Spatiotemporal Vegetation Degradation Simulation and Inversion in Inner Mongolia Autonomous Region

He Yang [1], Yongjun Zhang [1,2], Yongxiang Yao [1,2*], Duanning Chen [1], Yi Wan [1,2], Bin Zhang [3]

[1] School of Remote Sensing and Information Engineering, Wuhan University, 430079 Wuhan Hubei Province, China - (yheyhe, zhangyj, yaoyongxiang, 2024302141175, yi.wan)@whu.edu.cn;

[2] Hubei Luojia Laboratory, 430079 Wuhan Hubei Province, China;

[3] China railway siyuan survey and design group CO., LTD, 430063 Wuhan Hubei Province, China)- (007981@crfsdi.com)

Commission III, WG III/4

KEY WORDS: Spatiotemporal Vegetation Degradation, Cellular Automata, Random Forest, Improved Genetic Algorithm, Google Earth Engine

ABSTRACT:

Under climate and human pressures, vegetation in Inner Mongolia exhibits complex fragmentation and degradation. Scientifically inverting its spatiotemporal dynamics is crucial for regional ecological restoration. To address the challenges faced by traditional cellular automata (CA) models in large-scale complex ecological transition zones—such as computing power bottlenecks and subjective transition rules—this study proposes a cloud-based vegetation degradation simulation and inversion framework (CA-VDS) via Google Earth Engine. By coupling Random Forest (RF) and an Improved Genetic Algorithm (IGA) with CA, the framework extracts nonlinear driving potentials and automates the optimization of bidirectional transition thresholds. Validation against the 2020 baseline shows CA-VDS effectively resolves manual parameter tuning limitations. Furthermore, it smooths the spectral fluctuations caused by short-term sporadic disturbances through the underlying spatial neighborhood mechanism, demonstrating its value in simulating potential ecological degradation risks and developmental trajectories. This work not only verifies the reliability of CA-VDS in analyzing complex nonlinear ecological processes, but also establishes a reliable parameter baseline and model paradigm for subsequent integration with CMIP6 and other multi-scenario data to conduct long-term future ecological predictions.

1. Introduction

The Inner Mongolia Autonomous Region spans arid, semi-arid, and semi-humid climate zones, serving as an important ecological security barrier in northern China (Tong, 2016). However, under the dual stress of global climate change and high-intensity human activities, the vegetation in this region exhibits complex spatiotemporal characteristics of coexisting fragmentation and degradation. Scientifically inverting the driving mechanisms of historical vegetation degradation hold crucial significance for the formulation of regional ecological restoration policies.

Currently, regional ecological assessments highly rely on high-precision static remote sensing classification products. Although such products can provide extremely fine instantaneous surface snapshots, limited by strong climate variability, they often record apparent noise caused by short-term precipitation fluctuations or episodic disturbances. Without the mechanistic constraints of dynamic evolution models, it is difficult to effectively filter out these instantaneous disturbances using static snapshots alone, which consequently limits a deep understanding of the long-term succession potential of the vegetation system.

Meanwhile, as a supplement to dynamic mechanisms, geographic cellular automata (CA) models also face significant technical bottlenecks in their application within ecological transition zones like Inner Mongolia. First, the fusion efficiency of multi-source heterogeneous data and large-scale computational bottlenecks urgently need to be addressed; traditional simulation methods based on local computing power struggle to support the calculation of high-resolution ecological factors (Gorelick, 2017). Second, the ability to capture spatial heterogeneity and the subjective nature of transition rules severely restrict the model's reliability. The transition thresholds of traditional CA models mostly rely on expert knowledge (Wu, 1998) or empirical judgment methods (Clarke, 1997), lacking objective automated parameter inversion mechanisms, making it difficult to accurately reproduce the bidirectional nonlinear succession process of vegetation "degradation-recovery". The absence of combining dynamic and static approaches and objective optimization in existing assessment systems urgently requires the development of a novel simulation tool integrating cloud computing power and intelligent optimization algorithms.

