Publication Date: 30 June 2024
Archs Sci. (2024) Volume 74, Issue 3 Pages 216-225, Paper ID 2024334.
https://doi.org/10.62227/as/74334

Research on View Analysis and Spatial Optimisation of Landscape Design under XGBoost Model

Yanhui Sui[1,*]
[1]Department of Architecture, Weihai Vocational College, Weihai, Shandong, 264210, China.
Corresponding authors: Yanhui Sui (e-mail: rczhj0220@163.com).

Abstract In the process of urbanization, the landscape design of urban public space helps to improve the living quality and happiness of citizens. This paper takes the geographic data of Province G as the research object and preprocesses its landscape data. The agglomeration evolution characteristics of the regional landscape are studied by using the average nearest neighbor index, and the distribution pattern characteristics of the regional landscape are analyzed by the landscape pattern index, and the pattern characteristics of the regional landscape are also analyzed from the landscape level and the type level, respectively. The XGBoost algorithm was used to model the urban landscape design features, and the specific features affecting the urban landscape design were analyzed by combining with the SHAP model decoding method. It was found that the ANN index value of regional landscapes decreased by 16.05% between 2000 and 2022, and the AI index of urban landscapes decreased from 63.512 to 62.424.When the level of urbanization development was around 20%, the LPI index of woodland types was in the range of [0.5,1.5]. When the CONTAG index is between [0.2,0.8], the SHAP value of the urban landscape pattern index stabilizes between -0.2 and 0.2, and there is a significant decrease in the SHAP value after exceeding 0.8. Urban landscape design needs to be based on the level of urbanization development to meet the landscape needs of citizens through diverse landscape space combinations.

Index Terms ANN index, landscape pattern index, XGBoost model, SHAP model, landscape design

I. Introduction

In the past, the concepts and principles of ecological design have not been sufficiently emphasized in the teaching process of landscape design [1], [2]. Students often, in their studies, pursue unique forms and visual effects [3]. They pay more attention to the appearance and decoration of the landscape in their study and practice [4], and give relatively little consideration to the harmonization with the natural environment, ecological protection and sustainable development [5], [6]. This stems from the lack of in-depth study of ecological design concepts and principles in the teaching process [7]. In the previous teaching process, the teaching content and in-depth understanding of ecological design was insufficient, and students had insufficient knowledge of the structure, function and interrelationships of ecosystems [8], [9]. They lacked awareness of the importance of ecological restoration, biodiversity conservation, natural resource utilization and ecosystem services [10], and were unable to apply the principles and concepts of ecological design to actual landscape design [11]. Due to the lack of ecological awareness in the teaching and learning process [12], students tend to ignore the consideration of coordination and sustainability with the natural environment in their study and practice, and fail to recognize the interaction between landscape design and ecosystems.They may over-rely on artificial materials and neglect plant selection, ecological functions and ecosystem restoration [13]–[15].

Landscape includes all the environmental spaces with ornamental value formed by means of cultural propaganda and plant landscaping [16], while landscape space can be understood as the implementation of design around a certain environmental space, the flexible use of human factors, natural factors, and social and cultural elements, and the filling of various aesthetic, ecological, and cultural values in the environmental space, resulting in an external space with beautiful scenery, rich scenery, and certain functions [17]–[19]. Scientific landscape design helps to realize the harmonious coexistence of people and nature, and to bring into play the ecological value and use value of environmental space [20]. Based on landscape design, the application of the concept of sustainable development requires designers to take into account the cultural and ecological value of the landscape, in short, the use of the landscape to control environmental pollution, cater to the cultural and material needs of the masses, maintain the ecological environment and conserve natural resources, and promote local traditional culture [21],

[22].

This paper takes the geographically related data of Province G as the research object and resamples its landscape data for pre-processing. For the spatial optimization of landscape design, this paper chooses landscape pattern index, average nearest neighbor index, XGBoost model and SHAP model decoding method as research methods. For the evolution characteristics of regional landscape, this paper analyzes the clustering evolution characteristics of landscape space by using the average nearest neighbor index, and analyzes the morphological characteristics of landscape spatial distribution by landscape pattern index. The pattern characteristics of regional landscapes were also analyzed from landscape level and type level, respectively. In addition, the paper uses the XGBoost algorithm as the basis for modeling the landscape pattern characteristics related to landscape design, and selects the coefficient of determination, average absolute error and root mean square error as the evaluation indexes of the model. Aiming at the interpretability of the XGBoost model, this paper uses the SHAP model to explain and illustrate the importance degree of landscape design features, so as to clarify the spatial optimization path of landscape design.

II. Study Area and Methodology Selection

Since the reform and opening up, China's urbanization development has begun, and now it has entered the rapid development stage of urbanization, and the economy has also entered a period of comprehensive transformation. Reasonable urbanization process can effectively drive the adjustment and development of economy and industrial structure, however, excessive urbanization once caused unprecedented impact on urban landscape, breaking the balance of the landscape itself, leading to serious landscape conflicts, which is not conducive to the sustainable development of urban areas. Therefore, the study of landscape space optimization in the process of urbanization is of great significance.

