Environmental Science and Ecotechnology 22 (2024) 100479

Contents lists available at ScienceDirect

Environmental Science and Ecotechnology

journal homepage: www.journals.elsevier.com/environmental-science-and-ecotechnology/

Original Research

Trusted artificial intelligence for environmental assessments: An explainable high-precision model with multi-source big data

Haoli Xu [a, b, c], Xing Yang [a, c, *], Yihua Hu [a, c], Daqing Wang [d, **], Zhenyu Liang [a, c], Hua Mu [a, c], Yangyang Wang [a, c], Liang Shi [b], Haoqi Gao [a, c], Daoqing Song [e], Zijian Cheng [d], Zhao Lu [d], Xiaoning Zhao [d], Jun Lu [a, c], Bingwen Wang [a, c], Zhiyang Hu [f]

[a] State Key Laboratory of Pulsed Power Laser, College of Electronic Engineering, National University of Defense Technology, Hefei, 230037, China
[b] Jianghuai Advance Technology Center, Hefei, 230000, China
[c] Key Laboratory of Electronic Restriction of Anhui Province, Hefei, 230037, China
[d] Defense Engineering College, Army Engineering University of PLA, Nanjing, 210007, China
[e] International Studies College, National University of Defense Technology, Nanjing, 210000, China
[f] School of Electrical Engineering and Automation, Hefei University of Technology, Hefei, 230009, China

ARTICLE INFO

Article history:
Received 27 December 2023
Received in revised form
19 August 2024
Accepted 22 August 2024

Keywords:
Intelligent environmental assessment
Transformer
Multi-source data
Explainable AI

ABSTRACT

Environmental assessments are critical for ensuring the sustainable development of human civilization. The integration of artificial intelligence (AI) in these assessments has shown great promise, yet the "black box" nature of AI models often undermines trust due to the lack of transparency in their decision-making processes, even when these models demonstrate high accuracy. To address this challenge, we evaluated the performance of a transformer model against other AI approaches, utilizing extensive multivariate and spatiotemporal environmental datasets encompassing both natural and anthropogenic indicators. We further explored the application of saliency maps as a novel explainability tool in multi-source AI-driven environmental assessments, enabling the identification of individual indicators' contributions to the model's predictions. We find that the transformer model outperforms others, achieving an accuracy of about 98% and an area under the receiver operating characteristic curve (AUC) of 0.891. Regionally, the environmental assessment values are predominantly classified as level II or III in the central and southwestern study areas, level IV in the northern region, and level V in the western region. Through explainability analysis, we identify that water hardness, total dissolved solids, and arsenic concentrations are the most influential indicators in the model. Our AI-driven environmental assessment model is accurate and explainable, offering actionable insights for targeted environmental management. Furthermore, this study advances the application of AI in environmental science by presenting a robust, explainable model that bridges the gap between machine learning and environmental governance, enhancing both understanding and trust in AI-assisted environmental assessments.

1. Introduction

Rapid urbanisation and industrialisation threaten environmental and human health and have led to widespread environmental pollution [1]. Environmental assessments are an important tool for supporting sustainable development, and researchers are increasingly utilizing big data and artificial intelligence (AI) to conduct these assessments [2–6]. Although using AI to conduct environmental assessments is convenient and efficient, there are inherent risks in relying on AI models, such as transparency, interpretability, and trust. In environmental assessment, AI shows potential for providing high-efficiency, high-precision assessments; however, steps must be taken to ensure AI models' reliability, credibility, and explainability.

Excessive industrial emissions lead to the diffusion of pollutants

* Corresponding author. State Key Laboratory of Pulsed Power Laser, College of Electronic Engineering, National University of Defense Technology, Hefei, 230037, China.
** Corresponding author.
E-mail addresses: yangxing17@nudt.edu.cn (X. Yang), wangdaqing2020@126.com (D. Wang).

https://doi.org/10.1016/j.ese.2024.100479

into underground aquifers, causing widespread deterioration of the environment [7]. To more accurately and efficiently assess the environmental pollution and human health problems caused by industrial emissions, many scholars have attempted to analyse the groundwater flow, scope of pollution, and toxicology of pollutants in the study area by conducting hydrogeological surveys at local scales [8,9]. However, due to the complexity of environmental pollution, assessment models that can process big data are required. The evaluation of environmental pollution risk was developed based on the study of groundwater environmental vulnerability [10]. The development process can be divided into three stages: (1) evaluating groundwater pollution risk based on groundwater vulnerability [11,12], (2) evaluating groundwater pollution risk based on the superposition of groundwater vulnerability and pollution load [13,14], and (3) evaluating groundwater pollution risk based on the superposition of groundwater vulnerability, pollution load, and groundwater value [15,16]. The DRASTIC model is the most widely used coverage index method for evaluating the risk of environmental pollution [17–20]. However, since the index score and distribution weight are easily affected by subjective factors, researchers have continuously improved this model [21,22]. Kazakis developed two new methods, DRASTIC-PA and DRASTIC-PAN, by incorporating land use types into the original model [23]. When evaluating groundwater vulnerability in Jilin, China, Huan et al. replaced the two indices of aquifer medium and topographic slope with aquifer thickness and groundwater flow rate and added the evaluation index of land use type, thus constructing the DRSIHVL model [24]. In evaluating the environmental pollution risk in the northern suburbs of Yinchuan, Wu et al. proposed the DRTILSQ model, which improved upon the DRASTIC model in three aspects: groundwater inherent vulnerability, pollution load, and groundwater value [25]. The improved DRASTIC model has been widely used in many countries, including the United States [26,27], Canada [28,29], Europe [30,31], India [32,33], and China [34,35], to assess the pollution risks in diving or confined water regions. However, there is an ongoing need for improved environmental assessment models worldwide.

Many scholars have attempted to use a variety of indicators and algorithms to conduct a comprehensive, efficient, and high-precision environmental assessment. Many researchers have studied the impact of environmental pollution on human health through health risk assessments [36,37]. Human health risk assessments describe the potential effects of exposure to harmful substances. The assessment process determines current or future human health risks by evaluating the degree of human exposure to toxic or harmful substances. For example, the long-term use of nitrogen fertilisers in agriculture leads to excessive nitrate content in groundwater, and continuous consumption of nitrates can lead to conditions such as decreased blood oxygen-carrying capacity, blue baby syndrome, multiple sclerosis, and gastric cancer [38,39]. In addition, industrial waste discharge can lead to excessive fluoride or radioactive substances in groundwater. Long-term groundwater use with elevated fluoride levels can lead to skeletal and dental fluorosis [40]. Similarly, the presence of radioactive materials such as uranium in drinking water may have adverse effects on the kidneys and bones [41,42]. In the 21st century, human health risk assessments have attracted the attention of scholars worldwide, and many studies have been conducted in specific research areas on the degree of risk to human health caused by pollutants in water bodies. Li et al. conducted in-depth research on the uncertainty in health risks and analysed the uncertainty using the Monte Carlo method combined with random and fuzzy theories [43]. Muhammad et al. measured the arsenic (As) content in the water bodies of the Kohistan region and found that the intake of As through drinking water had a low non-carcinogenic risk and a moderate carcinogenic risk [44]. Siddique et al. calculated the health risk of trihalomethanes in drinking water in the Karachi urban industrial area of Pakistan and found that the chloroform content in some areas was unacceptable [45]. Considering the uncertainty of the health risk model parameters, Yang et al. analysed the health risk value of shallow groundwater in the Yinchuan Plain using fuzzy theory and calculated the health risks to adults and children [46]. In many cases, environmental and human health risk assessments focus on the pollutant content in groundwater. To improve the accuracy of assessment results, many scholars have adopted machine learning methods, such as the logistic regression model [27], artificial neural network (ANN) model, support vector machine (SVM) model [47,48], random forest (RF)-assisted ANN method, SVM-RF method, SVM-ANN method [49], and the use of large, multivariate, geospatial-temporal datasets [50]. Jaydhar envisioned that deep learning methods will contribute to the prediction and vulnerability analysis of groundwater parameters combined with hydrogeochemistry and the environment [51]. Although big data and AI-based environmental assessments are increasing in popularity [2–6], the assessment accuracy and efficiency of existing environmental assessment algorithms require further improvement. Moreover, few studies have been conducted on convolutional neural network (CNN) environmental big data models, and studies on transformer environmental big data models have not yet been published. In the field of intelligent algorithms, the transformer model has gained popularity and has the advantages of high precision and good interpretability [52–54]. For example, many popular underlying models of ChatGPT are transformer models [55]. Hence, this study aimed to use deep learning models and multi-source big data to achieve a comprehensive, efficient, and high-precision environmental assessment.

