeamayrat E M, Genovese P V. Monitoring urban expansion and urban green spaces change in Addis Ababa: directional and zonal analysis integrated with landscape expansion index. *Forests*, 2021, 12(4): 389.
- [19] Theres L, Radhakrishnan S, Rahman A. Simulating Urban Growth Using the Cellular Automata Markov Chain Model in the Context of Spatiotemporal Influences for Salem and Its Peripherals, India. *Earth*, 2023, 4(2): 296-314.