water

MDPI

Article

Health Risk Assessment of Groundwater in Cold Regions Based on Kernel Density Estimation–Trapezoidal Fuzzy Number–Monte Carlo Simulation Model: A Case Study of the Black Soil Region in Central Songnen Plain

Jiani Li [1,2], Yu Wang [1,2], Jianmin Bian [1,2,*], Xiaoqing Sun [1,2] and Xingrui Feng [1,2]

[1] Key Laboratory of Groundwater Resources and Environment, Ministry of Education, Jilin University, Changchun 130021, China; jnli24@mails.jlu.edu.cn (J.L.); wangyu1224@jlu.edu.cn (Y.W.); sunxq13@jlu.edu.cn (X.S.); fengxr24@mails.jlu.edu.cn (X.F.)
[2] Jilin Provincial Key Laboratory of Water Resources and Environment, Jilin University, Changchun 130021, China
* Correspondence: bianjm@jlu.edu.cn

Abstract

The quality of groundwater, a crucial freshwater resource in cold regions, directly affects human health. This study used groundwater quality monitoring data collected in the central Songnen Plain in 2014 and 2022 as a case study. The improved DRASTICL model was used to assess the vulnerability index, while water quality indicators were selected using a random forest algorithm and combined with the entropy-weighted groundwater quality index (E-GQI) approach to realize water quality assessment. Furthermore, self-organizing maps (SOM) were used for pollutant source analysis. Finally, the study identified the synergistic migration mechanism of NH_4^+ and Cl^-, as well as the activation trend of As in reducing environments. The uncertainty inherent to health risk assessment was considered by developing a kernel density estimation–trapezoidal fuzzy number–Monte Carlo simulation (KDE-TFN-MCSS) model that reduced the distribution mis-specification risks and high-risk misjudgment rates associated with conventional assessment methods. The results indicated that: (1) The water chemistry type in the study area was predominantly HCO_3^-–Ca^{2+} with moderately to weakly alkaline water, and the primary and nitrogen pollution indicators were elevated, with the average NH_4^+ concentration significantly increasing from 0.06 mg/L in 2014 to 1.26 mg/L in 2022, exceeding the Class III limit of 1.0 mg/L. (2) The groundwater quality in the central Songnen Plain was poor in 2014, comprising predominantly Classes IV and V; by 2022, it comprised mostly Classes I–IV following a banded distribution, but declined in some central and northern areas. (3) The results of the SOM analysis revealed that the principal hardness component shifted from Ca^{2+} in 2014 to Ca^{2+}–Mg^{2+} synergy in 2022. Local high values of As and NH_4^+ were determined to reflect geogenic origin and diffuse agricultural pollution, whereas the Cl^- distribution reflected the influence of de-icing agents and urbanization. (4) Through drinking water exposure, a deterministic evaluation conducted using the conventional four-step method indicated that the non-carcinogenic risk (*HI*) in the central and eastern areas significantly exceeded the threshold (*HI* > 1) in 2014, with the high-*HI* area expanding westward to the central and western regions in 2022; local areas in the north also exhibited carcinogenic risk (*CR*) values exceeding the threshold (*CR* > 0.0001). The results of a probabilistic evaluation conducted using the proposed simulation model indicated that, except for children's *CR* in 2022, both *HI* and *CR* exceeded acceptable thresholds with 95% probability. Therefore, the proposed assessment method can provide a basis for improved groundwater pollution zoning and control decisions in cold regions.

check for updates

Academic Editor: Paolo Madonia

Received: 25 August 2025
Revised: 13 October 2025
Accepted: 14 October 2025
Published: 16 October 2025

Citation: Li, J.; Wang, Y.; Bian, J.; Sun, X.; Feng, X. Health Risk Assessment of Groundwater in Cold Regions Based on Kernel Density Estimation–Trapezoidal Fuzzy Number–Monte Carlo Simulation Model: A Case Study of the Black Soil Region in Central Songnen Plain. *Water* **2025**, *17*, 2984. https://doi.org/10.3390/w17202984

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Keywords: health risk assessment; kernel density estimation–trapezoidal fuzzy number–Monte Carlo simulation; self-organizing map; water quality assessment; groundwater vulnerability

1. Introduction

Groundwater plays a critical role in environmental ecology. In recent years, human activities have increasingly impacted groundwater systems as accelerated economic development worsens environmental degradation. Groundwater, a core component sustaining regional water cycles and ecological equilibrium, is particularly critical in the black soil region of central Songnen Plain. Characterized as a typical cold region, this area experiences distinct freeze–thaw cycles that alter aquifer permeability, accelerate pollutant migration, and trigger seasonal solute release, including the sudden input of de-icing agents in winter and the concentrated leaching of agricultural pollutants during spring. This significantly increases the complexity of groundwater chemical evolution as well as the uncertainty of associated health risk assessments [1]. Groundwater quality evaluation and health risk assessment are foundational aspects of groundwater protection, and assessing contamination susceptibility under hydrogeological and anthropogenic conditions is essential for pollution prevention.

Groundwater quality is a key determinant of a safe drinking water supply and essential to maintaining agricultural productivity. Groundwater quality evaluation techniques have developed considerably in recent years. The entropy-weighted groundwater quality index (E-GQI) provides a reliable tool for quickly assessing the suitability of groundwater for drinking purposes [2]. This method uses weights based on raw data to avoid the subjective biases of traditional methods, making it more objective and intuitive. However, E-GQI evaluation considers numerous water quality parameters that must be selected subjectively. Therefore, this study selected key water quality parameters based on feature importance rankings obtained using a random forest (RF) algorithm and incorporated a vulnerability index to provide a critical dimension of water quality evaluation [3]. This approach can more comprehensively identify potential areas of high-vulnerability, low-quality groundwater risk. Groundwater vulnerability measures the sensitivity of a groundwater system to human activities or natural factors, such as geology, hydrogeology, and pollutant discharge conditions and physicochemical properties. The DRASTIC model is a widely used and mature model for groundwater vulnerability assessment [4]. In China, its application can be divided into several stages spanning from the mid-1990s, when it was based on localized parameter calibration, to the late 1990s, when fuzzy mathematics membership functions were introduced to construct a fuzzy DRASTIC model that addressed the issue of discontinuous representation of geological boundary transition zones [5], to the 21st century, when further improvements integrated land use type indicators and applied ArcGIS for statistical analysis to overcome subjectivity [6,7].