In addition, using AI models presents certain safety risks, especially in research fields where the outcome may be life-threatening, such as medical treatment, human health, and environmental ecology. Therefore, AI should not be used in these fields without taking steps to mitigate the risks. Additionally, the level of explainability and credibility of the AI model must be considered. Although deep learning models are generally regarded as important tools, the inherent lack of transparency and interpretability associated with black-box technology can foster a lack of trust in AI models. Therefore, measures should be implemented to improve the trustworthiness of model outcomes, such as through counterfactual interpretation. This is particularly important for the current study's application environment. In this manner, one can also become familiar with the unknown process by understanding the assumed input conditions associated with the change in results [56]. Therefore, this problem requires attention in the field of environmental assessments [56]. Although AI environmental assessment models are highly efficient and accurate, the level of trust in their outcomes is low due to the black box phenomenon. Ongoing research is required to improve the explainability and credibility of deep learning models. Model explainability remains an open issue in studying the operational mechanisms of deep learning models from multiple perspectives. Generally, research on this topic is divided into three dimensions: active or ex-post explainability, local or global explainability, and explained objects [57]. Among them, one classic local method is based on the explainability of the input, to attribute the prediction results of the model to different input features. This method masks or disturbs the original input's features to measure each feature's influence on the output and the contribution of each feature to the output. The basic principle of the perturbation-based explainability method is to evaluate the input features one at a time by constructing many perturbation samples and ranking the importance of each input feature. Therefore, another goal of this study was to examine the

explainability of a multi-source AI environmental assessment model.

In summary, environmental assessments require multi-source index data. There is still room for improvement in the accuracy of environmental assessment models, and the results of large models, such as the transformer environmental model, have not yet been reported. Applying the "AI + Environment" model also requires consideration of its reliability and explainability. These are urgent problems that must be addressed. Trusted and high-precision AI serving multi-source environmental assessment is based on deep learning Transformer model and multi-source environmental big data (Fig. 1). In this study, we utilised multivariate indicators, long time-series observation data, and other spatiotemporal big data [58]. Regarding big data high-precision model methods, we focused on transformer and CNN models and developed a new environmental assessment model based on multi-source big data. We analysed the efficiency and explainability of this deep learning model for multi-source data environmental assessment to ensure that the model was reliable and trustworthy. This explainable AI environmental assessment model can be used in "machine teaching" to guide humans to govern the environment based on clear and focused indicators.

2. Methods and materials

2.1. Explainable deep learning method of multi-source environmental risk assessment

This study primarily utilised the popular transformer model with an encoder–decoder structure. The encoder is stacked using an input embedding layer and multiple encoder layers. Each encoder layer contains a multi-head attention layer and a feed-forward layer, forming a residual connection between the two layers. The multi-head self-attention and feed-forward neural network layers were constructed by layer normalisation. The decoder was stacked using multiple decoder layers. In contrast to the encoder layer, a cross-attention layer was added to each decoder layer to fuse information from the encoder. In 2017, the transformer model was developed based on a mathematical technique known as self-attention [52]. Subsequently, based on the transformer architecture, the BERT model was ranked first among the major natural language processing models and continued to approach or even exceed human performance. The Vision Transformer, proposed by Google Brain researchers, applied the transformer model to the computer vision field with minimal changes, showing strong feature-learning capabilities [54]. Fig. 1 shows the network framework for deep learning (transformers).

In this study, we focused on introducing self-attention and multi-head attention mechanisms. The principle of the self-attention mechanism is that three different sequence vectors of images are obtained through a linear transformation: query (Q), key (K), and value (V). These three vectors are obtained through input vector transformation. The attention layer aggregates Q, K, and V, updates the output vector and uniformly expresses the entire process as follows:

$$\mathrm{Attention}(Q,K,V)=\mathrm{Softmax}\left(\frac{QK^{\mathrm{T}}}{\sqrt{d_k}}\right)V \tag{1}$$

The attention weight is generated by the dot-product operation between Q and K. To stabilise the gradient, the scaling factor $\sqrt{d_k}$ is used, where d_k is the dimension of Q and K. Then we used the Softmax function to normalise the attention weight.

Multi-head attention is a mechanism that can be used to improve the performance of the general self-attention layer. Different attention heads use different Q, K, and V matrices. Owing to the random initialisation, these matrices can project the trained input vector into different representation subspaces and be processed in parallel using multiple independent attention heads (layers). The resulting vector is aggregated and mapped to the final output. The multi-head attention calculation process is as follows:

$$\mathrm{MultiHead}(Q,K,V)=\mathrm{Concat}(Z_1,Z_2,\ldots,Z_i)W^{\mathrm{O}} \tag{2}$$

where Z_i represents the output matrix of each head, and W^{o} represents the output projection matrix.

Similarly, although we accept the validity of human decision-making despite our limited understanding of the brain's workings, we can extend comparable credibility to AI models. These models are based on neural networks and algorithms that emulate the structure and processes of the human brain. Therefore, although the deep-learning model of AI is considered an end-to-end black box, an in-depth study of its explainability and results can enhance its credibility. This study reverses the importance of model index characteristics through data-driven deep-learning model results. It combines the original attributes of indicators to study the explainability of the deep-learning model. The input is disturbed by one or a set of input feature deviations to obtain the difference between the output after the disturbance and the original output. The degree of influence of different inputs on the result is obtained to determine the importance of the index feature. The perturbation-based method can directly estimate the importance of features, and it is simple and universal. Examples of the applications of this method include occlusion experiments [59] and saliency-mapping methods [60,61]. However, the input data in the multi-source environmental assessment in this study were slightly different from ordinary image data, as we utilised multi-dimensional and multi-channel image data. The explainability of the deep learning model was consistent, and it was still necessary to determine the influence of each index feature on the entire model. The environmental risk assessment indicators are as follows:

$$\{Y_1,Y_2,Y_3,\ldots,Y_n\} \tag{3}$$

In this study, by removing or increasing the disturbance ΔY of index feature y, re-input it into the trained model, the accuracy of the environmental risk assessment results, that is, the increased total error Δe, is calculated. Here, Y is an index layer, where the pixels of the specific multi-dimensional data contain [X_1, X_2, X_3, …, X_n]. The multi-dimensional pixels correspond to different levels of environmental risk.

Similar to the calculation methods used with a saliency map [60,61], the change in each index is set to ΔY, and the error change of each index change to the overall environmental assessment result is Δe. The importance of the index characteristics can be expressed as:

$$s=\frac{\Delta Y}{\Delta e}=\frac{\partial Y}{\partial e} \tag{4}$$

To determine the importance of each index in the AI model, we can calculate its S value and match it with the meaning of the index feature. This provides an explanation of the AI environmental risk assessment and increases trust in the model.

In addition, we conducted the study in the Guanzhong Plain (107°40′–109°49′ E, 33°39′–34°45′ N), within the central region of the Yellow River Basin in China, north of the Weihe River, covering a total area of 10,180 km^2 (Fig. 2). The pollution originated from sewage discharge by the nearby industrial park (Fig. 2). During the

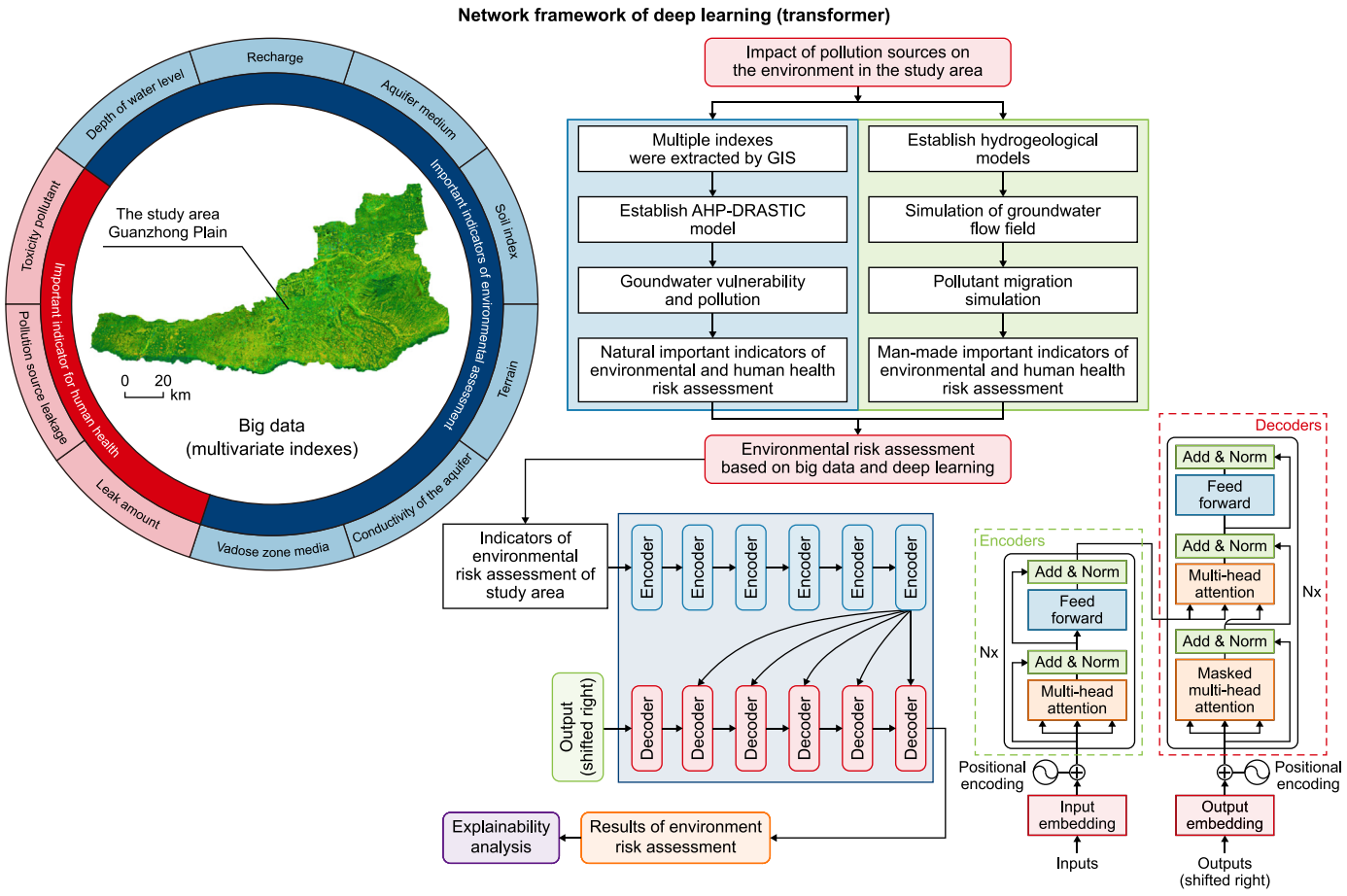

Fig. 1. The technical flowchart of explainable deep learning (transformer) and big data for environmental risk assessment.