A. Data collection and organization

1) Study area data

The landscape data required for this study are soil attribute, vegetation, meteorology and land use data, and the resolution of all raster data is uniformly 100m*100m after resampling.In this study, the soil attribute, vegetation and meteorology data are the landscape data of Province G from 2000 to 2022, and the data sources and spatial distribution are as follows:

1) The terrain factor includes elevation, slope and terrain roughness, the elevation and slope data come from the geospatial data cloud platform, and the terrain roughness is extracted based on the elevation data using the elevation mean square error algorithm.
2) Meteorological data are derived from the historical weather platform.
3) Socio-economic factors include land use, population density, distance from roads, and distance from residential land, distance from industrial and mining land (mining and smelting land), distance from general factories, distance from agricultural land, distance from forest land and grassland generated based on land use. The land use data were generated based on satellite images and concurrent field surveys, with satellite image data from the Bigemap GIS Office platform and population density data from the WorldPop platform.

2) Landscape data preprocessing

1) DEM
   By resampling the original precision DEM data from 15m resolution to 5m precision, after that the 5m precision DEM data was further analyzed to calculate the raster data such as slope direction, slope gradient, contour lines, etc. required for this study by using the 3D Analyst tool - Raster Surface Tool in ArcGIS Pro software. In addition this DEM data is also the source data used for elevation extraction for the Urban Landscape Pattern Index.
2) Land use classification data and land cover classification data
   According to the research purpose of this paper, the land use classification raster data is mainly used for urban landscape research from the perspective of natural zoning. The land use classification raster is used to interpret and classify high-resolution Worldview-2/QuickBird images, such as roads, residences, industries, and commercials through the eCognition software, and is combined with auxiliary data, such as the land use permit data, POI data, and historical maps, POI data, and historical maps.
3) POI Data
   POI mainly refers to point information containing names, categories, and spatial information in network maps, which become vector point data that can be used for further management, measurement, and statistics through geographic information data processing.

B. Selection of research methodology

1) Landscape pattern index

Landscape size, shape, quantity, type and spatial pattern have important impacts on ecological security, and due to certain correlations between various landscape indices, the comprehensive evaluation of landscape pattern tends to adopt a smaller number of indices to characterize the spatial distribution of landscape. The landscape pattern indices selected in this paper are mainly as follows:

1) Patch density
   Patch density refers to the ratio of the number of patches of a certain landscape to the total area of patches, reflecting the density of patches in the landscape.

$$PD = \frac{NP}{A}, \tag{1}$$

where NP is the number of patches and $A\left(hm^2\right)$ is the total area of the patches.

2) Number of patches

$$NP = N, \tag{2}$$

where N is the total number of all patches in the landscape. NP takes a range of values $NP{\sim}1$ with no upper limit. The size of NP value is positively correlated with landscape fragmentation.

3) Average patch area

$$MPS = \frac{A}{N} \times 10^6, \tag{3}$$

where A is the total area of the landscape and N is the total number of patches of each type.MPS can characterize landscape fragmentation, and at the landscape level, landscapes with smaller MPS are more fragmented than landscapes with larger MPS.

4) Maximum patch index

$$LPI = \frac{Max\left(a_1, \cdots a_n\right)}{A} \times 100, \tag{4}$$

where a_i is the area of patch i in the landscape, and A is the total area of the landscape. The value of LPI ranges from $0 < LPI\}100$. LPI is the proportion of the area of the largest patch in the whole landscape. When the largest patch in the landscape is getting smaller and smaller, LPI tends to be close to 0.

5) Landscape aggregation

$$AI = \left[\sum_{i=1}^{m}\left(\frac{g_{ii}}{\max g_{ii}}\right) P_i\right] \times 100, \tag{5}$$

where g_{ii} is the number of neighbors between pixels of the same patch type i, $\max g_{ii}$ is the maximum number of neighbors between pixels of patch type i, and P_i is the area ratio occupied by patch type i in the landscape. The value of AI ranges from $0\}AI\}100$. AI indicates the non-randomness probability or degree of aggregation of different patch types in the landscape that appear adjacent to each other in the landscape. When AI=0, it indicates that the degree of fragmentation of patch types is maximized and the degree of aggregation is minimized. the larger the value of AI, the greater the degree of aggregation of landscape patches, and when AI=100, it indicates that the whole landscape consists of only one patch.

6) Modified Simpson's diversity index

$$MSI = -\ln \sum_{i=1}^{m} P_i^2, \tag{6}$$

where P_i is the area proportion of patch type i in the landscape, and m is the number of patch types in the landscape. The value range of MSI is MSDI ≥ 0. When the landscape consists of only one patch, i.e., there is no diversity, MSIDI = 0, and MSIDI increases with the increase in the number of patch types, and, in addition, when the area proportion is more balanced among the patch types, the MSDI increases as well.

In addition, the patch area (CA) as well as the mean patch fractional dimension number (MPFD) were chosen to evaluate the changes of urban landscape space in this paper, in order to have a more intuitive understanding of the trend of changes in urban landscape space, and to provide a reference for optimizing the design of urban landscape space.

2) Average Nearest Neighbor Index

In order to globally determine whether the clusters of building points in the landscape area are clustered or not, this paper first uses the average nearest neighbor distance method to analyze the point pattern of building type points in the study area.