Drinking contaminated groundwater can seriously damage human health, and quantifying pollutant risk levels is a key step in assessing exposure probability and managing potential health risks. The United States Environmental Protection Agency's (USEPA's) classic "four-step method" (FSM) is a commonly used deterministic method for human health risk assessment (HHRA) that effectively distinguishes between carcinogenic and non-carcinogenic risks under various pollution scenarios and exposure pathways. It has been widely adopted owing to its simplicity and comprehensive evaluation capabilities [8]. However, the FSM is subject to limitations when dealing with uncertainty, primarily owing to the randomness of exposure parameters as well as limited sample sizes that fail to capture

variability in the risk assessment process, resulting in significant discrepancies between the calculated results and actual risks [9]. Previous stochastic health risk assessments have typically described the uncertainty in pollutant concentrations and exposure parameters using triangular stochastic simulation (TSS). Although the membership function associated with a TSS has the mathematical advantage of fixed endpoints, it struggles to reflect the gradual distribution characteristics of parameters. Currently, the use of improved trapezoidal fuzzy numbers (TFNs) adjusts the distribution boundaries to improve precision when describing the core intervals and extreme value ranges of parameters. However, the determination of the core interval still relies on subjective judgment or empirical assignment, which cannot represent multi-modal, skewed, or heavy-tailed features in actual data [10,11]. Therefore, this study proposed a kernel density estimation–trapezoidal fuzzy number–Monte Carlo simulation (KDE-TFN-MCSS) health risk assessment model. Note that the accurate estimation of the probability distribution function (PDF) for each variable in the model is necessary for the application of a Monte Carlo simulation (MCSS). Critically, the non-parametric kernel density estimation (KDE) method can evaluate data distribution characteristics entirely from samples and estimate the density functions of arbitrary shapes, reducing potential errors caused by the incorrect selection of probability distribution models, known as distribution mis-specification risks. The proposed KDE-TFN-MCS model significantly improves the probability density function estimation accuracy, thereby reducing the misjudgment of high-risk areas and minimizing the impact of parameter uncertainty. As a result, KDE provides the MCSS with more robust risk representation capabilities that are especially well-suited to estimating the probability distributions of complex random variables [12].

The acceleration of urbanization has introduced widespread ammonia nitrogen contamination in the groundwater of the central Songnen Plain owing to intensive agricultural activities, yet groundwater remains an important water resource in the region as an irreplaceable resource for drinking water, irrigation, and industry, Indeed, groundwater is often the primary source of drinking water and a key guarantee for agricultural irrigation, especially in rural areas. However, existing studies on cold-region groundwater quality have primarily evaluated the impact of natural freeze–thaw processes on water chemistry and have not conducted systematic analyses of the coupled effects of anthropogenic pollution (e.g., widespread ammonia nitrogen and rising Cl^- levels) and natural freeze–thaw cycles to determine the specific health risks posed by inherent geogenic contaminants, such as As. Therefore, the groundwater pollution characteristics and health risk evolution mechanisms in the Songnen Plain, a typical cold region, require further investigation. This study used groundwater quality monitoring data collected in the central Songnen Plain in 2014 and 2022 to analyze the water chemistry characteristics and their causes. First, an RF algorithm was used to determine water quality evaluation factors, which were combined with vulnerability indices and entropy-weighted indices to realize groundwater quality assessment. Next, self-organizing maps (SOMs) were employed to identify key risk factors [13], then the results obtained using the FSM and improved KDE-TFN-MCSS models to conduct groundwater health risk assessment were compared. The proposed KDE-TFN-MCSS model was shown to reduce distribution mis-specification risks in the evaluation process and improve the accuracy of the results for complex regional groundwater systems, thereby providing multi-dimensional support for groundwater risk management in cold regions. The research approach applied in this study is detailed in Figure 1.

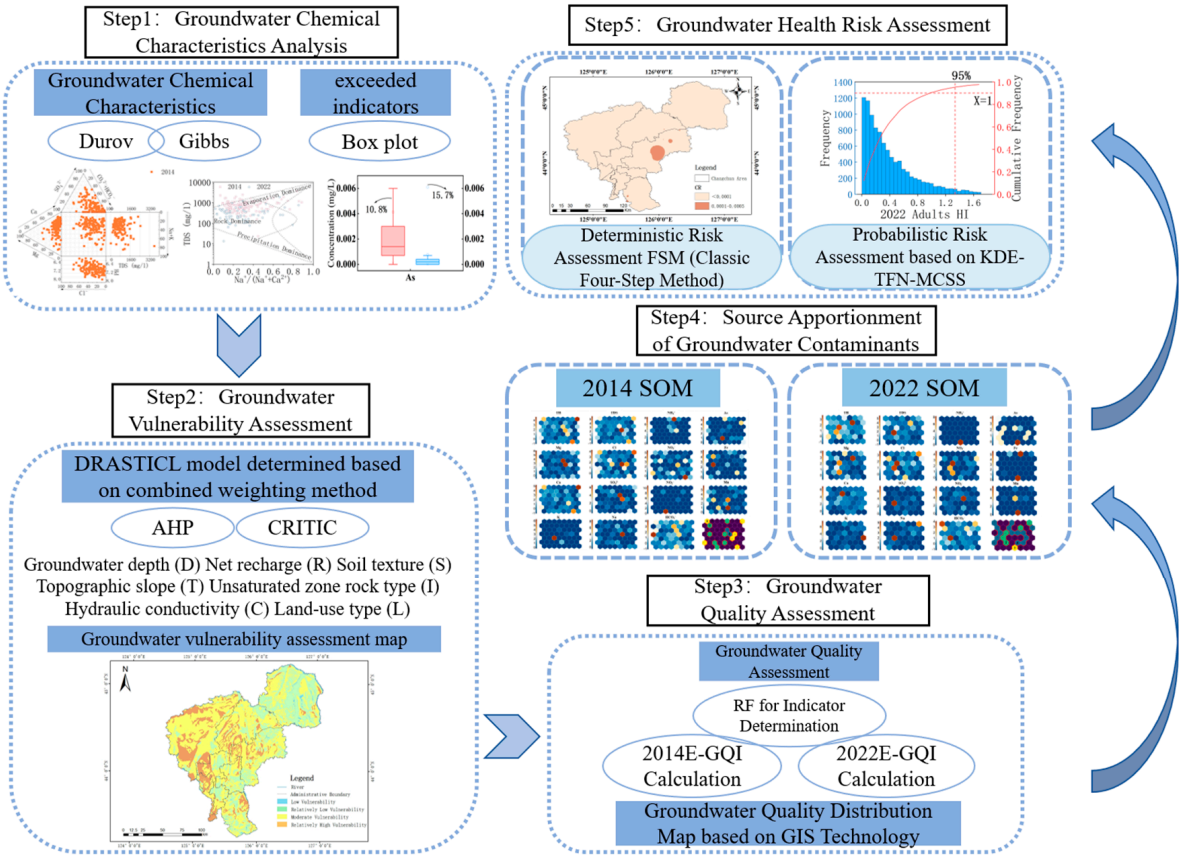


Figure 1. Research approach.

