



OPEN Integrated source apportionment, co-enrichment mechanisms analysis, and risk assessment of arsenic and fluoride in groundwater of large-scale irrigation districts in semi-arid regions

Qiyang Zhang¹✉, Aihui Wei¹, Panpan Xu^{2,3}, Hui Qian^{2,3} & Jianxi Ren¹

To address the widespread concern of arsenic (As) and fluoride (F) contamination in irrigated areas, this study analyzed 51 groundwater samples from the Jiaokou Irrigation District, China. An integrated approach of isotope tracing (δD and $\delta^{18}O$) and self-organizing map (SOM) was used to reveal the spatial distribution and enrichment mechanisms of As and F. Health risks were further evaluated via Monte Carlo simulation. The results indicated that groundwater in the study area was weakly alkaline brackish water. The groundwater was classified into three hydrochemical types: $Cl-SO_4-Na$, HCO_3-Na , and HCO_3-Ca by SOM cluster analysis. Weakly alkaline conditions inhibited As adsorption onto minerals (hematite, goethite) and fluorite precipitation, while promoting OH^- substitution for F^- in F-bearing minerals (biotite, amphibole), thereby facilitating the co-enrichment of As and F. Dissolution of feldspathic silicates increased water-rock interaction time and further enhanced alkalinity. Meanwhile, strong evaporation-concentration processes and agricultural activities contributed to As-F co-enrichment in groundwater. Overall, F posed a greater non-carcinogenic risk than As, with children being more susceptible. For carcinogenic risk, adults were the more sensitive group to arsenic. Specifically, F was the major contributor to non-carcinogenic risks for both adults (62.2%) and children (59.6%). As, however, showed higher sensitivity for carcinogenic risk in adults (92.6%). Management priorities should include safeguarding drinking water for children in high-risk areas and controlling groundwater fluoride within agricultural water strategies. These findings provide a scientific foundation for mitigating health risks and ensuring sustainable groundwater use in semi-arid irrigation districts.

Keywords arsenic-fluoride groundwater, integrated source apportionment, co-enrichment mechanisms, self-organizing map (SOM), health risk assessment

Groundwater, one of the most widely distributed freshwater resources globally, plays an irreplaceable role in ensuring safe drinking water, supporting regional economic development, and advancing ecological sustainability. In arid and semi-arid regions where surface water is scarce, groundwater frequently serves as the sole water source for domestic and economic use¹⁻⁴. However, the combined impacts from industrialization, agriculture, and other anthropogenic activities have compromised groundwater quality, presenting serious risks to water safety and public health^{5,6}. Among these threats, arsenic (As) and fluoride (F) pollution in groundwater has become increasingly prominent and is now recognized as a widespread global environmental problem⁷. Current estimates suggest that approximately 150 million people are directly affected by As in drinking water globally, while up to 260 million are exposed to high-F environments^{8,9}. With the continued identification of

¹School of Architecture and Civil Engineering, Xi'an University of Science and Technology, Xi'an 710054, Shaanxi, China. ²School of Water and Environment, Chang'an University, Xi'an 710054, Shaanxi, China. ³Key Laboratory of Subsurface Hydrology and Ecological Effects in Arid Region of the Ministry of Education, Chang'an University, Xi'an 710054, Shaanxi, China. ✉email: zhangqiyang@xust.edu.cn

contaminated areas and the availability of improved monitoring data, the affected population continues to grow. Several countries, including India¹⁰, Pakistan¹¹, China¹², and Mexico¹³, are currently facing severe As and F groundwater contamination, which poses a significant health threat to local populations. In China, elevated As and F in groundwater are concentrated in arid and semi-arid regions such as the Guanzhong Basin^{6,14}, Datong Basin¹⁵, Tumochuan Plain¹⁶, and Weining Plain⁵. The enrichment of As and F in groundwater is controlled by both the natural environment and human activities. Regional geology (lithology, mineralogy), climate, and groundwater residence time are the primary controls on As and F migration and enrichment in natural groundwater^{17–19}. Despite extensive research on the factors controlling groundwater arsenic and fluoride, significant gaps persist in understanding their co-occurrence mechanisms and anthropogenic drivers, particularly within human-altered, large-scale irrigation systems. Therefore, clarifying the enrichment mechanisms of As and F contamination in groundwater is of great significance for ensuring the safety of the water environment and sustainable development.

To identify pollutant sources and elucidate enrichment mechanisms, integrated approaches combining hydrochemical analysis (e.g., Gibbs diagrams, ion end-member diagrams, and PHREEQC simulations), multivariate statistics, and isotope tracing technology have been widely applied^{3,14,19–23}. However, traditional multivariate statistical methods such as principal component analysis (PCA) and cluster analysis (CA) can only provide preliminary insights into differences between reference vectors, demonstrating certain limitations²⁴. In the investigation and analysis of arsenic and fluoride contamination in groundwater, the abundance of samples and complexity of data structure require methods capable of analyzing multidimensional groundwater quality data while providing visual results to support decision-making²⁵. The Self-Organizing Map (SOM), an efficient unsupervised machine learning neural model, is widely applied for data visualization and attribute exploration²⁶. By organizing data in a grid topology, SOM can visually represent relationships between groundwater samples and identify distinct water quality patterns, which aids in tracing pollutant sources and the dominant geochemical processes^{27,28}. When integrated with K-means to further cluster SOM neurons, it can be effectively utilized for evaluating and classifying complex hydrogeochemical data. Therefore, groundwater samples with similar hydrochemical characteristics can be effectively classified by combining the SOM with the K-means method, and the correlation between arsenic and fluoride in groundwater can be discerned.

Moreover, the impact of As and F pollution in groundwater on human health has also attracted widespread attention. Long-term consumption of As-contaminated groundwater may lead to skin lesions and damage to the respiratory and nervous systems, and may also induce cancers (e.g., kidney and liver cancer)^{16,29}. Insufficient F intake increases the risk of dental caries and osteoporosis, whereas excessive intake may lead to dental and skeletal fluorosis³⁰. The World Health Organization (WHO) guideline for drinking water quality, established to protect human health, sets a limit of 0.01 mg/L for arsenic and recommends a daily fluoride intake of 0.5–1.5 mg/L. Therefore, health risk assessment models are commonly used to evaluate potential impacts of groundwater pollutants contaminants on human health^{14,31–33}. Nonetheless, these traditional deterministic models are sensitive to sample limitations and individual variability, leading to randomness and uncertainty in key parameters, which may result in deviations in assessment results³⁴. Monte Carlo simulation offers distinct advantages: (1) it inherently handles multi-source random variables without requiring discretization to resolve probability distributions, with computational accuracy independent of system state dimensionality; (2) through the implementation of reduction factors, it enables dynamic adaptation to scenario changes, thereby significantly enhancing the efficiency of risk prediction. Therefore, in the present study, a predictive model that combines Monte Carlo simulations with the traditional deterministic model is used to quantify the potential impacts of groundwater pollutants on human health via different exposure pathways (drinking water ingestion and dermal contact)^{35–37}. In addition, the health risk distribution map was drawn by using the deterministic model and GIS technology to visualize the regional differences of health risk. The coupling model can provide a robust and accurate prediction of population-level risks, and the results provide an important scientific basis for ensuring drinking water safety and protecting human health.

The Jiaokou Irrigation District, one of the ten major irrigation districts in the Guanzhong Basin, has been operational for over 60 years since its completion in 1960 and is an important grain production area in the Guanzhong Plain^{2,38}. Groundwater is the primary drinking water source in this district; its quality directly affects the health of local residents and socio-economic development. Monitoring indicates the presence of As and F contamination in the groundwater in the study area, with potential implications for both groundwater quality and public health. Although previous studies have indicated severe F contamination in the groundwater of the district^{39–41}, systematic studies on As-F co-contamination mechanisms and associated health risks are lacking.

Therefore, based on Self-organizing map (SOM), K-means clustering, isotope tracing, and human health risk assessment models coupled with Monte Carlo simulations, this study aims to: (1) analyze the spatial distribution of As and F in groundwater in the Jiaokou Irrigation District; (2) elucidate the correlations between As and F and their co-enrichment mechanisms; and (3) assess exposure risks of As and F contamination in groundwater to adults and children. The findings will help decision-makers implement targeted measures to mitigate As and F contamination in groundwater and ensure the safety of drinking water for local communities.

Materials and methods

Study area

The Jiaokou Irrigation District, which extends from 34°30'7"N to 34°52'37"N in latitude and from 109°12'40"E to 110°10'1"E in longitude, is situated in the eastern part of the Guanzhong Basin, Shaanxi Province, China. It is a large, multi-level, dam-free, electric-pumping irrigation area that uses the Wei River as its primary water source. The irrigation area is bounded to the east and west by the Luo and Shichuan rivers, and to the north and south by the Wei River and Pucheng Luyang Lake, respectively. Its total area is 1024.36 km², extending approximately 48.6 km east–west and 31.9 km north–south [Fig. 1(a)]⁴².

Over more than 60 years of irrigation development, the district has developed a dense network of ditches and canals. Land use is dominated by cultivated land (84.6%), followed by orchards (8.3%) [Fig. 1(b)]. The area has a warm-temperate semi-arid monsoon climate, with an average annual temperature of 14.12 °C, annual sunshine duration of 1842 h, and average annual precipitation and evaporation of 527.2 mm and 1220.9 mm, respectively^{38,42}.

Topographically, the area is primarily composed of the Wei River floodplain, first terrace, second terrace, and loess tableland, forming a “dustpan” shape sloping from the northwest, north, and northeast towards the central and southern parts of the irrigation area. Controlled by the Qinling zonal tectonic system, the main aquifers are Quaternary terrestrial sediments that are widely distributed but unevenly thick, including lacustrine (Q^l), alluvial (Q^{al}), and aeolian (Q^{eol}) deposits. Typical loess deposits occur in the north due to aeolian processes [Fig. 1(c)]^{2,42,43}. The aquifer in the study area is continuous and forms a complete system, with groundwater flowing from northwest to southeast. Recharge occurs primarily from precipitation, irrigation infiltration, and canal leakage. Influenced by the Kouzhen-Guanchi fault zone, groundwater occurs mainly as aeolian loess pore-fissure water in the northern loess plateau and loose-rock pore water in the alluvial layer³¹.

Sample collection and analysis

This research conducted a survey in the Jiaokou Irrigation District during the summer (August) of 2023. The sampling was designed to cover the entire irrigation district, considering factors such as geomorphology, groundwater flow dynamics, and human activities. A total of 51 water samples were collected, consisting of 48 from phreatic wells (including both domestic and irrigation wells, with depths ranging from 5 to 50 m) and 3 from drainage canals [Fig. 1(a)]. In this irrigation district, the primary function of the drainage canals is agricultural drainage, with their water mainly sourced from the discharge of shallow groundwater. Consequently, the hydrochemical characteristics of these canal waters largely reflect the properties of the local groundwater. Therefore, the drainage canal samples were analyzed together with the phreatic well samples in this study. Sampling locations were recorded using a handheld GPS device. Prior to collection, all wells were purged by continuous pumping for 10–15 min to remove stagnant water. A portable water-quality meter (AZ86031) was employed for in-situ measurements of pH, electrical conductivity (EC), and temperature until parameters stabilized, ensuring the collected samples were representative of the aquifer and actual usage conditions. Samples were then collected in pre-cleaned 1.5 L polyethylene bottles, which had been rinsed three times with distilled water followed by at

