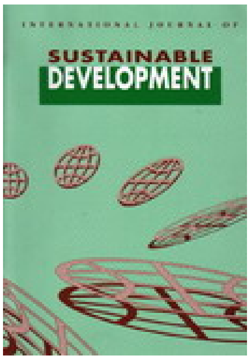


International Journal of Sustainable Development

ISSN online: 1741-5268 - ISSN print: 0960-1406
https://www.inderscience.com/ijsd

Assessment of environmental pollution in rural tourism based on random forest algorithm

Yi Zhao

DOI: 10.1504/IJSD.2024.10062823

Article History:

Received:	10 August 2023
Last revised:	18 October 2023
Accepted:	08 January 2024
Published online:	15 July 2024

Assessment of environmental pollution in rural tourism based on random forest algorithm

Yi Zhao

Business School,
Zhengzhou Railway Vocational and Technical College,
Zheng'zhou, 451460, China
Email: zhaoyihello@163.com

Abstract: In order to overcome the problems of low accuracy, poor recall rate, and low evaluation time in rural tourism environmental pollution assessment, this paper proposes a rural tourism environmental pollution assessment method based on the random forest algorithm. Firstly, web crawler technology is used to collect data on rural tourism environmental pollution, mainly including garbage and waste discharge data, water pollution data, and air pollution data. Secondly, based on the collected pollution data, construct an environmental pollution assessment index system. Finally, the random forest algorithm is used to calculate the feature importance of the indicators, select the feature with the highest accuracy, and use the linear weighting method to calculate the degree of environmental pollution to obtain the evaluation results. The experimental results show that the evaluation accuracy of the method proposed in this paper is the highest at 99.6%, and the pollution level evaluation takes up to 3.5 s.

Keywords: random forest algorithm; rural tourism environment; pollution level assessment; linear weighting method; low accuracy; poor recall rate; web crawler technology.

Reference to this paper should be made as follows: Zhao, Y. (2024) 'Assessment of environmental pollution in rural tourism based on random forest algorithm', *Int. J. Sustainable Development*, Vol. 27, No. 3, pp.311–324.

Biographical notes: Yi Zhao received his Master's degree in Business Administration from Gui'lin University of Technology in 2017. Currently, he is a Lecturer in the Business School of Zhengzhou Railway Vocational and Technical College. His research interests include tourism planning, ecotourism, and rural tourism.

1 Introduction

The rapid development of rural tourism may have negative impacts on the natural environment and ecosystem in rural areas. By assessing the pollution level of rural tourism environment, environmental problems can be identified and solved early, environmental quality can be protected and restored in rural areas, and sustainable development can be achieved. Rural tourism is an important way to promote economic growth and increase the income of rural residents. However, excessive tourism activities may lead to problems such as resource depletion, excessive pressure on scenic spots, and environmental pollution (Feng, 2022). By assessing the degree of pollution, it is possible to scientifically manage and plan the development of rural tourism, promote sustainable development of the tourism industry, and achieve coordinated development of economy, society, and environment. The management of rural tourism needs to rely on scientific data and evaluation results (Huang et al., 2021). By assessing the pollution level of rural tourism environment, scientific basis can be provided for the government and decision-makers, reasonable policies and management measures can be formulated, and the planning and development level of rural tourism can be improved. Rural tourism involves multiple stakeholders, including tourism practitioners, residents, and tourists (Zhao, 2021). Assessing the degree of environmental pollution can enhance public awareness and participation in rural tourism environmental issues, stimulate public awareness of environmental protection, promote public opinion supervision, and promote green development of the tourism industry (Mao, 2022). In summary, the research significance of assessing the degree of environmental pollution in rural tourism lies in protecting the environment, promoting sustainable development of the tourism industry, providing scientific basis for decision-making, and enhancing public participation and public opinion supervision, in order to achieve coordinated development between the tourism industry and the environment.