investigation, pollution indicators in the water samples exceeded regulatory standards, with each of the four factories discharging different substances above these thresholds. Table 1 lists the specific substances that exceeded the standards. Additional data can be found in the Supplementary Material.

The overall terrain pattern of the study area is southeast high, northwest low, and dustpan-shaped. The terrain is steeply inclined to the north, with an average altitude of more than 2000 m. The strata are mainly igneous rocks, followed by metamorphic rocks, and many turbulent rivers, which are significant to the water resources conservation of Xi'an [62]. The study area is characterised by a warm temperate, semi-arid, semi-humid continental monsoon climate. Quaternary loose sediments are widely distributed in the study area, providing an excellent location for groundwater storage. The aquifer within 300 m is Xi'an's primary source for urban water extraction. The aquifer group can be divided into phreatic and confined water according to the differences in landform, geological conditions, and hydrodynamic characteristics. The phreatic aquifer consists of Quaternary Holocene and Middle Pleistocene, mainly composed of sand, sandy gravel with clay, sandy soil, and loess,

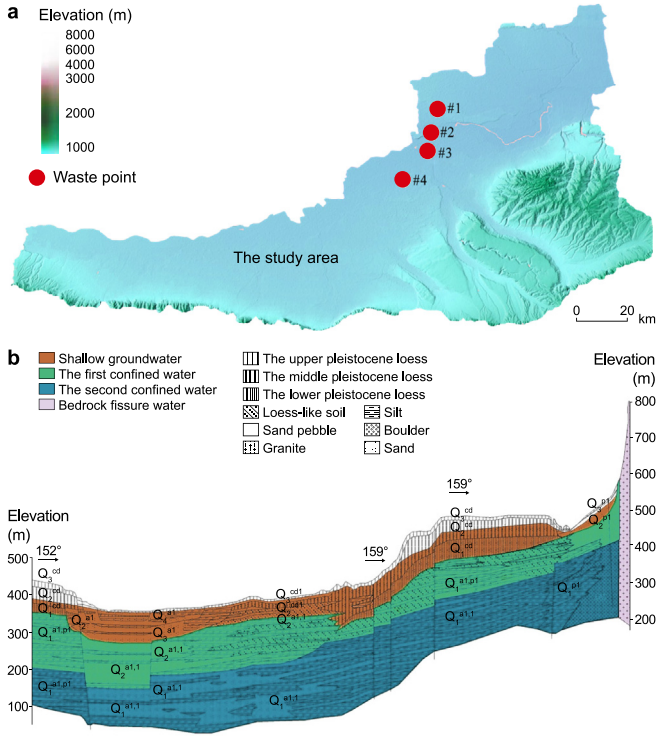

Fig. 2. The study area (**a**) and its hydrogeological section map (**b**).

Table 1
Excessive pollution index in industrial parks.

Pollution sources (unit)	Source location			
	#1	#2	#3	#4
Total hardness of water (THW, mg L^{-1})	506.0	599.0	0.0	0.0
Total dissolved solids (TDS, mg L^{-1})	1310.0	1441.0	0.0	0.0
As (mg L^{-1})	0.0	0.0	3.9	0.0
NH_3-N (mg L^{-1})	0.0	0.0	0.0	16.4

with thickness ranging from 10 to 70 m. This aquifer serves as the main mining layer of irrigation and suburban water supply, with mining depths generally ranging from 5 to 40 m. The confined aquifer, comprising the Middle and Lower Pleistocene of the Quaternary system, features interbedded layers with different thicknesses of sand, sand gravel, clay, and sand. This aquifer has a thickness ranging from 50 to 180 m and serves as the main mining layer in suburban self-provided well areas, with mining depths generally ranging from 40 to 300 m. The hydrogeological survey of the study area is as follows.

(1) Alluvial gravel pore water-bearing rock group

This comprises fast-circulating alluvial or diluvial coarse sand and gravel pore water-bearing rock groups. It has sufficient recharge sources and abundant water. The aquifer thickness is 20–80 m. The sediment layer is thick near the river and thin farther away. The water level is shallow, and the single well water output generally exceeds 2000 m^3 d^{-1}, distributed in the floodplain, terrace, ancient river depression, and other landforms of the Weihe River valley.

(2) Flood sand gravel pore water-bearing rock group

This is distributed in the Piedmont zone of the northern slope of the Qinling Mountains. The buried depth of the water level varies, and the thickness is inconsistent. Closer to the Piedmont, the groundwater level depth increases, reaching up to 30–40 m, while at the front edge of the alluvial plain, the water level is shallower. The buried depth of the water level is shallow. The main characteristics are transport, fast circulation, and uneven water richness. In the Yukou region and on both sides of the aquifer, the composition is primarily boulders, gravel, and coarse particles, resulting in substantial thickness and a permeability coefficient of 20–50 m d^{-1}, which supports a good water yield. The aquifer between the fans and the front edge has poor sorting, high mud content, small thickness, a permeability coefficient of 3–20 m d^{-1}, and poor water richness.

(3) The loess pore fissure phreatic aquifer rock group

This mainly comprises middle and lower Pleistocene loess and paleosol pore fissure phreatic aquifers. Its thickness generally ranges from 20 to 80 m, with water depths of 20–50 m, and a unit water inflow of approximately 5–1000 m^3 d^{-1}. In the vertical direction, the pores are developed, providing good connectivity and fast circulation and migration speeds. In the horizontal direction, the runoff is slow, with the water level depth increasing from the center of the tableland to its edges, water richness weakening from strong to weak. The average annual rainfall is 507.7–719.8 mm. The rainfall infiltration coefficient α is mainly affected by the lithology of the unsaturated zone, topographic slope, water level depth, and other hydrogeological conditions.

(4) Chemical characteristics of diving

The alluvial slope plain in the Piedmont of Luo Mountain is mainly composed of gravel pebbles and boulders, providing favorable conditions for groundwater recharge and runoff. Under long-term severe leaching, a single HCO_3-Ca water layer is formed, and the salinity is less than 0.5 g L^{-1}. In the loess tableland area, the phreatic water is mainly recharged by atmospheric precipitation and irrigation water, draining into the river valley by horizontal runoff. The soluble salt in the water alternating zone is mostly removed by leaching. Due to the substantial buried depth of the water level, evaporation and concentration are weak. The chemical composition of the water is mainly HCO_3-Ca·Mg type water, and the salinity is 0.5–1.0 g L^{-1}. From the front edge of the loess tableland to the vast plain area of the Weihe River valley terrace, the aquifer is mostly weathered granite and metamorphic rock debris. With the increase of underground runoff distance, the alternating adsorption and mixing effect are enhanced, and the salt gradually accumulates. The anions in the groundwater appear as NaCa, NaMg, or NaMgCa combination types with HCO_3^- cations, and the salinity is 0.7–1.6 g L^{-1}. In the main urban area and the northern suburbs, due to the discharge of many anthropogenic pollutants such as waste discharge and sewage irrigation, Cl-HCO_3, HCO_3-Cl, and HCO_3-SO_4 waters are formed, with salinity levels between 1 and 3 g L^{-1}. Bahe River, Fenghe River, and other major river riparian zones are types of dilution and desalination of submersible recharge rivers, which are HCO_3-Ca type water with salinity less than 0.5 g L^{-1}.

(5) Hydrochemical characteristics of confined water

The groundwater in this region is characterized by its significant depth, slow runoff, low salinity, and uniform hydrochemical type. Moving from south to north, the hydrochemical type transitions from HCO_3-Ca water in the Piedmont area to HCO_3-Ca·Na water in the loess tableland and valley terrace. HCO_3-Na-type water appears on the bank of the Weihe River, and the salinity increases gradually. Due to the variation of the hydrochemical environment, there are more complex hydrochemical types in the shallow geothermal anomaly area and the man-made pollution area in the suburbs.