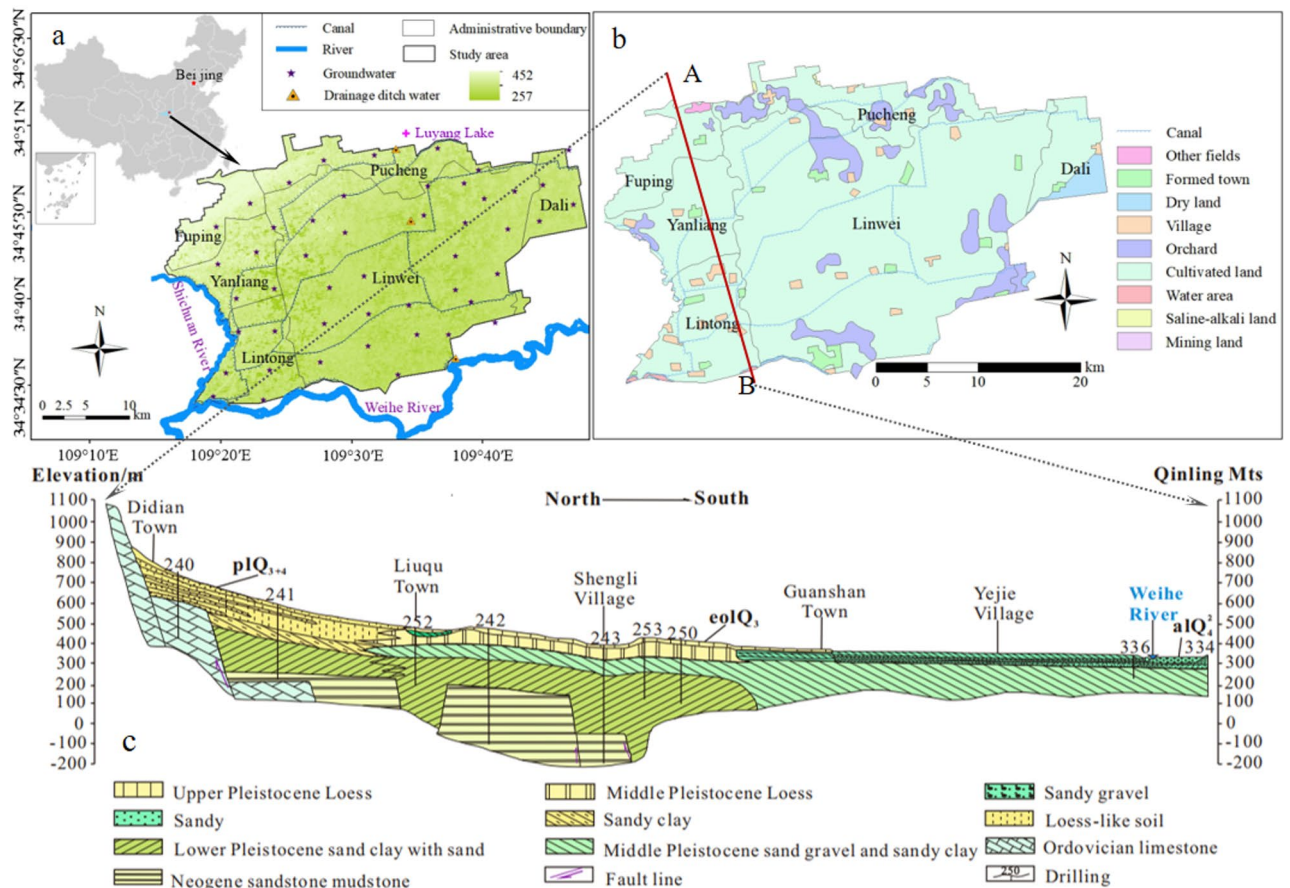


Fig. 1. Maps of the study area showing: (a) geographical location and distribution of groundwater sampling sites, (b) land-use map, (c) stratigraphic profile. Panels (a) and (b): ArcGIS 10.8.1 (<http://www.esri.com>), Panel (c): CorelDRAW X8 18.1.0.661 (<https://www.coreldraw.com>).

least three rinses with the sample water. The water samples intended for arsenic determination were acidified with 1% hydrochloric acid. The bottles were completely filled to minimize air bubbles, immediately sealed with caps, and labeled. All samples were stored at 4 °C and transported to the laboratory within one week for analysis.

All major cations, anions, and trace elements (As, F) were determined by Shaanxi Geology and Mineral 2nd Engineering Survey Institute Inspection Detection Co., Ltd. K^+ and Na^+ were determined by flame atomic absorption spectrometry (detection limits: 0.132 mg/L and 0.354 mg/L, respectively). Ca^{2+} , Mg^{2+} , SO_4^{2-} , CO_3^{2-} , and HCO_3^- were measured by titration, while Cl^- was determined by silver titration, all with a detection limit of 1.0 mg/L. NO_3^- was analyzed by spectrophotometry (detection limit: 0.2 mg/L). Arsenic was measured using atomic fluorescence spectrometry (AFS-922; detection limit: 0.001 mg/L), and F by ion chromatography (detection limit: 0.2 mg/L). Total dissolved solids (TDS) were determined using the gravimetric method. Stable hydrogen and oxygen isotopes (δD and $\delta^{18}O$) were analyzed with a gas isotope ratio mass spectrometer (Thermo Scientific 253 Plus) coupled with an elemental analyzer (Thermo Fisher Flash 2000HT) via high-temperature pyrolysis and reduction. The analytical accuracy was 1.0‰ for δD and 0.2‰ for $\delta^{18}O$.

Analytical quality was assessed using the charge-balance error (CBE) [Eq. (1)]^{3,44}. The CBE values for all samples in this study were within $\pm 5\%$ (range: -2.91% to 2.99%), indicating reliable results. CBE was calculated as follows:

$$CBE(\%) = \frac{\sum N_c - \sum N_a}{\sum N_c + \sum N_a} \times 100 \quad (1)$$

where N_c and N_a are the milligram equivalent concentrations of cations and anions (meq/L), respectively.

Self-organizing map and K-means clustering algorithms

The Self-organizing map (SOM), initially proposed by the Finnish scholar Kohonen⁴⁵, is an efficient unsupervised machine learning neural model. Through unsupervised training, SOM can be widely applied to complex data tasks such as classification, estimation, and prediction⁴⁶. By mapping high-dimensional data onto a two-dimensional hexagonal neuron grid, SOM achieves dimensionality reduction while preserving the topological structure of the input data, thereby enhancing data visualization and analytical capabilities^{24,26}.

K-means is a classic clustering method that partitions the data space into k clusters through an iterative process, and can be used to further classify the neurons after SOM training⁴⁷. The optimal number of clusters is determined based on the minimum Davies-Bouldin index (DBI), where a smaller DBI value indicates greater distinctions between different categories and better clustering performance^{47,48}. The combination of SOM and K-means helps reduce data complexity and enables superior visual analysis.

In this study, the SOM model was implemented using Matlab R2024a to classify groundwater samples. The optimal number of output neurons is determined by the formula $5\sqrt{n}$ (where n is the number of samples)⁴⁹. Based on this, a neural network was trained using a 5×7 map size, and the accuracy of the SOM was evaluated using the quantization error (QE) and topographic error (TE)^{27,28}. The key steps for clustering in the SOM neural network are as follows:

1. Input matrix and data standardization, x is the input variable.

$$X = (x_1, x_2, x_3, \dots, x_n)^T \quad (2)$$

2. The competitive learning is performed based on the Euclidean distance.

$$D_j(x) = \sqrt{\sum_{i=1}^n (x_i - w_{ij})^2} \quad (3)$$

Where $D_j(x)$ is the distance between x_i and the node neuron j ; n is the dimensionality of the input variable; w_{ij} represents the connection weight of node neuron j .

3. Calculating the spatial location of the winning neuron.

$$T_{j,I(x)} = \exp\left(-\frac{S_{j,I(x)}^2}{2\sigma^2}\right) \quad (4)$$

Where $T_{j,I(x)}$ denotes the updated weight coefficient; $S_{j,I(x)}$ is the distance between node neuron j and $I(x)$; $I(x)$ represents the winning node; σ is the neighborhood size.

4. Weight adjustment.

$$w_{ij}(t+1) = w_{ij}(t) + \beta_i h_i(t)(x_i - w_{ij}) \quad (5)$$

Where t is training times; β_i is the learning rate, which decreases exponentially with the number of training times.

5. K-means clustering.

$$DBI = \frac{1}{N} \sum_{i=1, i \neq j}^N \max \frac{\bar{S}_i + \bar{S}_j}{\|w_i - w_j\|_2} \quad (6)$$

Where N is the number of clusters; \bar{S}_i is the average Euclidean distance from samples in the i -th cluster to their centroid; $\|w_i - w_j\|_2$ is the Euclidean distance between the centroids of cluster i and cluster j .

Saturation index (SI)

The saturation index (SI) is a primary method for determining and investigating the saturation state of minerals in groundwater and is widely applied in hydrogeochemical research²⁷. It is defined as:

$$SI = \lg \frac{IAP}{K} \quad (7)$$

Where IAP is the ion activity product of the relevant mineral; K is the equilibrium constant of the mineral dissolution reaction. In this study, the PHREEQC software was employed to calculate the saturation indices of various minerals in the groundwater samples. Minerals with SI between -0.2 and 0.2 are in quasi-equilibrium⁴².

Health risk assessment

Humans are exposed to harmful substances from groundwater primarily via ingestion (i.e., drinking) and dermal contact. Prolonged excessive intake of As and F can cause various adverse health effects²⁹. This study applied the USEPA health risk assessment model to evaluate potential health risks posed by As and F to both children and adults^{1,5}. The chronic daily absorbed dose via the two pathways was calculated as follows:

$$CDI = \frac{C_w \times IR \times EF \times ED}{BW \times AT} \quad (8)$$

$$CDD = \frac{DA \times SA \times EV \times EF \times ED}{BW \times AT} \quad (9)$$

$$DA = K_p \times C_w \times T \times 0.001 \quad (10)$$

$$SA = 239 \times H^{0.417} \times BW^{0.517} \quad (11)$$

where CDI and CDD are the chronic daily absorbed doses for oral intake and skin contact, respectively (mg/kg-d); DA is the absorbed dose via skin contact (mg/cm²); C_w is pollutant concentration in groundwater (mg/L); and SA is the exposed skin surface area (cm²).

The total non-carcinogenic risk is the sum of the non-carcinogenic risk hazard coefficients of the two pathways of exposure to groundwater contamination, calculated as follows:

$$HQ_{\text{oral}} = \frac{CDI}{RfD} \quad (12)$$

$$HQ_{\text{dermal}} = \frac{CDD}{RfD} \quad (13)$$

$$HQ_i = HQ_{\text{oral}} + HQ_{\text{dermal}} \quad (14)$$

$$THQ = \sum_{i=1}^n HQ_i \quad (15)$$

where HQ is the non-carcinogenic hazard factor; RfD is the reference dose for both oral intake and skin contact routes (mg/kg-d).

Additionally, the carcinogenic risk to the human body may be caused by As. The formulas for calculating the carcinogenic risk through oral ingestion and skin absorption contact are as follows:

$$CR_{\text{oral}} = CDI \times SF \quad (16)$$

$$CR_{\text{dermal}} = CDD \times SF \quad (17)$$

$$TCR = CR_{\text{oral}} + CR_{\text{dermal}} \quad (18)$$

Where, CR represents carcinogenic risk; SF is the slope factor for carcinogenic pollutants (mg/kg-d)⁻¹.

Based on Guidelines for Health Risk Assessment of Groundwater Pollution in China, the RfD values of As and F are 3.0×10^{-4} and 0.04, respectively, and the SF value for As is 1.5. Other parameters are shown in Table S1 (in the Supplementary Materials). HQ values greater than 1 indicate a high non-carcinogenic risk, and CR values greater than 1×10^{-4} indicate that the carcinogenic risk is unacceptable.

Monte carlo model

Deterministic health risk models assume constant parameter values; however, substantial variability in population-related parameters during health risk characterization can introduce significant uncertainty in the assessment results³⁴. To quantify this uncertainty, a Monte Carlo simulation with 10,000 iterations was

performed using Crystal Ball software, with the random seed set to 2026. The relevant data are summarized in Table S2. The probability distributions of risk values were obtained after multiple iterations of the simulation, enabling a more robust assessment of As and F health risks in groundwater^{35,36,50}. Additionally, to identify the impact of parameters on the results of health risk assessment, a sensitivity analysis was conducted based on the Spearman rank correlation coefficient³⁶. The larger the absolute value of the correlation coefficient, the greater the impact of the variable on the health risk¹⁴.

The research routes of this study was illustrated in Fig. 2.

Results and discussion

Groundwater hydrochemical characteristics

The statistical results of groundwater hydrochemical indicators in the Jiaokou Irrigation District were presented in Table S3 (51 in total, JK1-JK51). Groundwater pH values ranged from 7.80 to 8.63, with a mean of 8.19, indicating weak alkalinity. Based on TDS, groundwater salinity could be classified as freshwater (TDS < 1000 mg/L), brackish water (1000 mg/L ≤ TDS ≤ 3000 mg/L), or saline water (TDS > 3000 mg/L)⁴². TDS values in the study area ranged from 988 to 11,802 mg/L, with an average of 3185.88 mg/L. Brackish and saline waters accounted for 54.90% and 43.14% respectively, which collectively indicate a significant groundwater salinization issue in the study area.