Li et al. (2022) proposed a method for evaluating the degree of rural tourism environmental pollution based on the pollution footprint model, selecting appropriate evaluation indicators based on the characteristics of rural tourism environmental pollution. Collect data on the number of tourists, number of tourist vehicles, air quality, water quality monitoring, etc. to support the evaluation. Based on the collected data, construct a rural tourism environmental pollution footprint model. This model can consider the direct and indirect impacts of rural tourism activities on the environment and convert them into comparable pollution footprint indicators. Evaluate the pollution level of rural tourism environment using the constructed pollution footprint model. Based on the evaluation results, analyse the main sources and influencing factors of environmental pollution, and propose corresponding improvement suggestions. This method can effectively improve the accuracy of the evaluation, but the evaluation efficiency is poor. Liu Xin et al. proposed a method for evaluating the degree of rural tourism environmental pollution based on the PSR model (Liu, 2022), and selected evaluation indicators suitable for rural tourism environmental assessment based on the PSR model. The PSR model includes three elements: pressure, state, response. Pressure refers to the impact of tourism activities on the environment, state refers to the current status and quality of the environment, and response refers to the management and protection measures of the government and society. Based on the collected data, construct an assessment model for the degree of environmental pollution in rural tourism. The model can quantify the three elements of stress, state, and response according to the framework of the PSR model, and convert them into comparable evaluation indicators through appropriate methods. Using the constructed evaluation model, different indicators are weighted and summarised to calculate a comprehensive pollution level index or rating, and the pollution level of rural tourism environment is evaluated. This method can effectively shorten the evaluation time, but the accuracy of the evaluation is poor. Liu et al. (2023) proposed a risk assessment method for site groundwater environmental pollution based on the game theory comprehensive weighting method. Based on the characteristics of site groundwater environmental pollution risk, evaluation indicators such as soil and groundwater quality, potential pollution sources, land use types, and pollution prevention and control facilities are selected to obtain information about potential pollution sources related to the site. Based on the game theory comprehensive weighting method, build a risk assessment model for groundwater pollution in the site. This model should consider the interaction relationship between different factors, use the concepts and methods of game theory to analyse the decision-making behaviour among different stakeholders, and determine the weights of each factor. By using the constructed evaluation model and calculating the comprehensive weights of various factors, the groundwater environmental pollution risk of the site is quantitatively expressed, and the groundwater environmental pollution risk of the site is evaluated to obtain corresponding evaluation results. This method can effectively improve the accuracy of risk assessment, but the evaluation efficiency is not high.

To address the above issues, this paper proposes a rural tourism environmental pollution assessment method based on the random forest algorithm. The technical route of this method is as follows:

1 Use web crawler technology to collect data on rural tourism environmental pollution, including garbage and waste discharge data, water pollution data, air pollution data, noise pollution, ecological environment damage, and diversity data. Based on the collected pollution data and the selection principle of evaluation indicators, a rural tourism environmental pollution evaluation index system should be constructed, which should comprehensively consider multiple pollution factors.

2 Use the random forest algorithm to analyse and model the collected data, evaluate the feature importance of each indicator, that is, their contribution to the assessment of environmental pollution level.

3 After completing the calculation of indicator importance, select the features with the highest accuracy, namely those that have a decisive impact on accurately assessing the degree of rural tourism environmental pollution. Using linear weighting method to calculate the degree of environmental pollution in rural tourism, the selected features are weighted and summed, and this method is used to evaluate the degree of environmental pollution.

2 Assessment of environmental pollution in rural tourism

2.1 Collection of environmental pollution data for rural tourism

In order to obtain real and objective information on rural tourism environmental pollution, including air quality, water quality, soil pollution, etc., data on rural tourism environmental pollution are obtained through web crawler technology (Chen et al., 2021). The data on environmental pollution in rural tourism mainly include:

1 *Garbage and waste discharge*: covering situations such as garbage accumulation and improper waste disposal, providing data on the quantity, type, and disposal method of garbage.

2 *Water quality and pollution*: including the water quality status of the water source, water pollution sources and their impact range, monitoring parameters such as suspended solids, chemicals, heavy metals, etc.