The average nearest neighbor distance is the mean value of the nearest distance between points. This analysis method determines the spatial pattern of the nearest neighbor point pairs by comparing the average distance of the calculated pairs of nearest neighbor points with the average distance of the nearest neighbor point pairs in the random distribution pattern. If the mean distance of the analyzed element is equal to the mean distance of the assumed random distribution, the element is considered random. If it is less than the mean distance of the assumed random distribution, the element is considered as a clustered element. Conversely, if it is greater than the mean distance of the assumed random distribution, the element is considered to be dispersed. Then:

$$ANN = \frac{D_A}{D_B} \tag{7}$$

$$D_A = \frac{\sum_{i=1}^{n} \min\left(d_{ij}\right)}{n} \tag{8}$$

$$D_B = \frac{0.5}{\sqrt{n/A}} \tag{9}$$

In the above equations, ANN is the nearest-neighbor ratio, D_A is the observed average distance between each building point and its nearest-neighbor building point, D_B is the expected average distance of building points in the stochastic model, and n is the total number of building points. d_{ij} is the distance from building point i to building point j, $\min\left(d_{ij}\right)$ is the distance from building point i to its nearest neighbor, and A is the area of the study area.

3) The XGBoost model

XGBoost is an extreme gradient boosting tree algorithm that combines supervised learning and integrated learning methods.XGBoost is first and foremost a supervised learning method, where supervised learning denotes a process of finding the mapping relationship between inputs and outputs by means of training samples containing labels. Its main implementation is to adjust each parameter by iterating the input-output pairs to finally obtain the optimal model. Among them, the main method to judge the optimality of a model is to use the objective function. The objective function can usually be denoted as:

$$Obj(X) = L(X) + \Omega(X), \tag{10}$$

where $Obj(X)$ denotes the objective function, $L(X)$ denotes the error function, and $\Omega(X)$ denotes the regular term.

In order to obtain the optimized model, the XGBoost algorithm uses the classical integrated learning approach. Integration learning denotes the construction of a stronger learning model by building and combining multiple weak learners. The XGBoost algorithm used in this paper is essentially a gradient boosting regression tree, and the goal of constructing the XGBoost model is to build a series of small regression trees, and to obtain a strongest regression tree to minimize the objective function by means of series decision making.

XGBoost model in the construction process, first of all, each round will add a new regression tree as the base learner, and in the last round of iteration results of the residuals of the basis for multiple iterations of training. Namely:

$$\begin{cases} \hat{y}_i^{(0)} = 0 \\ \hat{y}_i^{(1)} = \hat{y}_i^{(0)} + f_1(x_i) \\ \hat{y}_i^{(2)} = \hat{y}_i{}^{(1)} + f_2(x_i) \\ \cdots\cdots \\ \hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i) \end{cases} \tag{11}$$

$\hat{y}_i{}^{(t)}$ denotes the prediction result of round t, and $f_t(x_i)$ denotes the newly added regression tree of round t.

The objective function of the judgment model is as follows:

$$Obj^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i{}^{(t-1)} + f_t(x_i)\right) + \gamma^T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 + C, \tag{12}$$

where C is a constant, γ^T denotes the number of leaves, and $f_t(x)$ can be further defined as the tree structure of a weak regression tree q with leaf weights w_j, i.e. $f_t(x) = w_{q(x)}, w \in R, q :\in R$.

Compared with the traditional GBRT algorithm, a major improvement of XGBoost is that it is solved based on the second-order Taylor expansion of the objective function fitting instead of the underlying traditional first-order Taylor expansion. Omitting the inference step here, the objective function of the tth tree can finally be obtained which can be expressed as:

$$Obj^{(t)} \approx \sum_{j=1}^{T} \left[\left(\sum_{i \in I_j} \partial_{\hat{y}_i{}^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)\right) w_j\,, \tag{13}$$

where I_j denotes the set of samples falling on each leaf, $I_j = \{i | q(x_i) = j\}$.

In order to obtain the optimal solution, we derive Eq. (13) to obtain the optimal solution of w and the maximum gain obtained by its corresponding objective function. The result of partial derivation for w is as follows:

$$\tilde{w_j^*} = -\frac{\sum_{i \in I_j} \partial_{\hat{y}_i(t-1)} l\left(y_i, \hat{y}_i^{(t-1)}\right)}{\sum_{i \in I_j} \partial_{\hat{y}_i(t-1)}{}^2 l\left(y_i, \hat{y}_i^{(t-1)}\right) + \lambda}. \tag{14}$$

Substituting the optimal solution into Eq. (13) yields:

$$Obj = -\frac{1}{2} \sum_{j=1}^{T} \frac{\sum_{i \in I_j} \partial_{\hat{y}_i{}^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)^2}{\sum_{i \in I_j} \partial_{\hat{y}_i(t-1)}{}^2 l\left(y_i, \hat{y}_i{}^{(t-1)}\right) + \lambda} + \gamma^T. \tag{15}$$

The formula can be simplified by making the first-order derivative term $\sum_{i \in I_j} \partial_{\hat{y}_i{}^{(t-1)}} l\left(y_i, \hat{y}_i{}^{(t-1)}\right)$ in Eq. $G_j =$ and the second-order derivative term $\sum_{i \in I_j} \partial_{\hat{y}_i(t-1)}{}^2 l\big(y_i, \hat{y}_i{}^{(t-1)}\big) + \lambda\big)$ in Eq. (16):

$$Obj = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma^T. \tag{16}$$

Based on this objective function the core problem of XGBoost model construction can be solved, i.e., optimizing the splitting point. For the regression tree, maximizing the error reduction is the goal of optimization when finding the split. At this point, Eq. (16) can be used as a structure score to score the current regression tree structure. Therefore, the gain Gain can be defined as the difference between the sum of the left subtree scores and right subtree scores after splitting and the scores without splitting, considering the complexity cost of introducing new leaf nodes, which is used to score the current tree structure for evaluation. The gain formula is expressed as:

$$\text{Gain} = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma, \tag{17}$$

where G_L, G_R represents the sum of the first-order derivative terms of the left and right subsets and H_L, H_R represents the sum of the second-order derivative terms of the left and right subsets.