2. Materials and Methods

2.1. Study Area

The central Songnen Plain is an alluvial core area connecting the eastern high plain and western low plain, serving as a key transition zone controlling water/salt migration. Changchun is situated in the central Songnen Plain at the coordinates 43°05′~45°15′ N, 124°18′~127°05′ E and was selected as a representative location in this study. This region covers an area of 24.744 km^2 with terrain that primarily consists of plains and terraces decreasing in elevation from east to west. It has a semi-arid, semi-humid continental climate influenced by the monsoons of the northern temperate zone. The average annual temperature is approximately 4.6 °C and exhibits considerable diurnal variation. Winters are cold and dry, and spring temperatures rise quickly with little precipitation and high evaporation rates. The average annual precipitation of 577.3 mm is concentrated from May to September, and the potential evaporation is 752.8 mm, decreasing from the northwest to southeast.

The primary water system of the target region is a part of the Drinking Horse River basin in the Songhua River basin. The primary water-bearing layer is the Quaternary aquifer, which can be classified into unconfined and confined aquifers based on burial conditions. The thickness of the Quaternary aquifer increases gradually from the southeast to northwest of the region, with lithology primarily consisting of gravel and sand with local deposits of clay and silt. The Quaternary unconfined aquifer is widely distributed with a thickness generally between 15 and 50 m, though it may be thicker in some areas. The aquifer's sediment granularity is coarser in the Yitong and Drinking Horse River basins due to fluvial deposition, which improves permeability. The hydraulic gradient of the unconfined aquifer decreases with the water table depth, which generally ranges from 1.5 to

15 m in most areas. Critically, this type of aquifer is easily affected by surface pollution infiltration. The geographical location of the study area and sampling point distribution within are shown in Figure 2.

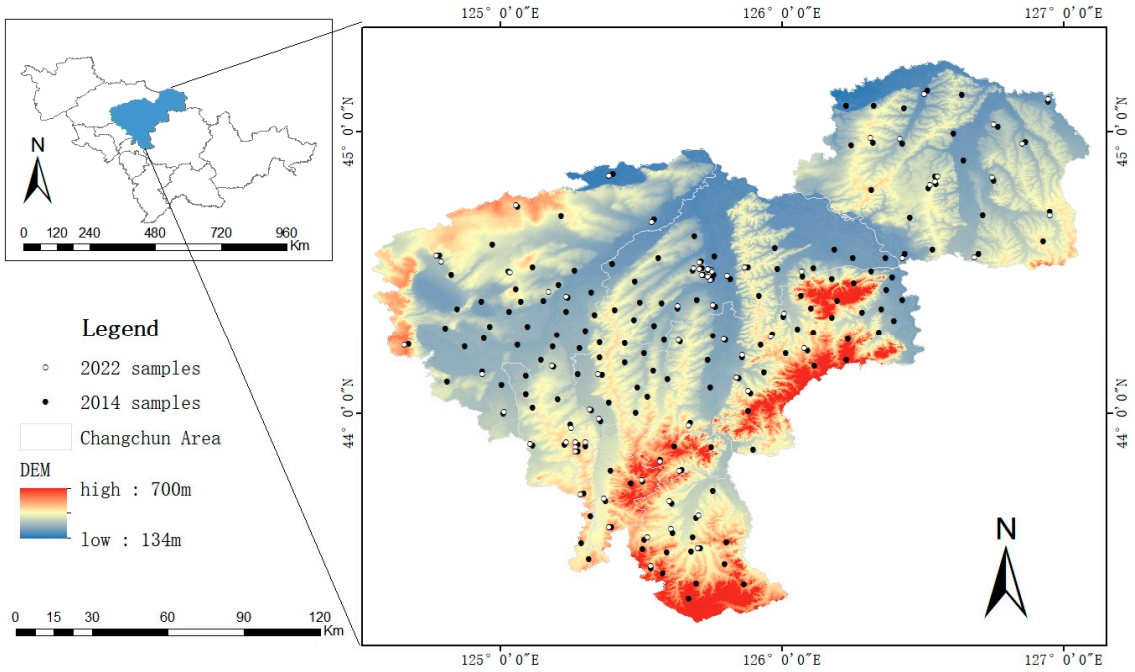


Figure 2. Location of the study area and distribution of sampling points.

2.2. Data Source

This study used groundwater test data collected by our research group during previous work in the study area in 2014 and 2022. All groundwater samples were collected during the wet season (July and August) from similar distribution depths; the sampling sites used in 2022 were a subset of those used in 2014. Data were collected from 209 unconfined sampling points of Quaternary aquifer, and 23 water quality indicators were evaluated comprising routine indicators (pH, total hardness (TH), and total dissolved solids (TDS)), eight major ions (K^+, Ca^{2+}, Na^+, Mg^{2+}, HCO_3^-, CO_3^{2-}, Cl^-, and SO_4^{2-}), three forms of nitrogen (NO_3^-, NO_2^-, and NH_4^+), geogenic indicators (F^-, Fe, Mn, and As), and heavy metals (Hg, Cr^{6+}, Pb, S, and Al). The methods used to obtain these indicators are detailed in Table 1.

Table 1. Water indicators and associated test methods.

