Screening of urban environmental vulnerability indicators based on coefficient of variation and anti-image correlation matrix method

Xuefeng Wu [a,b], Xing Huang [a,*]

[a] School of Economics and Management, Southwest University of Science and Technology, Sichuan, Mianyang 621010, China
[b] School of Environment and Resource, Southwest University of Science and Technology, Sichuan, Mianyang 621010, China

A B S T R A C T

A scientific evaluation indicator system is critical for assessing and predicting urban environmental vulnerability. This study establishes an indicator system of urban environmental vulnerability based on ecological sensitivity, resilience, and pressure and proposes a comprehensive screening method using the coefficient of variation and anti-image correlation matrix to improve the evaluation result of urban environmental vulnerability. The indicator screening method improves the reliability of the evaluation results. The proposed method is compared with Pearson correlation analysis to demonstrate the advantages of our method for calculating the information interpretation intensity and the cumulative contribution rate. The calculated results are compared with the threshold values to verify the suitability of the constructed index system. The results show that the information interpretation intensity of the selected indicators system is 1.061, and the cumulative information contribution rate is 89.78%. The selected indicators system accurately reflects urban environmental vulnerability.

1. Introduction

Due to urbanization and economic development in recent years, both population density and industrialization level have increased in cities, causing extensive pollution and threatening urban ecosystems (Shi et al., 2022). The degree of environmental vulnerability reflects the environmental quality and is a critical indicator of regional environmental health (Kang et al., 2018). It is increasingly vital to research environmental vulnerability. A scientific understanding and assessment of urban environmental vulnerability are critical for diagnosing and predicting urban environmental quality and laying the foundation for the scientific management of the urban environment (Huo et al., 2022). However, the selection of the evaluation indicators for urban environmental vulnerability may not be objective.

Most studies have focused on evaluation indicator systems and indicator screening methods. Several institutions have proposed evaluation indicator systems. For example, the National Research Council (1999) proposed a national environmental indicator system (Charles and Sasser, 2002); the H. John Heinz III Center for Science, Economics, and the Environment (2002) developed indicators for evaluating the ecological status of the United States (Tufford, 2003); and the Yale Center for Environmental Law and Policy (2005) put forward

environmental sustainability indicators (Esty et al., 2005). Chinese experts and scholars have developed indicator systems from different dimensions. For example, Dai et al. (2021) established an environmental vulnerability indicator system for Panzhihua City based on ecological sensitivity, resilience, and pressure; Ren and Zhao (2020) evaluated the environmental vulnerability in Guiyang City by assessing the exposure levels, ecological sensitivity, and adaptive capacity; Li et al. (2021) selected indicators of pollutant exposure, ecological sensitivity, and ecological resilience to determine spatial and temporal changes in ecological vulnerability in the Kashgar region; Jiang et al. (2022) constructed an ecological vulnerability indicator system for the Tibet Autonomous Regin, considering ecological pressure, state, and response; Yu et al. (2022) used remote sensing technology to construct remote sensing evaluation indicators for urban ecological livability and analyzed the ecological livability of Wuhan city; Hu and Xu (2018) evaluated the ecological quality of Fuzhou City based on the ecological index of remote sensing; Ru and Ma (2022) proposed an evaluation indicator system that considered natural and anthropogenic factors to assess ecological vulnerability in the Yellow River basin; and Zhang et al. (2021a) assessed the ecological risk in southwest Guangxi-Beibu, taking into account driving pressure state impact and response. The Existing studies mainly consider the ecological environment

vulnerability of the areas under investigation in terms of natural and human factors and the selected evaluation indicators can reflect the level of ecological environment vulnerability in the areas to a certain extent. but they have ignored that urban ecological environment vulnerability is more affected by social and economic factors and have not fully considered systematically selecting the indicators according to specific urban ecological environment vulnerability. Consequently, they are unable to evaluate urban ecological environment vulnerability in an objective manner.

Studies on indicator screening methods for assessing urban ecological vulnerability have primarily focused on the contribution of indicator information, less on the redundancy of indicator information. Methods include the analytic hierarchy process (Mahapatra, 2015; Yang et al., 2022; Hu et al., 2021; Tang et al., 2021), entropy weight method (Zhang et al., 2014), and principal component analysis (Zou and Yoshino, 2017; Guo et al., 2022; Li et al., 2006). For example, Zou et al. (2021) used the analytic hierarchy process to analyze the spatial and temporal evolution of Jilin province's environmental vulnerability; Thirumurthy et al. (2022) used the hierarchy process to analyze multi-criteria coastal environmental vulnerability; Boori and Choudary (2021) investigated the ecological vulnerability of the Samara region Russia using principal component analysis; Li and Fan (2014) studied the ecological vulnerability of cities in the Xijiang Economic Zone in Guangxi using the entropy weight method. However, the indicator information contribution and redundancy have been assessed using other methods in different fields, producing good results. The Delphi method (He and Zhang, 2009), factor analysis (Zhu and Yu, 2021), coefficient of variation (Chen et al., 2018; Ge et al., 2021), and information entropy (Zhang et al., 2015) have been used for determining indicator contribution. For example, Zhou et al. (2016) used factor analysis to select indicators with high information content to evaluate green industries; Sun et al. (2017) used the coefficient of variation to screen indicators of urban vulnerability in Urumqi. Long et al. (2022) used the coefficient of variation to screen indicators of Urban water environment carrying capacity; and Das et al. (2020) used entropy weights in mapping socio-environmental vulnerability to climate change in different altitude zones in the Indian Himalayas. Indicator information redundancy was evaluated with the Pearson correlation coefficient (Yi et al., 2017) and partial correlation coefficient (Meng et al., 2018; Chen, 2021). Huang and Tang (2022) used the Pearson correlation coefficient to screen indicators for information redundancy to evaluate a water safety system in an urban area; and Chen (2019) proposed pathological indicator cycle analysis for information redundancy screening. At present, the research on indicator screening methods mainly considers the screening between two indicators. Although the problem of information redundancy and information contribution between two indicators has been considered, the impact of the indicator system on the evaluation results has not been taken into account, resulting in inadequate rationality of the indicator selection process. The existing screening methods are also defective for over-screening indicators to a certain extent.

Existing research is insufficient. Firstly, most studies on environmental vulnerability focus on indicators related to the natural environment. These indicators are inadequate to assess urban environmental vulnerability. It is difficult to comprehensively and scientifically portray the objective situation of urban environmental vulnerability. Secondly, existing indicator screening methods focus on information redundancy and interpretability of the indicators. Existing studies ignore the degree of contribution and closeness of the indicator system to the evaluation results. In addition, problems with existing studies include information loss and excessive screening of indicators.