The average concentrations of major ions follow the order: $SO_4^{2-} > HCO_3^- > Cl^- > NO_3^- > CO_3^{2-} > F^-$ for anions and $Na^+ > Mg^{2+} > Ca^{2+} > K^+$ for cations. Box plots of the concentration distribution for various groundwater indicators are shown in Fig. S1 (in the Supplementary Materials). The average As concentration in groundwater is 0.0047 mg/L, with a maximum of 0.017 mg/L. As concentrations in 11.8% of samples exceed the threshold (0.01 mg/L). For F, the maximum concentration is 4.63 mg/L, with an average of 1.58 mg/L. F concentrations are below 0.5 mg/L in only 1.96% of samples, whereas they exceed 1.5 mg/L in 43.75% of samples. According to the Chinese Standards for Irrigation Water Quality (GB 5084–2021), the recommended limits for As and F in irrigation water for dryland crops are 0.1 and 3.0 mg/L, respectively. Fluoride concentrations in the study area were found to be well above the permitted level, indicating that the groundwater is therefore unsuitable for agricultural irrigation, and thereby threatens local food security and human health.

The coefficient of variation (CV) is a statistical measure used to assess spatial heterogeneity in ion concentrations. According to CV thresholds, variability can be categorized as weak (< 10%, uniform distribution), moderate (10–100%, significant spatial variation), or strong (> 100%, highly discrete distribution)⁴². In this study, the CV values for As and F in groundwater were 80.56% and 53.01%, respectively, both falling within the moderate variability range. This indicates that groundwater in the Jiaokou Irrigation District is likely influenced by anthropogenic agricultural activities, leading to local enrichment of As and F and resulting in their spatially heterogeneous distribution⁵¹.

Cluster analysis based on SOM

Statistical results in Table S3 reveal significant differences in anions and cations within the groundwater. Based on this, SOM and K-means methods were employed to perform unsupervised machine learning classification on 13 variables from 51 groundwater samples²⁷. The quantization error (QE) and topographic error (TE) were

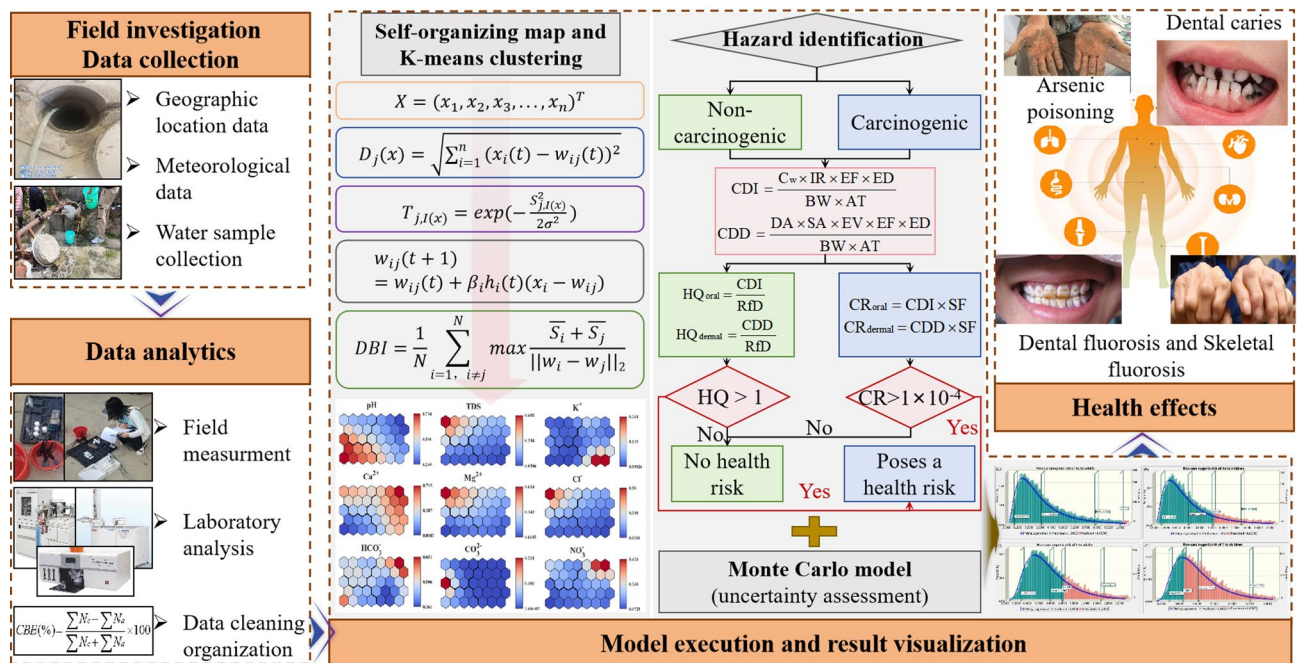


Fig. 2. The framework of research routes in the present study.

0.368 and 0.187, respectively, confirming that the model reliably preserved the topological structure of the data and thereby effectively illustrating the disparities among the samples. As shown in Fig. 3(a) (red and blue shades represent high and low values, respectively), correlations among variables can be identified by comparing color gradients across the component planes. Similar colors suggest a positive correlation, with higher similarity indicating a stronger relationship²⁶. Based on the minimum DBI, the groundwater samples are categorized into three distinct groups [Fig. 3(b)]. Cluster 1 comprises 12 samples characterized by elevated concentrations of TDS, Na^+ , Mg^{2+} , Cl^- , and SO_4^{2-} (Table S4), and is classified as the $\text{Cl}\cdot\text{SO}_4\text{-Na}$ type [Fig. 3(c)]. Cluster 2 includes 14 samples showing higher levels of CO_3^{2-} , F^- , As, and pH (Table S4), belonging to the $\text{HCO}_3\text{-Na}$ type [Fig. 3(c)]. Cluster 3 consists of 25 samples with notable concentrations of K^+ , Ca^{2+} , HCO_3^- , and NO_3^- (Table S4), categorized as the $\text{HCO}_3\text{-Ca}$ type [Fig. 3(c)].

Isotopic characteristics

Stable isotopes of hydrogen (δD) and oxygen ($\delta^{18}\text{O}$) were used to identify groundwater recharge sources⁵¹. δD values ranged from -71.92‰ to -55.98‰ (mean: -63.58‰), and $\delta^{18}\text{O}$ values ranged from -9.37‰ to -7.80‰ (mean: -8.59‰) (Table S5). The CV values for δD and $\delta^{18}\text{O}$ were -0.07 and -0.05 , indicating minimal variability in hydrogen and oxygen.

The $\delta^{18}\text{O}$ vs. δD relationship (Fig. 4) yielded a fitting line of hydrogen and oxygen that was nearly parallel to the local meteoric water line (LMWL: $\delta\text{D} = 7.49 \delta^{18}\text{O} + 6.13$), with most groundwater samples plotting below the

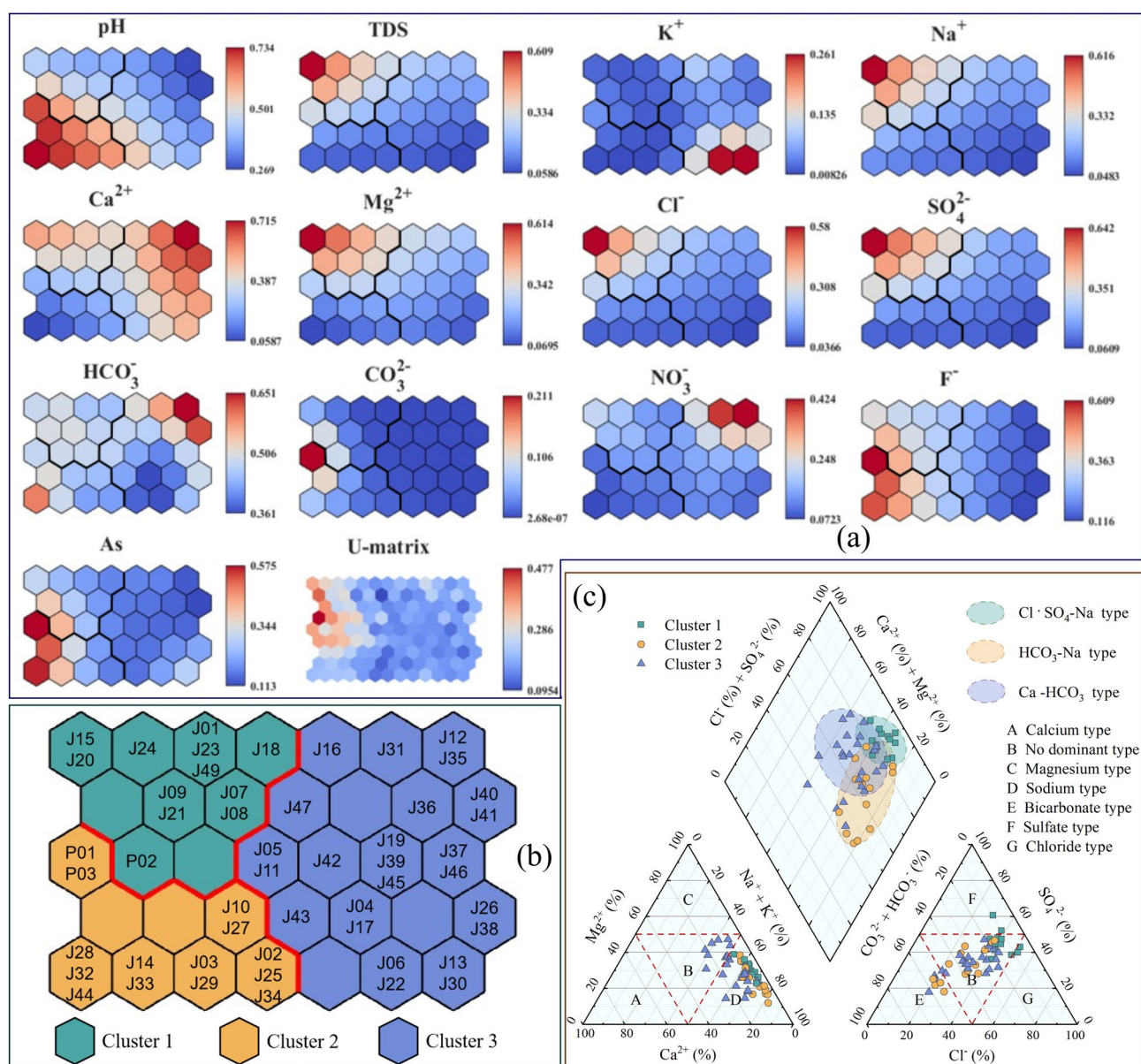


Fig. 3. Clustering results of groundwater samples: U-matrix and component planes of the 13 groundwater variables in the SOM (a), distribution map of the three groundwater clusters (b), and the Piper diagram (c).

LMWL. This pattern suggested that groundwater was primarily recharged by atmospheric precipitation and that evaporation contributed to the enrichment of δD and $\delta^{18}O$ in groundwater²⁰.

Spatial distribution of As and F in groundwater

The spatial distributions of As and F are shown in Fig. S2. Elevated As concentrations occurred mainly in the northern and southeastern parts of Linwei District and the central part of Pucheng County, with low concentrations distributed in other areas. The area where As exceeded regulatory limits was 36.99 km², accounting for 3.61% of the study area. Low-F groundwater (<0.5 mg/L) occurred in a small area in Yanliang District in the western part of the irrigation area (1.70 km²). High-F groundwater (>1.5 mg/L) mainly occurred in Linwei District, Pucheng County, and Dali County, covering 557.53 km² (54.43% of the total area), indicating that more than half of the groundwater was contaminated with fluoride. Overall, the spatial distributions of As and F were broadly consistent, suggesting that they had similar geochemical occurrence mechanisms. Furthermore, elevated concentrations of arsenic and fluoride are predominantly distributed within cultivated land [Figs. 5(a) and (b)].

Causes of As and F enrichment in groundwater

Correlation analysis based on SOM

The SOM clustering results [Fig. 3(a)] indicate strong correlations among As, F⁻, CO₃²⁻, and pH in the groundwater of the study area, while TDS exhibited a strong positive correlation with Na⁺, Mg²⁺, Cl⁻, and SO₄²⁻. As and F showed a moderate correlation ($r=0.63$) and were both moderately correlated with pH and CO₃²⁻. This indicating that they may be influenced by similar environmental factors²⁶. The Pearson correlation analysis also confirms these results (Fig. S3).