2.2. Lon simulation model of groundwater environmental pollution

In environmental assessment, the simulation of groundwater flow field and pollutant migration is usually used to establish the distribution of groundwater in the natural environment and the pollution discharge due to human factors. Models such as the migration numerical model of As and ammonia nitrogen can be used to predict and analyse environmental water pollution. The problem of contaminant transport in groundwater involves two mathematical models: the mathematical model of groundwater flow and the mathematical model of contaminant transport. Additional basic data information can be found in the Supplementary Materials.

(1) The mathematical model of groundwater flow under natural conditions can be expressed as a three-dimensional unsteady flow mathematical model:

$$\begin{cases} \mu\dfrac{\partial H}{\partial t}(x,y,z)=\dfrac{\partial}{\partial x}\left(D_x\dfrac{\partial H}{\partial x}\right)\dfrac{\partial}{\partial y}\left(D_y\dfrac{\partial H}{\partial_y}\right)\dfrac{\partial}{\partial z}\left(D_z\dfrac{\partial H}{\partial z}\right)+\varepsilon_1 & x,y,z\in\Omega \\ H(x,y,z)|_{\sum_1}=H_1(x,y,z) & x,y,z\in\sum\nolimits_1 \\ q(x,y,z)|_{\sum_2}=0 & x,y,z\in\sum\nolimits_2 \end{cases} \tag{5}$$

where H is the groundwater head (m); D_x, D_y, D_z are the x, y, z direction permeability coefficients (m d^{-1}); H_1 is the boundary head of the first category of aquifers (m); ε is the source–sink intensity (including mining intensity, etc.) (d^{-1}); Σ_1 is the first boundary of the aquifer; and Σ_2 is the second boundary of the aquifer.

(2) The numerical model of pollutant migration is expressed as follows:

$$R_d\frac{\partial c}{\partial t}=\frac{\partial}{\partial x_i}\left(A_{ij}\frac{\partial c}{\partial x_j}\right)-\frac{\partial}{\partial x_i}(cv_i)+\frac{q_s}{\theta}c_s+\sum R_k \tag{6}$$

where R_d is the retardation factor ($R_d = 1 + \rho_b k_d/\theta$), in which ρ_b represents the skeleton density and k_d denotes the distribution coefficient; c is the pollutant concentration in the groundwater (mg L^{-1}); t represents time (d); x_i is the distance in all directions along the axis (m); A_{ij} is the hydrodynamic diffusion coefficient; v_i is the groundwater seepage velocity (m d^{-1}); q_s is the unit flow of sources and sinks (m^3 d^{-1}); c_s is the concentration of source and sink (mg L^{-1}); θ is the aquifer porosity; and $\sum R_k$ are the reaction terms.

2.3. Human health risk assessment models

Environmental pollution emissions harm the human body mainly through drinking water and skin contact. Health risks can be divided into carcinogenic risks and non-carcinogenic risks. In the study area, the primary pollutants — total hardness of water (THW), total dissolved solids (TDS), and NH_3-N — are non-carcinogenic substances. The study's affected population was divided into children and adults. The exposure dose was calculated as follows for non-carcinogenic risk [63].

$$ADD=\frac{PC\times AI\times EF\times ED}{AW\times AT} \tag{7}$$

$$HI=HQ_{\text{Water}}+HQ_{\text{Skin}}=\frac{ADD}{RfD}+HQ_{\text{Skin}} \tag{8}$$

where ADD is the average daily exposure dose of water intake (mg kg^{-1} d^{-1}); PC is the pollutant concentration (mg L^{-1}); AI is the average daily intake (L d^{-1}); EF is the exposure frequency (d a^{-1}); ED is the duration of exposure (a); AW is the average weight (kg); and AT is the average time of action (d). RfD is the reference dose of each pollutant in mg kg^{-1} d^{-1}, where As is 0.0003 mg kg^{-1} d^{-1}, total hardness is 0.0001 mg kg^{-1} d^{-1}, TDS is 0.02 mg kg^{-1} d^{-1}, and NH_3-N is 0.001 mg kg^{-1} d^{-1}. HQ_{water} is the non-carcinogenic risk entropy of drinking water, HQ_{Skin} is the non-carcinogenic risk entropy of skin contact, and HI is the total human health risk assessment or non-carcinogenic risk value of the population under the two exposure routes of water and skin. In this study area, HQ_{Skin} is negligible compared with HQ_{water}, in which, when $HI > 1$, the non-carcinogens have potential health risks to the human body [64]. The parameter table for human health risk assessment is shown in Table 2.

2.4. GIS-based groundwater risk assessment model (GRAM) + artificial intelligence

Environmental assessment tends to be multidimensional and multitemporal index-based [58,68]. The comprehensive assessment system of pollution risk of sewage discharge in industrial parks in the study area combines a groundwater vulnerability assessment, a pollutant load diffusion model, and an environmental remote sensing assessment method. In this study, the environmental assessment indicators were reorganised into naturally important indicators of environmental and human health risk assessment and man-made important indicators for the same purposes. Further information can be found in the Supplementary Materials.

According to the analytic hierarchy process [69–72], the comprehensive evaluation of the environmental evaluation index and index weight can be calculated by the following formula:

$$S=\sum_{i=1}^{n}Y_1\times W_1\ i=1,2,3\ldots,n \tag{9}$$

where Y represents the environment-related indicators, such as slope, rainfall, soil index, aeration zone medium, permeability coefficient, pollutant toxicity, and pollutant leakage; W is the weight of each index. The results can be obtained using geographic information system (GIS) technology to process each evaluation index.

Table 2
Parameter table for human health risk assessment.

Parameter	Children	Adults
Pollutant concentration (mg L^{-1})	Constant	Constant
Average daily intake (L d^{-1}) [63]	1.1	1.7
Exposure frequency (D a^{-1}) [65]	350.0	350.0
Duration of exposure (a) [66]	10.0	30.0
Average weight (kg) [67]	16.0	62.7
Average time of action (d) [63]	3650.0	10950.0

Since the development of environmental risk assessment, researchers have begun to downplay the concept of weights and believe that a certain human factor interferes with weights. They have begun to use object-oriented machine learning methods, such as neural networks, SVM, RF, and even deep learning. Each model can be simplified to the fitting relationship between input indicators and results:

$$A = X \times B \tag{10}$$

where X is the prediction training model (RF, SVM, deep learning, etc.), A is the prediction result, and B is the model input data (such as multiple indicator data and sample label data). This study uses deep learning, such as transformers and convolutional neural networks [73–75], to train the model and compare the results with RF, ANN, SVM, and other machine learning models.

Using the water environment status and monitoring data of Xi'an City's Ecological Environment Bureau (http://xaepb.xa.gov.cn/xxgk/hjzkgb/shjzlyb/628db482f8fd1c0bdc9a5b46.html), we constructed machine learning label data. A total of 31 monitoring points controlled by the city and above were monitored in Xi'an, including 22 monitoring sections controlled by the province and above, using a total of 21 assessment and monitoring indicators. The environmental monitoring sites monitored the pH, dissolved oxygen, permanganate index, five-day biochemical oxygen demand, ammonia nitrogen, petroleum, phenol, mercury, lead, cadmium, anionic surfactant, chromium (hexavalent), fluoride, total phosphorus, cyanide, sulfide, arsenic, chemical oxygen demand, copper, zinc, and selenium levels of the water environment. According to the Chinese Water Environment Quality Standard and Water Environment Quality Evaluation Method (trial), water environment quality can be divided into five levels, with levels I to III accounting for 90.3% of the total, level IV accounting for 9.7%, and level V being negligible. Sample data from the machine learning models used in this study are shown in Table 3.

3. Results

3.1. Explainable high-precision results of deep learning and environmental big data

This study used multi-source environmental data, divided into important natural and man-made environmental and human health risk assessment indicators. In the environmental assessment, references are made to groundwater vulnerability, environmental pollution, and environmental remote sensing assessment. For this study, we selected several natural common indicators [58], such as groundwater depth, recharge, aquifer media, soil indicator, terrain (slope), and aquifer permeability coefficient, and several man-made common indicators [68], such as pollution source leakage, toxicity effect distribution, and pollution diffusion prediction. The multi-source index data and label data were used as input data (Fig. 3a), then GMS [76], MapGIS [77], ENVI [78], and other professional software were used to preprocess the multi-source index image data and normalise them into TIFF image format. The RF, Maxlikelihood, ANN, ACE, SVM, CNN, and transformer models were selected as intelligent models. The transformer model utilises a deep network structure (Fig. 3a). Following several training iterations, the results for each model are shown in Table 4.