To address the above issues, relying on the Google Earth Engine (GEE) cloud platform, which integrates massive high-resolution remote sensing data resources (Zhang, 2025), this study proposes a vegetation degradation simulation and inversion framework (CA-VDS) for Inner Mongolia that couples Cellular Automata, Random Forest (RF) (Breiman, 2001), and an Improved Genetic Algorithm (IGA). The main innovations of this study are as follows. (1) Construction of a fully cloud-based bidirectional dynamic simulation framework. It breaks through the computational bottlenecks of local computing power and provides a highly integrated computational foundation. (2) Proposal of an objective parameter inversion method based on an improved genetic algorithm. Coupling RF to capture nonlinear spatial probabilities, this study focuses on adaptively improving the initial population determination mechanism and elite retention strategy of traditional genetic algorithms, achieving efficient automated global optimization of transition rules in complex transition zones.

* Corresponding author

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-695-2026

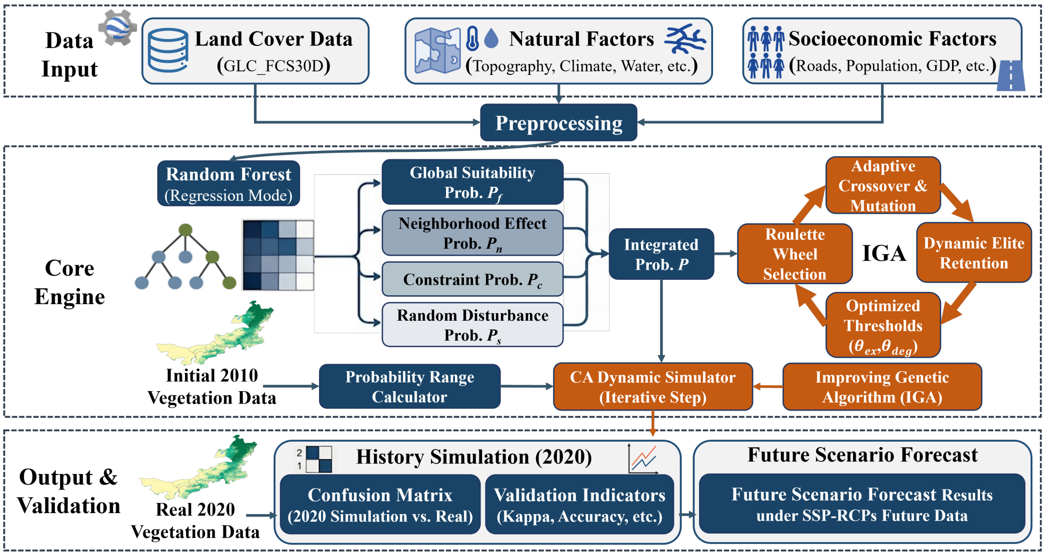

Figure 1. Overall architecture of the proposed CA-VDS.

2. Data and Methods

The proposed CA-VDS framework consists of four main components (Figure 1): (1) Multi-source data preprocessing, which integrates and standardizes heterogeneous physical and socioeconomic datasets on the GEE platform; (2) Parameter inversion and historical simulation, which couples RF, IGA, and CA to optimize bidirectional transition thresholds and invert baseline vegetation patterns; (3) Multi-scenario prediction extension, providing a built-in interface to access CMIP6 data for future projections; and (4) Accuracy assessment.

2.1 Study Area

The Inner Mongolia Autonomous Region is located in the northern frontier of China (37°24′N–53°23′N, 97°12′E–126°04′E), spanning semi-humid, semi-arid, and arid climate zones, and serves as a key ecological security barrier in northern China. Controlled by a precipitation gradient decreasing from east to west, the vegetation coverage in the region exhibits significant spatial differentiation patterns (Wang, 2020), as shown in Figure 2. The Greater Hinggan region in the northeast is dominated by forests, the vast central region is covered by typical steppes, and it gradually transitions into non-vegetated desert areas towards the west. As a typical ecological transition zone, this region has a low environmental carrying capacity. Particularly in the desert margin areas where vegetation and non-vegetation intersect, the risk of vegetation degradation is significant.

