International Journal of Environmental Technology and Management

ISSN online: 1741-511X - ISSN print: 1466-2132
https://www.inderscience.com/ijetm

Analysis on the spatial dynamic characteristics of land use in the urban agglomeration in central Yunnan based on random forest algorithm

Lede Niu, Jingzhi Lin, Lifang Zhou, Yan Zhou

DOI: [10.1504/IJETM.2023.10053934](https://doi.org/10.1504/IJETM.2023.10053934)

Article History:

Received:	01 October 2022
Last revised:	28 October 2022
Accepted:	07 December 2022
Published online:	18 December 2023

Analysis on the spatial dynamic characteristics of land use in the urban agglomeration in central Yunnan based on random forest algorithm

Lede Niu, Jingzhi Lin,
Lifang Zhou and Yan Zhou*

Faculty of Geography,
Yunnan Normal University,
Kunming 650500, China
Email: niulede@163.com
Email: 1085942305@qq.com
Email: 490628762@qq.com
Email: zhy555ylp@126.com
*Corresponding author

Abstract: In order to improve the accuracy of land use spatial analysis results, this paper takes the urban agglomeration in central Yunnan as an example, and proposes a land use spatial dynamic characteristics analysis method based on random forest algorithm. Using GIS technology to collect and process land remote sensing image data, we extract sparse description features of remote sensing images through the dictionary learning method, build a random forest classification model, classify land use space, and analyse dynamic features. The detailed analysis of land use change in the study area from 2005 to 2020 shows that the cultivated land area in this area has increased by 3,661 km^2, the dry land area has increased by 3,704 km^2, and the grassland area has decreased by 2,727 km^2, with the highest annual change rate of 0.62%.

Keywords: remote sensing image; land use; spatial dynamic feature; random forest algorithm; sparse description feature.

Reference to this paper should be made as follows: Niu, L., Lin, J., Zhou, L. and Zhou, Y. (2024) 'Analysis on the spatial dynamic characteristics of land use in the urban agglomeration in central Yunnan based on random forest algorithm', *Int. J. Environmental Technology and Management*, Vol. 27, Nos. 1/2, pp.92–109.

Biographical notes: Lede Niu received his Master's degree in the School of Geographical Sciences from Xi'nan University in 2007. He is currently an Associate Professor in the School of Tourism and Geographical Sciences of Yunnan Normal University. His research interests include regional economic development and management.

Jingzhi Lin received her Bachelor's in Science from the School of Resources and Environment of Beibu Gulf University in 2021. She is currently studying at the Faculty of Geography of Yunnan Normal University for her Master's in Geography. Her research interests include regional economic development and management.

Lifang Zhou received her Bachelor's in Tourism Management from the Faculty of Geography of Yunnan Normal University in 2021. She is currently studying in the Faculty of Geography of Yunnan Normal University for her Master's in Geography. Her research interests include regional economic development and management.

Yan Zhou received her PhD in Human Geography from the Nanjing University in 2018. She is currently a Lecturer in the College of Tourism and Geographical Sciences of Yunnan Normal University. Her research interests include land use planning and land use policy.

1 Introduction

Land is the most fundamental resource of a country and the basis for the survival of all people (Kalt et al., 2021). In nature, land is the link between human beings and the natural ecology, and with the continuous development of modern society, the total amount of urban construction land has also been greatly increased (Moroni et al., 2020). In the process of national development, the use of land is easily affected by different driving factors, which leads to great changes in the way and purpose of land use, and has a direct impact on the climate environment and biological reproduction in the region. Therefore, the analysis of the spatial dynamic characteristics of land use can be beneficial to the ecological development and environmental protection in the region (Tian et al., 2020).

Liu et al. (2021) adopted the stochastic frontier production function method in the process of studying the spatiotemporal characteristics of provincial industrial land use efficiency, and constructed a Translog-SFA model describing the relationship between industrial land productivity and land input level. The model evaluates the efficiency of industrial land use. After testing, it is proved that the method of using the SFA model for the measurement of land use efficiency is suitable, and it has a certain contribution to the analysis of the spatiotemporal characteristics of land use efficiency. From 2004 to 2017, the level of land use efficiency in China showed an upward trend. The eastern region has the highest average efficiency. In recent years, the land use efficiency in the west has gradually surpassed the land use efficiency in the northeast, and the development process is faster. Tansel and Inanloo (2019), based on land use characteristics, combined with atmospheric conditions, identified areas that might be affected by odours released from landfills, and estimated the number of people who might be affected. Shusheng et al. (2019) took the city wall area of Xi'an as the study area, and conducted a mathematical analysis of the land use change in the study area based on the plot level data. The results showed that the land use change in the study area has three characteristics. In terms of the rate of urban land use change, the period between 1995 and 2007 was the fastest, and the period between 1963 and 1995 was the slowest. Although the above-mentioned land use research can analyse the development of land use change, it does not classify the land use direction, and fails to analyse the land dynamic index, so the data results obtained may be inaccurate. The accuracy of the spatial analysis results of land use is improved, resulting in the lack of detailed analysis results and the lack of refined analysis results, which affects the rational planning of land use structure.

