IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 August 2023, accepted 25 September 2023, date of publication 23 October 2023, date of current version 1 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3327137

RESEARCH ARTICLE

Visualization and Synthesis of LCCC Space Based on RF-SVM Optimization

SHENMIN WANG, QIFANG MA, AND HANWEI LIANG
School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China

Corresponding author: Qifang Ma (001791@nuist.edu.cn)

This work was supported by the 2017 Jiangsu Provincial Land Resources Science and Technology Plan Project "Study on reclamation and utilization of industrial and mining wasteland" under Grant 2017048.

ABSTRACT The comprehensive carrying capacity of land is a key indicator for judging the sustainable development of nature and economy. Conducting comprehensive studies on this land capacity can provide the government with valuable information for decision-making. To divide and utilize land in a more reasonable and scientific manner, this study constructed an evaluation model based on the Random Forest and Support Vector Machine algorithms. The objective is to establish a more precise evaluation model and conduct a more comprehensive evaluation of the indicators. Based on the experimental results, it can be concluded that the carrying capacity of all five systems increased between 2014 and 2022. System Dexhibited the fastest increase in carrying capacity, with a growth rate of 58.43%, while system R recorded a growth rate of 46.29%. The difference in data before and after spatial processing in a certain region was analyzed: the average error of the five regions was relatively small, indicating that the processed values can accurately express spatial visualization information. The accuracy of Random forest support vector machine evaluation model is higher than that of support vector machine, which verifies the optimization effect of Random forest algorithm on support vector machine. In the comparative experiment of four models, each error evaluation parameter from the constructed evaluation module was superior to those from other models. These results illustrate the superior applicability of the proposed research model in terms of comprehensive land bearing capacity. The study optimized the evaluation model for comprehensive carrying capacity of land. This model proposes countermeasures and suggestions for the assessment of land carrying capacity and provides decision-making references for the protection and development of land resources.

INDEX TERMS Random forest, support vector machine, spatial visualization expression, comprehensive bearing capacity of land, evaluation model.

I. INTRODUCTION

Land resources are a rare and precious resource that has always been a necessary condition for human survival, while supporting the sustainable development of the ecological environment and social economy. Land Comprehensive Carrying Capacity (LCCC) refers to the population or production scale that land can carry under certain socio-economic conditions. It is an important basis for land use planning and regional sustainable development, and also an important indicator for evaluating land use and development potential [1]. The evaluation model for land carrying capacity was optimized in this study. The model suggests measures for assessing land carrying capacity and offers decision-making references for the protection and development of land resources. The traditional evaluation model is based on Support Vector Machine (SVM), but the accuracy of the traditional SVM model is not high and is not suitable for analyzing complex problems with large numbers of samples [3]. Random Forest (RF) is a machine learning algorithm that classifies data by constructing multiple decision trees. The core of RF is to establish a multi-level, overlapping forest model, with each model selected from different forest models through independence and correlation of the dataset.

The associate editor coordinating the review of this manuscript and approving it for publication was Shuihua Wang.

RF can effectively process high-dimensional data and has strong scalability [4]. Furthermore, the RF algorithm does not require data normalizationwhich saves time. Moreover, it has the ability to evaluate the importance of built-in indicators and estimate errors independently [5]. In view of this, this study utilizes RF to improve SVM and constructs an LCCC evaluation model based on RF-SVM. It is expected to establish a higher precision evaluation model to comprehensively evaluate the indicators.The significance of the research mainly includes the following points: firstly, it provides theoretical basis and data support for the comprehensive research of LCCC, secondly, it expands the application scope of RF and SVM, and finally, it provides theoretical basis for relevant government decisions to achieve the goal of reasonable land allocation and development, which helps to promote sustainable development of land.The content is mainly divided into four sections. The Part 1 is a brief introduction to the relevant research on RF, SVM, and LCCC; The second part designs the EIS for LCCC, and then improves the SVM evaluation model based on RF; The third part is to test and analyze the performance of the LCCC model based on RF-SVM; The fourth part is a summary and outlook of the research content.

II. RELATED WORKS

In the field of deep learning, research on RF has become relatively mature. Franklin E. B et al. proposed a new RF based identification method for the identification of components when studying the source and formation mechanism of aerosol substances.The method trains and tests samples with known components, predicts the corresponding factors of organic compounds in new environments based on the polarity space of volatility and the position in mass spectrometry, and achieves efficient quantification of new environmental substances. Quantitative experiments on the test set demonstrate that the new RF has high accuracy, small errors, and is significantly superior to other algorithms [6].In the composite material manufacturing, Lim J et al. found that in order to meet the physical properties of composite materials, repeated production of materials is required, which not only causes a lot of waste but also prolongs working hours. In view of this, they constructed a model for RF prediction of physical performance built on the dataset. It uses importance analysis to simulate the impact of numerical values on material properties. Then, it employs appropriate software to predict the physical properties of composite materials used in this industry. This model is highly accurate, effectively preventing material wastage and improving work efficiency. It also finds certain applications in the manufacture of commercial composite materials [7].To carry out quantitative analysis of heavy metals in soil more quickly, Mao Gang L.I et al. designed an analysis model based on Laser-induced breakdown spectroscopy and RF. First, the impact of various spectral preprocessing methods on the performance of the proposed model was investigated. Then, normalized spectral data was used to analyze heavy metal elements in soil quantitatively. The proposed model demonstrates high accuracy and fast speed, signifying that RF has certain applicability in heavy metal analysis field [8].

At the same time, SVM has also been applied in various fields. To study the sequence problem of genes, KangS et al. constructed a classifier based on SVM to distinguish the spectra of different gene sequences. The accuracy of the trained classifier can reach over 90%, and it can basically accurately identify each gene sequence. Compared to other methods, SVM has higher accuracy and stronger robustness [9]. Teng et al. raised a new detection method in the field of clinical diagnosis to address the slow detection speed of bacteria. This method detects bacteria by combining two spectral data technologies with SVM. First, spectral data is utilized to enhance SVM classifier performance and reduce analysis time. Then, feature lines are selected for input optimization. The method achieves high accuracy and reduces detection time [10]. Liu et al. designed a new equalizer based on SVM for blind equalization of signals. First, a group of Error function weighted by neighborhood symbols and augmented by generalized power factor are proposed as penalty terms in equalizer. In addition, they also proposed a new method to reduce the complexity of the Error function. Under the same complexity, the optimized SVM equalizer has lower interference and faster rate of convergence than the traditional equalizer [11]. Lots of scholars have conducted discussion on LCCC. A method based on LCCC was proposed by Tong et al. to address food nutrition issues by analyzing the supply and demand of food. The first step in constructing LCCC is to utilize data from local household surveys and land use rate surveys. The goal of the evaluation model is to understand the food supply and demand relationship in representative areas, such as grassland pastoral areas. The study selected three food indicators to analyze the tolerance of LCCC. This method can accurately predict the supply and demand of grain, providing practical reference value for the development of grain nutrition [12]. An evaluation model based on LCCC was established by Zhang H et al. The model combines national poverty alleviation policies to explore the impact of land resources on poverty. Geographical correlation coefficients were used in the study to discuss the significant relationship between poverty incidence and LCCC. Therefore, LCCC can be used to select regions for Targeted Poverty Alleviation, so as to help poor people achieve poverty alleviation and local sustainable development goals [13].