This paper proposes an evaluation indicator system for assessing urban environmental vulnerability based on ecological sensitivity, resilience, and pressure (SRP) to ensure that the evaluation indicators are representative. With the goal of optimal information redundancy screening, the coefficient of variation and the anti-image correlation matrix are used to determine the information contribution and

information redundancy of the indicators, respectively. The proposed method and Pearson correlation analysis are compared for the screening results, and the results are compared with the threshold values to verify the constructed indicator system's suitability.

2. Sample selection and data sources

2.1. Sample selection

The preliminary indicator data of 35 representative cities in China from 2001 to 2020 are used to establish the urban environmental vulnerability indicator system. For regions with different geographical distribution characteristics, select representative cities in each region as the research object. Changchun and Shenyang are selected as representative cities in the northeast region. In North China, Beijing, Tianjin, and Taiyuan are selected as representative cities. In East China, select Suzhou, Ningbo, Wuxi, Nanjing, Xuzhou, Hangzhou, Shaoxing, Changzhou, Wenzhou, Hefei, Fuzhou, Quanzhou, Nanchang, Jinan, and Weifang. In central China, Zhengzhou, Wuhan, and Changsha are selected; Select Guangzhou, Foshan, Nanning, Haikou, and Guiyang in South China; Select Chengdu, Chongqing, and Kunming in the southwest region. Select Xi'an, Lanzhou, Xining, and Yinchuan in the northwest region. During selection, factors such as the economic development level, climatic conditions, altitude, terrain and geomorphological characteristics of these cities have been comprehensively considered to ensure that the selected cities are representative. The location of 35 representative cities is shown in Fig. 1.

2.2. Indicator competitive and preliminary selection

The paper will proposal two stages including competitive and primary selection of indicators system. Frist of all, indicator competitive selection, as shown below.

Urban environmental vulnerability refers to the susceptibility of the urban environment to natural and anthropogenic disturbances. It is affected by SRP (Qi et al., 2017). The proposed SRP model is a comprehensive evaluation method to determine the ecological vulnerability of a region. Ecological sensitivity refers to the sensitivity of the ecosystem to external disturbances (Wang et al., 2019). Ecological resilience is the ability of an ecosystem to self-regulate and recover after disturbances (Zhang et al., 2021b), and ecological pressure refers to the effect of disturbances on an ecosystem (Li et al., 2006). These factors and their interactions influence urban environmental vulnerability. The higher the ecological sensitivity and ecological pressure, the greater the vulnerability; and the higher the ecological resilience, the lower the vulnerability. The combined effect of the three factors is analyzed to enable a comprehensive and objective evaluation of urban environmental vulnerability.

The keywords "urban environmental vulnerability", "ecological vulnerability", and "urban vulnerability" are used in this study to search the literature, and 268 representative core articles citing these keywords are downloaded, including studies conducted at different scales (counties, cities, provinces, and urban clusters) and in different landscapes (watersheds, reservoirs, arid areas, mountainous areas, hilly areas, northern agricultural and pastoral areas, plains, oases, loess plateau, grassland areas, and karst areas). With the evaluation indicators of ecological vulnerability used in these studies and proposed by institutions are summarized and indicators irrelevant to ecological vulnerability omitted, 68 indicators are made available after the merging of indicators identical or similar in meaning.

Then, the preliminary selection of indicators is as follows.

26 out of the 68 indicators are selected, considering a scientific basis, representativeness, dynamic evaluation ability, data availability, and comparability. and the number of the indicators is reduced to 20 after experts and scholars in related fields are consulted. The indicators of urban environmental vulnerability are listed in Table 1.
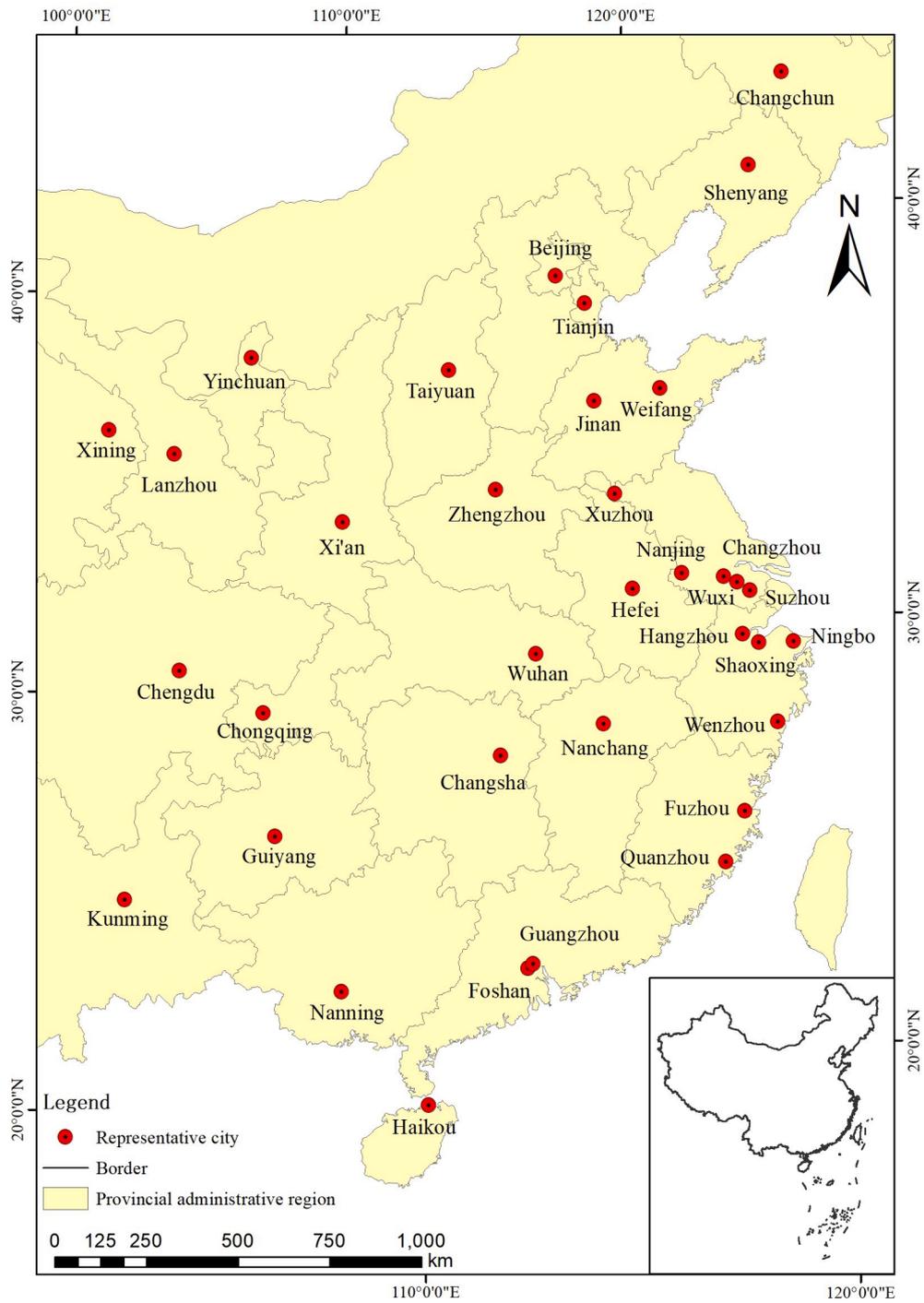
Fig. 1. The location of 35 representative cities.