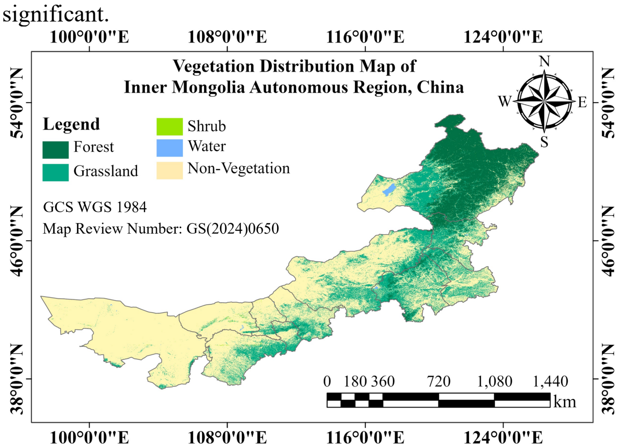

Figure 2. Vegetation distribution in Inner Mongolia Autonomous Region.

2.2 Data

To accurately invert the driving mechanisms of vegetation degradation in Inner Mongolia and support historical baseline simulations, this study constructed a multi-source heterogeneous driving factor dataset encompassing physical geography, ecological environment, and socioeconomic dimensions based on the GEE cloud platform (Table 1).

(1) Baseline vegetation data. The GLC_FCS30D global fine land-cover product with a 30m resolution was selected as the baseline data for validation. (2) Socioeconomic driving factors. Global road networks, nighttime lights, and spatialized GDP data were integrated to characterize the intensity of human activity disturbances; population density data utilized the 30m resolution HRSL dataset, with WorldPop data used to fill in missing areas. (3) Climate parameters. Based on strictly radiometrically corrected Landsat 8 OLI/TIRS imagery, a single-window algorithm was used to retrieve land surface temperature products, serving as a key indicator to measure surface hydrothermal stress. (4) Physical geography data. This includes global river distributions and topographic factors. Topographic data were primarily sourced from FABDEM, with NASADEM used to fill data gaps, which were then used to calculate elevation, slope, and aspect parameters. (5) Vegetation indices. The Normalized Difference Vegetation Index (NDVI) and Fractional Vegetation Cover (FVC) were retrieved based on Landsat 8 imagery to characterize the growth vitality of surface vegetation.

All raw data underwent standardization processing on the GEE platform. Vector data were converted into continuous Euclidean distance rasters via spatial distance functions; all raster data were uniformly resampled to a 30m resolution and projected to a unified coordinate system. Furthermore, all data were normalized to eliminate dimensional differences.

Factor	Dataset Name	Year	Scale	Source
Land Cover	GLC_FCS30D	2000-2020	30m	Zhang, 2024
Roads	GRIP	2018	vector	Meijer, 2018
Night Light	NOAA/VIIRS/DNB	2020	500m	Elvidge, 2021
Population	HRSL	2016	30m	Facebook, 2016
	WorldPop	2020	100m	Bondarenko, 2025
GDP	/	2019	1km	Chen, 2022
Temperature	Landsat8	2020	30m	Landsat8
Elevation	NASADEM	2000	30m	NASA, 2020
	FABDEM	2020	30m	Hawker, 2022
Slope	as above	2020	30m	as above
Aspect	as above	2020	30m	as above
Rivers	/	2019	vector	Yan, 2019

Table 1. Details of the multi-source heterogeneous datasets used in this study.

2.3 Method

The CA-VDS method proposed in this study aims to achieve accurate reconstruction and mechanism analysis of vegetation dynamics in complex ecological transition zones. The core logic lies in coupling the high-dimensional feature extraction capability of machine learning with the global optimization capability of heuristic algorithms.

2.3.1 Calculation of Transition Probabilities Based on Random Forest

The spatiotemporal succession of vegetation states is a complex process driven jointly by natural environmental constraints and human activity disturbances. It exhibits highly nonlinear and spatially heterogeneous response characteristics to driving factors such as climate, topography, and socioeconomics. To accurately invert this spatial heterogeneity and obtain global suitability probabilities, this study constructs a global probability inversion model based on RF. As an ensemble learning algorithm, RF can not only effectively overcome the defect of single classifiers being prone to overfitting, but also demonstrate strong robustness against multicollinearity when processing multi-source, high-dimensional geospatial data.