3 *Air pollution*: It involves air quality, and monitoring indicators include particulate matter (PM2.5, PM10), sulphur dioxide (SO2), nitrogen oxides (NOx), etc.

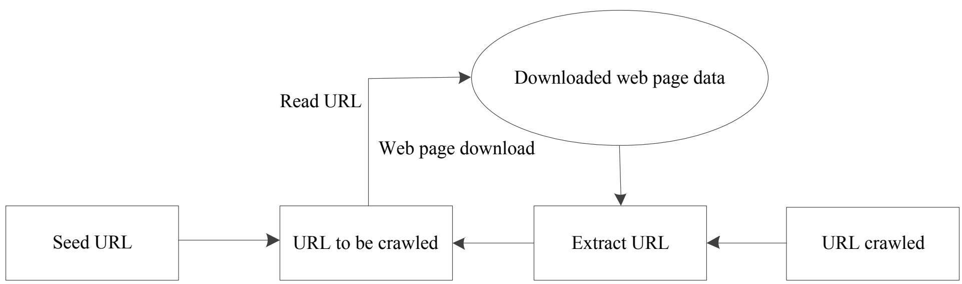
4 *Soil pollution*: including pesticide, fertiliser residue, heavy metal pollution, etc., involving soil quality and environmental health indicators of farmland.

5 *Noise pollution*: refers to the impact of noise, traffic noise, and other factors on local residents and the environment, usually measured in decibels.

6 *Ecological damage and biodiversity*: including data indicators such as deforestation, excessive cultivation, and damage to wildlife, reflecting the impact on local ecosystems.

In the era of big data on the internet, for massive amounts of data, by writing code, simulating the process of obtaining data through a browser, setting access time intervals, and obtaining the required data, the effect of automatic retrieval is achieved (Wang and Song, 2022). The crawler technology is shown in Figure 1.

Figure 1 Crawler workflow diagram



Collect data on rural tourism environmental pollution according to the above process as the raw data for evaluating the degree of rural tourism environmental pollution.

2.2 Selection of pollution level evaluation indicators

Generally speaking, the assessment of environmental pollution in rural tourism refers to the principle of scientificity, systematicity, indirectness, reliability, and comparability.

1 *Scientificity*. Selecting indicators based on scientific analysis cannot be based on subjective judgement. The selected indicators should be related to the degree of environmental pollution and have a rich basis (Zhao, 2021; Chen and Lei, 2021; Kurolap et al., 2021; Nikiforova et al., 2021).

2 *Systemic*. Selecting indicators from various aspects that affect the degree of environmental pollution can comprehensively reflect the degree of environmental pollution in rural tourism.

3 *Simplicity*. Indicators can represent the characteristics of a certain aspect and are interrelated, but do not include them to avoid increasing the difficulty of evaluation.

4 *Operability*. Indicators serve as quantitative analysis services for the degree of environmental pollution in rural tourism. The selected indicators should be operable and able to obtain data and information. Non operable indicators are meaningless.

5 *Comparability*. The indicators are meaningful only when they are horizontally comparable, and the evaluation objects have common influencing factors. Only when they have different states under different conditions can they be compared with each other, thus distinguishing the degree of pollution in rural tourism environment (Tang et al., 2021).

Based on the above choices, this study used methods such as expert consultation, frequency analysis, and comprehensive analysis to select 18 indicators for evaluating the degree of rural tourism environmental pollution, as shown in Table 1.

Table 1 Evaluation indicators

Primary indicators	*Secondary indicators*	*Third level indicators*
Assessment of Environmental Pollution in Rural Tourism	Air pollution	Ozone concentration
		Source distribution of fine particulate matter
		Emissions of toxic and harmful gases
	Water pollution	Heavy metal content
		Microbial pollution in water
		Pesticide and fertiliser usage
	Soil pollution	soil organic matter content
		Soil acidity and alkalinity
		Heavy metal content in soil
	Noise pollution	Noise intensity of surrounding traffic flow
		Mechanical equipment noise level
		Community noise assessment
	Waste disposal	Solid waste treatment rate
		Correct disposal rate of hazardous waste
		Wastewater and exhaust emission control
	Resource utilisation and protection	Renewable energy utilisation ratio
		Rationality of land use structure
		Number of ecological and environmental protection projects