Table 1

Preliminary indicators of urban environmental vulnerability.

Dimension	Indicator	Type	Calculation method and description	Data sources
Ecological sensitivity	X_1 Dryness	positive	Calculated cumulative temperature and precipitation	EPS-China Regional Economic Database(https://www.epsnet.com.cn)
	X_2 Elevation	positive	Average urban elevation	EPS-China City Database(https://www.epsnet.com.cn)
	X_3 Topographic relief	positive	Maximum elevation per unit area - minimum elevation per unit area	EPS-China City Database(https://www.epsnet.com.cn)
	X_4 Annual precipitation	negative	Calculated monthly average precipitation	Guotai'an Regional Economy Environment Database(https://www.gtafe.com)
	X_5 Average annual temperature	negative	Calculated from monthly average temperature	EPS-China Environment Database(https://www.epsnet.com.cn)
	X_6 Flood period precipitation	negative	Cumulative precipitation from June to September	EPS-China Regional Economic Database(https://www.epsnet.com.cn)

Dimension	Indicator	Type	Calculation method and description	Data sources
Ecological resilience	X_7 Normalized difference vegetation index	negative	IDL software processing results	Resource and Environmental Science Data Center (https://www.resdc.cn/)
	X_8 Net primary productivity of vegetation	negative	Derived from normalized difference vegetation index	Geospatial data cloud(https://www.gscloud.cn)
	X_9 Harmless disposal rate of domestic waste	negative	Ratio of harmless treatment amount of domestic waste to the total amount of generated domestic waste	Geospatial data cloud(https://www.gscloud.cn)
	X_{10} Vegetation cove	negative	Derived from normalized difference vegetation index	TRMM Satellite precipitation data(https://disc.gsfc.nasa)
	X_{11} Biological abundance	negative	Calculated based on land use data	National meteorological science data sharing service platform(https://data.cma.cn)
	X_{12} Proportion of investment in pollution control projects	negative	Ratio of investment in pollution control projects to annual GDP	National Ecological Science Data Center(https://www.nesdc.org.cn)
	X_{13} Sewage treatment rate	negative	Ratio of sewage treatment capacity to total sewage discharge	National Satellite Meteorological Center(https://satellite.nsmc.org.cn)
Ecological pressure	X_{14} Per capita GDP	positive	Ratio of GDP to annual average resident population	National Earth System Science Data Center(https://www.geodata.cn)
	X_{15} Population density	positive	Ratio of urban population to total urban area	EPS-China City Database(https://www.epsnet.com.cn)
	X_{16} Industrial waste water discharge amount	positive	Monitoring statistics	NASA Official website (https://www.nasa.gov)
	X_{17} Volume of industrial sulphur dioxide emission	positive	Monitoring statistics	EPS-China City Database(https://www.epsnet.com.cn)
	X_{18} Industrial smoke and dust emission	positive	Monitoring statistics	Resource and Environmental Science Data Center (https://www.resdc.cn/)

Dimension	Indicator	Type	Calculation method and description	Data sources
Ecological pressure	X_{19} Industrial waste gas emission	positive	Monitoring statistics	EPS-China Environment Database(https://www.epsnet.com.cn)
	X_{20} Highway network density	positive	Ratio of total length of regional roads to total area of the region	EPS-China Urban and Rural Construction Database(https://www.epsnet.com.cn)

2.3. Data sources and processing

The sample data are obtained from the Geospatial Data Cloud, National Satellite Meteorological Center, and other platforms, as summarized in Table .1. This study collected meteorological data, digital elevation data, vegetation data, and socioeconomic data for 35 representative cities in China from 2001 to 2020. The types of data, calculation methods, and data sources are shown in Table 1. The spatial resolution of annual average temperature and precipitation data is 1 km × 1 km, and the spatial resolution of elevation data is 90 m × 90 m. The spatial resolution of vegetation coverage and normalized vegetation index data is 250 m × 250 m. The spatial resolution of vegetation net primary productivity data is 500 m, and the spatial resolution of biological abundance data is 2 km. Topographic relief data is obtained by processing elevation data. Vegetation cover data is obtained by normalized vegetation index processing. The above data were resampled to a spatial resolution of 30 m × 30 m. The time resolution is one year. To ensure data integrity and coherence, missing data are interpolated. The missing annual average temperature data is obtained by interpolation. The annual average precipitation data is obtained by capturing. The elevation data is obtained by stitching and resampling. The topographic relief is first converted and then calculated by raster. The vegetation cover is processed by raster operation. Land type and vegetation net primary productivity data were reprojected, stitched, atmospherically corrected, and radiometrically calibrated. The biological abundance is obtained by raster calculation and assignment projection. The magnitude of digital elevation data, meteorological data, and vegetation data is eventually expressed as an annual average. Industrial wastewater discharge amount, the volume of industrial sulphur dioxide emission, industrial smoke, and dust emission, and Industrial waste gas emission are monitoring statistics. The data mainly comes from the statistical yearbook. Harmless disposal rate of domestic waste, per capita GDP, proportion of investment in pollution control projects, and population density are calculated from the monitoring data. The missing part of the socio-economic data is interpolated by multiple interpolations.