In model implementation, based on the GLC_FCS30D historical land cover data, this study first employs a stratified random sampling strategy to extract training samples of the "vegetation-non-vegetation" evolutionary trajectory, and maps the 30m resolution multi-dimensional driving factors to the sample feature space. Assuming the RF model contains K decision trees, for a cell x at any spatial location (i, j), the global suitability probability $P_{g,(i,j)}$ of vegetation degradation or recovery occurring can be expressed as the weighted average of the prediction results of all decision trees. Through this mechanism, the model can quantify the nonlinear contribution of each environmental variable to vegetation dynamic evolution, thereby generating a global probability surface that reflects the evolutionary potential of vegetation across Inner Mongolia.

However, relying solely on global environmental factors is insufficient to fully characterize complex ecological processes; the dynamic changes of vegetation are also constrained by local environments and random factors (Li, 2011). According to the First Law of Geography, the evolution of geographical feature states possesses spatial autocorrelation (Mathur, 2015), meaning the state change of a central cell is often drawn by the states of its surrounding neighborhood cells. Therefore, this study introduces the neighborhood effect mechanism of cellular automata and uses a 5×5 extended neighborhood window to calculate the local neighborhood probability $P_{n,(i,j)}$. This indicator characterizes the aggregation density of vegetation or non-vegetation within a local extent, and its calculation formula is:

$$P_{n,(i,j)} = \frac{\sum_{N \times N} con(S_{\text{neigh}} = \text{Target})}{N \times N - 1} \times w_n, \quad (1)$$

where S_{neigh} is the cell state within the neighborhood window, $Target$ is the target transition state, $N = 5$ is the window size, and w_n is the neighborhood weight coefficient. This mechanism promotes a more compact spatial morphology in the simulation results and effectively suppresses the random generation of isolated pixels, which aligns with the community aggregation characteristics of ecosystem succession.

In addition, ecological succession in nature inherently involves certain uncertainties, such as sudden natural disasters or accidental human interventions, which may cause reverse succession of vegetation in unsuitable areas. To simulate this non-deterministic process, this study introduces a random disturbance probability $P_{r,(i,j)}$, allowing the model to break free from the constraints of deterministic rules under small-probability conditions. Its mathematical expression is:

$$P_{r,(i,j)} = 1 + (-ln\gamma)^{\alpha}, \quad (2)$$

where γ is a random number in the range of $(0,1)$, and α is a parameter controlling the intensity of randomness. Meanwhile, to ensure the simulation results conform to objective geographical laws, the model superimposes an absolute spatial constraint probability $P_{c,(i,j)}$, setting areas where vegetation transition is impossible (such as permanent water bodies) as restricted zones with a value of 0.

In summary, for a cell (i, j) at any time t, its integrated probability of final state transition $P^{t}_{total,(i,j)}$ is jointly determined by the quadruple mechanisms of global suitability, neighborhood effect, random disturbance, and spatial constraints:

$$P^{t}_{total,(i,j)} = P_{g,(i,j)} \times P_{n,(i,j)} \times P_{r,(i,j)} \times P_{c,(i,j)}. \quad (3)$$

This integrated probability system not only considers the controlling effect of macroscopic environmental gradients but also incorporates micro-scale spatial self-organization mechanisms and random processes, providing a refined probability input baseline for the subsequent parameter inversion based on the genetic algorithm (Holland, 1975).

2.3.2 Automated Parameter Inversion Based on an Improved Genetic Algorithm

The dynamic evolutionary behavior of cellular automata models is highly sensitive to the setting of transition rule thresholds. To address the strong subjective uncertainty inherent in empirical trial-and-error methods used in traditional large-scale simulations of the Inner Mongolia ecological transition zone, this study transforms parameter calibration into a global optimization inversion problem with constraints, introducing the IGA to achieve the automated optimal configuration of thresholds.