2.3 *Assessment of environmental pollution characteristics in rural tourism based on random forest algorithm*

Random forest is an ensemble learning algorithm that combines the characteristics of decision trees and randomness to solve classification and regression problems. Random forests are composed of multiple decision trees, and by integrating the results of multiple decision trees, they can provide high-precision classification or regression results. Secondly, random forests have the ability to resist noise. By randomly selecting feature subsets to establish a decision tree, the impact of noise on the model is reduced, and the stability of the algorithm is improved. In addition, random forests can automatically process missing data and outliers, reducing the workload of data preprocessing. Most importantly, the random forest algorithm has good scalability for large-scale datasets and is suitable for handling complex problems. Therefore, the random forest algorithm is a powerful and effective tool in evaluating the degree of environmental pollution characteristics in rural tourism (Yang et al., 2022).

Decision tree is a tree structure that classifies or regresses data based on attributes. Taking classification as an example, its purpose is to separate the input data according to certain rules through a series of attribute divisions. The decision tree contains root nodes, internal nodes, and leaf nodes. The root node contains all sample data; The sample data of internal nodes is divided into sub nodes based on classification attributes; Leaf nodes are the classification results of decision trees for sample data.

The key to decision tree growth is to find the optimal partition attribute. In short, it is to find the optimal partition from the attribute set through certain criteria, so that the sub node data belongs to the same category as much as possible, that is, the higher the data purity, the better. Using 'Gini value' to represent the purity of node sample data:

$$Gini(D) = 1 - \sum_{i=1}^{n} p_i^2 \tag{1}$$

In the formula, D represents the dataset, n represents the number of sample data categories, and p_i represents the probability of two randomly selected samples from dataset D with inconsistent category labels. The smaller $Gini(D)$, the higher the purity of dataset D.

The Gini coefficient method is used to select partitioning attributes, which selects the attribute with the smallest Gini coefficient after partitioning as the optimal splitting attribute. If dataset D is divided into sub nodes with discrete attribute a and a has v values of $\{a^1, a^2, ..., a^v\}$, a total of v branches will be generated after division. The vth branch includes all samples with attribute a of D and a^v, denoted as D^v. The Gini coefficient is defined as:

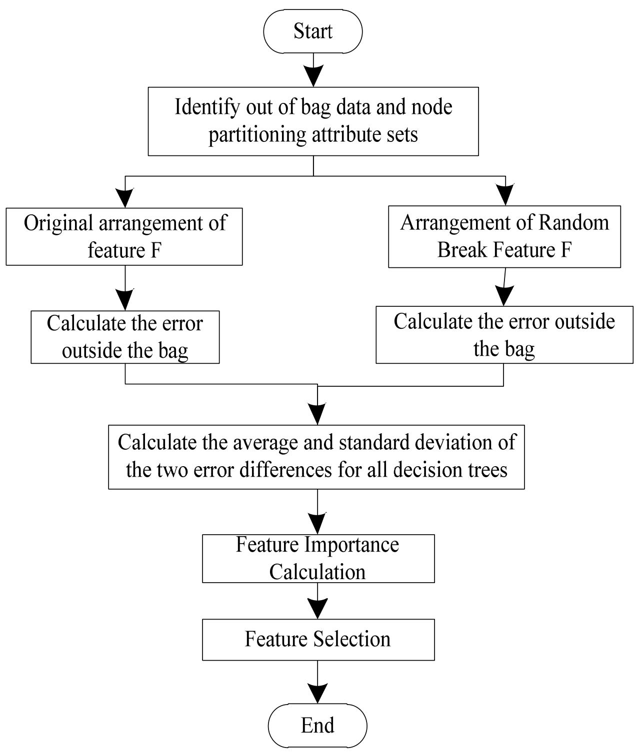
$$Gini\_index(D, a) = \sum_{v=1}^{|v|} \frac{|D^v|}{|D|} Gini(D^v) \tag{2}$$

The classification rules of a single decision tree model are complex and prone to overfitting. In order to overcome the shortcomings of decision trees, the idea of random forests is introduced. The first step is to construct a forest, which consists of independent and unrelated decision numbers as base classifiers. Each decision tree grows using its own input data to form its own results. Secondly, in order to ensure the diversity of the decision tree, randomly selected samples and random attribute selection are used to improve the growth mode of the decision tree, avoiding overfitting and reducing the complexity of decision rules. Finally, the output of the random forest is generated according to certain rules. The random forest algorithm has significantly improved performance compared to decision trees and is more widely used.