In summary, the study on LCCC is beneficial for the development of various aspects of the country. Tong S et al. can conduct land use efficiency surveys based on LCCC, construct evaluation models, and promote the resolution of food nutrition issues. Zhang Hai et al. were able to implement poverty alleviation policies based on LCCC and use the LCCC evaluation model for precision poverty alleviation. However, the models proposed are all based on traditional SVM models, which exhibit limited processing ability for complex sample data. The RF algorithm is advantageous in evaluating the significance of built-in indicators and does not demand data normalization, thus achieving desirable outcomes in various fields. To overcome the limitations of traditional models incapable of handling complex samples and construct a more reasonable and objective LCCC evaluation model, this study uses the RF-SVM algorithm to create an LCCC evaluation model with higher accuracy.

III. LCCC CONSTRUCTION BASED ON RF-SVM OPTIMIZATION

The study of LCCC can provide a theoretical basis for land use and planning, and objectively and scientifically evaluate the potential for land use and development. Establishing an evaluation model for LCCC is crucial in LCCC research. This study constructed RF-SVM-based LCCC. Firstly, the study comprehensively considers various factors of society and nature, establishes a criterion layer, and designs an evaluation system based on this. Then, two types of methods for visualizing evaluation indicators are introduced in detail. Next, the grading criteria for indicators are developed to establish SVM. Finally, the shortcomings of traditional SVM were discussed and improved with RF to construct LCCC optimized through RF-SVM.

A. DESIGN OF EVALUATION INDICATOR SYSTEM AND VISUALIZATION OF EVALUATION INDICATORS

To build a RF-SVM-based optimized LCCC, a comprehensive EIS must be designed beforehand. The system must comprehensively consider several factors, including ecological environment, society, and economy in order to design a comprehensive evaluation system [14].Simultaneously, the chosen indicators must precisely denotespecific relationships between critical issues within the system and interact with it to promote comprehensive management, thus achieving a more effective evaluation of LCCC's efficiency [15]. Therefore, this study constructs an EIS based on five criteria layers, as described in Fig.1.

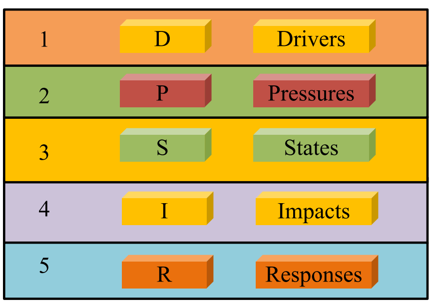


FIGURE 1. The meaning represented by 5 criteria layers.

As in Fig.1, D represents the driving force, which is the factor that causes changes in the land system, including natural and socio-economic driving forces; P is pressure, which refers to the pressure exerted by human activities on LCCC; S is the state, which is the direct result of the interaction between D and P, presenting the true state of LCCC; I is the final LCCC situation; R represents a response and an improvement proposal for the LCCC situation. This study constructs an LCCC EIS, as shown in Figure 2, using the aforementioned five levels as guidelines and adhering to principles such as scientific objectivity, comprehensiveness, and independent selection.

In Fig.2, the driving force criterion selects the ultimate capacity of agricultural machinery, GDP, and urbanization rate; The pressure criteria include population density, fertilizer and pesticide application/unit, and per capita water supply; The state criteria include land use intensity, per capita arable land area, and per capita grain yield; The selection of impact criteria includes temperature, water conservation, and habitat quality; The response criterion selects the proportion of wastewaterdisposal rate and pro-environment investment to GDP. This evaluation model has a total of 14 evaluation indicators, which can accurately reflect various influencing factors and the relationships between them. After the construction of the evaluation system is completed, a detailed understanding of the evaluation methods is also needed. This study chooses to visualize the evaluation index space. The evaluation indicator, Spatial Visualization (SV), describes the process of converting a dataset into a spatial grid, establishing a database, and displaying the data's characteristics in a more intuitive format [16]. SV allows for identifying significant features in the dataset, comprehending and analyzing the data, and ultimately enhancing the model's evaluation ability and accuracy. The SV processing used in this study adopts two methods, and their characteristics are listed in Fig.3.

Figure 3 shows that Method 1 is an SV processing technique which is primarily based on basic data. This method spatializes the indicators by assigning spatial weights. The processing of data using SV is a scientific and efficient method of spatialization. This method allocates data from a specific scale and maps it to a spatial area with a particular resolution, in order to model the statistical data distribution in geographical space [17].Method 2 applies spatial interpolation and ecological models to simulate and compute indicators. The regional grid mean of statistical data is used, which is divided by the number of existing grids for SV. Land use intensity, habitat quality, and water conservation models can be simulated using ecological models. Additionally, spatial interpolation is capable of handling temperature.

Data based SV processing requires a spatialization formula for each indicator that corresponds to it. The formula for spatialization of urbanization rate is Eq.(1).

$$Q = q \times \frac{(S_1 + \ldots + S_j)}{j} \tag{1}$$

In equation (1), the spatialized grid cell values are represented by Q. The weight of land-usagekinds is S. The statistical values of the indicators are expressed in q. The quantity of land-kinds is j. The spatial formula for the application amount of fertilizers and pesticides per unit area is Eq.(2).