First, a population initialization mechanism based on prior probability constraints is designed. The algorithm encodes the vegetation expansion threshold θ_{ex} and degradation threshold θ_{deg} to be inverted as real-number chromosomes. To eliminate invalid redundant solution spaces, this study introduces a prior-knowledge-guided strategy. Based on the characteristics of vegetation probability distribution during the historical period, an effective confidence interval $[\theta_{2\%}, \theta_{98\%}]$ for the thresholds is delineated. The population is forced to be initialized within this interval, thereby significantly improving the initial global search efficiency. The algorithm aims to find an optimal parameter combination $\Theta^* = (\theta_{ex}^*, \theta_{deg}^*)$ that maximizes the consistency between the model simulation results Y_{sim} and the actual observation data Y_{obs}. The mathematical expression of this optimization problem is as follows:

$$\text{Maximize } J(\Theta) = \text{Kappa}(\text{CA}(P_{total}, \Theta), Y_{obs}), \tag{4}$$

where $J(\Theta)$ is the fitness function, employing the Kappa coefficient to measure simulation accuracy. $CA(\cdot)$ represents the forward evolution process of the cellular automata.

Second, an adaptive co-evolutionary mechanism with a “high baseline-dynamic gain” is designed. On the basis of roulette wheel selection ensuring the “survival of the fittest,” to overcome the late-stage evolutionary stagnation easily caused by fixed mutation rates, the algorithm dynamically adjusts the crossover probability P_c and mutation probability P_m according to the dispersion degree of population fitness. The specific calculation formulas are:

$$\begin{cases} P_c = \max\left(0.6,\ 1.6 \times \dfrac{f_{max} - f_{avg}}{f_{max}}\right) \\ P_m = \max\left(0.6,\ 0.4 \times \dfrac{f_{max} - f_{avg}}{f_{max}}\right) \end{cases}, \tag{5}$$

where f_{max} and f_{avg} are the maximum and average fitness of the population, respectively. When the difference in population fitness is significant, the algorithm automatically increases the probabilities to accelerate the recombination of superior patterns; when the population tends to converge, the probabilities decrease accordingly but are still maintained at a high baseline of 0.6. This approach not only utilizes the exploration advantages during the period of diversity but also copes with the trap of premature convergence through a mandatory high-mutation baseline.

Finally, a triple dynamic elite strategy of "elimination-injection-retention" is designed. Traditional elite strategies usually only retain the optimal individual from the previous generation, making it difficult to cope with oscillations during parameter inversion. This study designs an enhanced mechanism: (1) At the end of each generation's evolution, several inferior individuals with the lowest fitness are eliminated to purify the population gene pool. (2) Randomly generated new chromosomes are introduced to fill the vacancies, maintaining the dynamic diversity of the population through the introduction of exogenous disturbances to prevent algorithm stagnation. (3) The historical optimal individual is unconditionally inherited to ensure that fitness is monotonically non-decreasing.

This strategy can effectively balance the stable convergence and dynamic diversity of the algorithm, supporting the acquisition of the optimal parameter combination through multiple iterations.

2.3.3 Dynamic Simulation of Cellular Automata

After obtaining the optimal parameter combination $(\theta_{ex}^*, \theta_{deg}^*)$ inverted by the IGA, this study constructed a dynamic simulation engine for cellular automata based on bidirectional transition rules. This engine takes the 30m resolution binary vegetation distribution map as the initial state S_0 and drives the spatiotemporal evolution of vegetation transition by integrating the aforementioned integrated probability P_{total}^t.

During the simulation process, any spatiotemporal node (i, j) follows a bidirectional asymmetric succession mechanism. The evolutionary state $S_{i,j}^{t+1}$ at time $t+1$ is expressed as follows:

$$S_{i,j}^{t+1} = \begin{cases} 1, & \text{if } S_{i,j}^t = 0 \text{ and } P_{total,(i,j)}^t > \theta_{ex}^* \\ 0, & \text{if } S_{i,j}^t = 1 \text{ and } P_{total,(i,j)}^t < \theta_{deg}^* \\ S_{i,j}^t, & \text{otherwise} \end{cases}, \tag{6}$$

Where the state values 1 and 0 represent vegetation and non-vegetation coverage, respectively. The simulation process employs a 10-year time step for closed-loop iteration, gradually realizing the continuous transmission of spatiotemporal dynamics. Before each step's output, morphological closing operations are introduced for post-processing to eliminate salt-and-pepper noise and maintain the spatial continuity of the fragmented landscape.