In the random forest algorithm, 'indicators' are referred to as 'features', features that can affect the degree of environmental pollution in rural tourism are referred to as 'related features', features that do not affect or have little effect are referred to as 'irrelevant features', and features that have a high correlation with other features are referred to as 'redundant features'. The key role of evaluation features in assessing pollution levels cannot be determined by the quantity, and quantitative analysis needs to be combined with the characteristics of the research area. Therefore, a feature selection process based on random forest algorithm was designed in this study, which is a crucial data preprocessing step. This can first propose irrelevant and redundant features, reducing model dimensions and training time costs. Instead, it can avoid repeated interpretations between variables caused by overfitting, in order to improve evaluation accuracy.

There are multiple methods for calculating the importance of features in the random forest algorithm, and the most commonly used one is based on the idea of 'mean de precision accuracy'. That is, if the data arrangement of a certain feature is randomly disrupted, and the accuracy of the out of pocket data is significantly reduced, it indicates that this feature has a significant impact on the results of sample classification or regression, which means its importance is relatively high. The calculation process is shown in Figure 2.

Figure 2 Calculation process of feature importance



The specific operation is as follows:

1 Calculate the error of the decision tree for each lesson using out of pocket data by ε_1.

2 Randomly shuffle the order of a feature F in the sample, resulting in an out of pocket error of ε_2 in the second calculation. Calculate the difference in error between the two locations:

$$d = \varepsilon_1 - \varepsilon_2 \tag{3}$$

3 Calculate the average value $\bar{d}$ and standard deviation σ of all decision tree differences d

4 Calculate the importance of feature F:

$$\text{importance}(F) = \frac{\bar{d}}{\sigma} \tag{4}$$

If a feature F is highly correlated with the degree of rural tourism environmental pollution and has a high degree of importance, the accuracy of the out of pocket data will be significantly reduced after random arrangement. Therefore, this method can be used to measure the importance of features.

After completing the calculation of feature importance, proceed with feature selection. The so-called feature selection refers to the process of ranking after obtaining the importance of features, and selecting a subset of relevant features that are more important. This paper is based on a recursive elimination strategy for feature selection. In short, it starts from the completed feature set, each time one of the least important features remains until only one feature is left, and finally selects the feature set with the highest accuracy. The specific steps are as follows:

1 Assuming the initial population feature set is $\{F_1, F_2, ..., F_n\}$, this sample set is used as the first round feature candidate, and the importance of each feature is calculated and ranked based on the random forest algorithm. The model is evaluated using out of pocket data.

2 Delete the least important features from the current feature set, and repeat the aforementioned operation for the remaining $n-1$ features as the second round of candidates.

3 Repeat steps (1) and (2) until the last feature.

4 Assuming the error of the feature subset in the *i*th round is the smallest, the optimal feature subset in this set will be output and the remaining features will be deleted.

The normalised result of feature importance $\text{importance}(F)$ is used as the weight of the indicator. This weight is obtained through quantitative analysis and does not involve subjective experience defects. The second is to convert the importance into a relative weight without changing the relationship between them. At the same time, by performing operations similar to normalisation, the weight can be avoided from being too discrete, thus achieving better data convergence. The weight calculation formula for indicator F_j is as follows:

$$w(F_j) = \frac{\text{importance}(F_j)}{\sum_{i=1}^{n} F_i} \tag{5}$$

Calculate the environmental pollution level of area i using linear weighting method r:

$$r(i) = \sum_{i-1}^{n} w(F_j)\, x(F_i) \tag{6}$$

In the formula, $x(F_i)$ is the value at indicator F_i.