$$Q = \frac{q}{n} \times \frac{(S_1 + \ldots + S_j)}{j} \tag{2}$$

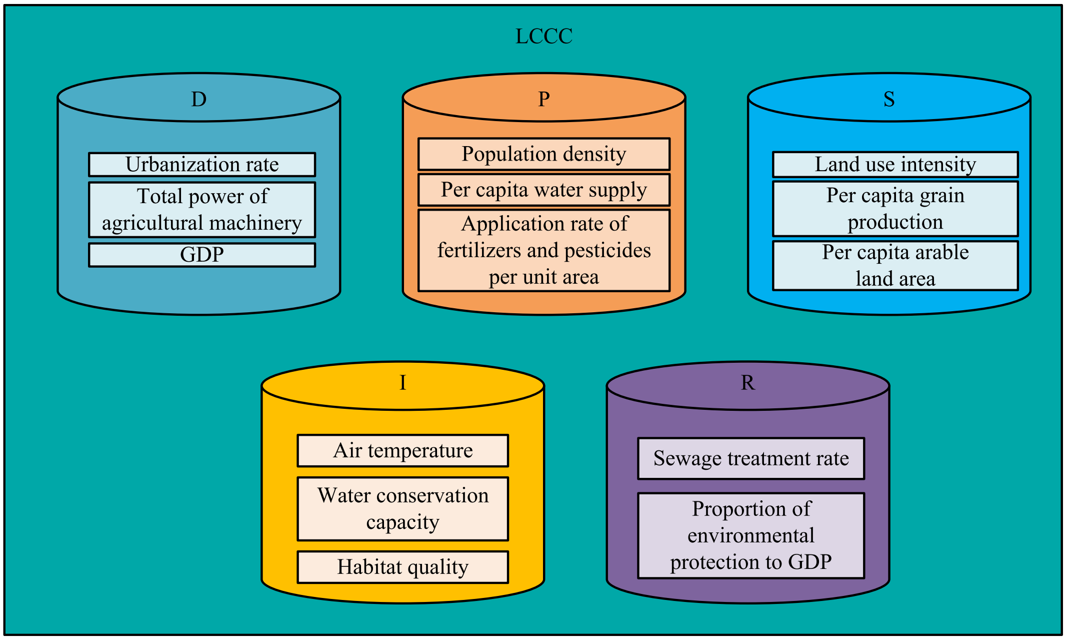


FIGURE 2. **Construction of LCCC index evaluation system.**

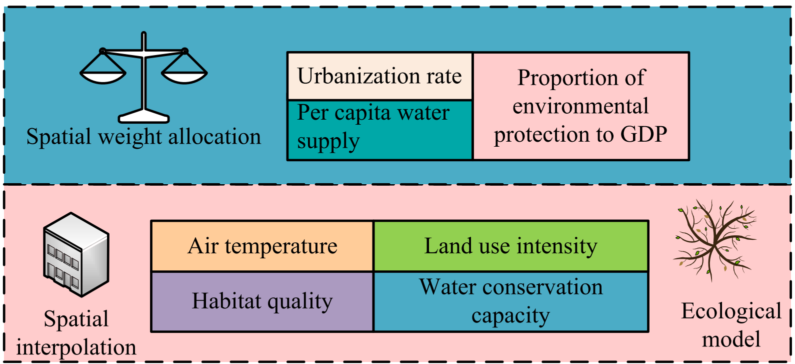


FIGURE 3. **Visualization processing of evaluation index space.**

n in equation (2) represents the sum of grids in the region. The formula for per capita water supply is Eq.(3).

$$Q = \frac{q}{n \times POP} \tag{3}$$

In Eq.(3), POP represents the population density. The formula for the ratio of pro-environmental investment to GDP is Eq.(4).

$$EPI = \frac{q}{n \times GDP} \tag{4}$$

In equation (4), EPI represents the spatialized grid cell value. GDP is the grid cell value. In the second type of method, the calculation formula for water source conservation is Eq.(5).

$$Y_{xj} = (1 - \frac{AET_x}{P_x}) \times p_x \tag{5}$$

In equation (5), Y_{xj} represents the depth of water production in grid unit x of land type j; AET_x is the annual actual increase and emission of x in the j region; p_x represents the annual precipitation of x. The expression for AET_x is Eq.(6).

$$\frac{AET_{xj}}{P_x} = \frac{1 + \omega_x R_{xj}}{1 + \omega_x R_{xj} + \frac{1}{R_{xj}}} \tag{6}$$

In equation (6), R_{xj} represents the ratio of potential evaporation to precipitation; ω_x is a physical parameter related to soil. The expression related to ω_x is Eq.(7).

$$\omega_x = Z \times \frac{AWC_x}{p_x} \tag{7}$$

In equation (7), AWC_x represents the available water volume of plants, and Z is a coefficient representing regional

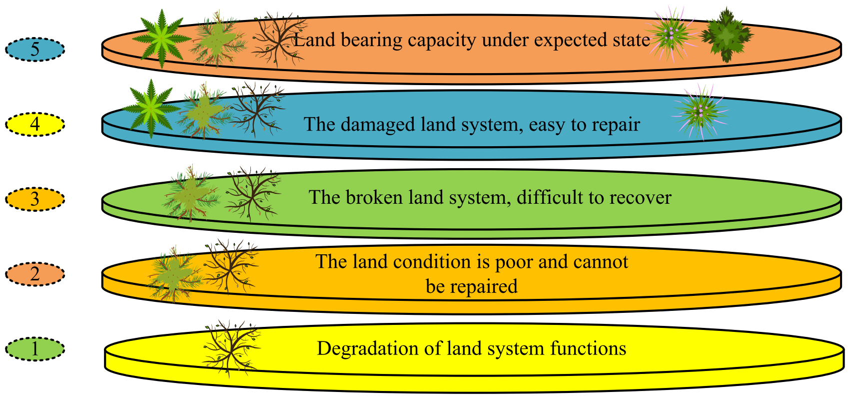


FIGURE 4. Five grading standards representing the bearing capacity of each indicator.

precipitation. The expression for R_{xj} is Eq.(8).

$$R_{xj} = \frac{K_{xj} \times ET_{ox}}{p_x} \tag{8}$$

In equation (8), ET_{ox} represents the potential evaporation of x, and K_{xj} is the evaporation coefficient of plants. Habitat quality can reflect the degree of rarity, such as the degree of degradation, and its expression is Eq.(9).

$$Q_{xj} = H_j \left[1 - \left(\frac{D_{xj}^z}{D_{xj}^z + k^z} \right) \right] \tag{9}$$

In equation (9), Q_{xj} represents the habitat quality of x in land-usage and land cover j. D_{xj} is the level of stress, k is a constant, H_j represents habitat suitability, z is a normalized constant, and the value is taken. The calculation formula for D_{xj} is Eq.(10).

$$D_{xj} = \sum\nolimits_{r=1}^{R} \sum\nolimits_{y=1}^{Y_R} \left(\frac{W_r}{\sum_{r=1}^{R} W_r} \right) r_y i_{rxy} \beta_y S_{jr} \tag{10}$$

In equation (10), R is the threat factor; r refers to the grid layer; y is thegrids numberin r; Y_R represents the amount of grids occupied by R; W_r is the weight of R; The threat factor r_y represents y is 0 or 1; i_{rxy} represents the threat level of r_y to x; β_y is the accessibility level of x, from the range of $0 \sim 1$, and its value 1 represents the level that is easily achievable; S_{jr} represents the sensitivity of j to r, and the expression for i_{rxy} is Eq.(11).

$$i_{rxy} = 1 - \left(\frac{d_{xy}}{d_{r\max}} \right) \tag{11}$$

In equation (11), d_{xy} is the linear distance between grid x and y; $d_{r\max}$ is the max influence distance of r.