Indicator	Test Method	Indicator	Test Method
TH	Ethylenediaminetetraacetic acid titration [14]	HCO_3^-, CO_3^{2-}	Acid–base titration [15]
TDS	Gravimetric method [16]	NH_4^+	Nessler's reagent spectrophotometry [17]
pH	Electrode method [18]	NO_2^-	N-(1-Naphthyl)ethylenediamine dihydrochloride spectrophotometry [19]
Ca^{2+}, Mg^{2+}, K^+, Na^+, Cl^-, SO_4^{2-}, NO_3^-, F^-	Ion chromatography [20,21]	Fe, Mn, As	Inductively coupled plasma optical emission spectrometry [22]
Hg	Atomic fluorescence spectrometry [23]	Pb	Graphite furnace atomic absorption spectrometry [24]
Cr^{6+}	Diphenylcarbazide spectrophotometry [25]	S^{2-}	Methylene blue spectrophotometry [26]
Al	Spectrophotometry (chrome azurol S) [27]		

2.3. Groundwater Evaluation Methods

2.3.1. Improved DRASTICL Model

The DRASTIC model, which is widely used for groundwater vulnerability assessment, is based on seven key evaluation parameters: groundwater depth (D), net recharge (R) directly related to climate and rainfall, aquifer rock type (A), soil texture (S) and topographic slope (T) surface conditions, unsaturated zone rock type (I), and hydraulic conductivity (C). Because the study area lies in a transition zone between mountain and plain regions, the variation in aquifer thickness is not a major factor influencing groundwater vulnerability; therefore, the analysis of aquifer thickness (A) was excluded from this case study. The dynamic effects of freeze–thaw cycles in cold regions on aquifer permeability were considered in this study by adding a land use type (L) parameter to quantify human disturbances such as de-icing agents and agricultural activities, creating a more regionally adaptive DRASTICL model [28,29] defined as follows:

$$DRASTICLindex = D_r D_w + R_r R_w + S_r S_w + T_r T_w + I_r I_w + C_r C_w + L_r L_w \quad (1)$$

where the subscript r represents the rating value of each parameter and the subscript w represents the weight of each parameter.

Thus, the *DRASTICL* vulnerability index can be obtained by calculating and summing the products of the rating values and weights; the higher the index value, the greater the groundwater vulnerability and risk of contamination.

2.3.2. Random Forest Method

The RF algorithm provides an efficient ensemble learning method that comprises dual random sampling of the observations and feature variables in the modeling dataset using the "bootstrap" resampling technique, which generates multiple new sample sets by repeatedly sampling with replacements from the original data [30]. This approach typically creates subsets of the same size as the original dataset, with some observations potentially appearing multiple times and others potentially omitted. This process is generally repeated 1000 to 10,000 times to estimate the uncertainty or distribution of statistical metrics and thereby randomly extract multiple subsets from the original dataset to build a series of decision trees. During the decision tree splitting process, the mean decreased Gini coefficient (*MDG*) is calculated to assess the quality of feature split points as follows:

$$Gini(i) = 1 - \sum_{j=1}^{|y|} \left(X(i,j) / \sum_{j=1}^{n} X(i,j) \right)^2, i \in d \quad (2)$$

where y denotes the total number of distinct categories or classes in the target variable, $X(i,j)$ represents the i-th index of the j-th variable and d represents the node number of the regression tree [31].

Each decision tree iterates through splitting until it generates a classification result that matches the tree properties. Finally, the RF algorithm integrates the classification wisdom of all decision trees and uses a voting mechanism to determine the final classification result, improving classification accuracy and reliability.

2.3.3. Entropy-Weighted Groundwater Quality Index (E-GQI)

The E-GQI uses water quality indicators to represent the influence of geological and anthropogenic activities on overall water quality. It assesses water quality based on standard allowable limits and has gained widespread application among scholars owing to its reliability and objectivity. However, the E-GQI considers many different water quality indicators, the selection of which can be subjective. Therefore, this study selected the key indicators based on the importance ranking of features from the RF algorithm and

coupled them with a vulnerability index to form a comprehensive evaluation system. The weight of each indicator was set using the entropy weighting method to quantify its dispersion through information entropy. This weighting method relies entirely on the statistical distribution of the water quality monitoring dataset, avoiding the risk of subjective bias inherent to conventional methods. The E-GQI approach applied in this study can be expressed as follows:

$$X_i = \frac{w_i}{\sum_i^k w_i} \tag{3}$$

$$M_i = \frac{m_i \times 100}{A_i} \tag{4}$$

$$GQ_{\mathrm{i}} = X_i \times M_i \tag{5}$$

$$GQI = \sum_{i=1}^{k} GQ_i \tag{6}$$

where w_i denotes the weight of indicator i, X_i denotes its relative value, M_i denotes its evaluated quality level, m_i denotes its measured value, A_i denotes its permissible standard value, GQ_i denotes its quality index, and k denotes the number of evaluated indicators. This study classified GQI values of <50, 50–100, 100–200, 200–300, and >300 as corresponding to water quality categories of "Good," "Fair," "Moderate," "Poor," and "Very Poor," respectively [32]. The water quality indicators and their classification limits adopted in this study are summarized in Table 2.

Table 2. Water indicators and classification limits [33].

Indicator \ Lmits	Class I	Class II	Class III	Indicator \ Lmits	Class I	Class II	Class III
TH	≤150	≤300	≤450	SO_4^{2-}	≤50	≤150	≤250
TDS	≤300	≤500	≤1000	HCO_3^-		-	
pH		6.5 ≤ pH ≤ 8.5		NH_4^+	≤0.02	≤0.1	≤0.5
Ca^{2+}		-		NO_3-	≤2	≤5	≤20
Mg^{2+}		-		NO_2^-	≤0.01	≤0.1	≤1
K^+		-		F^-	≤1	≤1	≤1
Na^+	≤100	≤150	≤200	Fe	≤0.1	≤0.2	≤0.3
Cl^-	≤50	≤150	≤250	Mn	≤0.05	≤0.05	≤0.1
				As	≤0.001	≤0.001	≤0.01

2.3.4. Self-Organizing Map (SOM)

The SOM is an unsupervised learning artificial neural network that learns from data in the input space and maps high-dimensional input data to a low-dimensional discrete space (typically a two-dimensional grid) to provide data clustering and visualization. The SOM network structure consists of input and output layers. Continuous iterative training with a competitive learning algorithm gradually adjusts the weights of the output layer neurons until they accurately reflect the data distribution in the output layer [34]. The structure of a classic SOM network is shown in Figure 3. This study integrated the SOM with the K-means clustering method to overcome the limitations of traditional clustering techniques when identifying water chemistry features. In this application, K-means clustering optimizes the classification boundaries based on the SOM outputs to enhance the interpretability of the water chemistry dataset and provide critical risk factor selection criteria for the subsequent HHRA. Before generating an SOM, three key parameters must be set: (1) the number of neurons, (2) the error quantification parameter for evaluating the topology, and (3) the optimal number of clusters determined by silhouette scoring [35,36].