3 Experiment

3.1 Experimental data

In order to accurately evaluate the degree of environmental pollution in rural tourism, it is necessary to fully collect environmental data of rural tourism, and collect data related to tourism environmental pollution in a certain rural tourism destination. The collected sample data is shown in Table 2.

Table 2 Sample data

Primary indicators	*Secondary indicators*	*Third level indicators*	*Data*
Assessment of Environmental Pollution in Rural Tourism	Air pollution	Ozone concentration	0.03 ppm
		Source distribution of fine particulate matter	Industrial emissions: 40%, vehicle exhaust: 30%, agricultural emissions: 20%, others: 10%
		Emissions of toxic and harmful gases	5000 tons/year
	Water pollution	Heavy metal content	Lead: 0.05 mg/L, cadmium: 0.02 mg/L, mercury: 0.001 mg/L
		Microbial pollution in water	Total E. coli population: 100 CFU/100 mL, Salmonella: 0 CFU/100 mL, Cryptosporidium: 5/L
		Pesticide and fertiliser usage	Pesticide: 1000 kg/year, fertiliser: 2000 kg/year
	Soil pollution	Soil organic matter content	2%
		Soil acidity and alkalinity	PH value is 6.5
		Heavy metal content in soil	Lead: 10 mg/kg, cadmium: 5 mg/kg, mercury: 0.1 mg/kg
	Noise pollution	Noise intensity of surrounding traffic flow	70 dB(A)
		Mechanical equipment noise level	65 dB(A)
		Community noise assessment	Low noise areas
	Waste disposal	Solid waste treatment rate	80%
		Correct disposal rate of hazardous waste	90%
		Wastewater and exhaust emission control	Compliance with standard emissions

Table 2 Sample data (continued)

Primary indicators	*Secondary indicators*	*Third level indicators*	*Data*
	Resource utilisation and protection	Renewable energy utilisation ratio	30%
		Rationality of land use structure	Farmland: 60%, forest land: 20%, construction land: 20%
		Number of ecological and environmental protection projects	5

3.2 Experimental plan and indicators

In order to fully validate the performance of the method proposed in this paper, the accuracy, time consumption, and recall rate of rural tourism environmental pollution assessment were used as indicators to compare and validate the method proposed in this paper with the methods in Li et al. (2022) and Liu (2022).

1 *Comparison of evaluation accuracy*: Conduct experiments on the method proposed in this paper and the methods in Li et al. (2022) and Liu (2022), calculate the evaluation accuracy separately, and conduct comparative analysis. Real datasets can be used for experiments to evaluate accuracy by measuring differences from the real situation. The results may indicate the superiority of our method in evaluating accuracy.

2 *Comparison of evaluation time*: Conduct time-consuming tests on the methods in this paper, Li et al. (2022), and Liu (2022), and calculate the time required for the evaluation process in the same hardware environment. Through comparative analysis, it can be concluded whether this method has advantages in evaluating efficiency.

3 *Comparison of assessment of recall rate*: For important pollution factors in the assessment of rural tourism environmental pollution degree, whether the method in this paper can better capture and identify them is tested to see whether the method in this paper has a higher recall rate by comparing with the methods in Li et al. (2022) and Liu (2022).

3.3 Experimental result

3.3.1 Accuracy of environmental pollution assessment in rural tourism

In order to verify the effectiveness of the method used in this paper in assessing the pollution level of rural tourism environment, Li et al. (2022) method, Liu (2022) method, and the method used in this paper were used to verify the accuracy of the pollution level assessment. The results are shown in Table 3.