The transformer environmental assessment model's accuracy was 97.68%, followed by the CNN and SVM models. The data analysis of the ROC curves of the different models is shown in Table 4. In these results, the area under the curve (AUC) of the transformer model was 0.891, followed by the AUC value of the CNN model (0.833) and the SVM model (0.722). The AUC value of the ACE model was 0.465, the Maxlikelihood model was 0.396, and the AUC value of the RFV model was 0.389. The closer the AUC of the ROC curve is to 1, the better the classification effect of the obtained model is. The standard error of the transformer and CNN results was also relatively reduced to 0.083. The standard errors of other results were 0.099 for RF and Maxlikelihood, 0.100 for ACE, 0.101 for SVM, and 0.104 for ANN. Deep learning and environmental multi-source big data results were divided into five levels (I, II, III, IV, and V) corresponding to the classification in the Chinese International Water Environment Standard document (Fig. 3b). According to the classification results and ROC curves, we focused on the more accurate and less error-prone results from the deep learning model. The analysis revealed that the water environment in the central and southwestern parts of the study area was mostly level II and level III; the northern part was mainly level IV; and the western part had a level V water environment. Thus, the proposed deep learning assessment model based on big environmental data can provide an environmental risk assessment with high precision. The explainability of the complex model is provided in Fig. 3c.

3.2. Explainability analysis results of deep learning and environmental big data

A high-precision environmental assessment model was developed using big data and deep learning. However, due to the difficulties in summarising the nonlinear attributes of the deep learning

Table 3
Location of sample area and levels of environmental assessment.

Sampling point	Position coordinates of center point	Environmental assessment level
Bahekou	34°4′2.04″ N,109°55′55.24″ E	II
Jiajiatang	34°13′31.22″ N,109°52′10.66″ E	III
Sanliqiao	33°48′12.18″ N,108°53′23.94″ E	III
Tianyukou	33°47′41.59″ N,108°43′43.94″ E	II
Xitongqiao	34°4′4.49″ N,109°47′43.72″ E	II
Sanlangcun	34°12′43.26″ N,109°53′28.62″ E	III
Madongcun	34°18′8.21″ N,109°59′37.78″ E	II
Xixinglong	33°46′14.68″ N,108°44′24.50″ E	III
Liangjiaqiao	33°50′21.98″ N,109°30′50.62″ E	III
Wenrao Road	33°48′46.72″ N,109°8′35.41″ E	III
Weihehengqiao	34°10′24.47″ N,109°38′40.53″ E	II
Gengzhengqiao	34°17′10.09″ N,110°2′5.16″ E	III
Farm West Station	34°7′47.18″ N,109°36′6.29″ E	V
Xinfengzhengdaqiao	34°16′19.46″ N,110°16′11.56″ E	III
Shichuan River into Wei Plain	34°37′35.15″ N,110°12′4.69″ E	IV
Rao River flows into the Wei River	33°43′39.66″ N,109°13′2.59″ E	II
Heihe River flows into the Wei River	33°51′44.44″ N,108°51′21.66″ E	II

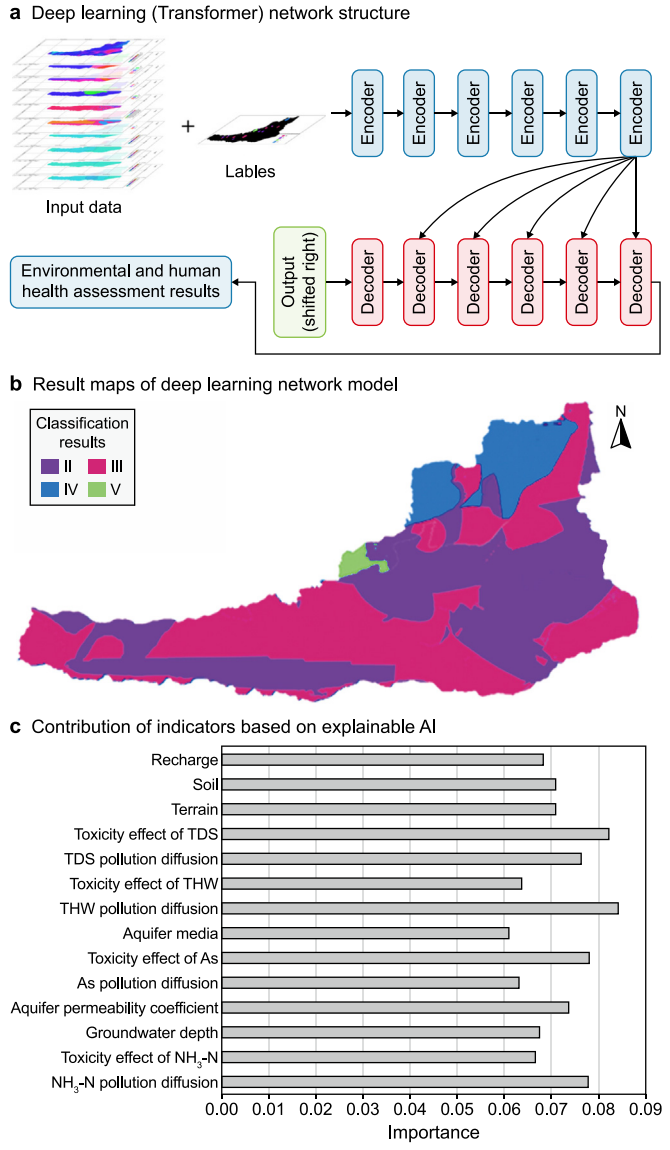

Fig. 3. High-precision and explainability analysis results of deep learning and environmental multi-source big data. **a**, Workflow network structure of environmental risk assessment based on deep learning. **b**, The results of the environment risk assessment using deep learning (transformer). **c**, The contribution of indicators (or importance of indicators) of the model based on explainable AI.

model and its attention mechanism using correlation analysis, the deep learning model remains a "black box". Despite this limitation, the results obtained from the model are credible. Therefore, an explainability analysis of the model was carried out to ensure the explainability and reliability of the high-precision environmental assessment model. Based on the saliency map [60,61], this study proposed a multi-source big data model explainability method to calculate the contribution of each index in the multi-dimensional intelligent model. It analysed the explainability of the deep learning environmental assessment model. The multi-source model inputs the multi-dimensional index $\{Y_1, Y_2, Y_3, \ldots, Y_n\}$ of the data, removes or increases the disturbance (ΔY) of one of the dimension data, re-enters it into the trained model, calculates the accuracy of the environmental assessment results, and increases the error (Δe). Here, y is the one-dimensional environmental data index layer, where the pixel of the specific multi-dimensional data contains [X_1, X_2, X_3, …, X_n], corresponding to different environmental assessment rating labels. Similar to the calculation idea of the saliency map [60,61], the change of each index is set to ΔY, the error change of each index to the overall environmental assessment result is Δe, and the importance of each dimension index feature can be expressed as $s = \Delta Y/\Delta e = \partial Y/\partial e$. The importance of a single-dimension index in the model can be obtained by calculating the S value of each dimension index. With the meaning of the index feature itself, the intelligent environment evaluation of the multi-source index can be explained so that the model can be credible. The index contribution of the deep learning model of environmental big data based on multi-source explainability in this study is shown in Fig. 3c. Through in-depth analysis of the essence of the high-efficiency multi-source intelligent environmental assessment model obtained, the toxicity effects of THW, TDS, and As were found to be high contribution indices in the nonlinear intelligent model, playing an important role, followed by the NH_3-N and TDS pollution diffusion indices. The big data-driven and intelligent depth model results can help humans take corrective actions, improve the model's credibility, and realise explainable intelligent environment evaluation.

The results of the single index distribution are shown in Fig. 4. Our explainability analysis of the deep learning model shows that the toxicity effects of THW, TDS, and As, along with the NH_3-N and TDS pollution diffusion indices, made a greater contribution to the deep learning model and should be our main focus. The distribution of these indicators is shown in Fig. 4i, j, 4m, 4p, and 4l. Based on these explainable results, the "Explainable artificial intelligence (XAI) + Environment" model shows the potential to guide humans to focus on these key indicators and achieve an explainable, reliable, and trusted high-precision environmental assessment.

Based on the explainability results of the deep learning model, it is crucial to prioritize the toxicity effect of the THW index. The Chinese national standard for the THW is 450 mg L^{-1}. Excessive THW can cause cardiovascular diseases such as hypertension and atherosclerotic heart disease. Both excessively hard and excessively soft water are detrimental to human health over prolonged periods. Our analysis found that THW levels exceeded the standard in the

Table 4
The ROC curve data analysis table shows different model results.

Test model	Area under curve	Standard error	Asymptotic significance level	Toward a 95% confidence interval	
				Lower bound	Upper bound
RF	0.389	0.099	0.270	0.195	0.583
Maxlikelihood	0.396	0.099	0.301	0.202	0.590
ANN	0.500	0.104	1.000	0.296	0.704
ACE	0.465	0.100	0.730	0.269	0.662
CNN	0.833	0.083	0.001	0.672	0.995
SVM	0.722	0.101	0.027	0.524	0.921
Transformer	0.891	0.083	0.001	0.695	0.996

Test variable results: At least one of the RF, Maxlikelihood, ANN, ACE, CNN, SVM, or transformer models were between the positive and negative actual state groups, and the statistics can be skewed.