The XGBoost algorithm has done the following three optimizations, viz:

1) Optimized the algorithm itself, with the part of the error function in the objective function raised to Taylor's second order.
2) Improved the efficiency of the algorithm by using parallel computation in the construction of the weak regression tree, which improves the computational efficiency by means of parallel CPU acceleration.
3) The robustness of the algorithm is optimized, with special treatment of missing values, and the addition of L1 and L2 regularization part improves the generalization ability and stability of the algorithm.

4) SHAP model decoding method

SHAP is a method of interpreting the output of a machine learning model based on game theory, and its core is Shapley value, which is an additive interpretation model. In this paper, the SHAP algorithm is used to explain and analyze the relevant factors affecting the landscape space in the XGBoost model to increase the interpretability of the model.The role of SHAP value is to quantify the influence of each feature in that sample on the final output of the predictive model. In

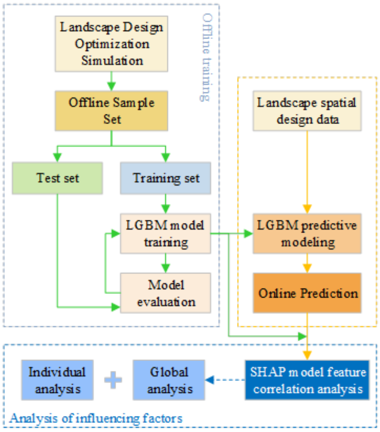

Figure 1: An interpretative framework based on SHAP

practice, the output values are attributed to the SHAP value of each feature value and used to measure its influence.Tasks that can be performed by SHAP are debugging the model, doing feature work, guiding the direction of data collection, and guiding decision making.

Assuming the ist sample x_i and its krd feature value is x_{ik}, the marginal contribution of the feature is m_{ij}, and the SHAP value of the edge with weight w_i, x_{ik} is $f(x_{ik})$, the kth feature of the ith sample x_i contributes to the predicted value y_i. Then the SHAP value corresponding to this feature is:

$$f(x_{ik}) = m_{ik}w_1 + m_{ik}w_2 + \cdots + m_{ik}w_n. \quad (18)$$

The baseline for the entire model, typically the mean of all sample target variables is:

$$y_i = y_{base} + f(x_{i1}) + \cdots + f(x_{ik}). \quad (19)$$

And the SHAP value also satisfies (19). When $f(x_{ik}) > 0$, it means that the feature has a positive and enhancing effect on the predicted value. On the contrary, it is the opposite and lowering effect. The biggest feature is that it not only reflects the influence of feature inputs in each sample, but also illustrates the positive and negative nature of the influence.

Aiming at the relevant influencing factors existing in the process of landscape design, this paper combines the SHAP model to construct an interpretable framework for the XGBoost model. The interpretable framework of XGBoost model based on SHAP is shown in Figure 1, which mainly analyzes the global and local characteristics of landscape design, so as to clarify the key direction of optimization in the process of landscape design, and to promote the urban landscape design more in line with the development trend of urbanization.

III. Evolutionary Characteristics of the Regional Landscape

Landscapes are geospatially heterogeneous, socio-economically driven, regional human-environmental systems, and land use is an important indicator for quantifying the interactions between human activities and the environment in a landscape.

Year	ANN index	Z score	P value
2000	0.673	-234.16	0.000
2004	0.651	-241.83	0.001
2008	0.638	-249.62	0.000
2012	0.612	-255.74	0.003
2016	0.594	-262.48	0.002
2020	0.576	-267.57	0.000
2022	0.565	-273.29	0.001

Table 1: Spatial distribution cluster evolution characteristics

Landscape indices, on the other hand, can be used to understand and quantify the composition and configuration of the landscape. When conducting spatial analysis of a landscape, understanding its spatial evolution characteristics is crucial for landscape design optimization. Therefore, based on the relevant data and research methods given in the previous section, this chapter conducts a quantitative analysis for the evolutionary characteristics of regional landscapes, aiming to explore the spatial optimization path of landscape design.

A. Spatial distribution characteristics

1) Characteristics of agglomeration evolution

The average nearest neighbor analysis of the data related to the urban landscape in Province G from 2000 to 2022 yielded the results of the distributional agglomeration evolutionary characteristics of the urban landscape space in Province G as shown in Table 1. During the study period, the ANN index values are all less than 1, with little change more constant, decreasing from 0.673 in 2000 to 0.565 in 2022, with an overall decrease of 16.05%. It indicates that the clustering distribution pattern of architectural landscape in province G is obvious, and there is only 1% or less possibility that this clustering pattern is the result of random generation. From the trend of ANN value, the ANN value shows a decreasing trend, indicating that the randomness of regional architectural landscapes is gradually becoming smaller. In addition, this paper also analyzes the clustering density of architectural landscapes through kernel density, and overall the evolution pattern of architectural landscapes in Province G has not shown large and obvious changes, but the overall density is continuously increasing. This also shows that the process of urbanization development has influenced the evolution of architectural landscape to a certain extent, making it closer and closer to the daily life of residents.

2) Distributional and morphological characteristics

The landscape pattern indices MSI, MPFD, AI and MPS selected in the previous paper characterize the morphological features of spatial distribution of architectural landscape in Province G. Using the formula of each index to bring in the relevant data for calculation, the morphological features of spatial distribution of landscape pattern for the period of 2000-2022 were derived as shown in Table 2.