The intensity of land use is one of the important indicators reflecting the impact of human activities on the ecological environment. The situation after adjusting the forest land intensity formula in this study is Eq.(12).

$$L = \sum\nolimits_{i=1}^{n} G_i \times S_i \tag{12}$$

In equation (12), G_i is the intensity level value; S_i represents the ratio to total area.

B. BUILDING LCCC BASED ON RF-SVM

In the evaluation process of LCCC, it is necessary to establish a grading standard, and this study constructs five grading standards that can fully reflect the bearing capacity of each indicator. The specific situation is displayed in Fig.4.

In Fig.4, the fifth level represents the expected land bearing capacity, which is strong; The fourth level represents that the land has suffered slight damage but is easy to repair, with strong bearing capacity and great development value; The third level represents the destruction of the land system, which can maintain basic transportation and provide basic services; The second level represents poor land conditions that cannot be repaired and barely maintain some services; At level one, the land condition is assessed to be extremely poor, the system function is severely degraded, and recovery is deemed to be very difficult. Additionally, it does not have the social carrying capacity.

The commonly used LCCC models include the SVM evaluation model. SVM is a machine learning technology that classifies data by establishing an optimal Hyperplane for classification. Its purpose is to establish a nonlinear transformation to map the dataset into a high-dimensional space and find an SVM classification surface in this high-dimensional space, thereby achieving classification of new data [18]. One advantage of SVM is their ability to handle small samples and high-dimensional data effectively, which allows for strong generalizability. Fig.5 shows the running steps of SVM.

In Fig.5, the first step is to preprocess the data, input indicators, and then normalize it to keep the numerical range between [0, 1], and select the range method as the normalization method; The second step involves creating an SVM model using the appropriate software's "svm()" function, with the radial kernel function selected as the basis kernel function. Create original samples, followed by training and testing samples, and then use them to predict samples. The step3 is to find the optimal values of relevant parameters, and set the parameters of the control kernel function and cost parameters when fitting the training samples. Set parameter intervals for each function and generate an SVM for each

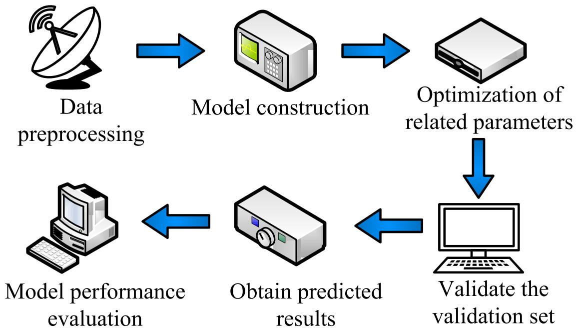


FIGURE 5. Steps for running SVM.

parameter value to find the most suitable parameter. The step4 is validation set testing, which involves substituting validation samples into the trained model to test the classification error of the samples. If the accuracy does not meet the standard, continue to return to the previous step and search for the most suitable parameters again until the standard is met; The fifth step is to substitute the data into the model for prediction, and obtain the final prediction result; The step6 is to test and compare the performance of the model. However, the accuracy of traditional SVM is not high and is not suitable for analyzing problems with large sample sizes [19]. RF is very good at handling complex samples and does not require data preprocessing, which can save time. RF can also evaluate the importance of built-in indicators. This process requires strict consideration of the internal laws of indicators and objective use of data to obtain the contribution value of the number of indicators. The RF algorithm includes a verification process for estimating accuracy that does not require validation experiments [20]. Considering this, the study uses RF to enhance SVM and then creates LCCC for RF-SVM. Sample RF randomly using the cloth bag method, and place it back uniformly. The number of samples in the sampler is equal to the number of generated samples, which to some extent reduces the error rate and has strong interference. The final classification result is Eq.(13).

$$G(x) = \arg marx \sum\nolimits_{i=1}^{k} I(g_i(x) = Y) \tag{13}$$

In equation (13), the ensemble classifier composed of multiple decision trees is represented by $G(x)$, g_i represents the i-th classifier in the ensemble classifier, the output variable is represented by Y, and $I(g_i(x) = Y)$ is the corresponding function of the classification rule. When RF uses the cloth bag method for regular sampling, the $1/3$ samples presentin each group are not extracted. The data that has not been extracted is out-of-bag (OOB), and the classification results of this type of data can be used for error estimation. The error of RF can be calculated through the error rate of OOB, taking the average of each error rate. Compared to the error analysis of other algorithms, the calculation of OOB error rate has a higher accuracy. RF has more advantages compared to decision tree algorithms.Firstly, RF is an integrated algorithm of decision trees; Secondly, RF effectively avoids the drawbacks of overfitting; Finally, in specific classification cases, it can use the Gini coefficient impurity of the index to evaluate the importance of the index, and its expression is Formula (14).

$$G(t) = 1 - \sum\nolimits_{k=1}^{Q} p^2(\frac{k}{t}) \tag{14}$$

In Formula (14), Q is the total number required for classification, t is the node, k is the quantity of classifications, and $p^2(\frac{k}{t})$ is the Conditional probability of Class k in t. It is determined by the purity of the index Gini coefficient of each decision tree.

The relationship between a feature and its target value can be represented by partial dependencies, which can be presented through partial dependency graphs. The Random Forest algorithm presents the partial dependency graph by fitting the model first and then expressing the functional expression of the partial dependency graph as an equation (15).

$$f(X_1) = \frac{1}{n} \sum\nolimits_{k=1}^{n} RF(X_1, X_2^k, X_3^k, \ldots, X_n^k) \tag{15}$$

As equation (15), X_n^k represents the n-th feature of the k-th sample in the training set. $f(X_1)$ represents the value of the partial dependency graph in X_1, that is, the evaluation value predicted by the previous model after converting the first variable in the training set into X_1. Based on different values of X_1 and $f(X_1)$, they are trained as a discount, which is a partial dependency graph. At this point, the horizontal axis is X_1, and the vertical axis is the partial dependency value. The operation of the RF-SVM model is exhibited in Fig.6.

Fig.6 first generates the original samples, selects 200,000 samples, and the ratio of the training set, test set, and prediction set is 6:2:2; Based on the five grading criteria, 100,000 samples were extracted from each interval, resulting in a total of 500,000 original samples; Randomly select 50% of 500,000 as training specimens and 50% as validation specimens. 14 indicators were input as variables into the model, and five grading criteria were adopted as output-variables. They were tested using the RF-SVM. Optimize model parameters. The training of the function requires setting the variables and purpose parameters of the decision tree during the node partitioning process. Afterwards, select the optimal parameters based on the results to train the algorithm. The OBB is used to evaluate the error of the model and evaluate the importance of the indicator: the larger the Gini coefficient of the indicator, the more important the indicator is. Finally, verify the validation set and output the results, substituting the samples into the trained RF-SVM model to test the classification error rate of the samples. If it does not meet the standards, go back to the previous step and reselect the optimal parameters until the standards are met. Substitute grid data into RF-SVM for prediction and obtain the final prediction result.