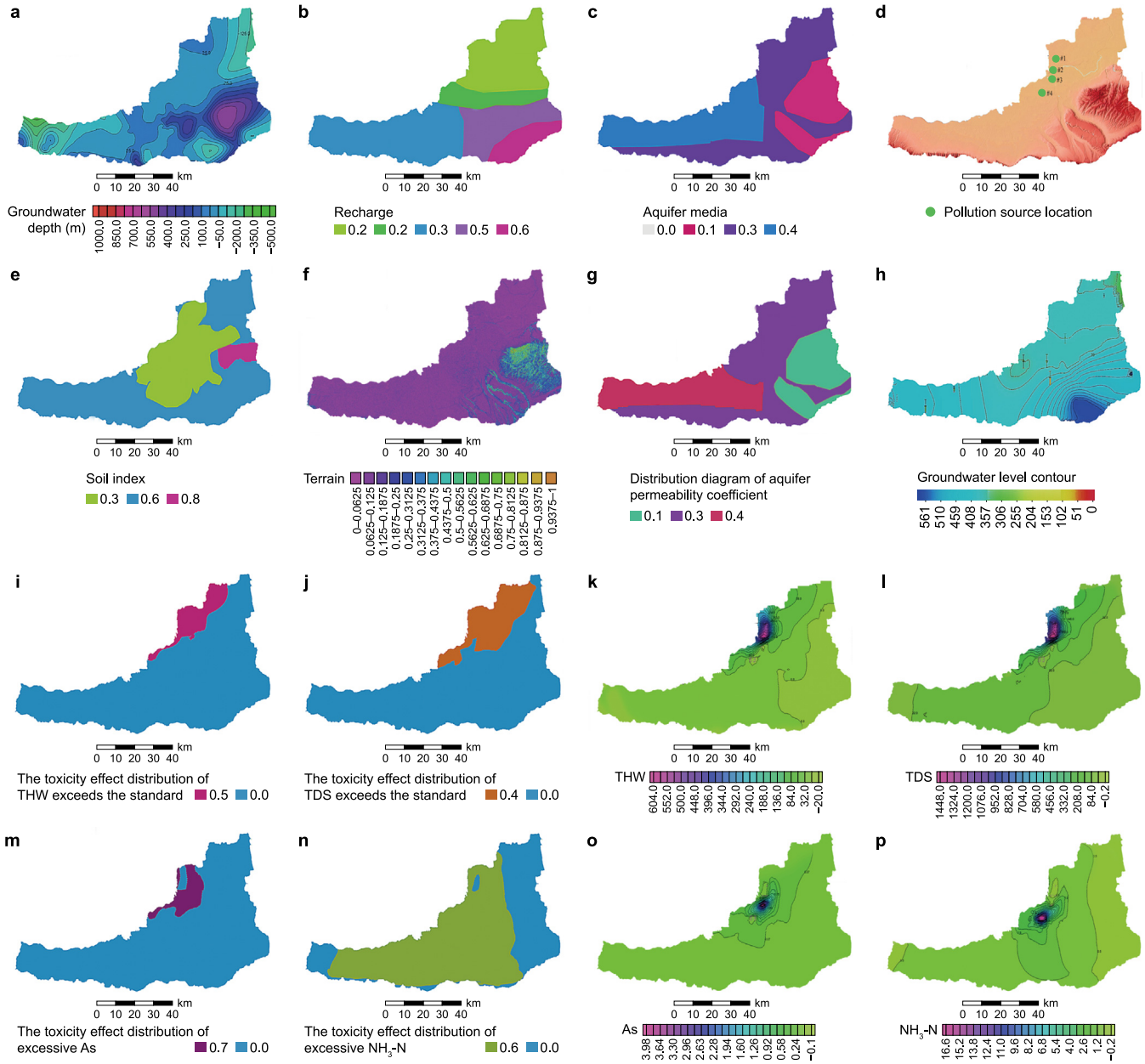

Fig. 4. Assessment results of natural and man-made important indicators. **a**, Groundwater depth. **b**, Recharge. **c**, Aquifer media. **d**, Pollution source location. **e**, Soil index. **f**, Terrain. **g**, Distribution diagram of aquifer permeability coefficient. **h**, Groundwater level contour. **i**, The toxicity effect distribution of total water hardness (THW) exceeds the standard. **j**, The toxicity effect distribution of total dissolved solids (TDS) exceeds the standard. **k**, THW pollution diffusion prediction. **l**, TDS pollution diffusion prediction. **m**, The toxicity effect distribution of excessive As. **n**, The toxicity effect distribution of exceeding NH_3-N. **o**, As pollution diffusion prediction. **p**, NH_3-N pollution diffusion prediction.

northern and western parts of the study area (Fig. 4i). Follow-up work is planned to remediate the THW toxicity distribution area.

According to the explainable results of the deep learning model, we need to pay attention to the toxicity effect of the TDS index. The national standard for total dissolved solids (TDS) is less than 1000 mg L^{-1}. While trace elements such as calcium and magnesium are essential for human health, other soluble substances in excess can be harmful. When the body lacks sufficient trace elements, it may absorb non-essential elements like lead, leading to dysfunctions in proteins or enzymes. TDS exceeded the standard in the western part of the northern region of the study area (Fig. 4j). Follow-up work will focus on examining the distribution area of TDS toxicity. The main reason for the excessive TDS is the discharge of industrial sewage, necessitating stronger supervision and regulation of sewage disposal practices.

The explainability results of the deep learning model indicate the need to address the toxicity effect of the arsenic (As) index. The national standard of As is less than 0.01 mg L^{-1}. Excess intake of As can cause ankylosing spondylitis in humans, affecting their families and having implications for sexual health. As shown in Fig. 4m, the northwest area of the study area was within the toxic influence range of As, and the supervision of sewage discharge needs to be strengthened.

In summary, the explainability analysis of the deep learning model for environmental big data highlights the need to pay close attention to the toxicity effects of THW, TDS, and As indices. The

toxic diffusion areas of pollutants shown in Fig. 4i, j, and 4m were mainly concentrated in the northern and western regions of the study area. The environmental assessment results of the study area were levels V and IV in the northern and western regions, respectively (Fig. 3b), aligning with the model explainability analysis. While a deep learning model without explainability analysis can produce good results, it lacks proper validation due to its "black box" nature, resulting in lower credibility. Meanwhile, using the environmental big data depth model with its explainability analysis provides a high-precision environmental assessment, yielding credible and reliable results. Additionally, this approach allows for identifying heavily polluted areas and prioritizing key governance indicators based on the explainability results.

4. Discussion and conclusion

In this study, we proposed an explainable, high-precision deep learning environmental assessment model based on multi-source environmental big data. While satisfying the requirements of high efficiency and precision, the XAI environmental assessment model can be used to explain relevant multi-source indicator data to environmental governance departments or personnel and guide personnel to focus on key indicators.

Although previous studies have used high-precision models to conduct environmental assessments [8–34,58,68], scant research exists on the use of transformer models in this capacity, and research focusing on the explainability of intelligent models is lacking. For example, Zhao [8] used the theory of groundwater environmental vulnerability, pollutant diffusion, and multi-source indicators. Based on the theory of groundwater environmental vulnerability and pollutant diffusion, Xu [58] used multi-source indicators and remote sensing cloud computing to assess multi-source environments over a long time series. Huang [68] combined the toxicity of pollutants and their impact on human carcinogenic risk and considered environmental multi-source index data. While these studies considered multi-source environmental indicators, they did not delve into efficient and high-precision models based on environmental multi-source big data as the current study does. In addition, previous studies have used models with explainability for environmental assessment. For example, Baek [79] used network and decision-tree-based micro-pollutant analysis of marine discharge. Their decision-tree model is an algorithm with its explainability. Based on the determined weights of the indicators in the decision tree, a pollution analysis of marine emissions was performed, and the method had a certain explainability. Sahani [80] used logistic regression, decision trees, and RF models to analyse and predict the sensitivity of a protected area and clarified the relative importance of trail susceptibility causal variables using an RF model.

The difference between our work and previous research lies in the use of a deep learning model based on the attention mechanism to assess the environment with high precision. Also, we used the explainability method to analyse the explainability of the deep network environmental assessment model in depth. Additionally, we clarified the contribution of multi-source indicators in the deep learning model. Most environmental assessment models still rely on the DRASTIC model, pollutant transport model, or pollutant toxicology analysis combined with traditional analytic hierarchy process, RP, decision tree, and SVM models. Previous research on applying multi-source deep learning environmental assessment models is still in its infancy. In this study, we used multi-source environmental data and compared the outcome of environmental assessments conducted using the RF, Maxlikelihood, ANN, ACE, SVM, CNN, and transformer models. We concluded that the deep learning transformer model had the highest accuracy. Therefore, it is suitable for conducting efficient and high-precision environmental assessments.

Existing multi-source AI environmental assessment models lack accuracy and precision. Furthermore, as deep learning models are often “black box” models, humans may be skeptical of the credibility of the resulting estimations or predictions. Moreover, although high-precision results have been obtained, they cannot provide the most critical guidance for future environmental governance because they cannot accurately identify the most critical indicators from multi-source data. Feature selection is a fundamental challenge in machine learning, yet features are the key to learning and understanding. When the impact of individual features is assessed and consolidated within the model, the results are more accurate and robust [81]. Thus, we added an explainability analysis to the high-precision deep learning model, and the results provide key indicators for future environmental assessment systems that can aid in human decision-making. According to the explainability results of the multi-source AI environmental assessment model, the toxicity effects of THW, TDS, and As and the diffusion of NH_3-N and TDS pollution ranked high among the contributing factors for environmental risk. Therefore, the model results indicate that these indicators should be the focus in environmental governance efforts.