3. Technological process of screening method for urban environmental vulnerability evaluation indicators system

3.1. Total ideas

Indicator screening for information contribution is performed before the screening of information redundancy. This strategy allows for the rapid evaluation of the indicator importance and improved the efficiency of indicator screening, resulting in indicators with low information redundancy and high information contribution. The coefficient of variation is used to perform information contribution screening to select indicators with the higher information contribution. This method eliminates the influence of outliers and can quickly and accurately
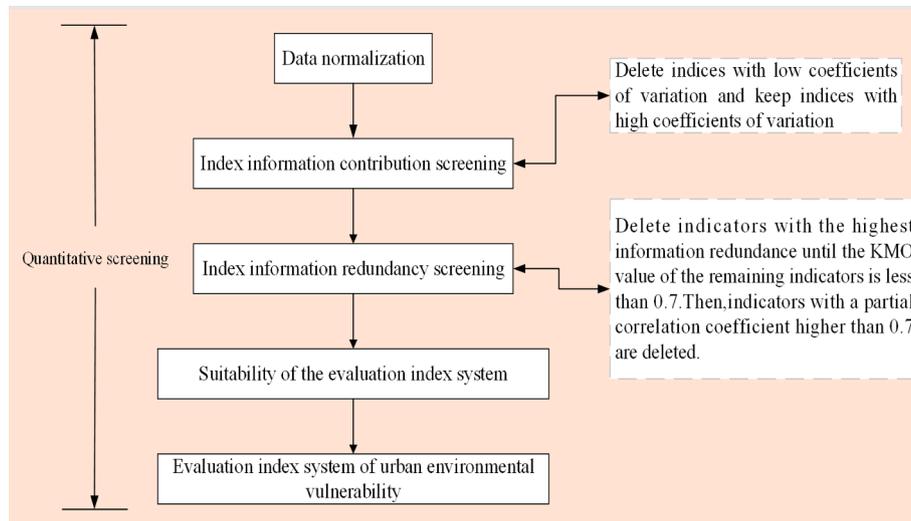
Fig. 2. Flowchart of quantitative indicator screening.

detect indicators contributing important information. The anti-image correlation matrix is adopted to determine indicator information redundancy. This method substantially reduces the level of information redundancy without losing information. Pearson correlation analysis is the most commonly used method to perform information redundancy screening of indicators. The Pearson correlation coefficients of the selected indicators are calculated to compare with the proposed method.

The flowchart of the quantitative indicator screening is shown in Fig. 2.

3.2. Data standardization

Since the urban ecological vulnerability indicators had different scales and units, Extreme Difference Standardization is performed. The method is used because the positive and negative aspects of these indicators are considered. In addition, the values of these indicators are non-negative numbers. The Extreme Difference Standardization is used to reduce these values to 0 to 1 to keep the original data characteristics unchanged.

Positive indicator:

$$X'_{ij} = \frac{X_{ij} - \min_{1 \leqslant i \leqslant n}(X_{ij})}{\max_{1 \leqslant i \leqslant n}(X_{ij}) - \min_{1 \leqslant i \leqslant n}(X_{ij})} \quad (1)$$

Negative indicator:

$$X'_{ij} = \frac{\max_{1 \leqslant i \leqslant n}(X_{ij}) - X_{ij}}{\max_{1 \leqslant i \leqslant n}(X_{ij}) - \min_{1 \leqslant i \leqslant n}(X_{ij})} \quad (2)$$

where j represents the indicator, i represents the sample data, X_{ij} is the original value of the indicator in the sample data, X'_{ij} is the standardized value of X_{ij}; the sample data are $i = 1, 2, …, n$, and the indicators are $j = 1, 2, …, k$.

3.3. Indicator screening based on information contribution

An indicator's coefficient of variation reflects its ability to discriminate information. The higher the coefficient of variation, the greater the discrimination ability of the indicator is, and the more significant the contribution of the indicator is. The threshold value of the coefficient of variation was selected to ensure that the remaining indicator information contribution met the selection requirements. The threshold value of the coefficient of variation for the evaluation indicators was 0.3(Chen

et al., 2018), and only evaluation indicators with coefficients of variation greater than 0.3 were retained.

$$V_j = \frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X'_{ij} - \frac{1}{n}\sum_{i=1}^{n}X'_{ij})^2}}{\frac{1}{n}\sum_{i=1}^{n}X'_{ij}} \quad (3).$$ where V_j is the coefficient of variation of indicator j, X'_{ij} is the standardized value of indicator X_{ij}, and n is the number of samples.

3.4. Indicator screening based on information redundancy

The anti-image correlation matrix was used to determine the level of information redundancy of an evaluation indicator by calculating the measure of sampling adequacy (MSA). The higher the MSA value, the higher the information redundancy of the indicator, indicating that it should be removed. The Kaiser-Meyer-Olkin (KMO) test was used to represent the information redundancy level of the evaluation indicator system. The higher the KMO value, the higher the information redundancy level of the indicator system is. The anti-image correlation matrix method was used to reduce the overall information redundancy before reducing the information redundancy of the local indicator system. When the KMO value of the indicator system was high, we removed the indicator with the highest MSA successively until the KMO value of the indicator system was sufficiently low. The threshold of the KMO values was 0.7(Chen, 2021) to achieve sufficiently low information redundancy. Reducing partial information redundancy was achieved by removing the indicator with a higher MSA value in a pairwise comparison of indicators using the partial correlation coefficient. The threshold value of the partial correlation coefficient was empirically set to 0.7 (Chi et al., 2016).

The KMO value of the indicators was calculated as follows.

$$KMO = \frac{\sum\sum_{j\neq g}r_{jg}^2}{\sum\sum_{j\neq g}r_{jg}^2 + \sum\sum_{j\neq g}p_{jg}^2} \geqslant 0.7 \quad (4)$$ where r_{jg} is the Pearson correlation coefficient of indicators X_j and X_g, and p_{jg} is the k-2 partial correlation coefficient of indicators X_j and Xg. It is the correlation between indicators X_j and Xg after excluding the influence of the remaining k-2 indicators of indicators X_j and Xg. The calculation formula is as follows.

$$p_{jg} = -\frac{c_{jg}}{\sqrt{c_{jj}c_{gg}}} \quad (5)$$ where c_{jg} is the element in the inverse matrix $C = (c_{jg})_{k\times k}$ of the correlation coefficient matrix of the indicator sets $X_1, X_2, …, X_k$.

The MSA_j of indicator j is calculated as follows.

$$MSA_j(k) = \frac{\sum_{j\neq g}r_{jg}^2}{\sum_{j\neq g}r_{jg}^2 + \sum_{j\neq g}p_{jg}^2} \quad (6)$$

The information redundancy screening steps are as follows:
Step 1: Necessity test for information redundancy screening of the

indicator system.

Step 2: The MSA_j values of each indicator were calculated using equation (6).

Step 3: Necessity test for screening the information redundancy of the indicator system consisting of the remaining indicators.

Step 4: Calculated the MSA_j of the remaining indicators according to equation (6).

Step 5: Necessity test for screening the information redundancy of the indicator evaluation system consisting of the remaining indicators.

Step 6: Partial information redundancy indicator screening.

3.5. Suitability of the evaluation indicator system

In order to verify the proposed method of information redundancy screening is optimal, Pearson correlation analysis was conducted to screen the indicators. The results were compared with the proposed method to evaluate its suitability. The results mainly analyze the number of indicators, information interpretation intensity, and cumulative information contribution.