Moreover, among the 51 groundwater samples analyzed, six exceeded the permissible limit for As and 24 exceeded it for F. Five samples exceeded the standards for both As and F⁻ simultaneously [Fig. 6(a)], accounting

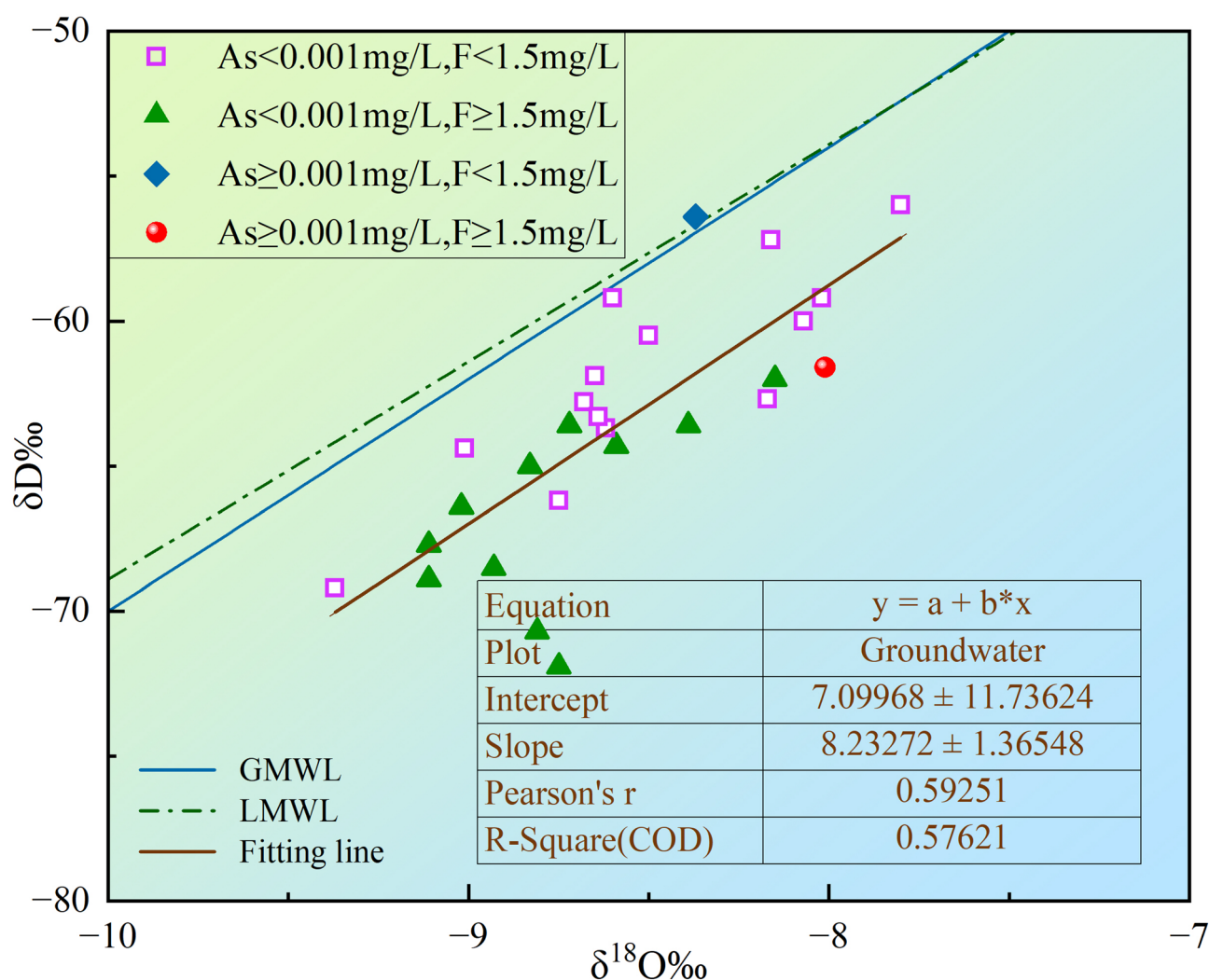


Fig. 4. The relationship of $\delta^{18}O$ vs. δD of groundwater.

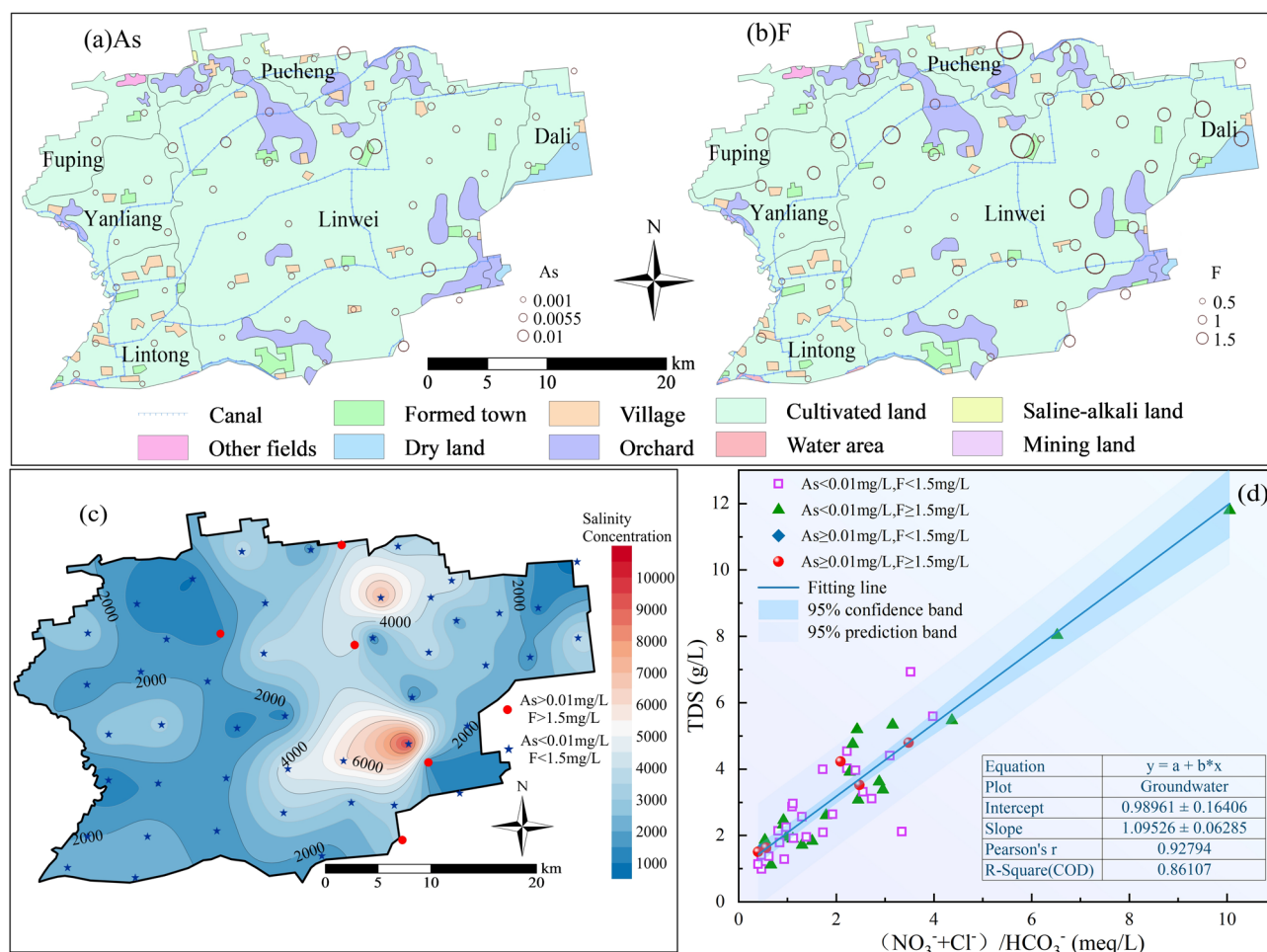


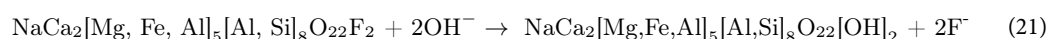
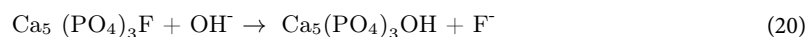
Fig. 5. Spatial variation of As and F values under different land use types (a) and (b), Map of As/F exceedance points over salinity contours (c), Relationship plot of TDS and $(\text{NO}_3^- + \text{Cl}^-) / \text{HCO}_3^-$ (d). Panels (a), (b), and (c) were prepared using Surfer 23 (<http://www.GoldenSoftware.com/products/surfer.com>).

for 20% of the samples that exceeded the standard threshold. This further confirmed the potential for co-occurrence of As and F in groundwater and their possible geochemical association.

Acid-base conditions

Groundwater in the Jiaokou Irrigation District is weakly alkaline, with an average pH of 8.19. As shown in Figs. 6 (b1) and (b2), pH exhibits a positive correlation with As and F concentrations; both increase gradually with rising pH. In groundwater with pH values between 7 and 9, As predominantly exists as arsenate or arsenite, which can be adsorbed by positively charged minerals and clay minerals, such as hematite, goethite, kaolinite, and illite, within the aquifer media^{16,52}. Higher pH reduces the positive surface charge of these minerals, weakening As adsorption and facilitating its enrichment into groundwater⁵³.

Fluoride is relatively more active in alkaline groundwater. Elevated pH reduces Ca^{2+} activity, inhibiting fluorite (CaF_2) precipitation and facilitating the enrichment of F in groundwater. Furthermore, due to the similar ionic radii of OH^- and F^- , weakly alkaline conditions promote the substitution of OH^- for F^- in fluorine-bearing minerals such as biotite, fluorapatite, and amphibole [Eqs. (19)–(21)], further contributing to fluoride enrichment^{54,55}. Thus, weak alkalinity favors the co-enrichment of As and F in groundwater.



Water-rock interaction

The Gibbs diagram distinguishes three hydrochemical evolution domains: evaporation-concentration, water-rock interaction, and atmospheric precipitation⁵⁶. As shown in Fig. S4, samples with co-enrichment of As and F are mainly distributed in the evaporation-concentration and water-rock interaction fields, indicating that