Random forest algorithm is an ensemble learning algorithm, which has the characteristics of high accuracy. When this algorithm is applied to the classification of land use directions, it is not easy to over-fit, and it can accurately complete the classification of land use directions. It can also accurately analyse the land dynamic index. Therefore, this paper uses the random forest algorithm to analyse the spatial dynamic characteristics of land use, which can realise the moderate development of land in the study area according to the characteristics of land use change in different areas.

2 Land use spatial classification method based on random forest algorithm

The previous land use spatial analysis methods directly conduct land use analysis after obtaining remote sensing images, which not only affects the analysis results of the land dynamic index, but also reduces the accuracy of the land use spatial analysis results. Therefore, this paper classifies the land use space based on the random forest algorithm.

2.1 Remote sensing image data collection

The land data used in this research process include 1:10,000 topographic map, 2005 multispectral scanner image, 2010, 2015 and 2020 TM remote sensing image data (Sui et al., 2020). Since the quality of remote sensing images is significantly affected by weather factors, the selected remote sensing images are all obtained in a cloudless environment with good weather conditions. For the study area, ARC/INFO geographic information system software was selected to implement 1:10,000 topographic map digitisation (Huovinen et al., 2019), and a 1:10,000 digital elevation model map was obtained.

In the process of collecting remote sensing images, the GIS geographic information system software is used to perform radiometric correction on the remote sensing images (Wang and Wang, 2002). Geometry correction, registration and standardisation are used to constrain the remote sensing image error of land use within one pixel. Through the division of county-level administrative divisions across the country, remote sensing images of sub-counties within the region are obtained. To maximise the use of geographical landforms, vegetation coverage, and meteorological environment and other information in different regions (Ye et al., 2022), implement hierarchical interactive interpretation of remote sensing images, and obtain 1:10,000 county land use vector maps in 2005, 2010, 2015 and 2020. Using GIS technology to process the remote sensing images of the study area, the existing vector maps of land conditions within the county-level area are combined into a provincial-level land use vector map, and finally the urban and rural areas can be obtained.

In the process of interpreting remote sensing images, field training is required (Li et al., 2022). Through the comparison of different regions and seasons, the perceptual awareness of remote sensing images is improved. In the process of random inspection of remote sensing image data (Pande et al., 2021), a quality inspection group was constructed by sampling with a ratio of 15%, three transects were set in different sampling counties, and remote sensing images that met the technical standards of the background database were defined as standard remote sensing images (Mandal, 2022). According to the random inspection results of all remote sensing images, the overall qualitative accuracy reached about 98.8%.

2.2 Feature extraction for sparse description of remote sensing images

After the remote sensing image data is collected, because there are many noises in the remote sensing image, the accuracy of the final analysis results will be affected if the image is used as the original data for land use spatial dynamic feature analysis. Therefore, it is necessary to extract the sparse description features of the remote sensing image, reduce the complexity of the remote sensing image, and increase the recognition ability of the design method on the remote sensing image.

In order to obtain the sparse description features of the remote sensing image, it is necessary to obtain the endmember construction dictionary in the remote sensing image. The dictionary atoms generated by the dictionary learning method can more accurately obtain the spectral features of the pixels of the land use remote sensing image (Wu et al., 2020), and the dictionary can describe the pixels more accurately.

The land use remote sensing image dataset and dictionary are represented by $X = \{x_i \mid x_i \in R^m, 1 \leq i \leq n\}$ and $Z \in R^{m \times P}$ (each column $Z_j \in R^m$ represents an atom), respectively. The main function of Z is to obtain the spectral features of the pixels, and at the same time, the atoms contained in it are used to sparse and linearly describe different pixels x_i. The solution of the dictionary learning problem can be described by solving the optimisation problem shown in equation (1):

$$\begin{cases} C_g + C_f \\ \text{s.t. } \left\| Z_j \right\|_2 \leq 1 \end{cases} \tag{1}$$

$$C_g = \min_{Z,A} \frac{1}{2} \|X - ZA\|_F^2 \tag{2}$$

$$C_f = \|\theta A\|_{1,1} \tag{3}$$

In equations (1), (2) and (3), $\min_{D,A} \frac{1}{2} \|X - DA\|_F^2$ represents the reconstruction error, and $A \in R^{P \times n}$ represents the coefficient matrix. $\|\theta A\|_{1,1}$ represents the sparse penalty function, where θ is a constant greater than 0.

In the process of dictionary learning, under the condition that dictionary Z and coefficient matrix A have certain variability, the above problem is not a convex optimisation problem (Guo et al., 2020). Under general conditions, the solution of equation (1) can be completed by iterative method, and the solution process is as follows:

1 Fix the dictionary Z, and convert formula (1) into a least square convex optimisation problem related to A and based on the $L1$ range.

2 Fixed coefficient matrix A, convert formula (1), solve the least square convex optimisation problem with quadratic constraints related to Z.