Table 3 Accuracy of environmental pollution assessment in rural tourism

Number of experiments/time	*Accuracy of environmental pollution assessment in rural tourism/%*		
	Li et al. (2022) Method	Liu (2022) Method	Proposed method
1000	68.2	66.1	99.6
2000	63.1	70.6	98.2
3000	65.2	72.9	97.6
4000	68.9	75.9	99.5
5000	70.0	78.2	99.6
6000	69.2	73.6	98.1

According to Table 3, when the number of experiments is 1000, the accuracy of the evaluation of rural tourism environmental pollution level using Li et al. (2022) method is 68.2%, the accuracy of the evaluation of rural tourism environmental pollution level using Liu (2022) method is 66.1%, and the accuracy of the evaluation of rural tourism environmental pollution level using this method is 99.6%; When the number of experiments is 6000, the accuracy of the evaluation of rural tourism environmental pollution level using Li et al. (2022) method is 69.2%, the accuracy of the evaluation of rural tourism environmental pollution level using Liu (2022) method is 73.6%, and the accuracy of the evaluation of rural tourism environmental pollution level using this method is 98.1%. This method can effectively improve the effectiveness of the evaluation of rural tourism environmental pollution level.

3.3.2 Time consumption for assessing the degree of environmental pollution in rural tourism

In order to verify the efficiency of the rural tourism environmental pollution level assessment method proposed in this paper, Li et al. (2022) method, Liu (2022) method, and the method proposed in this paper were used for pollution level assessment. The results are shown in Table 4.

Table 4 Time consumption for evaluating the degree of environmental pollution in rural tourism

Number of experiments/time	*Time consumption for assessing the degree of environmental pollution in rural tourism/s*		
	Li et al. (2022) method	*Liu (2022) method*	*proposed method*
1000	18.2	19.3	0.2
2000	38.9	39.6	0.9
3000	59.5	68.2	1.2
4000	98.1	110.8	1.8
5000	128.0	168.3	2.6
6000	156.3	219.6	3.5

According to Table 4, when the number of experiments is 1000, the evaluation time for rural tourism environmental pollution level using the method in Li et al. (2022) is 18.2s, the evaluation time for rural tourism environmental pollution level using the method in Liu (2022) is 19.3 s, and the evaluation time for rural tourism environmental pollution level using the method in this paper is 0.2 s; When the number of experiments is 6000, the evaluation time of rural tourism environmental pollution level using Li et al. (2022) method is 156.3 s, the evaluation time of rural tourism environmental pollution level using Liu (2022) method is 219.6 s, and the evaluation time of rural tourism environmental pollution level using this method is 3.5 s; The method proposed in this paper can effectively improve the efficiency of evaluating the degree of environmental pollution in rural tourism.

3.3.3 Recall rate for evaluating the degree of environmental pollution in rural tourism

In order to verify the effectiveness of the method used in this paper in assessing the pollution level of rural tourism environment, Li et al. (2022) method, Liu (2022) method, and the method used in this paper were used to verify the recall rate of pollution level assessment. The results are shown in Table 5.

Table 5 Recall rate of rural tourism environmental pollution assessment

Number of experiments/time	*Recall rate for evaluating the degree of environmental pollution in rural tourism/%*		
	Li et al. (2022) Method	*Liu (2022) method*	*Proposed method*
1000	68.3	66.3	98.3
2000	69.1	69.1	99.6
3000	70.6	73.6	99.2
4000	73.8	78.5	98.5
5000	72.9	79.1	99.8
6000	75.2	72.0	99.1

According to Table 5, when the number of experiments is 1000, the recall rate of rural tourism environmental pollution assessment using Li et al. (2022) method is 68.3%, the recall rate of rural tourism environmental pollution assessment using Liu (2022) method is 66.3%, and the recall rate of rural tourism environmental pollution assessment using this method is 98.3%; When the number of experiments is 5000, the recall rate of rural tourism environmental pollution assessment using Li et al. (2022) method is 72.9%, the recall rate of rural tourism environmental pollution assessment using Liu (2022) method is 79.1%, and the recall rate of rural tourism environmental pollution assessment using this method is 99.8%; The above results indicate that the method proposed in this paper can effectively improve the effectiveness of evaluating the degree of environmental pollution in rural tourism.