From the overall morphological characteristics, during the study period, the AI changed from 63.512 in 2000 to 62.424

Year	MSI	MPFD	AI	MPS
2000	1.104	1.065	63.512	16.042
2004	1.118	1.068	63.507	16.185
2008	1.127	1.072	63.482	16.273
2012	1.135	1.076	63.338	16.034
2016	1.146	1.084	63.241	15.861
2020	1.163	1.091	62.869	15.726
2022	1.182	1.083	62.424	15.697

Table 2: Spatial distribution characteristics

in 2022, with a decrease of 1.71%, and the overall change was not significant, indicating that the architectural landscape space in Province G did not show a large-scale aggregation and development.The MPS increased by 1.44% from 2000 to 2008, and then decreased by 3.54% from 2008 to 2022. The average patch area appears to grow first and then decline somewhat. In addition, from the viewpoint of landscape morphology characteristics, the patch shape index (SHAPE) continued to increase, indicating that the patch shape tends to be irregular, and the patch area index (FRAC) is used to quantitatively describe the size of the patch area and the curvature of its boundary line, with a value between [0,1], and the closer it is to 0 means that its shape is simpler, and the change of FRAC index during the study period is small and close to 0, which indicates that the architecture in Province G landscape spatial shape is more regular, and the influence of human factors on the spatial distribution pattern of architectural landscape in G province is larger.

B. Landscape pattern characteristics

1) Landscape level index

Urban landscape design and spatial optimization produce changes with the increase of urbanization level, this paper chooses the urbanization development level as the dependent variable of the change of urban landscape pattern, and selects the dispersion index (SPLIT), the maximum patch index (LPI), the patch density (PD), the perimeter-area fractional dimension (PAFD), the spreading degree index (CONTAG), the Shannon diversity Index (SHDI), Scatter and Juxtaposition Index (IJI), and Patch Cohesion Index (COHESION) were used as evaluation indexes. The trend of landscape pattern indices at the landscape level was calculated as shown in Figure 2, where Figures 2(a) to (h) show the changes of each index with the urbanization level, respectively.

Specifically, the urbanization development level in the interval of 0% to 25% shows a gradual increase in SHDI, SPLIT and PD at higher levels, and a gradual decrease in LPI, COHESION and CONTAG at lower levels. The region has a variety of landscape types such as cropland, woodland, grassland and unutilized land, so the landscape patches are scattered and the landscape types are diverse. The urbanization development level is located between 25% and 40%, which ushers in the inflection point of several indices, in which the landscape indices of PD, SPLIT and SHDI reach the peak, and on the contrary, the landscape indices of LPI, COHESION and CONTAG show a low valley, which reflects the dispersal of the developed and undeveloped land use in the urban-rural fringe area, which makes the various types of patches distributed in a scattered manner and the landscape is heterogeneity is strong. Areas with an urbanization development level greater than 40% are mainly concentrated in the central urban area, which shows that the indices of SHDI, SPLIT and PD continue to decrease, while LPI, COHESION and CONTAG gradually reach their peaks. The reason is that the area is dominated by construction land, and the shape of the patches develops in the direction of regularization, with an obvious tendency of spatial aggregation.PAFD reflects the complexity of the shape of the patches, and the fluctuation trend of the inverted U shape confirms the variability of the spatial order of the urban fringe areas.IJI shows a fluctuating downward tendency, which indicates that with the increase of the development level of the city, the type of patches is more and more complex. IJI shows a decreasing trend of fluctuation, indicating that with the increase of urbanization level, the separation between patch types is weakened, and the distribution relationship between landscape types tends to be simpler.

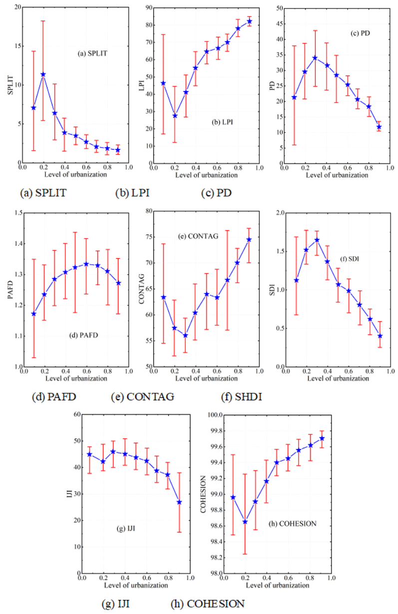

Figure 2: Landscape pattern index changes in landscape level

2) Type level index

At the type level, forest land, arable land and construction land, which have the most significant changes in land use types in the development of urbanization, were selected to analyze the trend of landscape pattern index changes in these three types of land. Figure 3 shows the trend of landscape pattern index changes at the type level, where Figures 3(a)∼(d) show the changes of each index with the urbanization level, respectively.