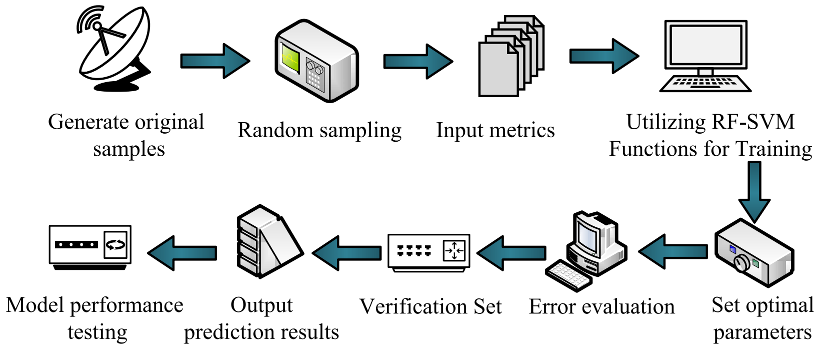


FIGURE 6. The operation of the RF-SVM model.

IV. LCCC EXPERIMENT BASED ON RF-SVM OPTIMIZATION

This study conducted performance testing on LCCC based on RF-SVM. In order to avoid experimental errors caused by different equipment, the study used the same computer for the experiment, with a CPU of Inter (R) Core (TM) i5-10210U, RAM of 20GB, operating system of Windows 10 Home, and memory of 8GB. Firstly, the performance of the algorithm is tested, and then the research method is compared and analyzed with common LCCC models to verify the applicability of the research method in LCCC evaluation.

A. PERFORMANCE TESTING OF LCCC EVALUATION MODEL BASED ON RF-SVM OPTIMIZATION

To determine the optimal parameter settings for model training, randomly selected samples from five levels were used as input to the model according to the grading criteria. This approach allowed us to obtain the relationship between the number of decision trees and the variation of errors, in addition to the relationship between the selected variables and errors during node partitioning. The relationship between the two types of changes is shown in Fig.7.

Fig.7 (a) shows the change curve of MSE under different of decision treenumbers: with the increase of it, the change law of sample mean square error at five levels tends to be consistent. The MSE first drops rapidly, then slowly, and finally at 100, it tends to be stable. 7(b) is the error variation curve for different candidate variables: as the number of candidate variables increases, the OOB error first rapidly decreases, then fluctuates slightly. When the number of candidate variables is 12, the curve tends to be horizontal. In summary, when the number of decision trees is 100 and there are 12 candidate variables, the parameters are optimal. To test the practicality of the model, real sample data from a certain region in 2014, 2016, 2018, 2020, and 2022 will be introduced, and RF-SVM models will be constructed through five criteria layers of D, P, S, I, and R indicator systems. From this, the bearing capacity results of the indicator system for the five criteria layers in this region are obtained. Fig.8 shows the changes in bearing capacity of indicator system D.

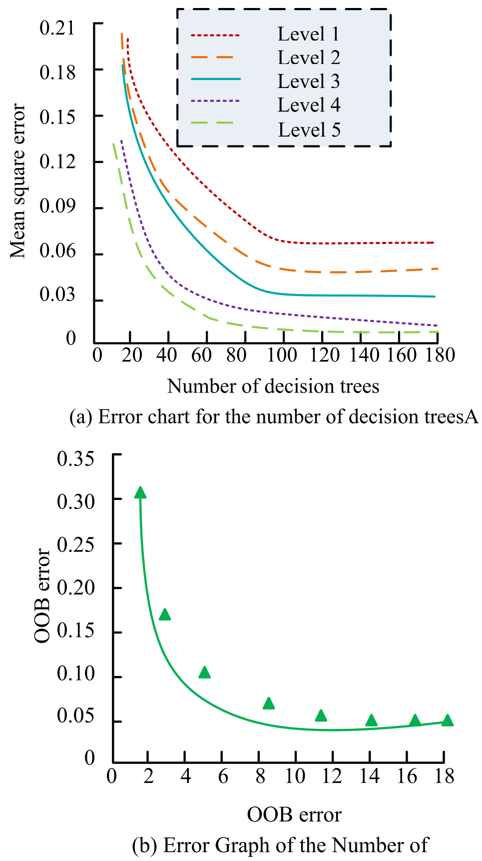


FIGURE 7. Changes in the decision tree amounts and error changes in candidate variables.

In Fig.8, between 2014 and 2022, the land area curve showed a downward trend at the 1st level, with an area reduction of 75.97%. The trend in the 2ndlevel is: descend-ascend-descend, with an area reduction of 72.47%; The 3rd level is: ascend-descend-ascend, with an area reduction of 37.48%. The second and third levels have shown an overall downward trend. The 4th level shows an overall upward trend, with an area increase of 97.87%; The overall trend of the 5th level shows a rapid increase, with an area increase of 97.84%. Fig.9 shows the changes of indicator system P.

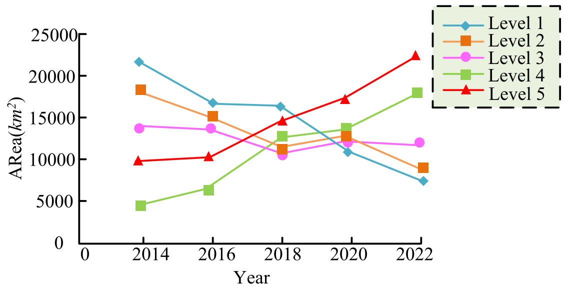


FIGURE 8. The variation of the bearing capacity of indicator system D over time.

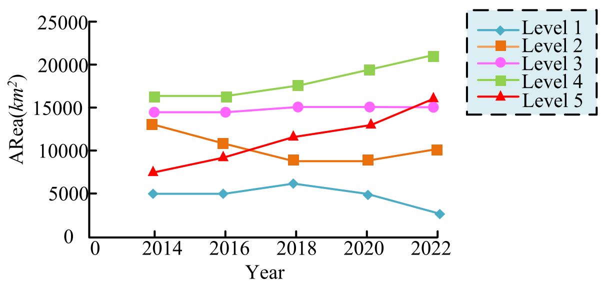


FIGURE 9. The variation of the bearing capacity of indicator system P over time.

Between 2014 and 2022, the land area curve at the first level showed an overall downward trend, with an area reduction of 9.31%; The second level is: downward upward, showing an overall downward trend, with an area reduction of 28.87%; The fluctuation range of the third level curve is relatively small, tends to be horizontal, and there is no significant change in area; The fourth level area shows an upward trend, 30.21%; The fifth level is a rapid increase, with an area increase of 99.25%. Fig.10 shows the relevant situation of system S.