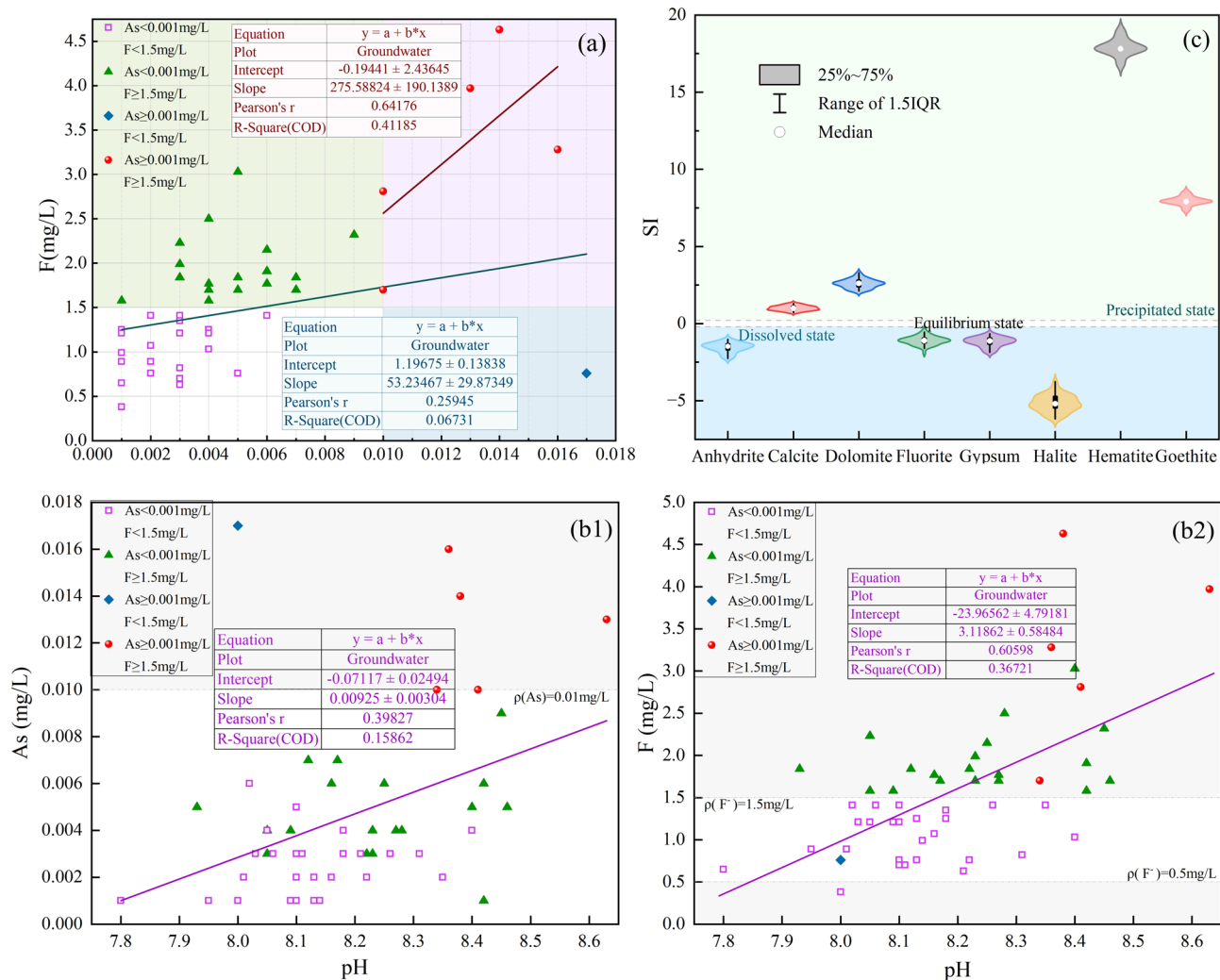
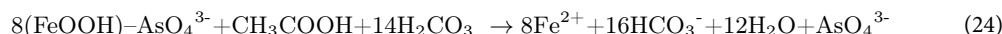
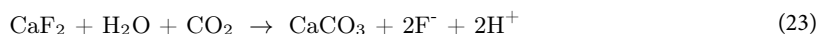
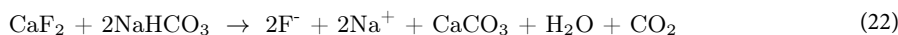


Fig. 6. Relationship between As and F concentrations in groundwater of the study area (a), Relationship between As and F concentrations with pH value (b1) and (b2), Saturation indices of minerals (c).

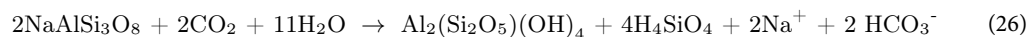
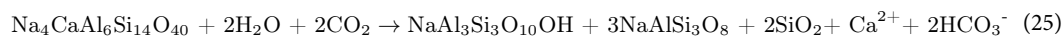
both processes contribute to the enrichment of As and F. High-F groundwater is more strongly influenced by evaporation, whereas high-As groundwater is mainly influenced by water-rock interactions. Additionally, anthropogenic activities may also exert a certain promoting effect on the enrichment of As and F.

Saturation index (SI) analysis, by elucidating the interactions between groundwater and minerals, serves as an effective method for identifying the dominant geochemical environment²⁷. Saturation indices (SI) for minerals including anhydrite, calcite, dolomite, gypsum, fluorite, halite, hematite, and goethite were calculated using PHREEQC to analyze the influence of water-rock interaction on groundwater chemical components. As shown in Fig. 6(c), calcite, dolomite, hematite, and goethite are in a saturated state (SI > 0.2), while other minerals are in an unsaturated state (SI < 0.2). Calcite and dolomite precipitation reduces the Ca²⁺ concentration in groundwater, thereby promoting the enrichment of F [Eqs. (22)–(23)]⁵⁴. Under alkaline conditions, As adsorbed on hematite and goethite undergoes desorption, promoting the enrichment of As in the groundwater. In addition, goethite (FeOOH) dissolves under specific reducing conditions, and the As in the mineral is released into groundwater in its free form [Eq. (24)]¹⁶.



To further identify the types of rock weathering sources in the study area, end-member plots of $\text{N}(\text{HCO}_3^- / \text{Na}^+)$, $\text{N}(\text{Mg}^{2+} / \text{Na}^+)$, and $\text{N}(\text{Ca}^{2+} / \text{Na}^+)$ were prepared⁵⁷. Fig. S5 shows that groundwater samples from the study area plot between the silicate and evaporite weathering fields, away from the carbonate field. This suggests that As and F enrichment is driven by the weathering of silicate and evaporite minerals. The dissolution of silicates (e.g., plagioclase, albite) in the aquifers under the action of CO₂ increases HCO₃⁻ concentrations [Eqs. (25)–

(26)], thereby making the groundwater alkaline and promoting As and F enrichment^{16,42,58}. The dissolution of feldspar-group minerals also produces fine-grained materials, slowing groundwater movement and prolonging water-rock interaction.



Evaporation

The relationship of Cl^- to δD and $\delta^{18}\text{O}$ was analyzed to further distinguish hydrogeochemical processes²⁰. An increase in δD and $\delta^{18}\text{O}$ with rising Cl^- indicates that evaporation is dominant; no change with increasing Cl^- indicates that evaporite dissolution is dominant; and a slight increase at low Cl^- concentrations indicates the mixed effect of both processes²⁰. As shown in Figs. S6a and b, As- and F-rich groundwater is influenced by both evaporation and evaporite dissolution. Recharge sources for groundwater in the study area include precipitation infiltration and agricultural irrigation (canal seepage, infiltration)². Long-term irrigation has raised groundwater levels over an area of approximately 350 km², reducing depths to less than the critical evaporation depth (5 m)⁴². Furthermore, it is observed that groundwater samples with As > 0.01 mg/L and F > 1.5 mg/L are invariably associated with a high soluble salt content of approximately 2000 mg/L [Fig. 5(c)]. This spatial coincidence points to a possible geochemical link between the co-enrichment of As and F and high salinity conditions. Under such conditions, intense evaporation promotes salinity buildup and the enrichment of As and F in groundwater.

Anthropogenic activities

The relationship between TDS and $(\text{NO}_3^- + \text{Cl}^-)/\text{HCO}_3^-$ was used to determine the potential anthropogenic influences on groundwater quality. As shown in Fig. 5(d), the strong positive correlation suggests substantial anthropogenic influence on groundwater quality⁵⁹, particularly on high-F concentrations. Furthermore, Figs. 5(a) and (b) show that high-As and high-F groundwater is primarily located in cultivated land zones.

To improve crop yields, large amounts of fertilizers and pesticides are applied in the irrigation district⁴². Common agricultural amendments, such as potash and NPK (nitrogen-phosphorus-potassium) fertilizers, contain substantial amounts of fluorides. Additionally, phosphate fertilizers, including superphosphate, contribute not only fluorides but also relatively high concentrations of arsenic. Pesticides including fluoroacetamide and sodium fluoroacetate are rich in F^{60,61}. Some insecticides used in agriculture may also be a potential source of As in groundwater⁶². Long-term use of pesticides and chemical fertilizers results in the accumulation of As and F in soils. These contaminants then enter groundwater through leaching and runoff, leading to increased concentrations in the aquifer. Therefore, the high concentrations of arsenic and fluoride in groundwater are driven by both natural factors and anthropogenic activities.

Climate conditions

The Jiaokou Irrigation District experiences a warm-temperate semi-arid monsoon climate, with a mean annual evaporation (1220.9 mm) significantly exceeding precipitation (527.2 mm) by approximately twofold^{38,42}. Groundwater samples exceeding arsenic and fluoride thresholds were predominantly located in the central part of the study area where the water-table depth was shallow (< 5 m, i.e., below the limiting evaporation depth). Intense evaporative concentration in these zones promoted the enrichment of arsenic and fluoride in groundwater. Furthermore, the arid climate necessitated a high dependence on irrigation for agriculture². Large-scale groundwater extraction and return flow created an enhanced pumping-irrigation-recharge cycle. This process not only reintroduced evaporatively concentrated water with elevated arsenic and fluoride into the aquifer but may also introduce exogenous arsenic and fluoride derived from fertilizers and pesticides, thereby further exacerbating arsenic and fluoride contamination in groundwater.

Conceptual model of As and F enrichment mechanisms

The synthesis of the aforementioned results led to the development of a conceptual model depicting the enrichment mechanisms of arsenic and fluoride in the study area's groundwater (Fig. 7). Groundwater in the study area is overall weakly alkaline, with alkalinity further enhanced by HCO_3^- generated from the dissolution of silicate rocks (e.g., plagioclase)^{42,58}. These conditions favor the enrichment of As and F. The alkaline environment promotes the accumulation of As by reducing its adsorption on mineral surfaces. In particular, the As adsorbed on the surfaces of hematite and goethite undergoes desorption under alkaline conditions¹⁶. Furthermore, fluoride (CaF_2) precipitation is inhibited, and the F^- in fluoride-bearing minerals (e.g., fluorapatite, biotite, amphibole) is easily replaced by OH^- , leading to the enrichment of F^- ⁵⁴. Thus, the interaction between groundwater and arsenic- or fluoride-bearing minerals is a key mechanism for both As and F enrichment. Additionally, long-term fertilizer and pesticide application have led to As and F accumulation in soils, which enter groundwater through leaching and infiltration. Collectively, these processes form a dual "natural-anthropogenic" enrichment pattern for As and F in the region.

Health risk assessment

Human health risk assessment

Non-carcinogenic risk The box plots illustrating health risks posed by arsenic and fluoride contamination in groundwater of the Jiaokou Irrigation District to adults and children are shown in Fig. 8. For adults, the non-carcinogenic risk values of arsenic ranged from 0.08 to 1.38 (average: 0.38), while for children, the values ranged from 0.16 to 2.65 (average: 0.72) (Table S6). Although average non-carcinogenic risk values for both adults and children were below the safety threshold, their maximum values exceeded this threshold in 7.84% and 17.65%

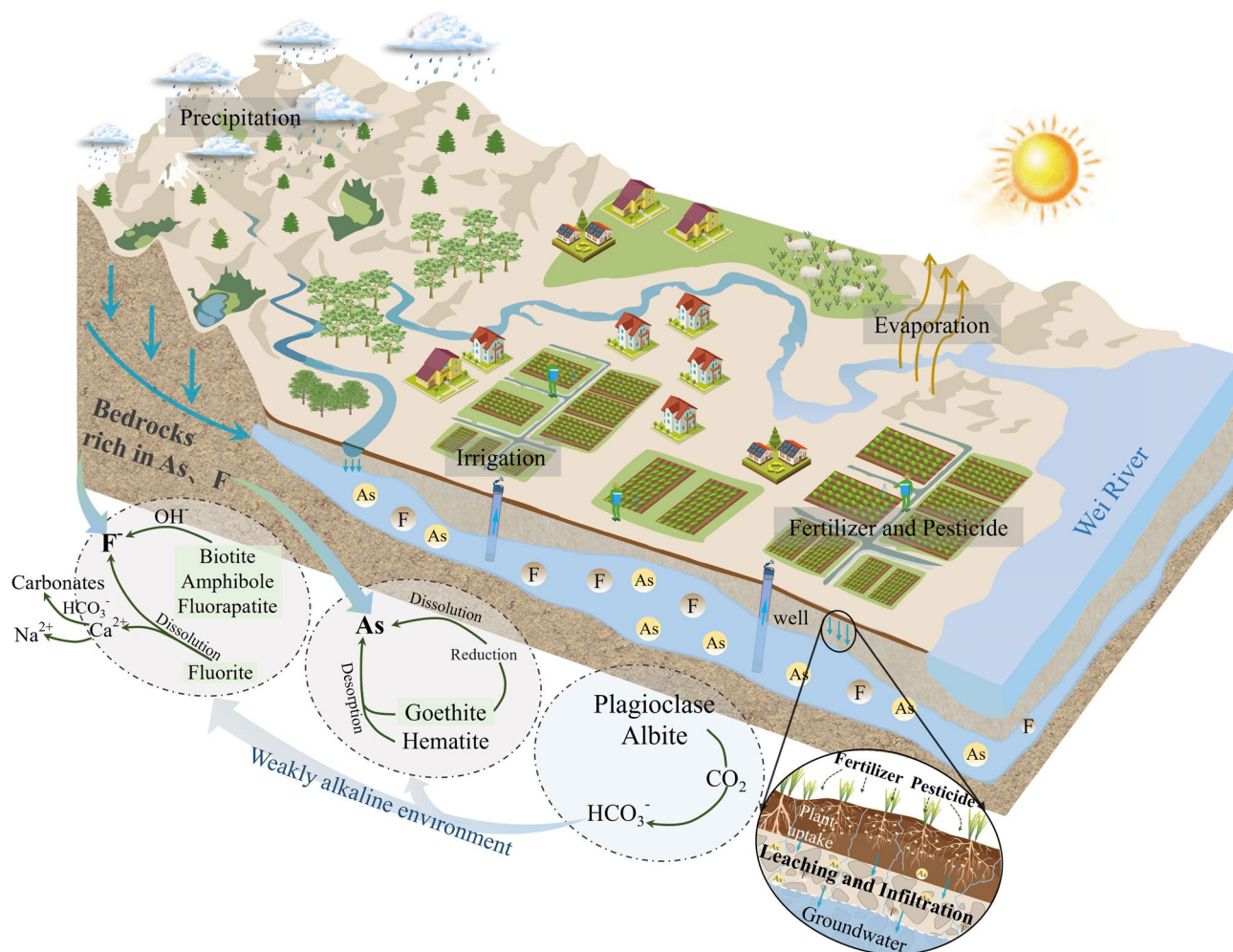


Fig. 7. Conceptual model of the enrichment mechanisms of As and F in groundwater.

of samples, respectively, indicating that arsenic pollution in groundwater can potentially pose hazards to human health. The average non-carcinogenic risk via oral ingestion is approximately 400 times higher than that via dermal contact, indicating that oral ingestion is the dominant exposure pathway. Furthermore, a significantly higher non-carcinogenic risk from arsenic was observed in children compared to adults.