4 Conclusion

The paper proposes a method for evaluating the degree of environmental pollution in rural tourism based on the random forest algorithm. Collect rural tourism environmental pollution data through web crawler technology, calculate the importance of rural tourism environmental pollution characteristics based on random forests, use XGBoost method to design a rural tourism environmental pollution degree evaluation function, and solve the function to obtain the final evaluation result of rural tourism environmental pollution degree. The experimental results indicate that:

1 When the number of experiments is 6000, the accuracy of the evaluation of rural tourism environmental pollution level using this method is the highest at 99.6%. This method can effectively improve the evaluation effect of rural tourism environmental pollution level.

2 When the number of experiments is 6000, the longest time for evaluating the degree of rural tourism environmental pollution in this method is 3.5 s. This method can effectively improve the efficiency of evaluating the degree of rural tourism environmental pollution.

References

Chen, C. and Lei, X. (2021) 'Heavy metal pollution characteristics, and ecological risk assessment of drinking water sources in Yunmeng Lake based on GIS', *Journal of Subtropical Resources and Environment*, Vol. 16, No. 03, pp.1–8.

Chen, M., Liao, Y., Zheng, X., Luo, H.L. and He, G. (2021) 'Heavy metal pollution assessment of Ganjiang river sediment based on multivariate statistical technology', *Environmental Chemistry*, Vol. 40, No. 12, pp.3861–3874.

Feng, L. (2022) 'Research on the impact of tourism industry agglomeration on environmental pollution', *Environmental Science and Management*, Vol. 47, No. 04, pp.39–44.

Huang, Z., He, Y., Xiao, Y., Tang, K. and Liu, L.F. (2021) 'Research progress on environmental risk assessment based on environmental pollution liability insurance', *China Environmental Management*, Vol. 13, No. 04, pp.57–63.

Kurolap, S.A., Petrosyan, V.S. and Klepikov, O.V. (2021) 'Environmental assessment of the impact of meteorological parameters on man-made pollution by carcinogenicly dangerous chemicals in the Voronezh air basin', *Ecology and Industry of Russia*, Vol. 25, No. 2, pp.60–65.

Li, J., Jiao, W., Min, Q., Mao, K. and Yao, L. (2022) 'Assessment of agricultural pollution in Huzhou city based on pollution footprint model', *Journal of Resources and Ecology*, Vol. 13, No. 01, pp.93–99.

Liu, J., Wang, Y., Tang, J., Chen, Z. and Du, R.W. (2023) 'Risk assessment of site groundwater environmental pollution based on comprehensive weight method of game theory', *Safety and Environmental Engineering*, Vol. 30, No. 1, pp.221–230.

Liu, X. (2022) 'Research on the evaluation index system of rural tourism environmental protection based on the PSR model', *Journal of Social Sciences of Jiamusi University*, Vol. 40, No. 05, pp.57–58+66.

Mao, M. (2022) 'Measures for air pollution control in urban environmental management', *Leather Production and Environmental Protection Technology*, Vol. 3, No. 14, pp.107–109.

Nikiforova, V., Lapina, S. and Vardanian, M. (2021) 'Environmental and health assessment of soil pollution impact on public health', *E3S Web of Conferences*, Vol. 295, No. 12, pp.13–25.

Tang, P., Jiang, Q. and Mi, L. (2021) 'One-vote veto: the threshold effect of environmental pollution in China's economic promotion tournament', *Ecological Economics*, Vol. 185, pp.107069–107076.

Wang, X. and Song, Y. (2022) 'Research on the impact of air pollution on inbound tourism in Chengdu based on gravity model', *Journal of Chengdu University (Social Science Edition)*, Vol. 2022, No. 03, pp.30–39.

Yang, J., Wang, Y., Sun, Q. and Wang, F. (2022) 'Strategic evaluation of ecotourism development of Qilian mountains national park based on SWOT-AHP model', *Qinghai Science and Technology*, Vol. 29, No. 3, pp.74–84.

Zhao, X. (2021) 'Tourism carrying capacity assessment of Zhangjiajie city based on AHP-FCE method', *Inner Mongolia Science and Technology and Economy*, Vol. 19, No. 15, pp.80–81.