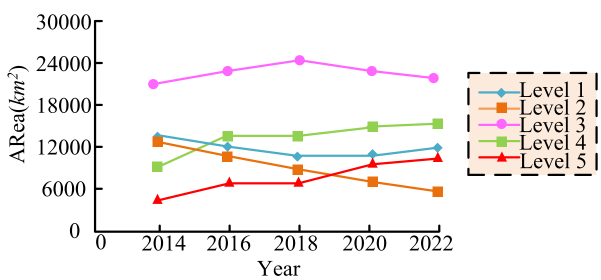


FIGURE 10. The variation of the bearing capacity of indicator system S over time.

During the 2014∼2022, the level1 first decreased and then increased, showing an overall downward trend, with an area reduction of 7.21%; The level2 overall decreased, with an area reduction of 53.96%; The level3 first rises and then decreases, with no significant difference between 2014 and 2022; The level4 and level5 showed an overall increase of 39.97% and 96.12%, respectively. Fig.11 shows the changes in System I.

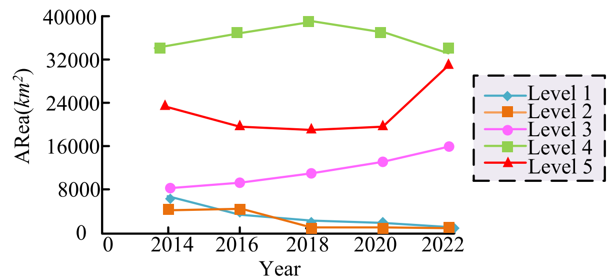


FIGURE 11. The variation of the bearing capacity of indicator system I over time.

During the 2014∼2022 in Fig.11, the 1st and 2nd levels showed a downward trend, with an area reduction of 96.16% and 57.12% respectively; The 3rd-level overall rose, with an area increase of 91.97%; The 4th-level first rises and then decreases, with relatively small changes in area over the years; The 5th-level first decreased and then rapidly increased, showing an overall upward trend, with an area increase of 34.97%. Fig.12 shows the bearing capacity of R.

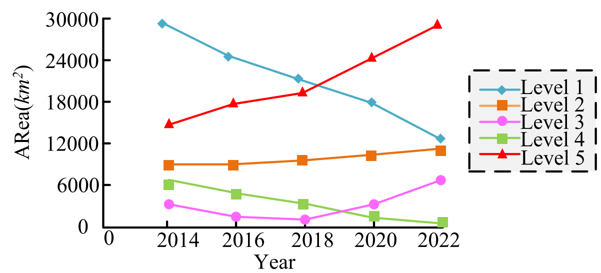


FIGURE 12. The variation of the bearing capacity of indicator system R over time.

In Fig.12, over the past five years, the land area curves at the first and second levels showed a rapid decline and an overall upward trend, with an area decrease of 41.48% and an increase of 5.21%, respectively. The land area of the third level increased by 85.22%, showing an overall upward trend: decreasing rising. The area at the fourth level decreased by 98.87%. The area of the fifth level increased by 135.19%, showing a rapid upward trend overall. The carrying capacity of the indicator system of the five criteria layers showed an upward trend from 2014 to 2022. System D has the fastest growth rate, with a growth rate of 58.43%, followed by 46.29% of R. Study the evaluation of LCCC in a certain area based on the grid of $500 \times 500m$ as the basic unit. There are a total of 227,968 grid data. To better integrate economic data with natural data, a calculation method combining spatial models was studied, and spatial weights were reassigned to the data, achieving visualization of LCCC space. The results are listed in Table 1.

In this study, the method of spatial weight redistribution of land index types is used to analyze the statistical Panel data SV. The evaluation and scoring of the importance level of indicators based on the type of land use can more accurately reflect the characteristics of the data. Table 1 shows the data

TABLE 1. Comparison of statistical panel values and spatially calculated values of fertilizer and pesticide application in a certain region.

Grade	Zone A	Zone B	Zone C	Zone D	Zone E
Statistical panel value (tons)	46298	69765	97543	47989	120855
Spatialization calculation value (tons)	45742	67742	83789	45734	108528
Error rate(%)	1.2%	2.9%	14.1%	4.7%	10.2%

on pesticide and fertilizer application rates for five regions A, B, C, D, and E in a certain region in 2022. Through specific analysis of the differences in data before and after spatialization processing, it can be seen that region C has the largest difference, reaching 14.1%; A is the smallest, only 1.2%. Overall, the error of this method is relatively small, with an average error of only 6.62%, indicating that the processed values can accurately express SV information.

B. COMPARATIVE ANALYSIS OF LCCC BASED ON RF-SVM OPTIMIZATION

To verify the improvement effect of LCCC optimized with RF-SVM on SVM, the evaluation accuracy of two methods was compared experimentally. See Fig. 13 for details.

Fig.13 shows a comparison of the accuracy verification between the two models. It determines the accuracy of the model by comparing the Area under the Curve (AUC) size, and the closer it is to 1, the higher the accuracy of the model. The AUC of SVM is 0.764, while the RF-SVM is 0.979, indicating that the accuracy of RF-SVM is higher than that of SVM. To further verify the superiority of RF-SVM performance, multiple tests were conducted on the same sample data. The common evaluation models for LCCC introduced in the experiment include SVM, Genetic Algorithm (GA), and Particle Swarm Optimization (PSO). Evaluation indicators: Mean Relative Error (MRE), Mean Squared Error (MSE), Mean Absolute Error (MAE), Root-mean-square Error (RMSE) and Coefficient of determination R2. Under the optimal preset parameters, the average index values of each parameter are displayed in Table 2.

Table 2 shows that the MRE of RF-SVM is 2.353%, MSE is 1.382%, MAE is 7.216%, RMSE is 0.1102, and R2 is 0.9998. Each error evaluation parameter is superior to other models. When the algorithms' optimal preset parameters were employed, the evaluation results ranked RF-SVM, RVM, GA, and PSO from best to worst. Overall, in the comparison of various model errors, RF-SVM achieved the best evaluation effect and the lowest error, demonstrating its higher accuracy and greater applicability for the comprehensive analysis of LCCC.