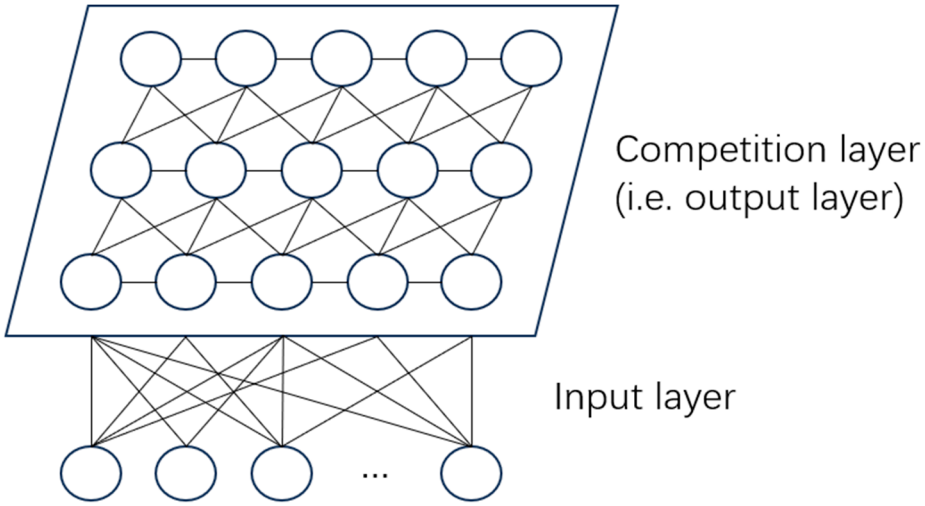


Figure 3. SOM network structure.

2.3.5. Groundwater Health Risk Assessment

The HHRA estimates the health risks of harmful pollutants entering the human body through various exposure pathways. Considering the primary uses of groundwater in the case study region as well as individual indicator variations in the collected data, this study analyzed the drinking water exposure pathway and used the KDE-TFN-MCSS and FSM to, respectively, perform probabilistic and deterministic HHRAs by calculating the non-carcinogenic risk (*HI*) and carcinogenic risk (*CR*) for children and adults [37,38]. The overall health risk from multiple pollutants was considered assuming that the risks from individual pollutants are independent and the total risk follows an additive relationship [39].

The distribution of mis-specification risks in TFN-MCSS was assessed in this study by coupling it with the KDE during groundwater HHRA. For a known sample size n, each sample in (X_1, X_2, …, X_n) was assumed to be independent of the others and follow the same continuous distribution $F(x)$. Under these conditions, the KDE for the unknown kernel density can be defined as follows:

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \tag{7}$$

where $\hat{f}_h(x)$ represents the function of the unknown kernel density; the constant h represents the smoothing coefficient of the smoothing curve, also known as the window width; and $K(\cdot)$ represents the kernel density given on the real number domain and is referred to as the kernel function for the unknown probability density.

The results of the KDE depend on the applied kernel function and window width. The specific choice of kernel function does not significantly affect the result; this study used the Gaussian kernel function. When the dataset is close to a normal distribution, the optimal $h = 1.06\sigma n^{-\frac{1}{5}}$ can be used; however, if the dataset is asymmetric or has multimodal features, the optimal h must be adjusted as follows:

$$h = 0.9\min\{\sigma, IQR/1.34\}\, n^{-\frac{1}{5}} \tag{8}$$

where σ represents the standard deviation of the dataset and IQR represents its interquartile range.

The PDFs of the relevant indicators were calculated based on the KDE results using the R 4.4.2 software to determine the boundaries of the core and support intervals, which were used to define the parameters a, b, c, and d for the TFN. Next, the obtained TFN data

were converted into interval values using the α-TST truncation technique. If $\forall a$ ($a \in [0, 1]$), TFN $\tilde{A}$ = (a, b, c, d) can be converted into interval values [40] as follows:

$$A^a = \{x|, \mu A(x) \geq a, x \in X\} \quad (9)$$

$$A^a = (a_L^a, a_R^a) = [(b - a)a + a, -(d - c)a + d] \quad (10)$$

in which the membership value a used in this calculation was 0.9.

The Oracle Crystal Ball 11.1.2.4 (OCB) software was used to define probability distributions for the input indicators leveraging the derived TFN parameters and their converted interval values. The software was configured to execute MCSS using a predefined number of trial iterations, typically 10,000, to propagate the uncertainties inherent in the model. Furthermore, OCB was instrumental in generating predictive contours that encompassed frequency distribution charts and cumulative probability plots for the obtained *HI* and *CR* outcomes.

The *HI* and *CR* for drinking water exposure were calculated as follows:

$$CDI_i = \frac{C_i \times IR \times EF \times ED}{BW \times AT} \quad (11)$$

$$CR_i = CDI_i \times SF \quad (12)$$

$$HQ_i = \frac{CDI_i}{RfD} \quad (13)$$

$$CR = \sum_{i=1}^{n} CR_i \quad (14)$$

$$HI = \sum_{i=1}^{m} HQ_i \quad (15)$$

where C_i is the concentration of groundwater contaminant i in mg/L, *IR* is the daily drinking water intake in L/d, *EF* is the exposure frequency, *ED* is the number of years of lifetime exposure to a toxic substance (a), *BW* is the body weight in kg, *AT* is the average exposure time in days, and *RfD* is the reference dose in mg/(kg·d). The specific values of the exposure parameters applied for the non-carcinogenic risk assessment are provided in Table 3.

Table 3. Exposure parameter values for non-carcinogenic risk.

Exposure Parameter	Reference Value	
	Adult	**Child**
IR (L/d)	[1.48, 1.72]	[0.96, 1.04]
BW (kg)	[61, 69]	[19, 23]
ED (years)	24	6
AT (days)	8760	2190
Rfd_{As} [mg/(kg·d)]	3×10^{-4}	3×10^{-4}
Rfd_{Cl^-} [mg/(kg·d)]	0.10	0.10
Rfd_{NH4^+} [mg/(kg·d)]	0.97	0.97

According to the USEPA health risk assessment standards, the reference threshold for *HI* is 1 and that for *CR* is 0.0001.