3 The dictionary Z can be obtained after several iterations.

Based on the dictionary Z obtained by the above process, according to the sparse description theory (Chen et al., 2021), the pixel $x \in X$ is described as a sparse linear combination of atoms in the dictionary Z, thereby solving equation (4):

$$\hat{a}(x) = \arg\min_{\alpha} \|x - Z\alpha\|_2^2 + \|\theta\alpha\|_1 \tag{4}$$

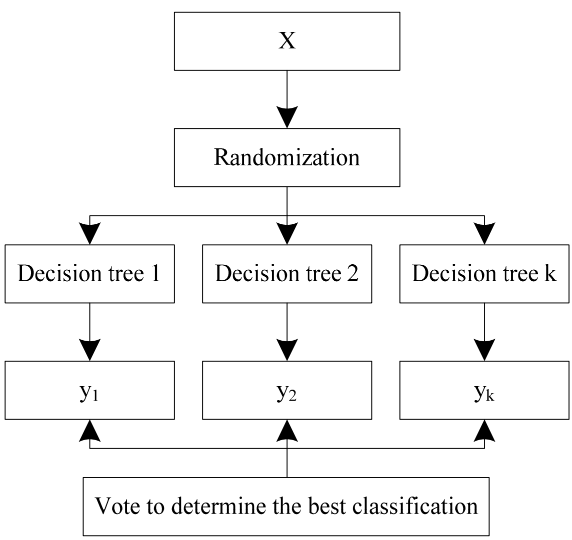
Equation (4) shows the convex l_1-regularised least square problem, and the $\hat{a}(x)$ obtained by solving it is a sparse vector, that is, there are only a few elements that are not 0.

Considering that the spectral curve structure of the remote sensing image pixels of the same type of land use has a high degree of consistency, the coefficients obtained by solving the dictionary Z also have a high degree of consistency. $\hat{a}(x)$ can be used as a new description form of the pixel x of the land use remote sensing image, that is, the sparse description feature of x on the dictionary Z. So far, the extraction of sparse features of remote sensing images has been realised.

2.3 Land use spatial classification

Stochastic forest algorithm is an integrated learning method based on bagging, which can deal with classification and regression problems well. It introduces randomness, and is not easy to over fit. It can effectively run on large datasets. It does not need to standardise the datasets, and the classification accuracy is high. Therefore, this paper uses random forest algorithm to classify land use space. In a random forest, there are several original classification trees $\{h(x, \Theta_k)\}$ without pruning operation. Θ_k represents the randomness vector in the classification tree that is not related to other distributions. The main advantage of the random forest classification algorithm is to reduce the probability of overfitting while making up for the defects of a single decision tree. Figure 1 shows the structure of the random forest algorithm.

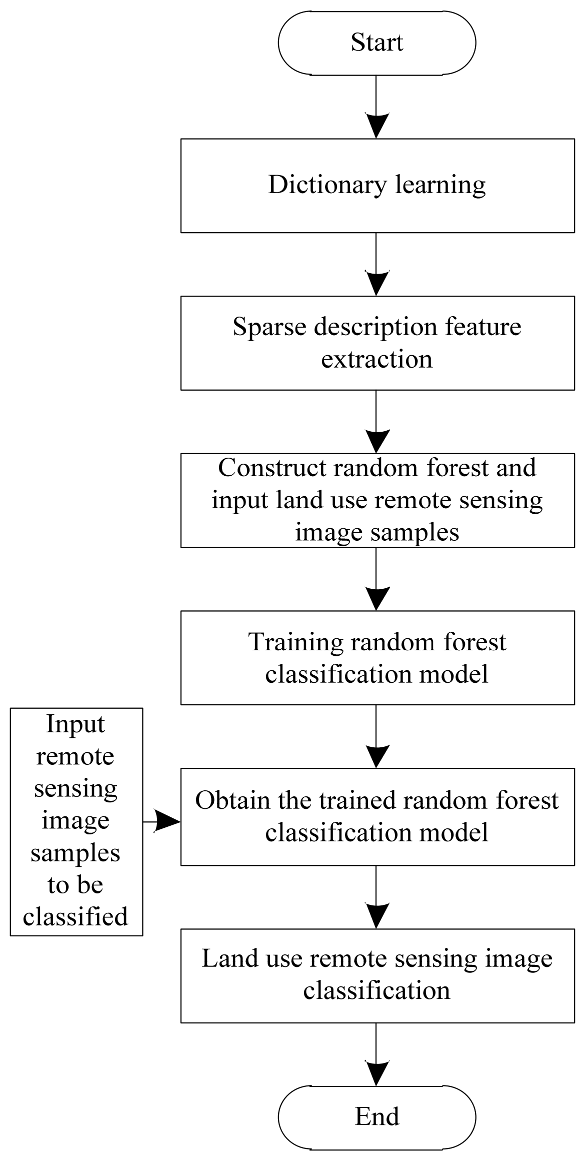
Figure 1 Structure of random forest algorithm



Prepare the remote sensing image samples that have extracted sparse description features. In order to improve the accuracy of remote sensing image classification, this paper combines the decision tree algorithm and adopts bootstrap sampling method to randomly construct the training set of each decision tree in the remote sensing image samples. Each time, a sample is randomly selected and put back into the training set until the training set and the initial dataset are the same size. According to the same method, multiple different training sets can be constructed, and each training set generates a decision tree.

Obtain k training sample sets (Di et al., 2019) in the overall training sample set X, and the training set is represented by $(x_1, x_2, \cdots, x_k)$ to build k decision trees.