V. DISCUSSION

Land resources are a crucial basis for human survival and development, as well as sustainable economic and social

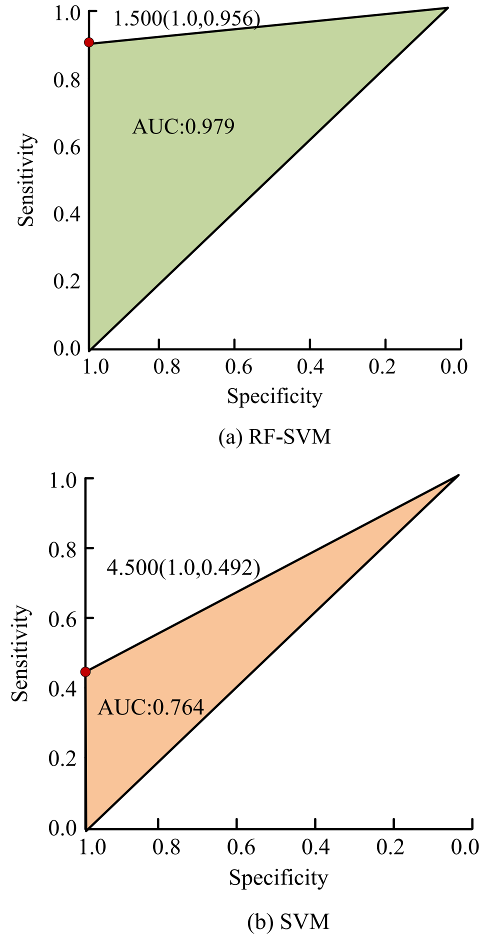


FIGURE 13. Comparison of ROC accuracy verification between two models.

progress [21]. LCCCis a significant factor for land use allocation and regional sustainable development, as well as a crucial assessment method for evaluating land use and development capabilities [22]. Research conducted by LCCC coordinates the relationship between the economy, construction, population, and environment. This provides decision-making references for land protection, planning, and restoration [23]. To create a more scientific and rational evaluation model for LCCC and overcome the limitations of traditional models. An LCCC evaluation model based on RF-SVM is constructed by introducing the RF algorithm to traditional SVM evaluation models. A comprehensive evaluation index system was constructed in this study, followed by visualizing the indicator space to more clearly reflect the characteristics of the data.Finally, an evaluation model for the LCCC evaluation model of RF-SVM is constructed based on the evaluation index system. The experimental results show that the accuracy of the research model is higher, and the judgment of LCCC is more accurate. It overcomes the limitations of its inability to process large-scale and complex data, saves

TABLE 2. Comparison of error rates between different evaluation models.

Performance evaluation index	GA	RVM	PSO	RF-RVM
MRE	6.409%	5.127%	8.219%	2.353%
MSE	3.612%	3.410%	6.618%	1.382%
MAE	14.16%	11.17%	24.31%	7.216%
RMSE	0.1925	0.1828	0.2627	0.1102
R2	0.9932	0.9953	0.9873	0.9998

running time, and is of great significance for the research of LCCC.

VI. CONCLUSION

To lift the accuracy of the model and provide a more comprehensive and objective evaluation of LCCC, this study constructed an LCCCon the basis of RF-SVM. When the decision trees are 100 and the candidate variables are 12, the performance parameters of the model are optimal. From 2014 to 2022, the bearing capacity of the five tested systems showed an upward trend, with System D experiencing the fastest growth rate, with a growth rate of 58.43% and R reaching 46.29%. In the analysis of 2022 pesticide and fertilizer application data in five regions A, B, C, D, and E, the difference in data before and after spatial processing shows that the average error of the five regions is small, only 6.62%. This indicates that the processed values can accurately express SV information. In the comparative experiment of the model, the accuracy of RF-SVM was higher than SVM, and the AUC of RF-SVM reached 0.979. This result verifies the optimization effect of RF on SVM. In the comparison of error rates, the MRE of RF-SVM is 2.353%, MSE is 1.382%, MAE is 7.216%, RMSE is 0.1102, and R2 is 0.9998. Each error evaluation parameter is superior to other models. The data verifies the excellent application value and performance of RF-SVM in evaluating LCCC indicators. Nevertheless, this study still has some limitations. The LCCC system solely comprises 14 indicators that fail to reflect all critical aspects of LCCC. This aspect will be enhanced in future research.

REFERENCES

[1] S. Tong, F. Zhiming, Y. Yanzhao, L. Yumei, and W. Yanjuan, "Research on land resource carrying capacity: Progress and prospects," *J. Resour. Ecol.*, vol. 9, no. 4, pp. 331–340, Jul. 2018, doi: 10.5814/j.issn.1674-764x.2018.04.001.

[2] P. Goswami and S. Nishad, "Quantification of regional and global sustainability based on combined resource criticality of land and water," *Curr. Sci., Fortnightly J. Res.*, vol. 114, no. 2, pp. 355–366, Jun. 2018, doi: 10.18520/cs/v114/i02/355-366.

[3] Y. Wu, X. Sun, B. Dai, P. Yang, and Z. Wang, "A transformer fault diagnosis method based on hybrid improved grey wolf optimization and least squares-support vector machine," *IET Gener., Transmiss. Distrib.*, vol. 16, no. 10, pp. 1950–1963, May 2022, doi: 10.1049/gtd2.12405.

[4] X. Wang, X. Gao, Y. Liu, and Y. Wang, "Stockwell-transform and random-forest based double-terminal fault diagnosis method for offshore wind farm transmission line," *IET Renew. Power Gener.*, vol. 15, no. 11, pp. 2368–2382, Aug. 2021, doi: 10.1049/rpg2.12170.

[5] S. Chen, Y. Yang, F. Deng, Y. Zhang, D. Liu, C. Liu, and Z. Gao, "A high-resolution monitoring approach of canopy urban heat island using a random forest model and multi-platform observations," *Atmos. Meas. Techn.*, vol. 15, no. 3, pp. 735–756, Feb. 2022, doi: 10.5194/amt-15-735-2022.

[6] E. B. Franklin, L. D. Yee, B. Aumont, R. J. Weber, P. Grigas, and A. H. Goldstein, "Ch3MS-RF: A random forest model for chemical characterization and improved quantification of unidentified atmospheric organics detected by chromatography–mass spectrometry techniques," *Atmos. Meas. Techn.*, vol. 15, no. 12, pp. 3779–3803, Jun. 2022, doi: 10.5194/amt-15-3779-2022.

[7] C. Joo, H. Park, J. Lim, H. Cho, and J. Kim, "Development of physical property prediction models for polypropylene composites with optimizing random forest hyperparameters," *Int. J. Intell. Syst.*, vol. 37, no. 6, pp. 3625–3653, Jun. 2022, doi: 10.1002/int.22700.

[8] M. G. Li, J. Liang, C. H. Yan, H. S. Tang, T. L. Zhang, and H. Li, "Rapid quantitative analysis of heavy metals in soil by laser induced breakdown spectroscopy combined with random forest algorithm," *Chin. J. Anal. Chem.*, vol. 49, no. 8, pp. 1410–1418, Oct. 2022, doi: 10.19756/j.issn.0253-3820.211067.

[9] S. Kang, I. Kim, and P. J. Vikesland, "Discriminatory detection of ssDNA by surface-enhanced Raman spectroscopy (SERS) and tree-based support vector machine (Tr-SVM)," *Anal. Chem.*, vol. 93, no. 27, pp. 9319–9328, Jul. 2021, doi: 10.1021/acs.analchem.0c04576.