3. Results and Discussion

3.1. Groundwater Chemical Characteristics

3.1.1. Water Chemistry Indicator Statistics

The statistical variations in the water quality indicators over time were explored using box plots to analyze the routine indicators (TH, TDS, pH), cations (K^+, Ca^{2+}, Na^+, Mg^{2+}),

anions (HCO_3^-, Cl^-, SO_4^{2-}), nitrogen compounds (NO_3^-, NO_2^-, NH_4^+), and geogenic indicators (Fe, F^-, Mn, As), as shown in Figure 4. According to the thresholds specified in the Standard for Groundwater Quality (GB/T 14848-2017) (Table 2) [33], the monitoring data indicated that 37% of the case study area had groundwater quality below Class III, with the primary pollutants comprising TH, NO_3^-, NO_2, NH_4^+, Fe, and Mn. In both sampling periods, the indicators exceeding the Class III threshold comprised TH, TDS, pH, Na^+, SO_4^{2-}, NO_3^-, NO_2^-, NH_4^+, F^-, Fe, Mn, and As, but the exceedance rates were lower in 2022 than in 2014, reflecting overall improvement in water quality. Specifically, the contents of the routine indicators, eight major ions, three nitrogen compounds, and geogenic indicators all decreased between 2014 and 2022, with a notable reduction in groundwater salinity primarily owing to increased precipitation in 2022, which accelerated groundwater renewal and dilution. However, the concentrations of geogenic indicators remained above the standard limits, and the exceedance rate of As even increased, indicating a high level of pollution in the area. Furthermore, the monitoring data from 2022 indicated that the thresholds for all three nitrogen compounds were exceeded at 15 groundwater monitoring points. This pollution exhibited point-source diffusion following a strip-like migration pattern, with the average NH_4^+ concentration increasing significantly from 0.06 to 1.26 mg/L to exceed the Class III limit of 1.0 mg/L, indicating an accelerating trend in NH_4^+ pollution in local areas that can be attributed to increased agricultural runoff. Notably, NH_4^+ can, under certain conditions, convert into nitrite in water. The consumption of nitrate over the long term can produce strong carcinogenic substances that pose severe health risks to humans and harm aquatic life.

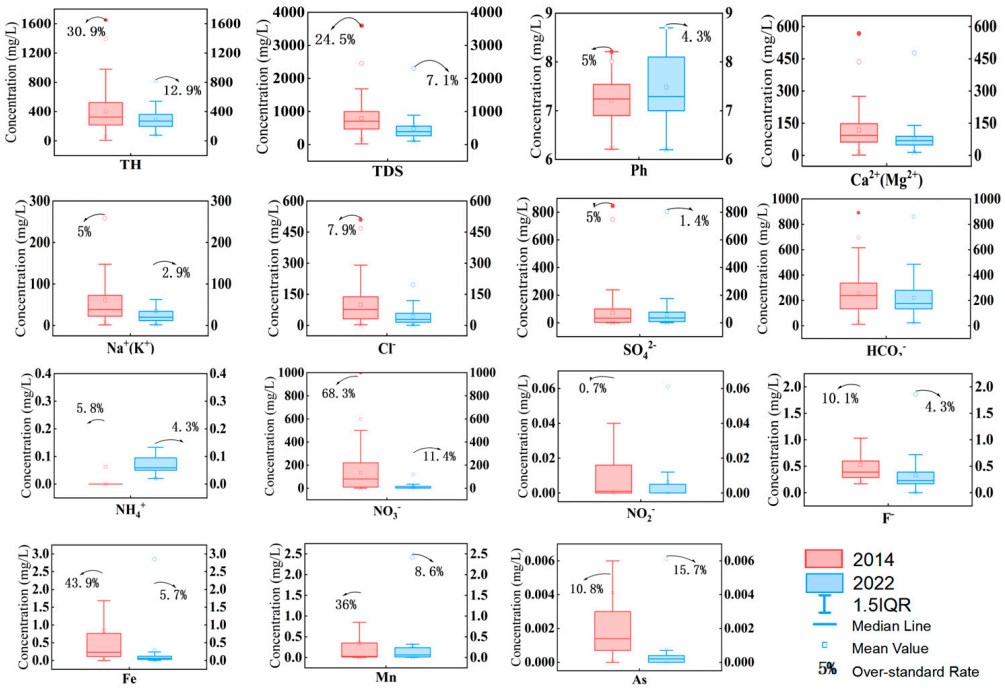


Figure 4. Statistics of groundwater chemical composition.

3.1.2. Water Chemistry Types

The Durov diagrams of the case study area's groundwater in 2014 and 2022, shown in Figure 5, were plotted using the Origin 2022 software to determine the groundwater chemistry types. The results indicated that the water chemistry type in this area in both years was primarily HCO_3^--Ca^{2+} with moderately weak alkaline water. The TH in the study area generally fell within the moderately high range, indicating hard water. Furthermore, the pH range shifted from 6.2–8.2 in 2014 to 6.2–8.8 in 2022, indicating increased alkalinity.

Finally, statistical analyses revealed that the groundwater TDS and salinity in 2022 were lower than those in 2014 owing to increased precipitation in the study area in 2022, which accelerated groundwater renewal and dilution.