Figure 2 Flowchart of land use spatial classification



Select m indicators from any n different nodes of the classification tree, and according to the minimum standard of node impurity, determine the optimal features in m subsequent indicators to classify and grow the nodes (Pastorino et al., 2021), so that the decision tree can grow to the maximum extent to different leaf nodes The impurity (Luo et al., 2021) achieves the minimum gini index, but does not perform decision tree pruning.

Repeatedly build k decision trees, combine many decision trees to form a random forest, build a random forest model, and train the model until the model accuracy meets the application requirements.

According to all the grown decision trees, the existing remote sensing image samples are input into the trained random forest model, which is divided into cultivated land, forest land, grassland, water area, urban and rural industrial and mining residential land and unused land. The classification results of the remote sensing image samples of land use to be classified are determined based on most of the voting results of all decision tree voting (Ba and Thai, 2020). The formula is described as follows:

$$f(x_t) = majorly\ vote\{h_i(x)\} \tag{5}$$

In equation (5), *majority vote* represents the majority vote result.

So far, the land use spatial classification based on random forest algorithm has been realised. The dictionary learning method is used to obtain the sparse description features of the learning dictionary Z and pixel $x \in X$ to realise the extraction of the sparse features of the remote sensing images; build a random forest model, set the number and variables of decision trees, and train the model until the model accuracy meets the application requirements; the remote sensing images of land use to be classified are input into the trained random forest classification model, thereby obtaining the corresponding classification results of the remote sensing land use images. The specific implementation process of the land use spatial classification method based on the random forest algorithm is shown in Figure 2.

The specific classification description is shown in Table 1.

Table 1 Classification of land use types

Class I land use type	*Class II land use type*
Cultivated land	Paddy field and dry land
Woodland	Woodland, dense woodland, open woodland, bushes and other woodlands
Grassland	High/medium/low coverage grassland
Waters	Canals, lakes, reservoirs, tidal flats, etc.
Urban and rural industrial and mining residential land	Urban construction land, facility land, residential area land, ecological landscape land, rail transit land, etc.
Unused land	Sand, Gobi, swamp and bare land

3 Spatial dynamic characteristics analysis of land use

After obtaining the classification results of land use remote sensing image samples through the trained random forest classification model, according to the spatial location of land use, according to the quantitative change of land use, regional dynamic change, land use type conversion tendency characteristics, land use level characteristics, to analyse the dynamic characteristics of land use space.

3.1 Spatial position conversion and quantity fluctuation characteristics of land use

The spatial position transformation of land use describes the corresponding changes in the spatial position of land use within the region. The actual usable area of the land of category j after spatial transformation and change is represented by S_j, and its calculation formula is described as follows:

$$S_j = 2 \times \min P_1 \tag{6}$$

$$P_1 = \left(P_{j+} - P_{jj}, P_{+j} - P_{jj}\right) \tag{7}$$

In equations (6) and (7), P_{j+} represents the area of the j type land before the location conversion, P_{jj} represents the unchanged area, and P_{+j} represents the j type land area after the location conversion.

Quantity change describes the change of different land use areas under the condition of land use type transition. The formula is described as follows:

$$Q_j = \max P_1 - \min P_1 \tag{8}$$

In formula (8), Q_j represents the quantitative change of the j type land area.

The overall change C_j of the land of category j can be determined by the sum of S_j and Q_j.

3.2 Land use dynamic index characteristics

The dynamic characteristics of land use describe the quantitative fluctuation characteristics of land use types in any time category, that is, the fluctuation of the quantity of land resources and the fluctuation of the combination of land use types. Land use dynamic index to illustrate.

A single land use dynamic index D can quantitatively analyse the fluctuation speed of a certain land use type in any time range in the region. The formula is described as follows:

$$D = \frac{\dfrac{\left(S_a - S_e\right) + \left(S_b - S_e\right)}{S_a}}{T} \times 100\% \tag{9}$$

In formula (9), S_a represents the area of the initial land type in the time category, S_e represents the area of the land type that generates fluctuations, S_b represents the area of the land type at the end of the time category, and T represents the time category.

The overall land use dynamic index D_s can describe the regional differences in the fluctuation speed of land use types, which reflects the overall impact of human behaviour on the fluctuation of regional land use types. The formula is described as follows:

$$D_s = \frac{\frac{\left(\sum_{i=1}^{n} \Delta S_{i-j}\right)}{2\sum_{i=1}^{n} S_i}}{T} \times 100\% \tag{10}$$

In formula (10), S_{i-j} represents the absolute value of the actual land use area after the transformation of the first type of land in the time category after the space utilisation is realised, S_i represents the initial area of the non-*i* type land use type in the time category, and S_j represents any land use in the final period type of area.

3.3 Characteristics of land use type conversion tendency

The land use type conversion tendency feature describes the level index of the data conversion of any two land use types, which can reflect the level of conversion between two different land use types. The value of the land use type conversion tendency characteristic is 0–1. The larger the value, the greater the trend and intensity of regional land use type transformation.