The results of non-carcinogenic risk assessment for F in groundwater are shown in Table S6. For adults, the non-carcinogenic risk values ranged from 0.23 to 2.82 (average: 0.96), while for children, these values were higher, ranging from 0.44 to 5.41 (average: 1.85). The proportions of samples exceeding the safety threshold were 41.18% for adults and 80.39% for children, indicating that fluoride pollution poses a more widespread and severe health threat to children. This is consistent with the findings of Zhang et al.⁴¹. Similar to arsenic, oral ingestion was the predominant pathway. Comparatively, the non-carcinogenic risk of fluoride for both adults and children was substantially higher than that of arsenic.

For combined As-F exposure, the average non-carcinogenic risk for adults was 1.34 (maximum: 3.96), with 64.71% of water samples exceeding the safety threshold ($HQ = 1$). For children, the average non-carcinogenic risk was 2.57 (maximum: 7.60), with 96.08% of the water samples above the safety threshold. These findings indicate that more than 60% of groundwater in the study area poses potential health threats to both adults and children, with children facing markedly higher risks from As-F co-contamination. In summary, As and F co-contamination constitutes a significant non-carcinogenic risk to both adults and children.

Carcinogenic risk of arsenic Carcinogenic risk assessment for As in the groundwater of the study area is presented in Table S7. In contrast to non-carcinogenic risk patterns, adults exhibited approximately threefold higher carcinogenic risk for arsenic than children, with oral ingestion as the dominant pathway. For children, carcinogenic risks via both oral ingestion and dermal contact were below the acceptable threshold (1×10^{-4}). For adults, the carcinogenic risk values from dermal contact (3.55×10^{-8} to 6.04×10^{-7}) were below the threshold, but the carcinogenic risk values from oral ingestion ranged from 1.44×10^{-5} to 2.45×10^{-4} , with 17.65% of samples exceeding the acceptable level. This indicates that drinking arsenic-contaminated groundwater may pose a carcinogenic risk to adults.

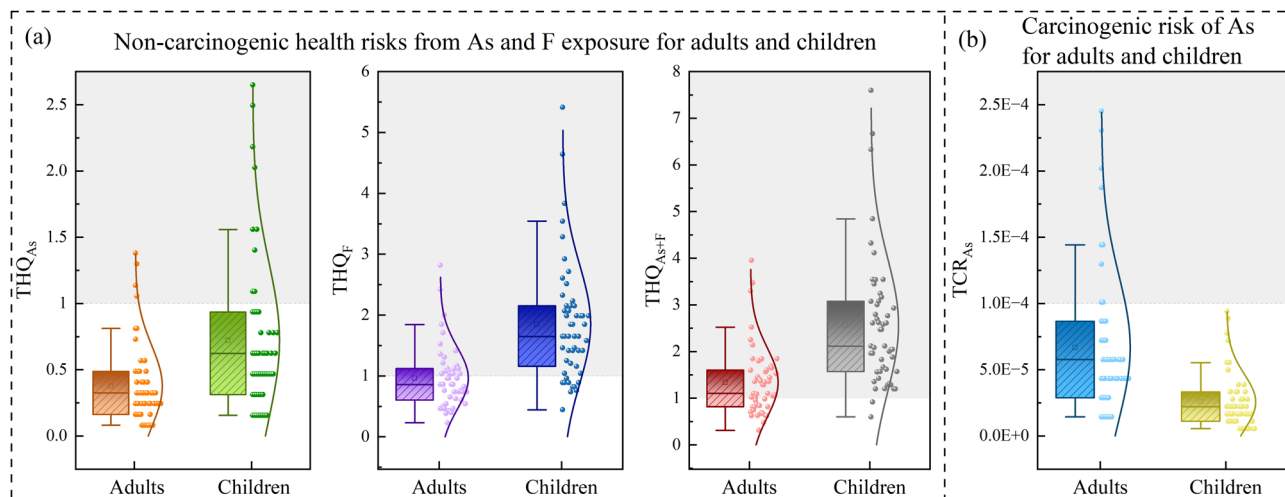


Fig. 8. Box plots showing (a) non-carcinogenic health risks from As and F exposure for adults and children; (b) carcinogenic risk of As for adults and children.

Spatial distribution of human health risks Spatial patterns of health risks were mapped using Surfer (Fig. S7). For adults, non-carcinogenic risks from arsenic were primarily concentrated in the northern and southeastern parts of Linwei District and the central part of Pucheng County (Fig. S7a). These zones occupied an area of 14.47 km², accounting for 1.41% of the study area. In addition to the aforementioned areas, carcinogenic risks extended to northwest Linwei District and portions of Yanliang District (Fig. S7e). These zones occupied a total area of 138.89 km², accounting for 13.56% of the study area. For children, except Lintong District and Dali County, other counties and districts had varying degrees of non-carcinogenic risks, covering an area of 187.41 km² (18.30% of the study area), with high-risk zones mainly distributed in Linwei District and Pucheng County (Fig. S7b). No carcinogenic risk for arsenic to children was detected in the study area (Fig. S7f).

Fluoride-related non-carcinogenic risks to adults occurred in all counties and districts except Lintong District (affected area: 446.87 km², accounting for 43.62% of the study area) (Fig. S7c), while for children, almost the entire study area had non-carcinogenic risks (affected area: 970.07 km², accounting for 94.70% of the study area) (Fig. S7d). The affected area of As-F non-carcinogenic risk is substantially larger for children than for adults, with high-risk areas mainly concentrated in Linwei District and Pucheng County.

Uncertainty assessment based on monte carlo simulation

Monte Carlo simulations (10,000 iterations; random seed = 2026) yielded health risk estimates at the 5% and 95% confidence levels for arsenic and fluoride pollution in groundwater of the study area (Fig. 9). The results show that both non-carcinogenic and carcinogenic risks posed by arsenic and fluoride pollution in groundwater to adults and children follow a log-normal distribution. For arsenic, probabilities of non-carcinogenic risk exceeding the safety threshold were 2.56% (adults) and 13.99% (children), with mean values of 0.3119 and 0.5976, respectively [Figs. 9(a) and (b)]. For carcinogenic risks, the probabilities of exceeding the safety threshold were 11.96% (adults) and 0.00% (children), with average values of 5.5410×10^{-5} and 2.1231×10^{-5} , respectively [Figs. 9(e) and (f)].

For fluoride, probabilities of non-carcinogenic risk exceeding the safety threshold were 24.21% (adults) and 68.90% (children), with mean values of 0.7874 and 1.5099, respectively [Figs. 9(c) and (d)]. These results confirm that fluoride poses a higher non-carcinogenic risk to humans than arsenic. Both adults and children in the study area face health risks from drinking groundwater contaminated with arsenic and fluoride, with children experiencing higher non-carcinogenic risks.

The Monte Carlo predictions are consistent with the aforementioned health risk assessment results. In addition, compared with traditional methods, uncertainty assessment using Monte Carlo simulation can present predicted results under different probability distributions, identify and reduce uncertainties in risk assessment, and provide a scientific basis for decision-making aimed at reducing health risks.

Sensitivity analysis

Sensitivity reflects the relationship between variance and risk, where the sign (positive/negative) indicates the direction of correlation. A greater absolute sensitivity value denotes a stronger influence on the risk assessment⁶³. The results of the sensitivity analysis for different population groups are presented in Fig. 10. Except for BW, all other parameters showed positive correlations with health risks in both adults and children. The total non-carcinogenic risks for adults and children were primarily influenced by F in groundwater, with sensitivity values of 62.2% and 59.6% (corresponding to correlation coefficients of 0.75 and 0.73), respectively. As, EF, IR, and BW had relatively smaller impacts. However, EF and IR were more sensitive to non-carcinogenic risk in children [Figs. 10(a) and (b)]. As shown in Figs. 10(c) and (d), arsenic in groundwater exhibited higher sensitivity for carcinogenic risk in adults. Compared with adults, children's IR showed greater sensitivity to carcinogenic risk.

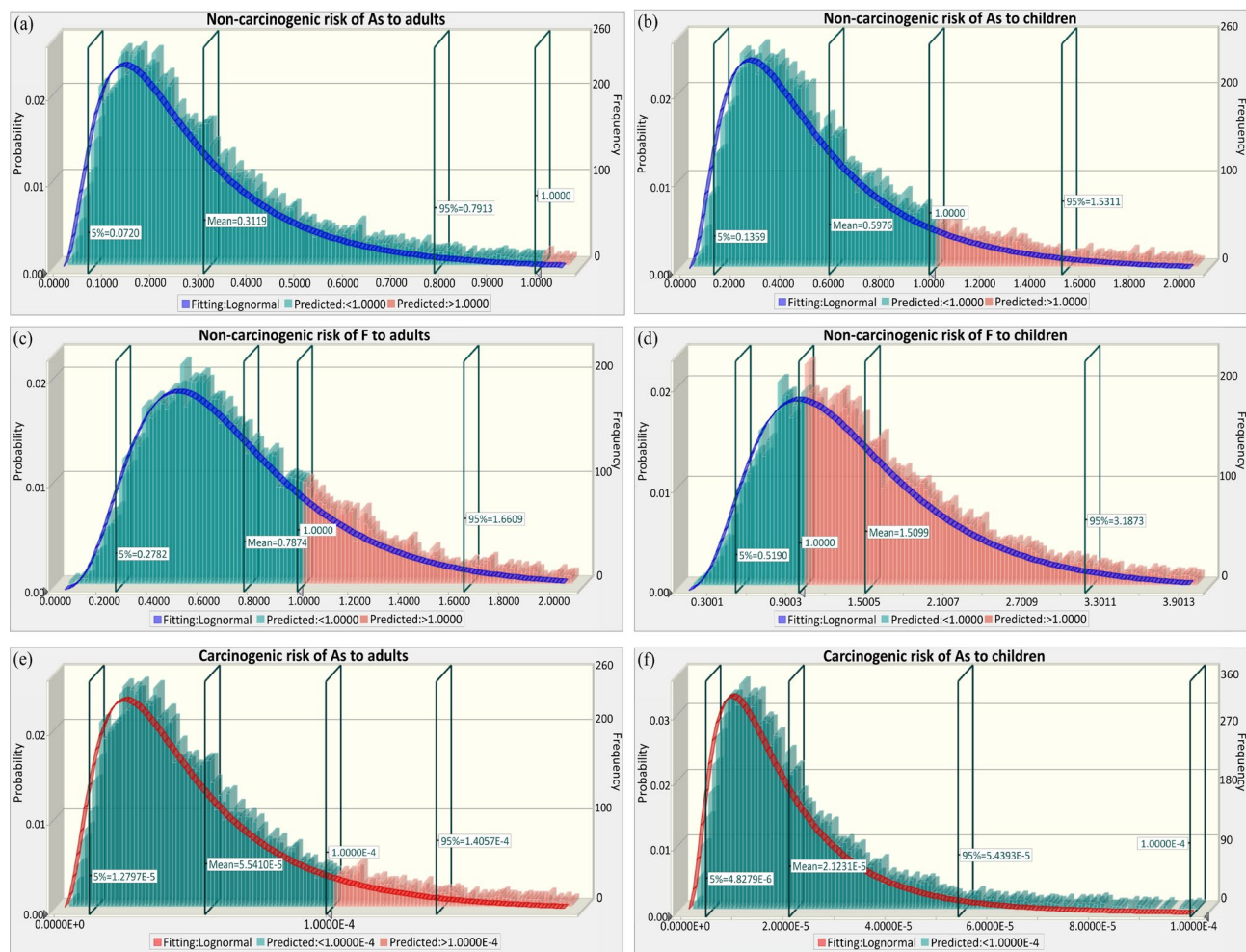


Fig. 9. Simulation plot of health risk uncertainty for adults and children.