As a forward-looking design of the CA-VDS, the simulation engine features a built-in multi-scenario dynamic coupling interface, supporting the integration of various SSP-RCP socioeconomic projection data. This allows the model to dynamically update the global probability layer P_g^t during projection and gradually accomplish future vegetation evolution predictions. Although this study focuses its core empirical research on the high-precision validation of the 2020 historical baseline, this collaborative design of “dynamic updating of environmental factors + static constraints of historical rules” has established a comprehensive computational foundation for a GEE-based vegetation degradation simulation framework to conduct long-term future ecological projections.

Finally, to comprehensively evaluate the reliability of the CA-VDS simulation framework, this study uses 2020 as the validation baseline year and compares the model's historical simulation output with the 30m high-resolution GLC_FCS30D real-observation land cover product for validation. Five indicators—overall accuracy, Kappa coefficient, Figure of Merit (FoM), Precision, and Recall—are employed to quantitatively assess the model's performance.

3. Results

Based on the CA-VDS model, this study completed a 30m resolution dynamic simulation of vegetation distribution in Inner Mongolia for 2020, and conducted a pixel-by-pixel overlay comparison between the simulation output and the GLC_FCS30D product, ultimately generating a vegetation spatial pattern map (Figure 3) and a confusion matrix table (Table 2).

GLC_FCS30 D (30m)	CA-VDS Simulation Results (30m)		
	Vege(%)	Non-Vege (%)	Total (%)
Vege(%)	37.45	1.20	38.65
Non-Vege(%)	3.02	58.33	61.35
Total(%)	40.47	59.53	100.00

Table 2. Confusion matrix of vegetation distribution in Inner Mongolia in 2020 ('Vege' and 'Non-vege' denote vegetation and non-vegetation areas, respectively)

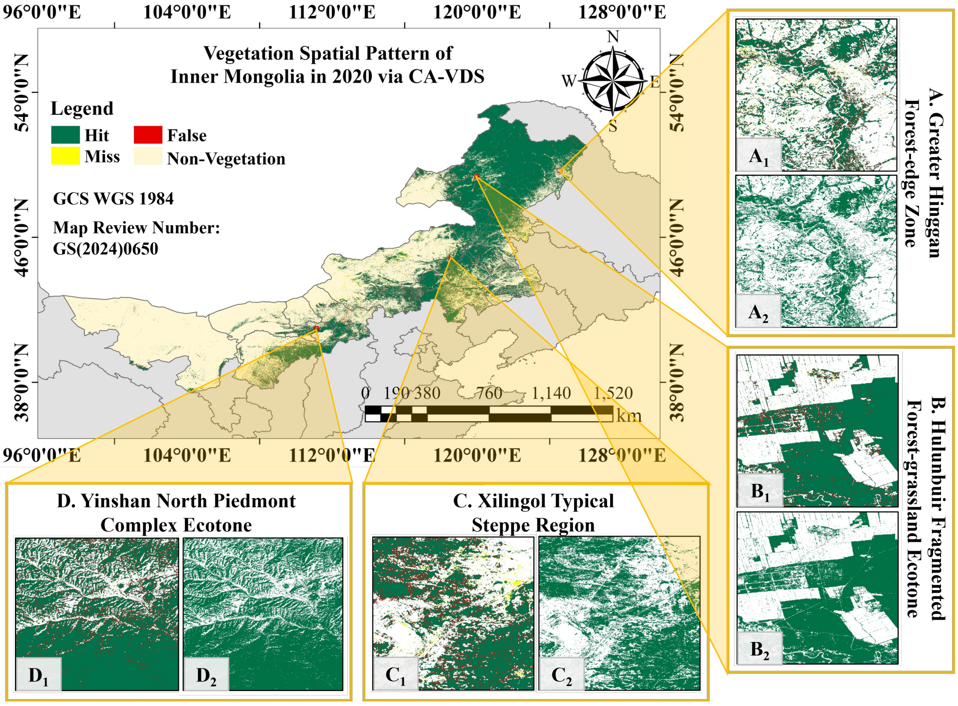

Figure 3. Historical simulation results of Inner Mongolia Autonomous Region in 2020.