Using equation (11), the land use type conversion probability matrix can be described:

$$P = \left(P_{ij}\right) \tag{11}$$

The cross-tabulation summary function is used to perform land type conversion analysis on the fluctuation of regional land use types, so that P can be obtained. On this basis, formula (12) is used to convert it into a land use type conversion probability matrix:

$$P_{ij} = \frac{S_{ij}}{\sum_{j=1}^{n} S_{ij}} \tag{12}$$

In equation (12), P_{ij} and S_{ij} respectively represent the probability of converting the i type land use type to the j type land use type and the area of the i type land use type to the j type land use type in the study time range.

The characteristics of the combination types of the existence of different types of patch types of land use spaces in the land landscape are described by the probability matrix of landscape symbiosis, as shown in formula (13):

$$X = \left(X_{ij}\right) \tag{13}$$

At the same time, the landscape co-occurrence probability matrix can also be used to describe the spatial correlation between different landscape types. The formula is described as follows:

$$X_{ij} = \frac{C_{ij}}{\sum_{j=1}^{n} C_{ij}} \tag{14}$$

In equation (14), X_{ij} and C_{ij} respectively represent the adjacent probability of patch types i and j in the landscape and the number of j type patch grids around i type patches, and n represents the overall number of patch types.

If land use type i and land use type j are adjacent, Q_{ij} can be used to describe the propensity index of land use type i to be converted to land use type j, and formula (15) is used to describe the matrix composed of P_{ij} / X_{ij}, then Q_{ij} can be obtained by formula (16) describes:

(15)

$$Q_{ij} = \begin{cases} \dfrac{Y_{ij}}{Y_{ji}} \ (j > 1) \\ Y_{ij} \ (j = 1) \end{cases} \tag{16}$$

When the value of Q_{ij} is greater than 1, it means that when land use type i and land use type j are adjacent to each other, the tendency of land use type i to be converted to land use type j and the probability of land use type j to be converted to land use type i. It is relatively large; the closer the Q_{ij} value is to 1, the more likely the land use type i is to be converted to the land use type j.

3.4 Land use level

The level of land use is based on the degree of regional land use development and reflects the overall impact of human behaviour on the fluctuation of regional land use. It can be described by equation (17):

$$I = \frac{(R - U)}{R} \times 100\% \tag{17}$$

In formula (17), R and U represent the total land area and the used land area, respectively. I can quantitatively describe the overall level and fluctuation trend of regional land use types.

4 Experimental study

In order to verify the application effect of the land use spatial dynamic characteristics analysis method based on random forest algorithm proposed in this paper in practice, this paper takes the central Yunnan area as the research object, and adopts the method designed in this paper to analyse the land use spatial dynamic characteristics within the research object area. The analysis is carried out mainly from three aspects: the fluctuation of land use quantity, the dynamic index of land use and the tendency of land use type transformation.

4.1 Overview of the study area

The central Yunnan urban agglomeration is taken as the main area of this research. Central Yunnan includes Kunming City, Qujing City, Yuxi City, Chuxiong Yi Autonomous Prefecture and Mengzi City, Gejiu City, Jianshui County, Kaiyuan City, Mile City, Luxi County, Shiping County in the northern part of Honghe Hani and Yi Autonomous Prefecture composed of seven counties and cities, it is the most economically developed area in Yunnan Province, with a total area of 114,600 km^2 and a population of 44.02% of the entire Yunnan Province.

The study area belongs to the eastern Yunnan Plateau Basin, which is dominated by mountains and inter-mountain basins. The soil type is mainly red soil. There are various types of vegetation, mostly secondary vegetation and artificial vegetation. The development intensity is low, and the available land resources have certain potential.

4.2 Experimental results and analysis

4.2.1 Fluctuation analysis of land use quantity

The method of this paper is used to analyse the fluctuation of the land use quantity of the research object. The analysis process starts from the two directions of the primary land use type and the secondary land use type. The results are as follows.

4.2.1.1 Types of primary land use

The method of this paper is used to analyse the fluctuation of the land use quantity of the primary land use type within the research object area, and the results are shown in Table 2 and Table 3.

From the analysis of Tables 2 and 3, it can be seen that among the first-level land use types in the research object area from 2005 to 2020, three types of land use: cultivated land, grassland and unused land occupy the main positions. The area of cultivated land, water area and urban and rural industrial, mining and residential land has shown a trend of continuous improvement. By 2020, the proportion of cultivated land has changed from 11.43% to 13.92%, the area has increased by 3,661 km^2, and the proportion of water has increased from 3.20%. The change was 3.30%, the area increased by 184 km^2, the proportion of urban and rural industrial and mining residential land changed from 1.13% to 1.64%, and the area increased by 832 km^2. Among the other utilisation types whose land area continued to decrease, grassland area decreased most significantly, with a decrease of 2,727 km^2, a decrease of 1.78%; followed by unused land, with an area of 1,757 km^2, a decrease of 1.17%; the decrease of woodland area was smaller, with a decrease of 1.17%. The area decreased by 177 km^2, a decrease of 0.15%. Specifically, after 2015, the area expansion efficiency of the research object has been significantly improved, and the impact on land use types has become more and more obvious. The cultivated land has increased by 128 km^2, the urban industrial and mining area has increased by 666 km^2, and the grassland area has been reduced by 1,197 km^2.