The COHESION index reflects the degree of aggregation of landscape patches, with construction land showing a trend

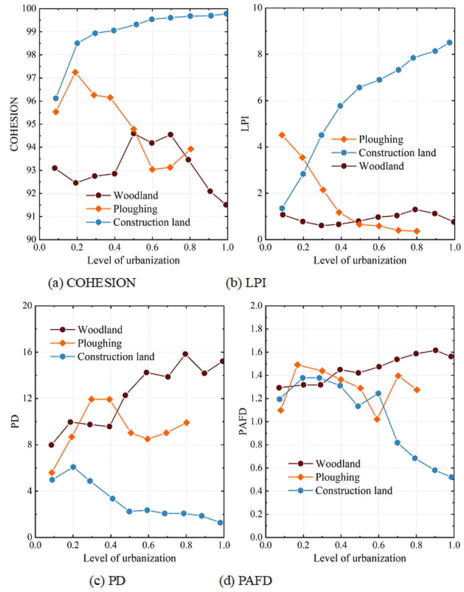

Figure 3: Landscape pattern index changes in type level

of smooth fluctuation, cropland showing an increase when the urbanization level is within 20% and starting to gradually decay after exceeding 20%, and forest land showing a trend of smooth fluctuation in change. In terms of LPI, with the increase of urbanization development level, the patch dominance of construction land gradually increases, cropland rapidly decreases, and forest land shows a fluctuating and decreasing trend at the level between [0.5,1.5]. In the case of PAFD, forest land showed a fluctuating increase with the increase in the level of urbanization and comprehensive development, indicating that the increase in the intensity of human disturbance leads to the increase in the complexity of forest land patches. Cultivated land and built-up land showed a trend of increasing and then decreasing, confirming the complexity of patch shape characteristics and the diversification of landscape patterns in the urban-rural intersection area. As far as PD is concerned, the patch density of forested land increased gradually and the degree of fragmentation increased with the influence of urbanization. The patch density of arable land showed a stepwise increase and then decrease, reaching a peak at about 20% of the integrated urbanization development level and then gradually decreasing, while the patch density of construction land showed a steady decreasing trend.

IV. Influences on Landscape Evolution

With the rising level of urbanization, the living conditions of human beings have been improved continuously, but also brought about serious ecological damage and air pollution, which triggered people's unprecedented attention to the built environment of scientific cities and the construction of healthy urban landscapes. The construction of urban green landscape pattern in the context of the development of the new era has an important role in promoting the construction of ecological civilization. Relevant studies show that urban green space

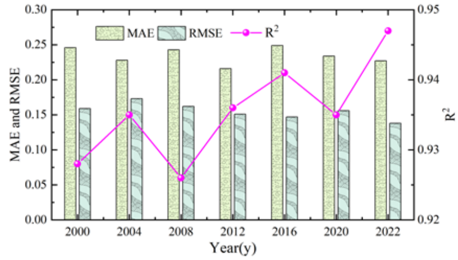

Figure 4: XGBoost model test results

landscape plays an important role in mitigating air pollution and improving public health, so it is of great significance to explore the influencing factors of urban green space landscape pattern to optimize the urban green space landscape pattern, protect the ecological environment and improve the urban air quality.

A. Model test results

In this paper, the XGBoost model is used to optimize the landscape design space type, and in order to avoid problems such as bias and overfitting of the XGBoost model during operation, the ten-fold cross-validation method and the three indexes, namely, the coefficient of determination (R2), the mean absolute error (MAE), and the root-mean-square error (RMSE), are utilized to optimize and evaluate the model performance. In this paper, the architectural landscape data of Province G from 2010 to 2022 is used as the training set, and the test results of the XGBoost model are obtained as shown in Figure 4.

According to the test results of the XGBoost model, the coefficient of determination (R2) for each period is more than 0.92, indicating that each landscape pattern index has good explanatory power for urban landscape design optimization. This paper utilizes the mean absolute error (MAE) to assess the accuracy of the prediction, and the smaller the value indicates that the model prediction is more accurate. The MAE of each period selected in this paper is lower than 0.25, indicating that the predicted landscape pattern index is more accurate. Root mean square error (RMSE) is the square root of the root mean square error between the fitted data and the real data of the corresponding sampling points, and the smaller the value indicates the better fitting effect. The RMSE for different periods is less than 0.2, indicating that the error between the fitted landscape pattern index and the original value is small. This indicates that the modeling process has obtained better results, which lays a good foundation for analyzing the relationship between landscape pattern index and landscape spatial optimization design.

B. Characteristic Importance

In order to optimize the design of urban landscape space, this paper uses the XGBoost model to model the optimization of landscape spatial design, and introduces several landscape pattern indices to analyze their importance using the SHAP model decomposition method, so as to explore which dimension is mainly used for spatial optimization in the process of landscape design, and to provide support for the spatial optimization of urban landscape. Patch Density (PD), Perimeter Area Fractional Dimension (PAFD), Spreading Index (SPLIT), Spreading Degree Index (CONTAG), Scattering and Juxtaposition Index (IJI), and Cohesion Index of Patches (COHESION) were selected as the indexes for spatial optimization of landscape design, which were categorized by using the XGBoost model, and analyzed in terms of their relevance to landscape design by combining them with the SHAP interpretable framework. Correlation with landscape design. Figure 5 shows the changes of each index with SHAP values, where Figures 5(a) to (f) are each landscape pattern index, respectively.