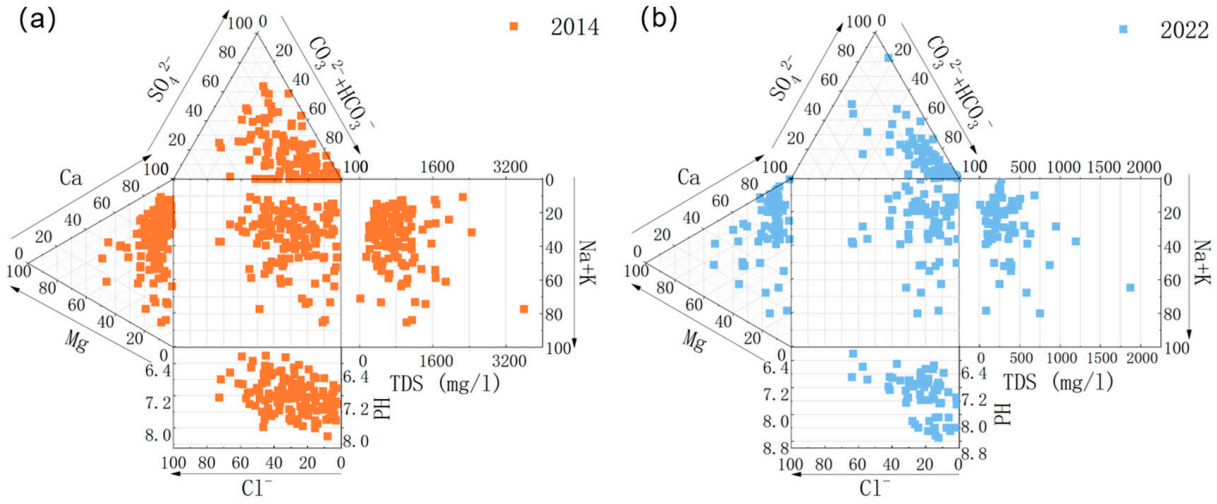


Figure 5. Groundwater Durov diagrams for (**a**) 2014 and (**b**) 2022.

3.1.3. Water Chemistry Origin Analysis

The Gibbs diagram can be used to determine the characteristics of groundwater evolution owing to evaporation crystallization, rock dissolution, and multi-factor influences [41], as shown in Figure 6. Rock dissolution and evaporation crystallization were the two primary factors influencing groundwater chemical evolution in the case study area. Ion ratio analysis indicated that groundwater chemistry was primarily controlled by silicate rock dissolution, with evaporation crystallization being a secondary factor. The observed decrease in groundwater salinity was partly attributed to reduced evaporation crystallization in 2022 compared to 2014.

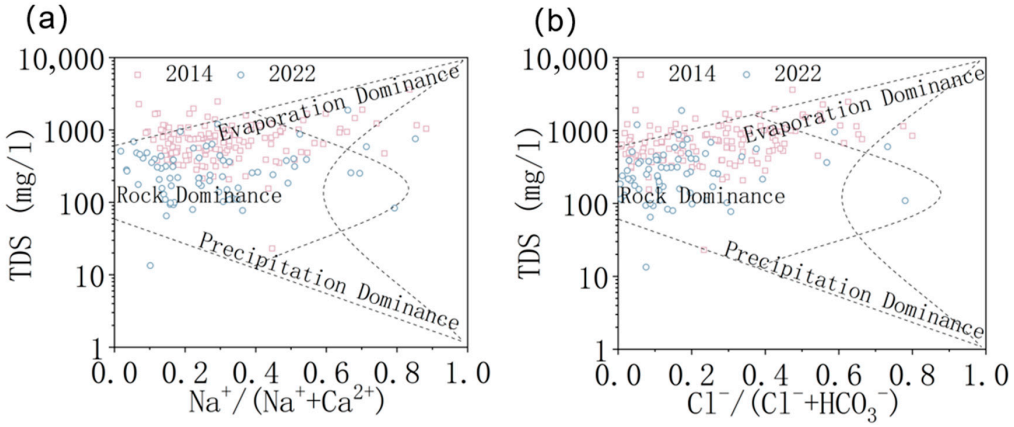


Figure 6. Groundwater Gibbs diagrams for (**a**) 2014 and (**b**) 2022.

3.2. Groundwater Quality Assessment

3.2.1. Groundwater Vulnerability

This study used hydrogeological surveys and field observation data in combination with regional geological and environmental characteristics to classify and assign values to groundwater vulnerability assessment parameters. The ArcGIS 10.8.2 geostatistical analysis module was employed to plot the rating map for each parameter, and a combined weighted sum approach was used to generate the groundwater vulnerability index distribution map using the DRASTICL model (Figure 7) [42]. The groundwater vulnerability in the study

area was classified into five levels from high to low. Higher and moderate vulnerability areas were the most common and widely distributed in the central and western regions; lower vulnerability areas were the second most common and primarily concentrated in the northeast; low vulnerability areas were the least common and only sporadically appeared in the northeast. The overall case study area did not exhibit any high-vulnerability regions. The results of this assessment were used as factors in the subsequent water quality evaluation.

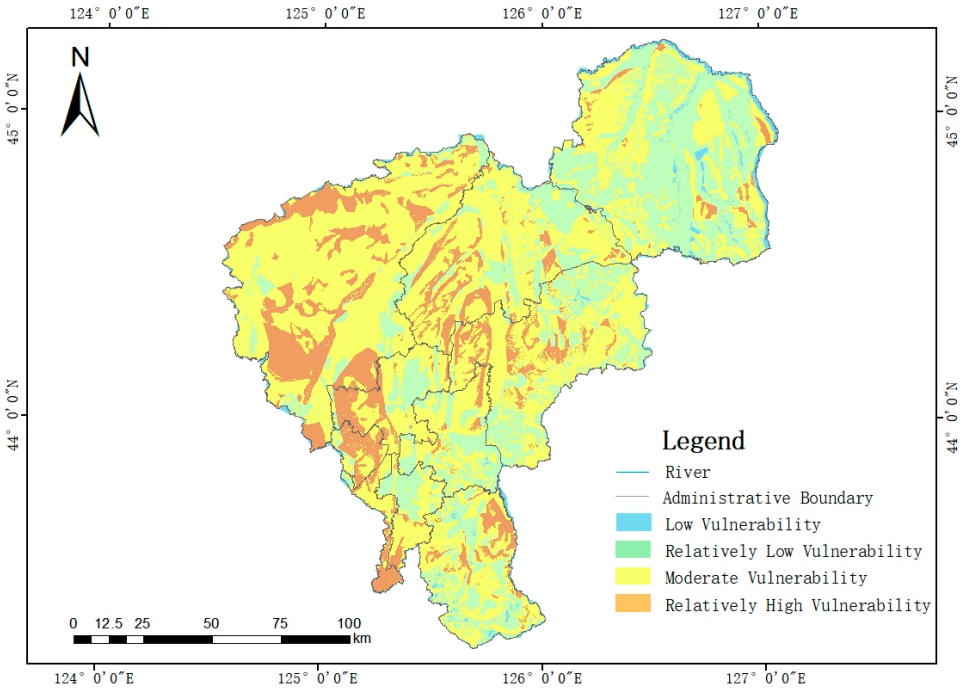


Figure 7. Groundwater vulnerability index distribution.