4.2.1.2 Types of secondary land use

The method of this paper is used to analyse the fluctuation of the land use quantity of the secondary land use type within the research object area, and the results are shown in Table 4 and Table 5.

Table 2 Fluctuation of land use/cover quantity in different years of first-class land use type within the study area

Particular year	*2005*		*2010*		*2015*		*2020*	
Fluctuation term	*The measure of area / km^2*	*Ratio / %*	*The measure of area / km^2*	*Ratio / %*	*The measure of area / km^2*	*Ratio / %*	*The measure of area / km^2*	*Ratio / %*
Cultivated land	17,623	11.43	18,787	12.21	19,756	12.89	21,284	13.92
Woodland	6,416	4.26	6,366	4.16	6,356	4.16	6,239	4.11
Grassland	64,089	41.96	63,121	41.30	62,559	40.96	61,362	40.18
Waters	4,865	3.20	4,989	3.36	5,072	3.32	5,049	3.30
Urban and rural industrial and mining residential land	1,679	1.13	1,828	1.25	1,845	1.22	2,511	1.64
Unused land	58,053	38.02	57,592	37.72	57,166	37.45	56,296	36.85

Table 3 Fluctuation of land use/cover quantity at different stages of the first-class land use type within the study area

Stage	*2005–2010*		*2010–2015*		*2015–2020*	
Fluctuation term	*Variation / km²*	*Annual change range / %*	*Variation / km²*	*Annual change range / %*	*Variation / km²*	*Annual change range / %*
Cultivated land	1,164	0.81	969	0.59	1,528	1.01
Woodland	–50	–0.03	–10	–0.01	–117	–0.07
Grassland	–968	–0.64	–562	–0.37	–1,197	–0.78
Waters	124	0.08	83	0.06	–23	–0.02
Urban and rural industrial and mining residential land	149	0.1	17	0.01	666	0.43
Unused land	–461	–0.3	–426	–0.29	–870	–0.56

Table 4 Fluctuation of land use/cover quantity in different years of secondary land use type within the study area (km^2)

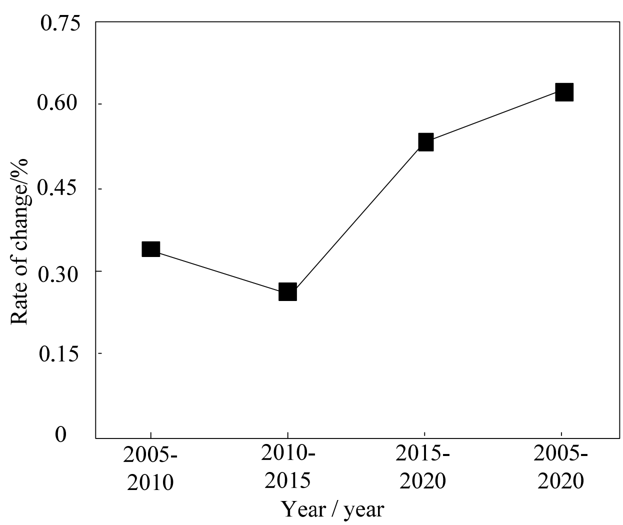
Type	*Particular year*				*Fluctuation from 2005 to 2020*
	2005	*2010*	*2015*	*2020*	
Paddy field	176	134	137	146	–30
Dry land	17,382	18,653	19,563	21,086	3,704
Woodland	4,562	4,573	4,563	4,543	–19
Shrubland	1,121	1,072	1,065	1,033	–88
Open woodland	714	692	690	633	–81
Other forest land	22	33	41	143	121
High coverage grassland	17,193	17,066	17,043	16,976	–217
Medium coverage grassland	16,264	16,118	16,023	15,789	–475
Low coverage grassland	30,635	29,932	29,495	28,590	–2,045
Canal	8	8	8	7	–1
Lake	1,182	1,385	1,389	1,359	177
Reservoir	2,983	2,983	2,983	2,983	0
Tidal flat	706	718	708	710	4
Urban land	562	611	721	935	373
Rural residential area	811	815	830	881	70
Other construction land	312	411	501	716	404
Sand	19,731	19,706	19,562	19,439	–292
Gobi	19,221	19,083	19,011	18,749	–472
Swamp	193	301	285	283	90
Bare land	18,921	18,515	18,316	17,842	–1,079

From the analysis of Table 4, it can be seen that among the secondary land use types within the research object area from 2005 to 2020, the land use fluctuations are the largest in dry land, low-coverage grassland and bare land. The use area of dry land increased by 3,704 km^2, and the growth trend was relatively strong; while the area of low-coverage grassland and bare land showed a continuous decreasing trend. The use area of low-coverage grassland decreased from 30,635 km^2 in 2005 to 28,590 km^2 in 2020, with a total decrease of 2,045 km^2. The area of bare land utilisation decreased from 18,921 km^2 in 2005 to 17,842 km^2 in 2020, a total decrease of 1,079 km^2. Although the area of low-coverage grassland has declined, it is still the core component of ecological environment quality.

4.3 Land use dynamic index analysis

The method of this paper is used to analyse the dynamic index of land use within the research object area, and the dynamic index of land use within the research object area in 2005, 2010, 2015 and 2020 is analysed by using equations (9) and (10), the results are shown in Figure 3.