[10] G. Teng, Q. Wang, X. Cui, K. Wei, W. Xiangli, and G. Chen, "Feature-level fusion of laser-induced breakdown spectroscopy and Raman spectroscopy for improving support vector machine in clinical bacteria identification," *J. Raman Spectrosc.*, vol. 52, no. 4, pp. 805–814, Apr. 2021, doi: 10.1002/jrs.6071.

[11] X. Liu, Y. L. Guan, and Q. Xu, "Support vector machine-based blind equalization for high-order QAM with short data length," *IEEE Signal Process. Lett.*, vol. 28, pp. 259–263, 2021, doi: 10.1109/LSP.2021.3050928.

[12] Y. Wanni, Z. Lin, and L. Qi, "Land resource carrying capacity in Xilin Gol grassland transects: A perspective on food nutritional demand," *IET Renew. Power Gener.*, vol. 10, no. 6, pp. 614–620, Sep. 2019, doi: 10.5814/j.issn.1674-764x.2019.06.006.

[13] H. Zhang, Z. Wang, J. Liu, J. Chai, and C. Wei, "Selection of targeted poverty alleviation policies from the perspective of land resources-environmental carrying capacity," *J. Rural Stud.*, vol. 93, pp. 318–325, Jul. 2022, doi: 10.1016/j.jrurstud.2019.02.011.

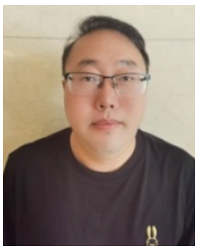
[14] H. Bai, W. Gao, D. Wang, Y. Chen, H. Zhang, Y. Zhao, K. Zhao, Y. Sun, and Z. Sun, "Allocating total emission pollutant control based on water environmental carrying capacity: Model establishment and case study," *Water Policy*, vol. 21, no. 6, pp. 1175–1192, Dec. 2019, doi: 10.2166/wp.2019.031.

[15] Y. Guo, Z. Mustafaoglu, and D. Koundal, "Spam detection using bidirectional transformers and machine learning classifier algorithms," *J. Comput. Cognit. Eng.*, vol. 2022, pp. 5–9, Apr. 2022, doi: 10.47852/bonviewJCCE2202192.

[16] Y. Yang and X. Song, "Research on face intelligent perception technology integrating deep learning under different illumination intensities," *J. Comput. Cognit. Eng.*, vol. 1, pp. 32–36, Jan. 2022, doi: 10.47852/bonviewJCCE19919.

[17] E. Nsugbe, "Toward a self-supervised architecture for semen quality prediction using environmental and lifestyle factors," *Artif. Intell. Appl.*, vol. 10, pp. 35–42, Jan. 2023, doi: 10.47852/bonviewAIA2202303.

[18] J. Zhu, S. Li, Q. Ma, B. He, and J. Song, "Support vector machine-based rapid magnitude estimation using transfer learning for the Sichuan–Yunnan region, China," *Bull. Seismological Soc. Amer.*, vol. 112, no. 2, pp. 894–904, Apr. 2022, doi: 10.1785/0120210232.
[19] S. Ding, Z. Zhang, Y. Sun, and S. Shi, "Multiple birth support vector machine based on dynamic quantum particle swarm optimization algorithm," *Neurocomputing*, vol. 480, pp. 146–156, Apr. 2022, doi: 10.1016/j.neucom.2022.01.012.
[20] J. Zhou, Q. Mao, J. Zhang, N. M. Lau, and J. Chen, "Selection of breast features for young women in northwestern China based on the random forest algorithm," *Textile Res. J.*, vol. 92, nos. 7–8, pp. 957–973, Apr. 2022, doi: 10.1177/00405175211040869.
[21] M. J. Santos, A. B. Smith, and S. C. Dekkerl, "The role of land use and land cover change in climate change vulnerability assessments of biodiversity: A systematic review," *Landscape Ecol.*, vol. 37, no. 1, pp. 367–371, Jun. 2022, doi: 10.1007/s10980-021-01276-w.
[22] J. Dong and Y. Lyu, "Appraisal of urban land ecological security and analysis of influencing factors: A case study of Hefei city, China," *Environ. Sci. Pollut. Res.*, vol. 29, no. 60, pp. 90803–90819, Dec. 2022, doi: 10.1007/s11356-022-22029-5.
[23] C. Fu, Z. Chen, G. Wang, X. Yu, and G. Yu, "A comprehensive framework for evaluating the impact of land use change and management on soil organic carbon stocks in global drylands," *Current Opinion Environ. Sustainability*, vol. 48, pp. 103–109, Feb. 2021, doi: 10.1016/j.cosust.2020.12.005.



SHENMIN WANG was born in Jinan, Shandong, in 1979. He received the bachelor's degree from Qufu Normal University, in 2003, and the master's and Ph.D. degrees in physical geography from Nanjing University, in 2008. In 2008, he joined the Nanjing University of Information Science and Technology, mainly engaged in teaching and research on land use planning, land space planning, land use change, and regional response. He is currently an Associate Professor. In terms of scientific research, he have led one sub project of the Science and Technology Support Program of the Ministry of Science and Technology, two collaborative projects with the China Land Survey and Planning Institute, two land and resources science and technology projects of the Jiangsu Provincial Department of Land and Resources, one land and resources science and technology project of Anhui Province, and more than ten other projects. He has published over 40 academic papers and seven invention patents in core and above journals both domestically and internationally. He received one "Second Prize for Land and Resources Science and Technology Progress in Anhui Province" (2/7). He is the Editor-in-Chief of three academic monographs and the Deputy Editor-in-Chief of one monograph.

QIFANG MA was born in Tianchang, Anhui, China, in 1979. She received the B.S. degree in land use plan from Anhui Agricultural University, in 1997, the M.S. degree in land management from Sichuan Agricultural University, in 2001, and the Ph.D. degree in physical geography from Nanjing University, in 2007. Since 2010, she has been an Associate Professor with the Department of Human Geography, Nanjing University of Information Science and Technology. From 2012 to 2014, she was a Visiting Scholar with the Centre for Sustainable Asian Cities, School of Design and Environment, National University of Singapore. Her research interests include land use planning, land use change and regional response, and natural resource utilization and management.

HANWEI LIANG was born in Inner Mongolia Autonomous Region, China, in 1987. She received the M.S. degree in resources science and technology from Beijing Normal University, China, in 2011, and the Ph.D. degree in civil engineering from Nagoya University, Japan, in 2014. Since 2015, she has been an Associate Professor with the School of Geographical Sciences, Nanjing University of Information Science and Technology. She is the author of two books and more than 30 articles. Her research interests include empirical analysis and policy implications, leveraging remote sensing data and GIS tools to comprehensively analyze and address resource, and environmental challenges. Her objective is to make contributions to the advancement of environmentally friendly and resilient cities. She was a recipient of the 6th Jiangsu Province Productivity Theory and Practice Outstanding Achievement Award, First Prize, in October 2020.