From the perspective of the global spatial pattern and macroscopic statistical indicators, the simulation results exhibit a high goodness of fit. In the figure, the "Hit" areas (green pixels) occupy an absolute spatial dominance, highly contiguously distributed in the core hinterlands of the Greater Hinggan forest region in the east and the typical steppe region in the central part. Correspondingly, the confusion matrix statistics show that the model's Accuracy reaches 95.78%, the Kappa coefficient is 0.91, and the macroscopic vegetation area proportions of the two are highly convergent (observed value 38.65%, simulated value 40.47%).

This high degree of consistency at the macroscopic scale indicates that, breaking away from the spectral dependence of traditional remote sensing classification, and relying solely on the dynamic coupling of core driving forces such as climate, topography, and human activities, the CA-VDS framework can identically anchor the steady-state fundamentals of the ecosystem dominated by the geographical baseline, achieving high-confidence regional vegetation mapping.

To analyze the model's micro-performance in complex habitats, four typical ecological windows (Figure 3 A-D) were selected. While the simulation exhibits high spatial consistency in core vegetation areas, moderate regional "overestimations" (red False pixels) gradually emerge in forest-grassland and agropastoral ecotones. This local divergence stems from the difference between dynamic evolution models and static spectral classification. The baseline GLC_FCS30D relies on instantaneous surface information, making it susceptible to short-term sporadic disturbances (e.g., precipitation fluctuations, overgrazing) that cause patch fragmentation. Conversely, the CA-VDS framework employs spatial neighborhood radiation effects to topologically smooth and fill fragmented degraded patches during iterative deductions. Consequently, the approximately 3% non-overlapping area largely reflects the system's developmental potential rather than mere calculation errors. And this structural deviation can delineate the high-risk degradation zones and key targets for future ecological restoration.

4. Discussion

Although this CA-VDS framework has achieved a high spatial goodness of fit in historical simulations, it still has certain limitations. At the data-driven level, the resolution of some socioeconomic factors is relatively low, and may introduce uncertainties at the local micro-scale. At the model mechanism level, the existing framework mainly derives evolution rules based on long-term constraints, such as meteorological and topographic, making it difficult to capture the instantaneous impacts of extreme climate events or sudden human disturbances on vegetation patterns. However, the purpose of quantitative remote sensing is not to replicate complex micro-reality with absolute precision, but to provide a macroscopic spatiotemporal map that offers insights into system operating mechanisms. The local deviations between the model and observations reveal the differences between short-term surface conditions and long-term vegetation trends. Analyzing these differences helps identify ecologically fragile areas and provides practical guidance for regional ecological restoration.

Future multi-scenario ecological projections under the background of global change will be the focus of subsequent research. Although the empirical study at this stage focuses on the spatial validation of historical baseline, this process has successfully inverted objective underlying transition thresholds via the IGA, establishing a reliable parameter baseline for long-term future evolutionary deductions. The built-in multi-scenario dynamic coupling interface of this framework can seamlessly integrate CMIP6 SSP-RCPs projection datasets. In subsequent research, relying on the mechanism of "dynamic updating of future environmental factors + static constraints of historical evolution rules", the model can continuously deduce vegetation response trends under different pathways. So future work will focus on quantifying uncertainties under multi-scenario cross-simulations and explore the introduction of extreme event disturbance modules to enhance the refined prediction capability.

5. Conclusion

To address computing bottlenecks and subjective transition rules in simulating vegetation succession, this study proposes a cloud-based spatiotemporal vegetation degradation simulation and inversion framework (CA-VDS) coupling RF, IGA, and CA on the GEE. By extracting nonlinear driving potentials via RF and optimizing bidirectional transition thresholds via IGA, it achieves objective, automated quantification of vegetation evolution rules in complex habitats.

Validation against the 2020 baseline demonstrates excellent performance (95.78% overall accuracy, 0.91 Kappa). Notably, the model's approximately 3% local "overestimation" is not merely a computational defect, but largely a structural deviation driven by spatial neighborhood effects. This mechanism helps to filter out short-term spectral noise caused by sporadic disturbances, stripping away instantaneous apparent fluctuations to better reveal the potential developmental capacity and high-risk degradation areas of the regional vegetation system. In a certain sense, this corresponds to the key observation areas of ecological protection research.