Removal methods of arsenic and fluorine

To address arsenic and fluoride contamination in irrigation district groundwater, reliable and low-cost treatment technologies suitable for local conditions, such as coagulation/filtration and bio-composite adsorption techniques, can be adopted. Coagulation/filtration is a widely used water treatment technique that has been implemented in multiple countries for the removal of arsenic and fluoride from water^{53,64}. Regarding bio-composite adsorption, biosorbents derived from orange and apple peels have demonstrated promising effectiveness in adsorbing both arsenic and fluoride⁶⁵. Compared to conventional adsorbents such as activated carbon and alumina, these biosorbents offer advantages not only in technical performance but also in economic feasibility⁶⁶. Promoting such economically viable technologies is of significant importance for ensuring residents' drinking water safety. Moreover, drinking water safety interventions should prioritize children in high-risk areas.

Conclusions

This study investigated As and F contamination in the groundwater of the Jiaokou Irrigation District, Shaanxi Province, China. The spatial distribution and enrichment mechanisms of these contaminants were systematically analyzed through Self-organizing map (SOM), K-means clustering, and isotope tracing (δD and $\delta^{18}O$). A health risk assessment model incorporating Monte Carlo simulation was developed to quantify the associated risks for adults and children. The key findings are summarized below.

The groundwater was characterized as weakly alkaline brackish water. SOM cluster analysis classified it into three hydrochemical types, including $Cl-SO_4-Na$, HCO_3-Na , and HCO_3-Ca . Co-contamination of As and F was prevalent, with exceedance rates for As, F, and both contaminants reaching 11.80%, 43.75%, and 9.80% of the samples, respectively. Spatially, elevated concentrations were primarily concentrated in Linwei District and Pucheng County, covering 3.61% and 54.43% of the total area, respectively. Integrated analysis confirmed the correlation between As and F, with both also linking to pH and CO_3^{2-} .

Groundwater was primarily recharged by atmospheric precipitation and irrigation infiltration. During infiltration, water-rock interactions in a weakly alkaline environment facilitate As and F enrichment. This environment reduced As adsorption onto minerals, inhibited fluorite precipitation, and promoted the replacement of F^- with OH^- in fluoride-bearing minerals. The dissolution of silicate minerals (e.g., plagioclase,

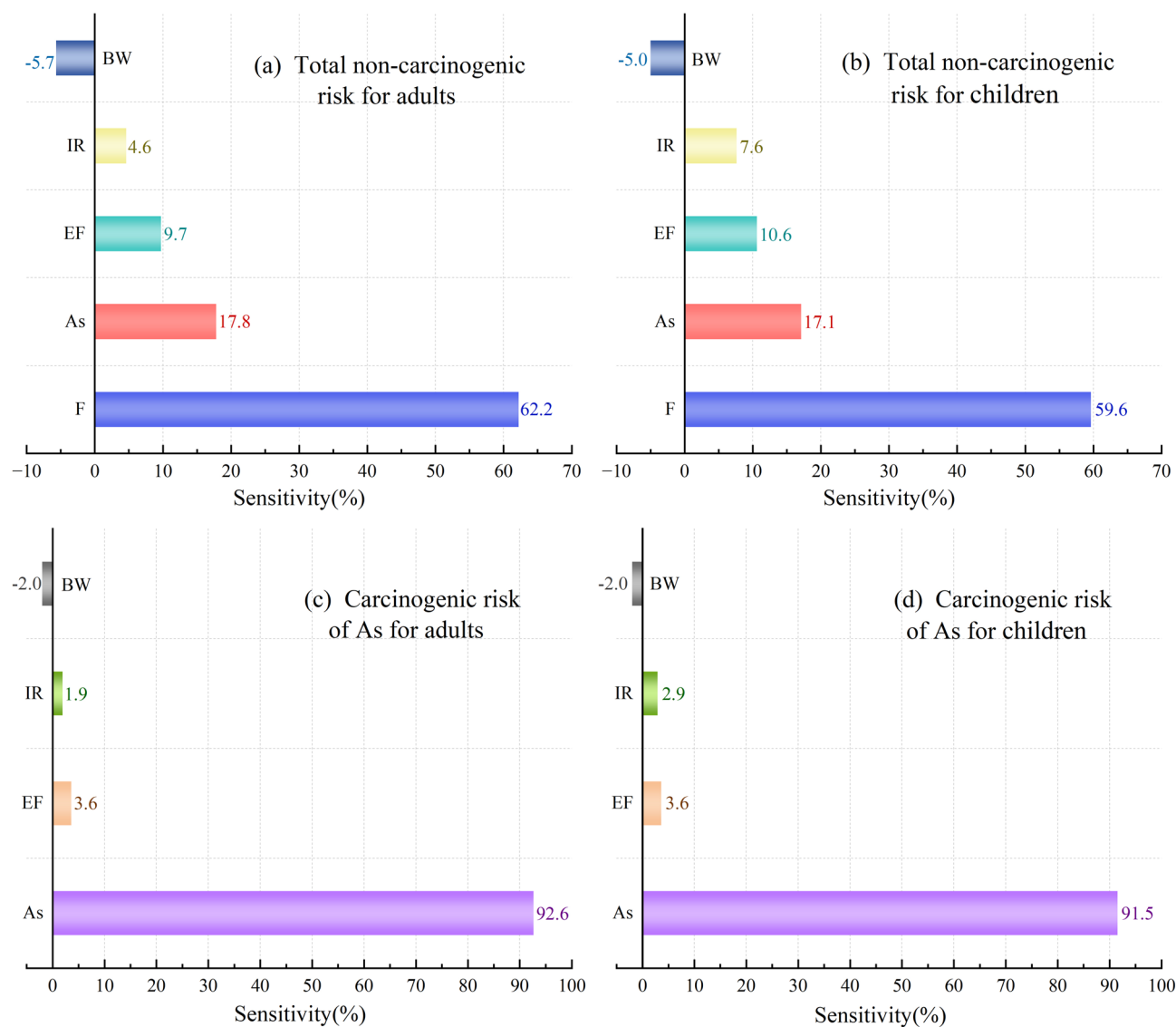


Fig. 10. Sensitivity analysis of (a) HQ_{Total} for adults; (b) HQ_{Total} for children; (c) CR_{As} for adults; (d) CR_{As} for children.

albite) further enhanced groundwater alkalinity and prolonged water-rock contact. Evaporation and the intensive use of fertilizers and pesticides also contributed to the enrichment of arsenic and fluoride.

The health risk assessment revealed that children in the study area were susceptible to non-carcinogenic risks from both As and F, whereas adults faced a higher carcinogenic risk from As. Oral ingestion was the primary exposure pathway for both risk types. F presented a substantially higher non-carcinogenic risk than As. Spatially, high-risk areas were primarily clustered in Linwei District and Pucheng County, which was highly consistent with the pollution distribution. Furthermore, sensitivity analysis identified F concentration as the dominant factor for non-carcinogenic risk, and As concentration as the key variable for its carcinogenic risk.

This work provides a comprehensive understanding of the distribution and enrichment mechanisms of arsenic–fluoride co-contamination in the groundwater of the Jiaokou Irrigation District, thereby informing future strategies for groundwater protection and public health risk mitigation in arid-semiarid irrigation regions. However, this study has limitations in the investigation of other heavy metals/trace metals within the research area. Future studies should incorporate XRD analysis to clarify the role of these elements in the enrichment processes of arsenic and fluoride.

Data availability

The datasets generated and/or analysed during the current study are not publicly available due to confidentiality but are available from the corresponding author on reasonable request. The 51 groundwater samples utilized in this study were collected with the permission of local residents. Specifically, samples were obtained from residential properties or privately owned irrigation wells. Consequently, the data constitutes original research material generated by the authors and is free from any associated disputes.