The mean of the variance sum of the final screened indicator data was compared with the mean of the variance sum of the primary indicators. The value reflects the information interpretation intensity of the indicator system; it is calculated as follows.

$$IN' = \frac{\frac{1}{w}\sum_{j=1}^{w}\frac{\sum_{i=1}^{n}(x'_{ij}-\overline{x'_{ij}})^2}{n-1}}{\frac{1}{k}\sum_{j=1}^{k}\frac{\sum_{i=1}^{n}(x'_{ij}-\overline{x'_{ij}})^2}{n-1}} \quad (7).$$ where IN' is the information interpretation

intensity of the indicator system. A value greater than 1 indicates that the information interpretation ability of the screened indicator system is higher than that of the system with the primary indicators. w denotes the number of final screened indicators, and k denotes the number of primary indicators.

The cumulative information contribution rates r_s of the remaining indicators were calculated based on the coefficient of variation of the remaining indicators. It is the ratio of the sum of the coefficients of variation of the remaining s indicators to the sum of the coefficients of

variation of the primary k indicators:

$$r_s = \frac{\sum_{j=1}^{s}\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x'_{ij}-\frac{1}{n}\sum_{i=1}^{n}x'_{ij})^2}}{\frac{1}{n}\sum_{i=1}^{n}x'_{ij}}}{\sum_{j=1}^{k}\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x'_{ij}-\frac{1}{n}\sum_{i=1}^{n}x'_{ij})^2}}{\frac{1}{n}\sum_{i=1}^{n}x'_{ij}}} \quad (8)$$ where r_s represents the cumulative

information contribution rate of the system with the remaining indicators, X'_{ij} represents the standardized value of the indicator X_{ij}, and n represents the number of samples. It is considered in principal component analysis that the principal components contain the majority of information if the information content exceeds 85 % (Chen, 2016). Thus, our indicator system is considered to be appropriate when the cumulative information contribution rate of the screened indicators is greater than 85 %.

4. Screening and application of urban environmental vulnerability evaluation indicators system

4.1. The standardized values of data

Data standardization results of 20 indicators of 35 representative cities from 2001 to 2020 are shown in Fig. 3. Considering the different scales and units of urban ecological environment vulnerability, this study adopts extreme difference standardization to process the collected indicator data. After standardization, the data of 20 indicators of 35 representative cities from 2001 to 2020 have been standardized to 0 to 1.

Fig. 3 shows that the data characteristics of the normalized index data are consistent with those of the untreated data, and the processed data can fully represent the data characteristics before processing.

4.2. Screening results based on information contribution screening

The standardized values of the samples data were substituted into equation (3) to calculate the coefficient of variation V_j for the 20 indicators; the results are listed in Fig. 4. Fig. 4 shows that the coefficients
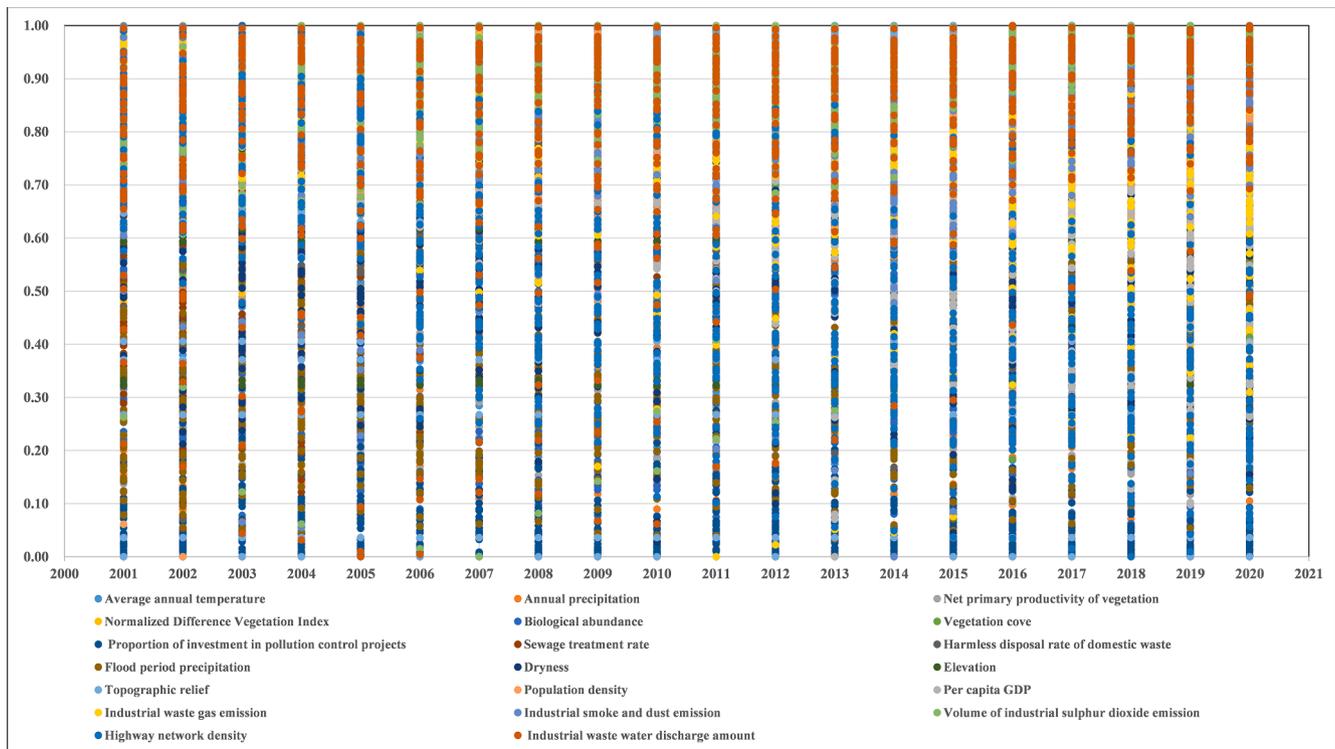