Acknowledgements

This work was supported by the Key Program of the National Key Research and Development Program of China Grant 2024YFB3909001, the Key Program of the National Natural Science Foundation of China under project number 42530106, the Program of the National Natural Science Foundation of China under Project 42401534 and Project 42471470, the LIESMARS Special Research Funding, the Key Laboratory of the Ministry of Education on Application of Artificial Intelligence in Equipment(2024-AAIEKF02-02), and the Hubei Provincial Natural Science Foundation of China under Grant 2025AFB426.

References

Bondarenko, M., Priyatikanto, R., Tejedor Garavito, N., Zhang, W., McKeen, T., Cunningham, A., Woods, T., Hilton, J., Cihan, D., Nosatiuk, B. and Brinkhoff, T., 2025. The spatial distribution of population in 2015-2030 at a resolution of 30 arc (approximately 1km at the equator) R2025A version v1.

Breiman, L., 2001. Random forests. Machine learning, 45(1), pp.5-32.

Clarke, K.C., Hoppen, S. and Gaydos, L., 1997. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. Environment and planning B: Planning and design, 24(2), pp.247-261.

Chen, J., Gao, M., Cheng, S., Hou, W., Song, M., Liu, X. and Liu, Y., 2022. Global 1 km× 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019 based on calibrated nighttime light data. Scientific Data, 9(1), p.202.

Elvidge, C.D., Zhizhin, M., Ghosh, T., Hsu, F.C. and Taneja, J., 2021. Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. Remote Sensing, 13(5), p.922.

Facebook Connectivity Lab and Center for International Earth Science Information Network-CIESIN-Columbia University, 2016. High resolution settlement layer (HRSL). Source imagery for HRSL© 2016 DigitalGlobe.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, *202*, pp.18-27.

Holland, J.H., 1975. Genetic Algorithms-John H. Holland. Computer programs that" evolve" in ways that resemble natural selection can solve complex problems even their creators do not fully understand, pp.1-4.

Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C. and Neal, J., 2022. A 30 m global map of elevation with forests and buildings removed. Environmental Research Letters, 17(2), p.024016.

Li, X., Shi, X., He, J. and Liu, X., 2011. Coupling simulation and optimization to solve planning problems in a fast-developing area. *Annals of the Association of American Geographers*, *101*(5), pp.1032-1048.

Mathur, M., 2015. Spatial autocorrelation analysis in plant population: An overview. Journal of Applied and Natural Science, 7(1), p.501.

Meijer, J.R., Huijbregts, M.A., Schotten, K.C. and Schipper, A.M., 2018. Global patterns of current and future road infrastructure. Environmental Research Letters, 13(6), p.064006.

NASA, J., 2020. NASADEM Merged DEM Global 1 arc second V001 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2020-12-30 from doi:10.5067/MEaSUREs/NASADEM/NASADEM_HGT.001

Tong, S., Zhang, J., Ha, S., Lai, Q. and Ma, Q., 2016. Dynamics of fractional vegetation coverage and its relationship with climate and human activities in Inner Mongolia, China. Remote Sensing, 8(9), p.776.

Wu, F. and Webster, C.J., 1998. Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B: Planning and design*, *25*(1), pp.103-126.

Wang, Y., Zhang, C., Meng, F.R., Bourque, C.P.A. and Zhang, C., 2020. Evaluation of the suitability of six drought indices in naturally growing, transitional vegetation zones in Inner Mongolia (China). PLoS One, 15(5), p.e0233525.

Yan, D., Wang, K., Qin, T., Weng, B., Wang, H., Bi, W., Li, X., Li, M., Lv, Z., Liu, F. and He, S., 2019. A data set of global river networks and corresponding water resources zones divisions. Scientific data, 6(1), p.219.

Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X. and Liu, L., 2024. GLC_FCS30D: the first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method. Earth System Science Data, 16(3), pp.1353-1381.

Zhang, A., Tariq, A., Quddoos, A., Naz, I., Aslam, R.W., Barboza, E., Ullah, S. and Abdullah-Al-Wadud, M., 2025. Spatio-temporal analysis of urban expansion and land use dynamics using google earth engine and predictive models. *Scientific Reports*, *15*(1), p.6993.