In conclusion, big data and deep learning technologies have been introduced to the field of environmental risk assessment, which can increase the accuracy of assessments and reduce assessment errors. Leveraging the local explainability analysis results in deep learning, humans can identify areas with poor environmental conditions according to the multi-source big data environmental assessment model. This enables a targeted focus on addressing specific indicators, aligning with the human-centred AI approach [82]. The model developed in this study delivered a comprehensive and high-precision environmental assessment, demonstrating its reliability and credibility by addressing the "black box" issue inherent in deep-learning environmental assessment models. This model facilitates AI-driven recommendations for environmental improvements, achieving a breakthrough from “machine learning” to “machine teaching” in the environmental field. This innovation can foster a greater understanding of and trust in AI environmental assessments. This study provides a foundation for future research on high-precision AI environmental assessments and elucidates the importance of performing explainability analysis on environmental assessment models.

Data availability

According to the water environment status and monitoring data of the Xi'an Ecological Environment Bureau (http://xaepb.xa.gov.cn/xxgk/hjzkgb/shjzlyb/628db482f8fd1c0bdc9a5b46.html), we made machine learning label data. The original satellite data can be getten from https://glovis.usgs.gov/or https://www.cheosgrid.org.cn/. And the hydrogeology and geological data can be obtained from the official website of Geological Archives of China https://www.ngac.cn/125cms/c/qggnew/index.htm. And the paper provides an annex describing the basic data of the study area. In addition, the groundwater simulation model, pollutant diffusion data or other data in this study can be available from the corresponding author (Xing Yang, Yihua Hu and Daqin Wang) upon request.

Code availability

The Python scripts and ArcMap & ENVI & MODFOLOW + MT3DMS models & explainable deep learning method developed and used in this study are available from the

corresponding author (Xing Yang, Yihua Hu and Daqin Wang) upon request.

CRediT authorship contribution statement

Haoli Xu: Conceptualization, Formal Analysis, Methodology, Writing - Original Draft. **Xing Yang:** Conceptualization, Supervision, Writing - Review & Editing. **Yihua Hu:** Project Administration, Supervision. **Daqing Wang:** Methodology, Project Administration, Resources, Writing - Original Draft. **Zhenyu Liang:** Project Administration. **Hua Mu:** Writing - Review & Editing. **Yangyang Wang:** Writing - Review & Editing. **Liang Shi:** Writing - Review & Editing. **Haoqi Gao:** Formal Analysis, Software. **Daoqing Song:** Formal Analysis, Supervision. **Zijian Cheng:** Project Administration, Supervision. **Zhao Lu:** Writing - Review & Editing. **Xiaoning Zhao:** Writing - Review & Editing. **Jun Lu:** Writing - Review & Editing. **Bingwen Wang:** Supervision. **Zhiyang Hu:** 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.

Acknowledgments

The authors deeply appreciate the great support from the Dreams Foundation of Jianghuai Advance Technology Center (No.2023-ZM01D006), the National Natural Science Foundation of China (No.62305389), the Scientific Research Project of National University of Defense Technology under Grant (22-ZZCX-07) and Hefei Comprehensive National Science Center. In addition, the authors thank Prof. Zhengdong Deng and Prof. Zhibin Ding from the Army Engineering University of PLA for their valuable knowledge of groundwater remote sensing assessment and water environment analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ese.2024.100479.

References

[1] V. Amiri, et al., Mercury pollution in the coastal Urmia aquifer in northwestern Iran: potential sources, mobility, and toxicity, Environ. Sci. Pollut. Res. 28 (2021) 17546–17562.
[2] F.H. He, et al., The synthetic geo-ecological environmental evaluation of a coastal coal-mining city using spatiotemporal big data: a case study in Longkou, China, J. Clean. Prod. 142 (2017) 854–866.
[3] N. Cho, et al., Construction of spatiotemporal big data using environmental impact assessment information, Korean Journal of Remote Sensing 36 (2020) 637–643.
[4] H.L. Xu, et al., Multivariate and spatio-temporal groundwater pollution risk assessment: a new long-time serial groundwater environmental impact assessment system, Environ. Pollut. 317 (2023) 120621.
[5] W.W. Liu, et al., Intelligent comprehensive evaluation system using artificial intelligence for environmental evaluation, Environ. Impact Assess. Rev. 86 (2021) 106495.
[6] S. Xu, et al., Carbon reduction assessment of public buildings based on Apriori algorithm and intelligent big data analysis, Soft Comput. 1 (2023) 1–10.
[7] B. Houria, et al., Hydrochemical characterisation of groundwater quality: merdja plain (Tebessa town, Algeria), Civil Engineering Journal 6 (2020) 318–325.
[8] X.N. Zhao, et al., Groundwater pollution risk assessment based on groundwater vulnerability and pollution load in an isolated island, Chemosphere 289 (2022) 133134.
[9] Z.Y. Zhang, et al., Assessment of global health risk of antibiotic resistance genes, Nat. Commun. 13 (2022) 1553.
[10] S.S. Singha, et al., A GIS-based modified DRASTIC approach for geospatial modeling of groundwater vulnerability and pollution risk mapping in Korba district, Central India, Environ. Earth Sci. 78 (2019) 628.
[11] L. Dong, et al., Evaluating ground water vulnerability in West Lake Watershed by using DRASTIC model, Chin. J. Appl. Ecol. 13 (2002) 217–220.
[12] H.H. Jiang, et al., Risk assessment on U(VI) to groundwater based on a modified DRASTIC model, GIS and AHP method, Environ. Monit. Assess. 3 (2015) 118–122+125.
[13] G. Busico, et al., A modified SINTACS method for groundwater vulnerability and pollution risk assessment in highly anthropized regions based on NO3- and SO42- concentrations, Sci. Total Environ. 609 (2017) 1512–1523.
[14] N. Kazakis, K.S. Voudouris, Groundwater vulnerability and pollution risk assessment of porous aquifers to nitrate: modifying the DRASTIC method using quantitative parameters, J. Hydrol. 525 (2015) 13–25.
[15] J. Wang, et al., Assessment of groundwater contamination risk using hazard quantification, a modified DRASTIC model and groundwater value, Beijing Plain, China, Sci. Total Environ. 432 (2012) 216–226.
[16] F. Li, et al., Assessment and uncertainty analysis of groundwater risk, Environ. Res. 160 (2018) 140–151.
[17] D.C. Jhariya, et al., Assessment of groundwater vulnerability to pollution by modified DRASTIC model and analytic hierarchy process, Environ. Earth Sci. 78 (2019) 610.
[18] S. Maleki, et al., Z-numbers based novel method for assessing groundwater specific vulnerability, Eng. Appl. Artif. Intell. 122 (2023) 106104.
[19] M. Kong, et al., Developing and validating intrinsic groundwater vulnerability maps in regions with limited data: a case study from Datong City in China using DRASTIC and Nemerow pollution indices, Environ. Earth Sci. 78 (2019) 262.
[20] A.A. Nadiri, et al., Groundwater DRASTIC vulnerability mapping by unsupervised and supervised techniques using a modelling strategy in two levels, J. Hydrol. 574 (2019) 744–759.
[21] A. Neshat, et al., Estimating groundwater vulnerability to pollution using a modified DRASTIC model in the Kerman agricultural area, Iran, Environ. Earth Sci. 71 (2014) 3119–3131.
[22] M. Bordbar, et al., A new hybrid framework for optimization and modification of groundwater vulnerability in coastal aquifer, Environ. Sci. Pollut. Res. 26 (2019) 21808–21827.
[23] N. Kazakis, K.S. Voudouris, Groundwater vulnerability and pollution risk assessment of porous aquifers to nitrate: modifying the DRASTIC method using quantitative parameters, J. Hydrol. 525 (2015) 13–25.
[24] H. Huan, et al., Assessment and validation of groundwater vulnerability to nitrate based on a modified DRASTIC model: a case study in Jilin City of northeast China, Sci. Total Environ. 440 (2012) 14–23.
[25] H. Wu, et al., A modified DRASTIC model for assessing contamination risk of groundwater in the northern suburb of Yinchuan, China, Environ. Earth Sci. 75 (2016) 483.
[26] K. Loague, Regional scale groundwater vulnerability estimates: impact of reducing data uncertainties for assessments in Hawaii, Ground Water 32 (2010) 605–616.
[27] A. Mair, A.I. El-Kadi, Logistic regression modeling to assess groundwater vulnerability to contamination in Hawaii, USA, J. Contam. Hydrol. 153 (2013) 1–23.
[28] J.E. Liggett, D.M. Allen, Evaluating the sensitivity of DRASTIC using different data sources, interpretations and mapping approaches, Environ. Earth Sci. 62 (2011) 1577–1595.
[29] A. Neshat, et al., Groundwater vulnerability assessment using an improved DRASTIC method in GIS, Resour. Conserv. Recycl. 86 (2014) 74–86.
[30] E. Senser, A. Davraz, Assessment of groundwater vulnerability based on a modified DRASTIC model, GIS and an analytic hierarchy process (AHP) method: the case of Egirdir Lake basin(Isparta, Turkey), Hydrogeol. J. 21 (2012) 701–714.
[31] K. Brindha, L. Elango, Cross comparis on of five popular groundwater pollution vulnerability index approaches, J. Hydrol. 524 (2015) 597–613.
[32] H. Rajput, et al., Modification and optimization of DRASTIC model for groundwater vulnerability and contamination risk assessment for Bhiwadi region of Rajasthan, India, Environ. Earth Sci. 79 (2020) 136.
[33] D.C. Jhariya, et al., Assessment of groundwater vulnerability to pollution by modified DRASTIC model and analytic hierarchy process, Environ. Earth Sci. 78 (2019) 610.
[34] W. Wu, et al., Groundwater vulnerability assessment and feasibility mapping under reclaimed water irrigation by a modified DRASTIC model, Water Resour. Manag. 28 (2014) 1219–1234.
[35] H. You, et al., Evaluation of groundwater vulnerability with improved DRASTIC method, Procedia Environmental Sciences 10 (2011) 2690–2695.
[36] M.B. Shakoor, et al., Human health implications, risk assessment and remediation of As-contaminated water: a critical review, Sci. Total Environ. 601–602 (2017) 756–769.
[37] R.C. Estoque, et al., Heat health risk assessment in Philippine cities using remotely sensed data and social-ecological indicators, Nat. Commun. 11 (2022) 1581.
[38] Y. Ji, et al., Seasonal variation of drinking water quality and human health risk assessment in Hancheng city of Guanzhong Plain, China, Expo. Heal. 12 (2020) 469–485.
[39] R. Tian, J. Wu, Groundwater quality appraisal by improved set pair analysis with game theory weightage and health risk estimation of contaminants for Xuecha drinking water source in a loess area in Northwest China, Hum. Ecol.