Fig. 3. Data standardization results of 20 indicators of 35 representative cities from 2001 to 2020.
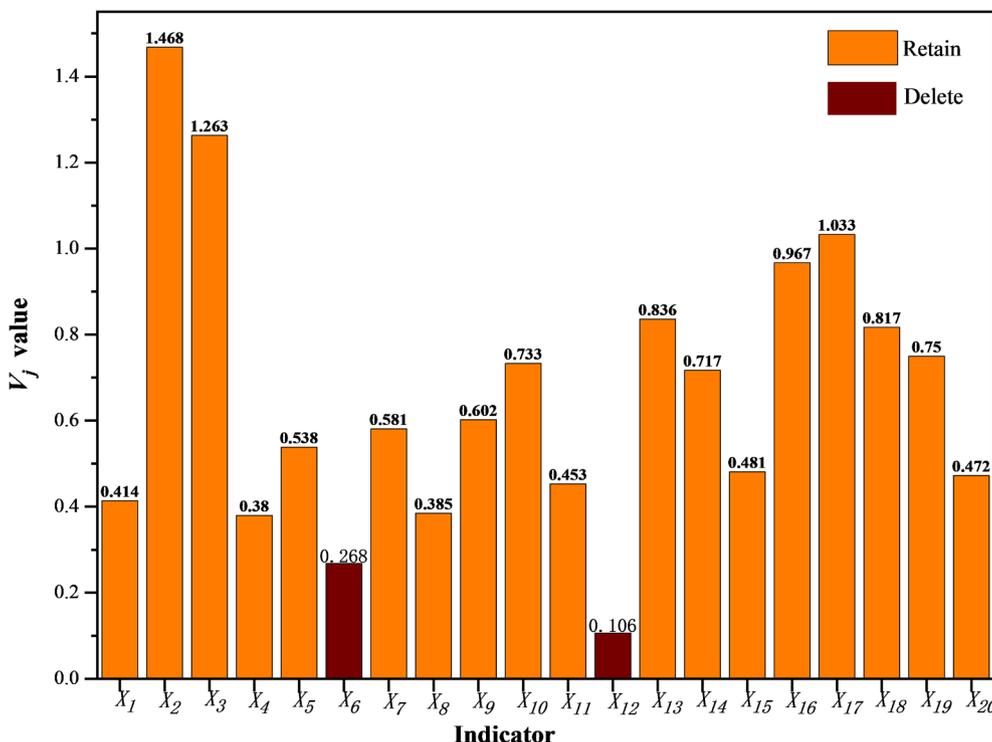
Fig. 4. Results of indicator information contribution screening.

Table 2
KMO Values

KMO		0.713
Bartlett's Test of Sphericity	approximate chi-square	8184.555
	degree of freedom	153
	Statistical significance	0.000

of variation of indicators X_6 and X_{12} are less than 0.3. These two evaluation indicators were deleted, resulting in 18 remaining evaluation indicators.

4.3. Screening results based on information redundancy screening

Necessity test results for information redundancy screening of the indicator system. The KMO value of the indicator system composed of the remaining 18 indicators is 0.713. The KMO values are shown in Table 2. A value greater than 0.7 indicates information redundancy in the system.

The MSA$_j$ values of each indicator are listed in column 3 of Fig. 5. X_9 has the highest MSA value of 0.916; thus, it was deleted.

Necessity test results for screening the information redundancy of the indicator system consisting of the remaining 17 indicators. The KMO of the new indicator system is 0.700. This value indicates information redundancy in the indicator system.

The MSA$_j$ values of the remaining 17 indicators are listed in column 5 of Fig. 5. X_4 has the highest MSA value of 0.766; thus, it was deleted.

Necessity test results for screening the information redundancy of the indicator evaluation system consisting of the remaining 16 indicators. The KMO value of the new indicator system is 0.665. Since this value is less than 0.700, the system meets the information redundancy requirements.

The partial correlation coefficients absolute values of the remaining 16 indicators are less than 0.7, indicating that the system meets the information redundancy requirements.

4.4. Compare to the screening method of urban environmental vulnerability indicators

The results of Pearson correlation analysis showed that the information redundancy levels of the indicators X_5, X_7, X_8, and X_{13} are too high; thus, they need to be removed. The information interpretation intensity of the indicator system derived from the proposed method and the Pearson correlation analysis was 1.061 and 0.983, and the cumulative information contribution rate was 89.78 % and 74.13 %, respectively (Fig. 6).

4.5. Result analysis

The results show that the coefficient of variation of the indicators can reflect the level of information contribution of the indicators. Calculate the coefficient of variation values of individual preliminary indicators. The results show that the values of indicator X_6 and indicator X_{12}, i.e., 0.268 and 0.106, respectively, are both less than the threshold value 0.3. According to the threshold principle, these two indicators should be deleted. Screen the indicators by information redundancy with the Pearson's method and four indicators, namely, annual average temperature, normalized difference vegetation index (NDVI), net primary productivity of vegetation and sewage treatment rate (STR), are deleted. NDVI and WTR should not be deleted, because the former can reflect the vegetation growth status of a city, i.e., the higher the NDVI value, the stronger the ability of water and soil conservation, the stronger the anti-interference ability, the faster the urban ecosystem recovery, and the stronger the ecological recovery ability, while STR can reflect urban wastewater treatment capacity, i.e., the higher the STR, the more stable the urban ecosystem and the stronger the ecological resilience. These two indicators are extremely important to reflect the vulnerability of urban ecological environment and, therefore, should be retained.