Received: 3 December 2025; Accepted: 25 February 2026

Published online: 05 March 2026

References

- Zhang, Q. Y., Xu, P. P. & Qian, H. Groundwater quality assessment using improved water quality index (WQI) and human health risk (HHR) evaluation in a semi-arid region of northwest China. *Expo Health*. **12** (3), 487–500 (2020).
- Zhang, Q. Y. et al. Salinization of shallow groundwater in the Jiaokou Irrigation District and associated secondary environmental challenges. *Sci. Total Environ.* **908**, 168445 (2024).
- Jia, H. et al. Alterations to groundwater chemistry due to modern water transfer for irrigation over decades. *Sci. Total Environ.* **717**, 137170 (2020).
- Famiglietti, S. J. The global groundwater crisis. *Nat. Clim. Chang.* **4** (11), 945–948 (2014).
- Li, P. Y., Li, X. Y., Meng, X. Y., Li, M. N. & Zhang, Y. T. Appraising Groundwater Quality and Health Risks from Contamination in a Semiarid Region of Northwest China. *Expo Health*. **8** (3), 361–379 (2016).
- Qiao, J. B. et al. Distributions of arsenic and other heavy metals, and health risk assessments for groundwater in the Guanzhong Plain region of China. *Environ. Res.* **181**, 108957 (2020).
- Padilla-Reyes, A. D. et al. Arsenic and fluoride in groundwater triggering a high risk: Probabilistic results using Monte Carlo Simulation and Species Sensitivity Distribution. *Chemosphere* **359**, 142305 (2024).
- Muñoz, O. M. et al. Geochemistry of naturally occurring arsenic in groundwater and surface-water in the southern part of the Poopó Lake basin, Bolivian Altiplano. *Groundw. Sustain. Dev.* **2**, 104–116 (2016).
- Wang, Y. X. et al. Genesis of geogenic contaminated groundwater: As, F and I. *Crit. Rev. Environ. Sci. Technol.* **51** (24), 2895–2933 (2021).
- Kumar, A. J., Subodh, P. C., Asish, S., Towfiqul, M. R. A. I. & Dipankar, R. Hydrogeochemical evaluation and corresponding health risk from elevated arsenic and fluoride contamination in recurrent coastal multi-aquifers of eastern India. *J. Clean. Prod.* **369**, 133150 (2022).
- Muhammad, S., Abdur, Y. M. R., Abida, F. & Naseem, R. M. Arsenic and fluoride co-exposure through drinking water and their impacts on intelligence and oxidative stress among rural school-aged children of Lahore and Kasur districts, Pakistan. *Environ. Geochem. Health.* **44** (11), 3929–3951 (2022).
- Wang, X. G. et al. Spatial distribution and influencing factors, and health risks of high fluoride and arsenic groundwater in Eastern Henan Plain, Northern China. *Environ. Geochem. Health.* **47** (10), 433 (2025).
- Alarcón-Herrera, M. T. et al. Co-occurrence, possible origin, and health-risk assessment of arsenic and fluoride in drinking water sources in Mexico: Geographical data visualization. *Sci. Total Environ.* **698**, 134168 (2020).
- Gao, Y. Y. et al. Cumulative health risk assessment of multiple chemicals in groundwater based on deterministic and Monte Carlo models in a large semiarid basin. *J. Clean. Prod.* **352**, 131567 (2022).
- He, X. D. et al. Poor groundwater quality and high potential health risks in the Datong Basin, northern China: research from published data. *Environ. Geochem. Health.* **43** (2), 791–812 (2021).
- Dong, S. G. et al. Hydro-geochemical control of high arsenic and fluoride groundwater in arid and semi-arid areas: A case study of Tumochuan Plain, China. *Chemosphere* **301**, 134657 (2022).
- Díaz, L. S. et al. Control factors of the spatial distribution of arsenic and other associated elements in loess soils and waters of the southern Pampa (Argentina). *Catena* **140**, 205–216 (2016).
- Fantong, W. Y. et al. Geochemical provenance and spatial distribution of fluoride in groundwater of Mayo Tsanaga River Basin, Far North Region, Cameroon: implications for incidence of fluorosis and optimal consumption dose. *Environ. Geochem. Health.* **32** (2), 147–163 (2010).
- Abiye, T., Bybee, G. & Leshomo, J. Fluoride concentrations in the arid Namaqualand and the Waterberg groundwater, South Africa: Understanding the controls of mobilization through hydrogeochemical and environmental isotopic approaches. *Groundw. Sustain. Dev.* **6**, 112–120 (2018).
- Su, H., Kang, W. D., Li, Y. R. & Li, Z. Fluoride and nitrate contamination of groundwater in the Loess Plateau, China: Sources and related human health risks. *Environ. Pollut.* **286**, 117287 (2021).
- Qiu, H. L., Gui, H. R., Xu, H. F., Cui, L. & Yu, H. Occurrence, controlling factors and noncarcinogenic risk assessment based on Monte Carlo simulation of fluoride in mid-layer groundwater of Huaibei mining area, North China. *Sci. Total Environ.* **856**, 159112 (2023).
- Li, Y., Bi, Y. H., Mi, W. J., Xie, S. L. & Ji, L. Land-use change caused by anthropogenic activities increase fluoride and arsenic pollution in groundwater and human health risk. *J. Hazard. Mater.* **406**, 124337 (2021).
- Salifu, A., Petrusovski, B., Ghebremichael, K., Buamah, R. & Amy, G. Multivariate statistical analysis for fluoride occurrence in groundwater in the Northern region of Ghana. *J. Contam. Hydrol.* **140**, 34–44 (2012).
- Yin, Z. Y., Luo, Q. K., Wu, J. F., Xu, S. H. & Wu, J. C. Identification of the long-term variations of groundwater and their governing factors based on hydrochemical and isotopic data in a river basin. *J. Hydrol.* **592**, 125604 (2021).
- Zeng, J. W. et al. Driving factor, source identification, and health risk of PFAS contamination in groundwater based on the self-organizing map. *Water Res.* **267**, 122458 (2024).
- Zhang, Y. B. et al. Hydrogeochemical analysis and groundwater pollution source identification based on self-organizing map at a contaminated site. *J. Hydrol.* **616**, 128839 (2023).
- Appukkuttan, A. et al. Exploring hydrochemical drivers of drinking water quality in a tropical river basin using self-organizing maps and explainable AI. *Water Res.* **284**, 123884 (2025).
- Appukkuttan, A. & Reghunath, R. Machine learning-based groundwater potential mapping and factor analysis in tropical lateritic terrains using self-organizing maps and random forest. *Environ. Monit. Assess.* **197** (12), 1340 (2025).
- Zhang, Y. et al. Human health risk assessment of groundwater arsenic contamination in Jinghui irrigation district. *China J Environ. Manag.* **237**, 163–169 (2019).
- WHO. *Guidelines for drinking water quality: fourth edition incorporating the first addendum* (World Health Organization, 2017).
- Zhang, Q. Y. et al. Effect of hydrogeological conditions on groundwater nitrate pollution and human health risk assessment of nitrate in Jiaokou Irrigation District. *J. Clean. Prod.* **298**, 126783 (2021).
- Xu, B., Zhang, Y. & Wang, J. F. Hydrogeochemistry and human health risks of groundwater fluoride in Jinhuai irrigation district of Wei river basin, China. *Hum. Ecol. Risk Assess.* **25** (1–2), 230–249 (2019).
- Zhu, Y. W. et al. A hydrochemical and isotopic approach for source identification and health risk assessment of groundwater arsenic pollution in the central Yinchuan basin. *Environ. Res.* **231**, 116153 (2023).
- Ganyaglo, Y. S. et al. Groundwater fluoride contamination and probabilistic health risk assessment in fluoride endemic areas of the Upper East Region, Ghana. *Chemosphere* **233**, 862–872 (2019).
- Jiang, C. L. et al. Distribution, source and health risk assessment based on the Monte Carlo method of heavy metals in shallow groundwater in an area affected by mining activities, China. *Ecotoxicol. Environ. Saf.* **224**, 112679 (2021).
- Kaur, L., Rishi, S. M. & Siddiqui, U. A. Deterministic and probabilistic health risk assessment techniques to evaluate non-carcinogenic human health risk (NHHR) due to fluoride and nitrate in groundwater of Panipat, Haryana, India. *Environ. Pollut.* **259**, 113711 (2020).

37. Yan, J. H., Chen, J. S. & Zhang, W. Q. A new probabilistic assessment process for human health risk (HHR) in groundwater with extensive fluoride and nitrate optimized by non parametric estimation method. *Water Res.* **243**, 120379 (2023).
38. Liu, R. *Study on the Rational Allocation of Water Resources in Jiaokou Irrigation District* (Chang'an University, 2014). (in Chinese).
39. Li, G. D., Qian, H. L., Zhang, X. G., Guo, S. H. & Zhang, M. Investigation and Analysis of Drinking Fluorosis in Weinan City. *Chin. J. Endemiol.* **28**, 68–69 (2009).
40. Liu, R. Study on Transference and Transform Simulation of Fluoride in Groundwater and Relation between Fluoride and Human Body Health in Dali Region, Guanzhong Basin. In: Chang'an University (in Chinese) (2009).
41. Zhang, Q. Y., Xu, P. P., Qian, H. & Yang, F. X. Hydrogeochemistry and fluoride contamination in Jiaokou Irrigation District, Central China: Assessment based on multivariate statistical approach and human health risk. *Sci. Total Environ.* **741**, 140460 (2020).
42. Zhang, Q. Y. *Study on Spatial-Temporal Evolution Analysis As Well as Early Warning and Response of Groundwater Environment in Jiaokou Irrigation District* (Chang'an University, 2022). (in Chinese).
43. Xu, P. P., Zhang, Q. Y., Qian, H., Li, M. N. & Hou, K. Characterization of geothermal water in the piedmont region of Qinling Mountains and Lantian-Bahe Group in Guanzhong Basin, China. *Environ. Earth Sci.* **78**, 1–17 (2019).
44. Chen, J. et al. Hydrogeochemical Characteristics and Quality Assessment of Groundwater in an Irrigated Region, Northwest China. *Water* **11** (1), 96 (2019).
45. Kohonen, T. Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **43** (1), 59–69 (1982).
46. Nakagawa, K., Yu, Z. Q., Berndtsson, R. & Hosono, T. Temporal characteristics of groundwater chemistry affected by the 2016 Kumamoto earthquake using self-organizing maps. *J. Hydrol.* **582**, 124519 (2020).
47. Santos, R. M., Roisenberg, A., Iwashita, F. & Roisenberg, M. Hydrogeochemical spatialization and controls of the Serra Geral Aquifer System in southern Brazil: a regional approach by self-organizing maps and k-means clustering. *J. Hydrol.* **591**, 125602 (2020).
48. Hou, K. et al. Evolutionary process of soil grain-size fractals in irrigation areas of northern China: a pedological analysis of ancient irrigation wisdom. *Catena* **257**, 109160 (2025).
49. Vesanto, J. & Alhoniemi, E. Clustering of the self-organizing map. *IEEE Trans. Neural Netw. Learn. Syst.* **11** (3), 586–600 (2000).
50. Kumari, M. & Kumar, A. Human health risk assessment of antibiotics in binary mixtures for finished drinking water. *Chemosphere* **240**, 124864 (2020).
51. Xu, P. P. et al. Hydrogeochemical genesis mechanism and secondary environmental challenges of groundwater in southern Jingyang, Shaanxi Province. *Geol. Bull. China.* **44** (5), 811–824 (2025).
52. Nicolli, H. B., Bundschuh, J., García, J. W., Falcón, C. M. & Jean, J. S. Sources and controls for the mobility of arsenic in oxidizing groundwaters from loess-type sediments in arid/semi-arid dry climates—Evidence from the Chaco-Pampean plain (Argentina). *Water Res.* **44** (19), 5589–5604 (2010).
53. Alarcón-Herrera, T. M. et al. Co-occurrence of arsenic and fluoride in groundwater of semi-arid regions in Latin America: Genesis, mobility and remediation. *J. Hazard. Mater.* **262**, 960–969 (2013).
54. Rashid, A. et al. Fluoride prevalence in groundwater around a fluorite mining area in the flood plain of the River Swat, Pakistan. *Sci. Total Environ.* **635**, 203–215 (2018).
55. Dey, R. K. et al. Hydrogeochemical processes controlling the high fluoride concentration in groundwater: a case study at the Boden block area, Orissa, India. *Environ. Monit. Assess.* **184** (5), 3279–3291 (2012).
56. Gibbs, R. J. Mechanisms controlling world water chemistry. *Science* **170** (3962), 1088–1090 (1970).
57. Qian, C. et al. Hydrogeochemical characterization and suitability assessment of groundwater in an agro-pastoral area, Ordos Basin, NW China. *Environ. Earth Sci.* **75**, 1–16 (2016).
58. Wang, J. X., Jiang, Y. T., Wang, H. M., Huang, Q. Q. & Deng, H. B. Groundwater irrigation and management in northern China: status, trends, and challenges. *Int. J. Water Resour. Dev.* **36** (4), 670–696 (2020).
59. Song, H., Li, P. Y., Su, F. M., Wang, D. & Ren, X. F. Identification and apportionment of shallow groundwater nitrate pollution in Weining Plain, northwest China, using hydrochemical indices, nitrate stable isotopes, and the new Bayesian stable isotope mixing model (MixSIAR). *Environ. Pollut.* **298**, 118852 (2022).
60. Jia, H. et al. Fluoride Occurrence and Human Health Risk in Drinking Water Wells from Southern Edge of Chinese Loess Plateau. *Int. J. Environ. Res. Public Health.* **16** (10), 1683 (2019).
61. Hartley, N. T., Macdonald, J. A., McGrath, P. S. & Zhao, F. J. Historical arsenic contamination of soil due to long-term phosphate fertiliser applications. *Environ. Pollut.* **180**, 259–264 (2013).
62. Shahid, M. et al. A meta-analysis of the distribution, sources and health risks of arsenic-contaminated groundwater in Pakistan. *Environ. Pollut.* **242**, 307–319 (2018).
63. Mohammad, M. et al. Temporospatial variation and health risk assessment of trihalomethanes (THMs) in drinking water (northwest Iran). *Environ. Sci. Pollut. Res.* **28** (7), 8168–8180 (2020).
64. Camacho, M. L., Gutiérrez, M., Alarcón-Herrera, T. M., Villalba, L. M. & Deng, S. G. Occurrence and treatment of arsenic in groundwater and soil in northern Mexico and southwestern USA. *Chemosphere* **83** (3), 211–225 (2011).
65. Robledo, P. A., López, G. M., Morales, A. C. G. & Reynoso, C. L. Arsenic and Fluoride in Groundwater, Prevalence and Alternative Removal Approach. *Processes* **9** (7), 1191 (2021).
66. Manna, S., Roy, D., Adhikari, B., Thomas, S. & Das, P. Biomass for water defluoridation and current understanding on biosorption mechanisms: A review. *Environ. Prog. Sustain. Energy.* **37** (5), 1560–1572 (2018).

Acknowledgements

The completion of this article was inseparable from the contributions of all authors. Their support is gratefully acknowledged. The authors would like to thank all the reviewers who participated in the review, as well as MJEditor (www.mjeditor.com) for providing English editing services during the preparation of this manuscript.

Author contributions

Qiyang Zhang: Conceptualization, Data curation, Formal analysis, Methodology, Funding acquisition, Investigation, Resources, Writing- Review & editing. Aihui Wei: Investigation, Methodology, Visualization, Writing-Original draft. Panpan Xu: Writing- Review & editing, Methodology. Hui Qian: Resources, Supervision. Jianxi Ren: Project administration.

Funding

This study was financially supported by the National Natural Science Foundation of China (Grant No. 42302292, 42572319), Outstanding Youth Program of Xi'an University of Science and Technology (No. 6310224014), Talent Introduction Project of Xi'an University of Science and Technology (No. 2050123005), and Open Fund of Key Laboratory of Subsurface Hydrology and Ecological Effect in Arid Region of the Ministry of Education, Chang'an University (No. 300102295508).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-026-42293-6>.

Correspondence and requests for materials should be addressed to Q.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026