Figure 3 Comprehensive dynamic index of land use of research object



From Figure 3, it can be seen that the rate of land use change in the research object area continued to increase from 2005 to 2020, and the annual change rate was up to 0.62%; from the analysis of the study period, it can be obtained: the annual change rate of land use from 2005 to 2010. The annual change rate from 2010 to 2015 was 0.26%, and the fastest rate of change from 2015 to 2020 was 0.55%. This shows that in this stage, the type of land use is more significantly affected by social development and human behaviour.

4.4 Analysis of land use type conversion tendency

Based on the transformation of land use types, the co-occurrence probability matrix of land use types is introduced, and the two are organically combined to analyse the tendency characteristics of land use type transformation, so as to analyse the land use changes in the research object area spatially. Table 5 shows the dominant tendency of conversion between different land use types in the non-self-adjacent state.

Table 5 Analysis of dominant tendency of land use type transformation

The type I land use type is converted to the type J land use type	*Dominance tendency index*
Conversion of grassland to woodland	1.46
Conversion of grassland into urban and rural industrial and mining residential land	2.66
Conversion of grassland to water	1.84
Conversion of grassland to arable land	1.33
Conversion of unused land to grassland	1.91
Conversion of unused land into forest land	24.56
Conversion of water area to forest land	3.52
The water area is converted into urban and rural industrial and mining residential land	5.63
Conversion of water area into unused land	2.10
Conversion of cultivated land to forest land	751.95
Conversion of cultivated land into unused land	1.62
Conversion of cultivated land into urban and rural industrial and mining residential land	3.73

From the analysis of Table 5, it can be seen that the propensity index for the conversion of grassland to other landscape types except unused land is > 1, which indicates that grassland has a tendency to convert to other land use types; when the unused land is adjacent to grassland and forest land, the conversion dominance propensity index of conversion to grassland and forest land are 1.91 and 24.56, respectively, which indicates that unused land has a tendency to convert to these two land use types; water area is a land use type with high stability, under the condition that it is adjacent to other different land use types, the propensity index value of its conversion to urban and rural industrial and mining residential land is the largest, reaching 5.63; the propensity index value of cultivated land being converted to forest land is the largest, reaching 751.95; at the same time, for unused land and urban and rural areas. There may be a corresponding conversion tendency in terms of industrial and mining residential land.

It can be seen that the application of the spatial dynamic characteristics analysis method of land use based on random forest algorithm proposed in this paper can effectively analyse the fluctuation of land use quantity, land use dynamic index and land use type conversion tendency in central Yunnan. The obtained analysis results are more detailed, which is helpful for rational adjustment of land use structure, and can promote the planning and management of land use in the study area.

5 Conclusions and discussion

The analysis of the spatial dynamic characteristics of land use is one of the main topics in the study of global environmental fluctuations. In order to better realise the land application, this paper takes the central Yunnan urban agglomeration as an example to study the spatial dynamic characteristics analysis method of land use based on the random forest algorithm:

1. After collecting remote sensing images of the study area and processing them with GIS technology, a random forest classification model was constructed, and the spatial dynamic characteristics of land use were analysed according to the spatial location of land use.
2. The experimental results show that, from 2005 to 2020, the area of grassland in the central Yunnan urban agglomeration decreased most significantly, with a decrease of 2,727 km^2, a decrease of 1.78%; the use area of low-coverage grassland decreased by 2,045 km^2, the unused land area decreased by 1,757 km^2, and the bare land use area decreased by 1,079 km^2. The area of cultivated land will increase by 3,661 km^2, the area of dry land will increase by 3,704 km^2, the area of urban and rural industrial and mining residential land will increase by 832 km^2, other construction land will increase by 404 km^2, and urban land will increase by 373 km^2.
3. The rate of land use change in the central Yunnan urban agglomeration continued to increase from 2005 to 2020, with an annual change rate of up to 0.62%. Grassland has a tendency to convert to other land use types, and its propensity index is > 1; unused land has a tendency to convert to grassland and forest land types, which are 1.91 and 24.56, respectively; the propensity value of water area to convert to urban and rural industrial, mining, and residential land is 5.63. The propensity value of cultivated land to be converted into forest land is 751.95.
4. It can be proved that the spatial dynamic characteristics analysis method of land use based on random forest algorithm proposed in this paper can effectively analyse the spatial dynamic characteristics of land use in the study area and help meet the urban development requirements of the study area. However, in future research, it is necessary to increase the analysis of climatic conditions and natural geological disasters in the area where the central Yunnan urban agglomeration is located, so that the method proposed in this paper can more comprehensively study the spatial dynamic characteristics of land use and improve urban land use. The intensive level of utilisation promotes the coordinated development of land use and ecological environment, and the harmonious coexistence of man and nature.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant Nos. 42261037 and 42101278), Yunnan Fundamental Research Project-General Program (Grant No. 202201AT070024 and No. 202201AT070039).

References

Ba, T.L. and Thai, H.A. (2020) 'Hyperspectral remote sensing image classification based on random average band selection and an ensemble kernel extreme learning machine', *Applied Optics*, Vol. 59, No. 13, pp.412–428.