In 2010∼2022, PD, PAFD, and SPLIT, which characterize the shape of landscape patches, reflect the shape of landscape patches and are positively correlated with landscape spatial design, i.e., as the complexity of the shape of individual landscape patches increases, the ability of landscape spatial design increases. This may be that the more complex the shape of ecological patches, the more frequent the exchange of material and energy information between patches and other patches, and the more favorable to the ecological function radiation of patches.CONTAG characterizes the spatial spreading degree of urban landscape patches, and its value is in the [0.2,0.8], the SHAP value stably stays in the range of [-0.2,0.2], but in the interval of [0.8,1.2], with the CONTAG values increase, the SHAP values show a decrease. The reason for this is that within the central urban area, green plants in the green space are the main body that plays the role of landscape, and the areas with higher CONTAG are large areas such as urban-rural junction, where the green space accounts for a disproportionate share of the landscape space, which leads to a lower level of design of its landscape space. The correlation trend between IJI and COHESION, which characterize the size of the plaque, and the level of landscape spatial design is similar, both of which show that the larger the value of the indicator, the higher the value of SHAP, which is positively correlated with landscape spatial design. That is, the larger the area of landscape design patches, the more conducive to the level of urban landscape spatial design. It is worth noting that when the values of COHESION and IJI are near 0, the corresponding SHAP value fluctuates more in the interval of [-0.2,0.1]. There may be two reasons for this, one is that differences in shape lead to differences in landscape patches of similar area sizes, and the other is that different plant species and community structures in landscape patches cause differences in landscape types under the same area. Therefore, for landscape patches with small scale in the city, the improvement of landscape spatial design ability should pay more attention to the regulation of patch shape and spatial distribution when the increase of area is limited.

Further, the same approach is chosen in countries where the province of a and the land is severely desertification, and the conclusion of the idea of G is also applicable to

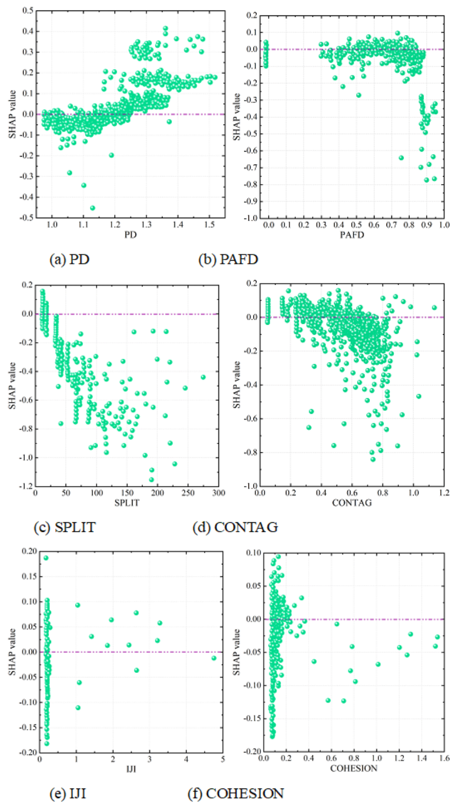

Figure 5: The important degree of the characteristics is sorted

other provinces or countries, showing the universality and universality of the research.

C. Policy principles of landscape planning

The planning and construction of urban landscape is becoming more and more concerned, and the urban landscape is becoming an indispensable part of people's lives. At present, most countries in the world have given high attention to the planning and construction of urban landscape, and formulate relevant laws and regulations and planning systems to regulate and control them. On the basis of the collection of different national data, the following points are mainly considered: first, the adequacy of the data, the country that focuses on choosing the urban landscape planning and construction ideas, and the comprehensive and comprehensive information can be used to make more detailed instructions and analysis, and the meaning is greater. The policy formulation, the planning guidance and the administrative management institutions are all different, but it is generally divided into distributed phase and independent types, and the distributed phase of which refers to the country's requirements for landscape planning to be dispersed in other planning and construction related laws and regulations and urban planning, which means that the country has developed more independent and complete landscape regulations and landscape planning.

V. Conclusion

In this paper, using the geographic data of Province G as the data source, the evolutionary characteristics of the regional landscape and the influencing factors of landscape evolution were analyzed by using the landscape pattern index, the XGBoost model, and the SHAP model interpretation method.

1) The agglomeration evolution of the regional landscape is analyzed in combination with the average nearest neighbor index, whose ANN index value decreases from 0.673 to 0.565 between 2000 and 2022, with an overall decrease of 16.05%. It reflects that the overall change of the urban landscape is small, and the trend of aggregation and distribution of the urban landscape is more obvious.
2) During the period of 2000-2022, the AI index of urban landscape decreases from 63.512 to 62.424, with a smaller overall decrease, the MPS index shows a trend of growth followed by a decrease, the change of SHAPE index continues to increase, and the change of FRAC index is smaller. It shows that overall the spatial shape of urban landscape is more regular, and it also shows that human factors under the development of urbanization are the main factors causing changes in the spatial distribution pattern of urban landscape.
3) Based on the change of urbanization level, it will have different impacts on the pattern index of urban landscape space at different stages. When the urbanization level is around 20%, the PD index of the cultivated land type reaches the peak, while the LPI index of the forest land type shows a fluctuating downward trend at the level between [0.5,1.5]. Urban landscape design and spatial optimization are built on the level of urbanization development, backed by urbanization development can provide more diverse options for landscape design.
4) The coefficients of determination (R2) obtained by the XGBoost model are all greater than 0.92, and the MAE is less than 0.25 and the RMSE is less than 0.2. Analyzing the characteristics of landscape pattern indexes related to the urban landscape design by using the XGBoost model can provide a reference for the spatial optimization of landscape.
5) In urban landscape design, the patch density (PD), perimeter area fractional dimension (PAFD), dispersion index (SPLIT), spreading degree index (CONTAG), scattering and juxtaposition index (IJI), and cohesion index of patches (COHESION) should be fully considered. Among them, when the CONTAG index is [0.2,0.8], the SHAP value stabilizes between [-0.2,0.2], while the SHAP value shows a significant decrease after exceeding 0.8.