Risk Assess. 25 (2019) 132–157.
[40] WHO, Guidelines for Drinking-Water Quality, fourth ed., World Health Organization, 2017. https://www.iwapublishing.com/template.cfm?name=isbn9781780400303.
[41] D. Lariviere, et al., Uranium bone content as an indicator of chronic environmental exposure from drinking water, J. Environ. Radioact. 121 (2013) 98–103.
[42] A.I. Selden, et al., Nephrotoxicity of uranium in drinking water from private drilled wells, Environ. Res. 109 (2009) 486–494.
[43] J. Li, et al., An integrated fuzzy-stochastic modeling approach for risk assessment of groundwater contamination, J. Environ. Manag. 82 (2007) 173–188.
[44] S. Muhammad, et al., Arsenic health risk assessment in drinking water and source apportionment using multivariate statistical techniques in Kohistan region, northern Pakistan, Food Chem. Toxicol. 48 (2010) 2855–2864.
[45] A. Siddique, et al., Multipathways human health risk assessment of trihalomethane exposure through drinking water, Ecotoxicol. Environ. Saf. 116 (2015) 129–136.
[46] S. Yang, et al., Health risk assessment of phreatic water based on triangular fuzzy theory in Yinchuan plain, Ecotoxicol. Environ. Saf. 164 (2018) 732–738.
[47] M. Wang, et al., Human health risk identification of petrochemical sites based on extreme gradient boosting, Ecotoxicol. Environ. Saf. 233 (2022) 113332.
[48] X.L. Li, et al., Rapid diagnosis of heavy metal pollution in lake sediments based on environmental magnetism and machine learning, J. Hazard Mater. 416 (2021) 126163.
[49] B. Ke, et al., Predicting the sorption efficiency of heavy metal based on the biochar characteristics, metal sources, and environmental conditions using various novel hybrid machine learning models, Chemosphere 276 (2021) 130204.
[50] K. White, et al., Exploration of E. coli contamination drivers in private drinking water wells: an application of machine learning to a large, multivariable, geospatio-temporal dataset, Water Res. 197 (2021) 117089.
[51] A.K. Jaydhar, et al., Hydrogeochemical evaluation and corresponding health risk from elevated arsenic and fluoride contamination in recurrent coastal multi-aquifers of eastern India, J. Clean. Prod. 369 (2022) 133150.
[52] A. Vaswani, et al., Attention is all you need, Comput. Lang. (2023) 1–15.
[53] H.P. Wu, et al., CvT: introducing convolutions to vision transformers, in: 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 1–10.
[54] Res-ViT, Residual vision transformers for image recognition tasks, IEEE 35th International Conference on Tools with Artificial Intelligence (ICTAI). 2023 (2023) 1–8.
[55] T.B. Brown, et al., Language models are few-shot learners, Neural Information Processing Systems (2020) 1–75.
[56] J. Del Ser, et al., On generating trustworthy counterfactual explanations, Inf. Sci. 655 (2024) 119898.
[57] D. Guning, et al., XAI-Explainable artificial intelligence, Sci. Robot. 4 (2019) 7120.
[58] B.H.M. Van der Velden, et al., Explainable artificial intelligence (XAI) in deep learning-based medical image analysis, Med. Image Anal. 79 (2022) 102470.
[59] H.L. Xu, et al., Multivariate and spatio-temporal groundwater pollution risk assessment: a new long-time serial groundwater environmental impact assessment system, Environ. Pollut. 317 (2023) 120621.
[60] M.D. Zeiler, R. Fergus, Visualizing and understanding convolutional networks, ECCV 2014 (2014) 818–833.
[61] K. Simonyan, et al., Deep inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps, ICIR, 2014. 2014, pp. 1–8.
[62] Y.M. Liu, et al., Explainable deep learning for insights in El Niño and river flows, Nat. Commun. 14 (2023) 339.
[63] Y.Y. Yao, et al., Evaluation of environmental engineering geology issues caused by rising groundwater levels in Xi'an, China, Eng. Geol. 294 (2021) 106350.
[64] H.H. Shi, et al., Source analysis and health risk assessment of heavy metals in groundwater of Leizhou Peninsula, Environ. Sci. J. Integr. Environ. Res. 4 (2021) 4246–4256.
[65] USEPA, Risk Assessment Guidance for Superfund, Human Health Evaluation Manual (Part A), Office of Emergency and Remedial Response, Washington DC, 1989. https://www.epa.gov/risk/risk-assessment-guidance-superfund-rags-part.
[66] Z.M. Dong, et al., Uncertainties in human health risk assessment of environmental contaminants: a review and perspective, Environ. Int. 85 (2015) 120–132.
[67] H. Yang, et al., Waste management, informal recycling, environmental pollution and public health, J. Epidemiol. Community 72 (2018) 237–243.
[68] C.J. Schenck, et al., Exploring the potential health risks faced by waste pickers on landfills in South Africa: a socio-ecological perspective, Int. J. Environ. Res. Publ. Health 16 (2019) 2059.
[69] Y. Huang, et al., Groundwater non-carcinogenic health risk comprehensive assessment based on DLAFVRT model in an island city, Ecol. Indicat. 148 (2023) 110043.
[70] Z.D. Deng, et al., Construction and investigation of groundwater remote sensing fuzzy assessment index, Chinese Journal of Geophysics-Chinese edition 56 (2013) 3908–3916.
[71] G.D. Tian, et al., Green decoration materials selection under interior environment characteristics: a grey-correlation based hybrid MCDM method, Renewable Sustainable Energy Rev. 81 (2018) 682–692.
[72] R. Ma, et al., Deficiency and excess of groundwater iodine and their health associations, Nat. Commun. 13 (2022) 7354.
[73] H.D. Guo, et al., Eeasuring and evaluating SDG indicators with big earth data, Sci. Bull. 67 (2022) 1792–1801.
[74] H.L. Xu, et al., Predicting groundwater potential assessment in water-deficient islands based on convolutional neural networks, The Egyptian Journal of Remote Sensing and Space Sciences 25 (2022) 1013–1023.
[75] Y.J. Ma, et al., Deep learning algorithm using fundus photographs for 10-year risk assessment of ischemic cardiovascular diseases in China, Sci. Bull. 67 (2022) 17–20.
[76] MODFLOW modeling with GMS. https://www.aquaveo.com/software/gms-groundwater-modeling-system-introduction, 2021.
[77] MapGIS Desktop, 2014. https://www.mapgis.com/.
[78] Learning Materials, 2024. https://envi.geoscene.cn/.
[79] S.S. Baek, et al., Analysis of micropollutants in a marine outfall using network analysis and decision tree, Sci. Total Environ. 806 (2022) 150938.
[80] N. Sahani, et al., GIS-based spatial prediction of recreational trail susceptibility in protected area of Sikkim Himalaya using logistic regression, decision tree and random forest model, Ecol. Inf. 64 (2021) 101352.
[81] B. Pfeifer, et al., Robust random forest-based all-relevant feature ranks for trustworthy AI, Stud. Health Technol. Inf. 294 (2022) 137–138.
[82] A. Holzinger, et al., Digital transformation in smart farm and forest operations needs Human-Centered AI: challenges and future directions, Sensors 22 (2022) 3043.