The information interpretation intensity and the cumulative information contribution rate derived from the proposed method are higher than that derived from the Pearson correlation analysis. Pearson correlation analysis suffers from over-screening of the indicators. The
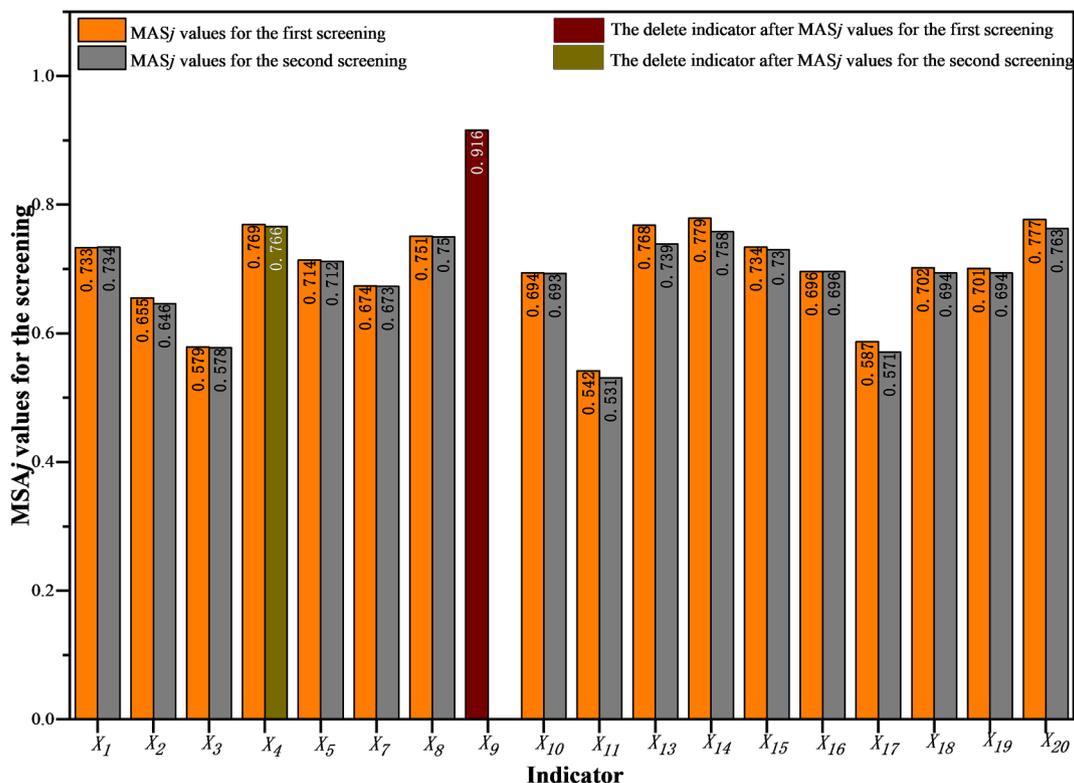
Fig. 5. Screening result of indicator information redundancy.
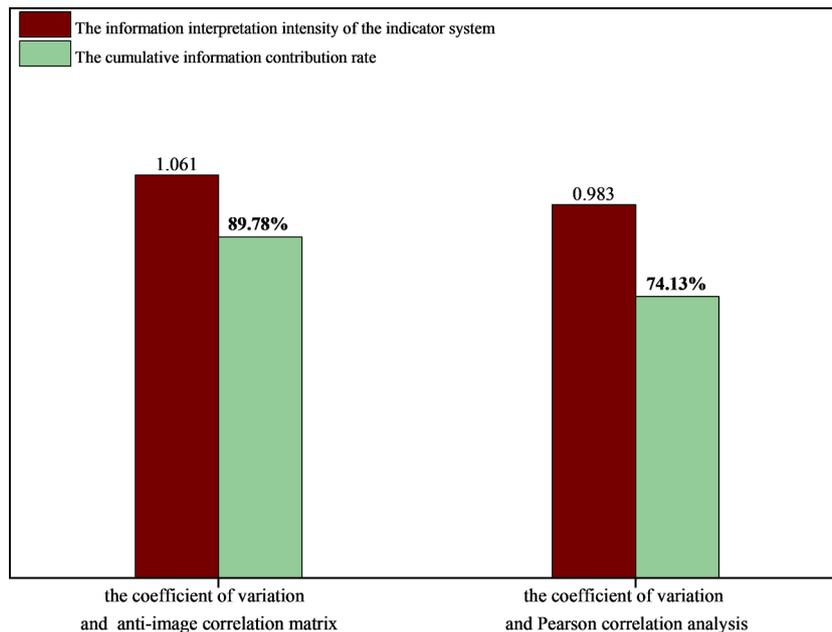

Fig. 6. Result comparison.

information interpretation intensity indicator screened using the proposed method is greater than 1, which means the information interpretation ability of the screened indicator system is more than that of the preliminary indicator system. The cumulative information contribution rates obtained from the proposed method exceed the threshold value of 85 %, indicating that the proposed environmental vulnerability evaluation indicator system is suitability.

5. Conclusions

This study proposed an evaluation indicator system for assessing urban environmental vulnerability. We selected 20 indicators related to ecological sensitivity, resilience, and pressure and used data from 35 cities. The coefficient of variation and anti-image correlation matrix were adopted to determine the information contribution and information redundancy of the indicators, respectively. We compared the information interpretation intensity and cumulative information

contribution rate derived from the proposed method and Pearson correlation analysis to verify the suitability of the urban environmental vulnerability indicator system. The conclusions are as follows:

(1) Sixteen indicators were ultimately selected from the 20 primary indicators. They were reliable indicators to reflect urban environmental vulnerability.

(2) The information interpretation intensity of the screened indicator system was 1.061, and the cumulative information contribution rate was 89.78 %, indicating that this indicator system had higher information interpretation ability, a higher cumulative information contribution rate, and less redundancy than the system derived from Pearson correlation analysis. The proposed screening method prevents information loss and excessive screening and ensures optimal information redundancy of the screened indicators.

CRediT authorship contribution statement

Xuefeng Wu: Data curation, Methodology, Writing – original draft. **Xing Huang:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgement

References

Boori, M., Choudary, K., 2021.Spatiotemporal ecological vulnerability analysis with statistical correlation based on satellite remote sensing in Samara, Russia-ScienceDirect.Journal of Environmental Management285,112138.

Charles E, Sasser. 2002.Ecological Indicators for the Nation. Wetlands22(3),634-635.

Chen, H., 2016. Research on evaluation index screening based on information substitutability. Statistics & Information Forum 31 (10), 17–22.

Chen, H., 2019. Research on evaluation indicator screening based on pathological indicator cycle analysis. China Management Science 27 (1), 184–193.

Chen, H., 2021. Research on evaluation indicator screening method based on Anti-image correlation matrix. China Management Science 1, 1–10.

Chen, R., He,L., Yang,Q.,Pan,A.,Ma,H.,2018.The screening method of urban ecological-social-economic comprehensive analysis indicator system. Journal of Wuhan University of Technology (Information and Management Engineering Edition)40 (2),151-157.

Chi, G., Zhang, Y., Shi, B., 2016. Probit-based debt credit rating model and empirical evidence for small enterprises. Journal of Management Science 19 (6), 136–156.

Dai, X., Gao,Y., He,X.,Liu,T.,Yao,Y.,2021.Spatial-temporal pattern evolution and driving force analysis of ecological environment vulnerability in Panzhihua City. Environmental Science and Pollution Research international28(6),7151-7166.

Esty, D., Levy, M., Srebotnjak, T., 2005. Environmental Sustainability Indicator: Benchmarking National Environmental Stewardship. Yale Center for Environmental Law & Policy Yale University.

Ge, S., Zeng, G., Hu, H., Cao, X., 2021. Evaluation of green innovation capacity and spatial characteristics of the Yangtze River Delta urban agglomeration. Yangtze River Basin Resources and Environment 30 (1), 1–10.

Guo, B., Zang, W., Luo, W., L.,, 2022. Spatial-temporal Shifts of Ecological Vulnerability of Karst Mountain Ecosystem impact of Global Change and Anthropogenic Interference. Sci. Total Environ. 741, 140256.

He, Y., Zhang, Y., 2009. Evaluation of ecological vulnerability in Yunnan Province. Geographical Research and Development 28 (2), 108–111.

Hu, X., Xu, H., 2018.A new remote sensing index for assessing the spatial heterogeneity in urban ecological quality: A case from Fuzhou City China. Ecological Indicators 89, 11–21.

Hu, X., Ma, C., Huang, P., Guo, X., 2021. Ecological vulnerability assessment based on AHP-PSR method and analysis of its single parameter sensitivity and spatial autocorrelation for ecological protection – A case of Weifang City China. Ecological Indicators 125, 107464.