In summary, the optimization of urban landscape design needs to pay full attention to the level of urbanization development, effectively enrich the spatial organization of the urban landscape, create a rich spatial sequence experience, and enhance the sense of participation of the citizens in the urban landscape through a more humanized experience of the nodes and facilities, so as to enhance the sense of well-being and satisfaction of the citizens.

References

[1] Kwak, Y., Deal, B., & Mosey, G. (2021). Landscape design toward urban resilience: Bridging science and physical design coupling sociohydrological modeling and design process. *Sustainability, 13*(9), 4666.

[2] Wei Lun Lee, A., Ying Chung, S., Shee Tan, Y., Mun Ho Koh, S., Feng Lu, W., & Sze Choong Low, J. (2023). Enhancing the environmental sustainability of product through ecodesign: a systematic review. *Journal of Engineering Design, 34*(10), 814-843.

[3] Dorado, I. G., Getreuer, P., Wronski, B., & Milanfar, P. (2020). Image Stylization: From predefined to personalized.

[4] Hendevari, M. S. , Bemani, M. , & Nikmehr, S. . (2020). A novel dual-band double-sided high-impedance surface (db-ds-his) with systematic design methodology. *Aeu-international Journal of Electronics and Communications, 115*, 153039.

[5] Chuang, T. F., & Chang, Y. H. (2021). A new design concept of an ecological corridor for frogs to improve ecological conservation. *Sustainability, 13*(20), 11175.

[6] Henderson, C. J., Stevens, T., Lee, S. Y., Gilby, B. L., Schlacher, T. A., Connolly, R. M., ... & Olds, A. D. (2019). Optimising seagrass conservation for ecological functions. *Ecosystems, 22,* 1368-1380.

[7] McGarigal, K., Compton, B. W., Plunkett, E. B., DeLuca, W. V., Grand, J., Ene, E., & Jackson, S. D. (2018). A landscape index of ecological integrity to inform landscape conservation. *Landscape Ecology, 33*, 1029-1048.

[8] Ji, X., Small, D. S., Leonard, C. E., & Hennessy, S. (2017). The trend-in-trend research design for causal inference. *Epidemiology, 28*(4), 529-536.

[9] Thai, N. T. T. , De Wever, B. , & Valcke, M. . (2017). The impact of a flipped classroom design on learning performance in higher education : looking for the best 'blend' of lectures and guiding questions with feedback. *Computers & Education, 107*(APR.), 113-126.

[10] Lionetto, M. G. , Caricato, R. , & Giordano, M. E. . (2021). Pollution biomarkers in the framework of marine biodiversity conservation: state of art and perspectives. *Water, 13*(13), 1847.

[11] Liu, Y. , Sun, Y. , Yang, A. , & Gao, J. . (2021). Digital twin-based ecogreen building design. *Complexity, 2021*(10), 1-10.

[12] Voicescu, S. A., Lane, J. F., Cooke, S. J., Higgs, E., Fisher, A. C., Rochefort, L., ... & Murphy, S. (2023). Awareness and use of the Society for Ecological Restoration's International Principles and Standards for the Practice of Ecological Restoration in Canada. *Restoration Ecology, 31*(1), e13789.

[13] Diran, D., Hoppe, T., Ubacht, J., Slob, A., & Blok, K. (2020). A data ecosystem for data-driven thermal energy transition: Reflection on current practice and suggestions for re-design. *Energies, 13*(2), 444.

[14] Zhou, Y., Sharma, A., Masud, M., Gaba, G. S., Dhiman, G., Ghafoor, K. Z., & AlZain, M. A. (2021). Urban rain flood ecosystem design planning and feasibility study for the enrichment of smart cities. *Sustainability, 13*(9), 5205.

[15] Karrasch, L., Maier, M., Kleyer, M., & Klenke, T. (2017). Collaborative landscape planning: Co-design of ecosystem-based land management scenarios. *Sustainability, 9*(9), 1668.

[16] Beisheim, K. , & Otte, A. . (2017). Evaluating the ornamental value of the caucasian flora in georgia. *Tuexenia, 37*(1), 333-354.

[17] Su, C., Xiao, Y., Shi, Y., & Peng, Y. (2020, August). Application of Humanized Design Concept in Industrial Design. In *Basic & Clinical Pharmacology & Toxicology* (Vol. 127, pp. 150-150). 111 RIVER ST, HOBOKEN 07030-5774, NJ USA: WILEY.

[18] Tu, S. (2022). (Retracted) Computer hand-painting of intelligent multimedia images in interior design major. *Journal of Electronic Imaging, 31*(5), 051418-051418.

[19] Wu, H. (2020, August). Urban Public Art Design from the Perspective of Humanized Elements Integration. In *Basic & Clinical Pharmacology & Toxicology (Vol. 127*, pp. 206-206). 111 RIVER ST, HOBOKEN 07030-5774, NJ USA: WILEY.

[20] Geng, J., & Liang, C. (2021). Analysis of the Internal Relationship between Ecological Value and Economic Value Based on the Forest Resources in China. *Sustainability, 13*(12), 6795.

[21] Jose, S., Walter, D., & Mohan Kumar, B. (2019). Ecological considerations in sustainable silvopasture design and management. *Agroforestry Systems, 93*(1), 317-331.

[22] Wan, C. K., & Lin, S. Y. (2022). Negotiating social value, time perspective, and development space in sustainable product design: A dialectics perspective. *Design Studies, 81*, 101121.