Chen, X., Liu, Y., Li, F. et al. (2021) 'Remote sensing image recommendation based on spatial-temporal embedding topic model', *Computers & Geosciences*, Vol. 157, No. 5, p.104935.

Di, S., Li, Z.L., Tang, R., Pan, X., Liu, H. and Niu, Y. (2019) 'Urban green space classification and water consumption analysis with remote-sensing technology: a case study in Beijing, China', *International Journal of Remote Sensing*, Vol. 40, Nos. 5–6, pp.1909–1929.

Guo, D., Shi, W., Hao, M. et al. (2020) 'FSDAF 2.0: improving the performance of retrieving land cover changes and preserving spatial details', *Remote Sensing of Environment*, Vol. 248, No. 53, p.111973.

Huovinen, P., Ramirez, J., Caputo, L. and Gomez, I. (2019) 'Mapping of spatial and temporal variation of water characteristics through satellite remote sensing in Lake Panguipulli, Chile', *The Science of the Total Environment*, 20 August, Vol. 679, No. 31, pp.196–208.

Kalt, G., Mayer, A., Haberl, H. et al. (2021) 'Exploring the option space for land system futures at regional to global scales: the diagnostic agro-food, land use and greenhouse gas emission model BioBaM-GHG 2.0', *Ecological Modelling*, Vol. 459, No. 41, p.421.

Li, W., Jia, S., He, W. et al. (2022) 'Analysis of the consequences of land-use changes and soil types on organic carbon storage in the Tarim River Basin from 2000 to 2020', *Agriculture, Ecosystems & Environment*, Vol. 327, No. 25, p.107824.

Liu, S.C., Lin, Y.B., Ye, Y.M. and Xiao, W. (2021) 'Spatial-temporal characteristics of industrial land use efficiency in provincial china based on a stochastic frontier production function approach', *Journal of Cleaner Production*, Vol. 295, No. 2, p.126432.

Luo, X., Du, H., Zhou, G., Li, X. and Huang, Z. (2021) 'A novel query strategy-based rank batch-mode active learning method for high-resolution remote sensing image classification', *Remote Sensing*, Vol. 13, No. 11, p.2234.

Mandal, A.K. (2022) 'The need for the spectral characterization of dominant salts and recommended methods of soil sampling and analysis for the proper spectral evaluation of salt affected soils using hyper -spectral remote sensing', *Remote Sensing Letters*, Vol. 13, No. 6, pp.588–598.

Moroni, A., Spagnolo, V., Crezzini, J. et al. (2020) 'Settlement, space organization and land-use of a small Middle Bronze Age community of central Italy. The case study of Gorgo del Ciliegio (Arezzo-Tuscany)', *Quaternary International*, Vol. 539, No. 15, p.379.

Pande, C.B., Moharir, K.N. and Khadri, S.F.R. (2021) 'Assessment of land-use and land-cover changes in Pangari Watershed Area (MS), India, based on the remote sensing and GIS techniques', *Applied Water Science*, Vol. 11, No. 6, pp.1–12.

Pastorino, M., Montaldo, A., Fronda, L., Hedhli, I. and Zerubia, J. (2021) 'Multisensor and multiresolution remote sensing image classification through a causal hierarchical Markov framework and decision tree ensembles', *Remote Sensing*, Vol. 13, No. 5, p.849.

Shusheng, W., Liang, Z.Z., Yuqian, X.U. and Xiaolong, L.I. (2019) 'Mathematical analysis of urban land use change in Xi'an City wall area by using parcel-level data', *Science China Technological Sciences*, Vol. 62, No. 4, pp.171–181.

Sui, L., Kang, J., Yang, X., Wang, Z. and Wang, J. (2020) 'Inconsistency distribution patterns of different remote sensing land-cover data from the perspective of ecological zoning', *Open Geosciences*, Vol. 12, No. 1, pp.324–341.

Tansel, B. and Inanloo, B. (2019) 'Odor impact zones around landfills: delineation based on atmospheric conditions and land use characteristics', *Waste Management*, April, Vol. 88, No. 28, pp.39–47.

Tian, F., Li, M., Han, X. et al. (2020) 'A production-living-ecological space model for land-use optimisation: a case study of the core Tumen River region in China', *Ecological Modelling*, Vol. 437, No. 36, pp.1231–1247.

Wang, Q. and Wang, H. (2022) 'An integrated approach of logistic-MCE-CA-Markov to predict the land use structure and their micro-spatial characteristics analysis in Wuhan metropolitan area, Central China', *Environmental Science and Pollution Research*, January, pp.20–29, DOI: 10.1007/s11356-021-17750-6.

Wu, M., Zhao, Y., Sun, L. et al. (2020) 'Remote sensing of spatial-temporal variation of chlorophyll-a in the Jiaozhou Bay using 32 years Landsat data', *Journal of Coastal Research*, Vol. 102, No. sp1, pp.276–289.

Ye, S., Ren, S., Song, C. et al. (2022) 'Spatial patterns of county-level arable land productive-capacity and its coordination with land-use intensity in Mainland China', *Agriculture, Ecosystems & Environment*, Vol. 326, No. 13, p.107757.