Huang, Y., Tang, D., 2022. Urban water security evaluation based on information substitutability indicator screening from the perspective of variable weights. China Rural Water Conservancy and Hydropower 3, 54–68.

Huo, T., Zhang, X.u., Zhou, Y., Chen, W., 2022. Spatial and temporal variability evaluation and correlation analysis of ecological vulnerability based on exposure-sensitivity-adaptation model–a case study of the Suzhou section of the Grand Canal in China. J. Ecol. 42 (6), 2281–2293.

Jiang, Y., Shi, B., Su, G., Lu, Y., Li, Q., Meng, J., Ding, Y., Song, S., Dai, L., 2022. Spatiotemporal analysis of ecological vulnerability in the Tibet Autonomous Region based on a pressure-state-response-management framework. Ecol. Ind. 130, 108054.

Kang, H., Tao, M., Chang, Y., Zhang, Y., Li, X., Chen, P., 2018.A. Feasible Method for the Division of Ecological Vulnerability and its Driving Forces in Southern Shaanxi. J. Clean. Prod. 205, 619–628.

Li, P., Fan, J., 2014. Regional Ecological Vulnerability Assessment of the Guangxi Xijiang River Economic Belt in Southwest China with VSD Mode. Journal of Resources and Ecology 5 (2), 163–170.

Li, L., Sun, G., Lu, H., Lu, H., Shi, H., 2021. Analysis of spatial and temporal variability and driving forces of ecological vulnerability in Kashgar region. Geography of Arid Regions 44 (1), 277–288.

Li, A., Wang, A., Liang, S., Zhou, W., 2006. Eco-environmental vulnerability evaluation in mountainous region using remote sensing and GIS—A case study in the upper reaches of Minjiang River, China. Ecological Modelling 192 (1), 175–187.

Long, X., Wu, S.G., Wang, J.Y., Wu, P.Z., Wang, Z., 2022. Urban water environment carrying capacity based on VPOSR-coefficient of variation-grey correlation model: A case of Beijing, China. Ecological Indicators 138 (5), 1–10.

Mahapatra, M., 2015. Coastal Vulnerability Assessment Using Analytical Hierarchical Process for South Gujarat Coast, India. Natural Hazards 76 (1), 1–21.

Meng, B., Kuang, H., Luo, J., 2018. Screening model and application of economic and social development evaluation indicators based on significant differences. Scientific Research Management 39 (11), 17–26.

Qi, S., Gong, J., Qian, C., Xie, Y., Zhang, Y., 2017. Ecological and environmental vulnerability assessment of the Bailong River basin in Gansu Province based on SRP model. Soil and Water Conservation Bulletin 37 (1), 224–228.

Ren, H., Zhao, Y., 2020. Evaluation of ecological vulnerability of rapidly urbanizing areas in karst mountains of Guiyang City. Ecological Science 39 (4), 252–258.

Ru, S., Ma, R., 2022. Evaluation, spatial Analysis, and prediction of ecological vulnerability in the Yellow River Basin. J. Nat. Resour. 37 (7), 1722–1734.

Shi, J., Shi, P., Wang. Zi., Wang, Y.,2022. Study on the influence of human disturbances on the dynamic evolution of ecological vulnerability in the Lanxi urban agglomeration. China Environmental Science (11), 1-13.

Sun, L., Wang, Y., Sheng, Y., Si, Q., 2017. Urban vulnerability evaluation of Urumqi based on variance coefficient-composite indicator method. Safety and Environmental Engineering 24 (6), 14–19.

Tang, Q., Wang, J., Jing, Z., 2021. Tempo-spatial Changes of Ecological Vulnerability in Resource-based Urban Based on Genetic Projection Pursuit Model. Ecol. Ind. 121, 107059.

Thirumurthy, S., Jayanthi, M., Samynathan, M., Duraisamy, M., Kabiraj, S., Anbazhahan, N., 2022. Multi-criteria coastal environmental vulnerability assessment using analytic hierarchy process based uncertainty analysis integrated into GIS. J. Environ. Manage. 313 (7), 1–9.

Tufford, D., 2003. The State of the Nation's Ecosystems: Measuring the Lands, Waters and Living Resources of the United States. Landsc. Res. 28 (2), 225–239.

Wang, B., Ding, M., Guan, Q., Ai, J., 2019.A. grid-based ecological vulnerability assessment of Nanchang City. J. Ecol. 39 (15), 5460–5472.

Yang, H., Zhai, G., Zhang, Y., 2022. Ecological vulnerability Assessment and Spatial Pattern Optimization of Resource-based Cities: A case Study of Huaibei City, China. Human and Ecological Assessment 11, 1–20.

Yi, P., Li, X., Zhou, Y., Li, W., 2017. Screening model of ecological city evaluation indicator and its application. Journal of Northeastern University (Natural Science Edition) 38 (8), 1211–1216.

Yu, J., Li, X., G, X., Shen, H.,2022. A remote sensing assessment index for urban ecological livability and its application. Geo-spatial Information Science (7),1-22.

Zhang, Z., Hu, B., Qiu, H., 2021a. Comprehensive assessment of ecological risk in southwest Guangxi-Beibu bay based on DPSIR model and OWA-GIS. Ecol. Ind. 132, 108334.

Zhang, Z., Hu, B., Qiu, H., Deng, Y., 2021b. Ecological and environmental vulnerability assessment of southwest of Guilin Beibu Gulf based on mountain-river-sea perspective and SRP model. Earth and Environment 49 (3), 297–306.

Zhang, X., Wang, C., Li, E., Xu, C., 2014. Assessment Model of Ecoenvironmental Vulnerability Based on Improved Entropy Weight Method. Scientific World Journal 5, 797814.

Zhang, H., Zhang, J., Hang, J., Wang, J., Wang, N., 2015.An. indicator system establishment method based on dynamic cyclic screening model. Firepower and Command and Control 40 (4), 41–44.

Zhou, Y., Wang, H., Chi, G., 2016. Model construction and empirical evidence of green industry evaluation indicator system based on factor analysis. Journal of System Management 25 (2), 338–352.

Zhu,J., Yu,L.,2021.Coupled synergistic measurement of science and technology innovation and financial development in China– screening based on VIF-variance coefficient. Journal of Shanghai University (Natural Science Edition)27(4),785-794.

Zou, T., Yoshino, K., 2017. Environmental Vulnerability Evaluation Using a Spatial Principal Components Approach in the Daxing'anling Region, China. Ecological Indicators 78 (5), 405–415.

Zou, T., Chang, Y., Chen, P., Liu, J., 2021. Spatial-temporal variations of ecological vulnerability in Jilin Province (China), 2000 to 2018. Ecol. Ind. 133, 108429.