3.2.2. Selection of Groundwater Quality Indicators

A feature importance analysis of water quality indicators was performed using the RF algorithm in the SPSSPRO software as described in Section 2.3.2, resulting in the ranking shown in Figure 8. The top 16 ranked indicators were Mg, Mn, NH_4^+, As, NO_3^-, TDS, Fe, Cl^-, TH, Ca, SO_4^{2-}, K, F^-, HCO_3^-, Na, and NO_2^-. Among these indicators, the geogenic and nitrogen indicators had relatively high importance rankings. These key indicators and the vulnerability index discussed in Section 3.2.1 were selected as groundwater quality evaluation factors [43].

3.2.3. Groundwater Quality Assessment Results

The top ten groundwater quality indicators ranked by vulnerability index and feature importance were selected as evaluation factors, and their weights were determined using the E-GQI discussed in Section 2.3.3.

The weighted quality indices were aggregated using Equations (3)–(6), and spatial interpolation was performed using Kriging in ArcGIS 10.8 to generate the groundwater quality distribution maps shown in Figure 9 for spatial pattern analysis. In 2014, groundwater quality in the case study area was generally poor and predominantly classified into Classes IV and V, with limited occurrences of Class III in the western and northeastern regions and no instances of Classes I or II. By 2022, groundwater quality had generally improved, predominantly falling into Classes I–IV following a banded distribution from southwest to northeast. However, slight deterioration was observed in some central and northern zones, where Class V persisted.

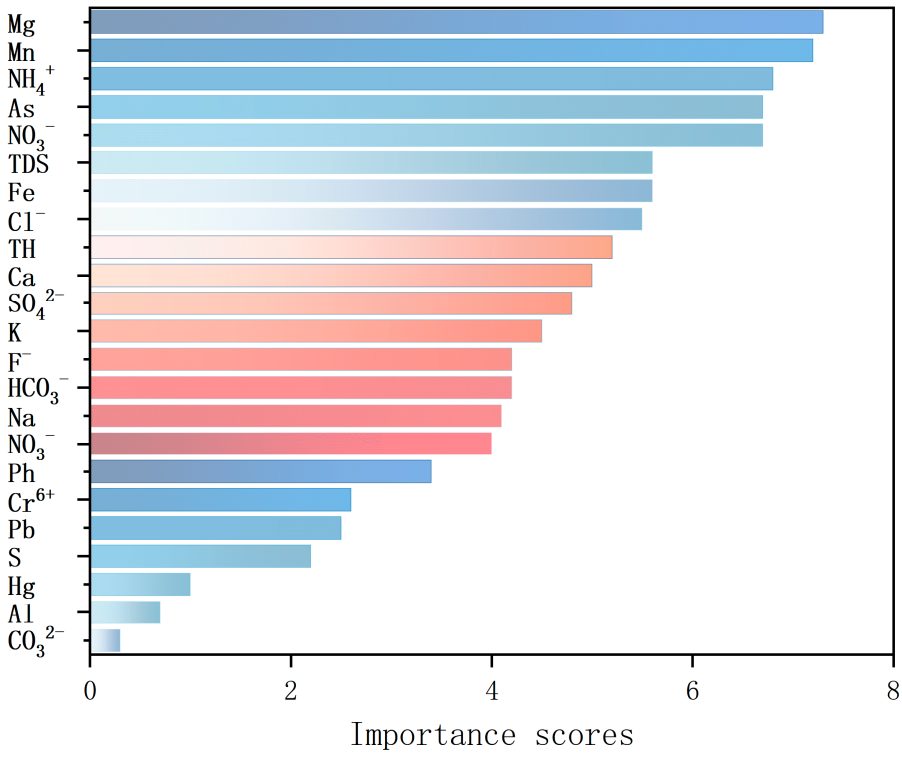


Figure 8. Feature importance ranking of water quality indicators.

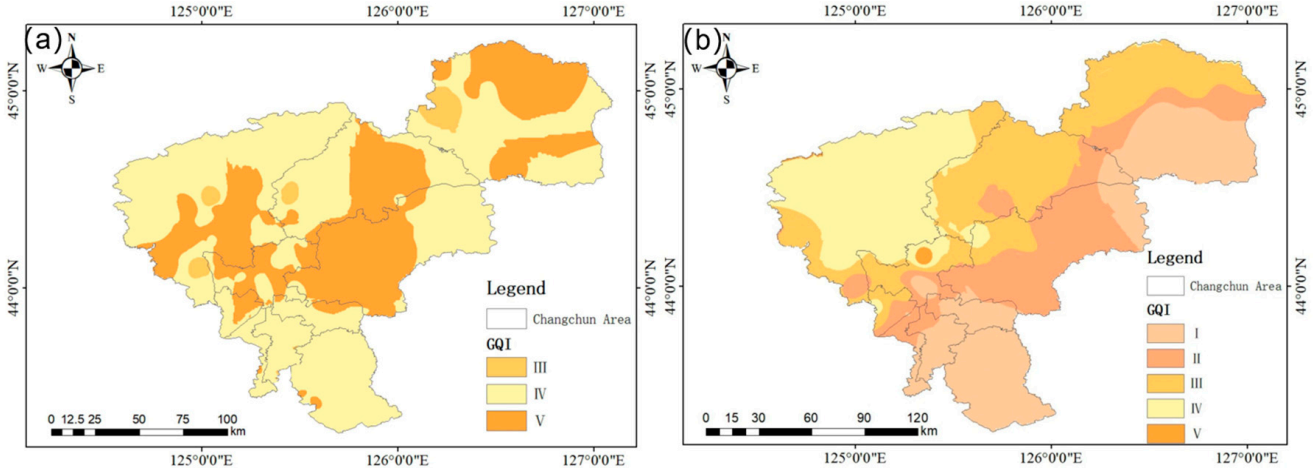


Figure 9. E-GQI water quality distribution maps for (**a**) 2014 and (**b**) 2022.

3.3. Groundwater Health Risk Assessment

3.3.1. Source Apportionment of Groundwater Contaminants

An SOM analysis was conducted using the R software with optimal neuron counts empirically determined by $5\sqrt{n}$. Topological structures comprising 7×8 and 6×7 were constructed using a learning rate of 0.7 with 10,000 iterations to generate neuron distribution maps for 2014 and 2022 (Figures 10 and 11).

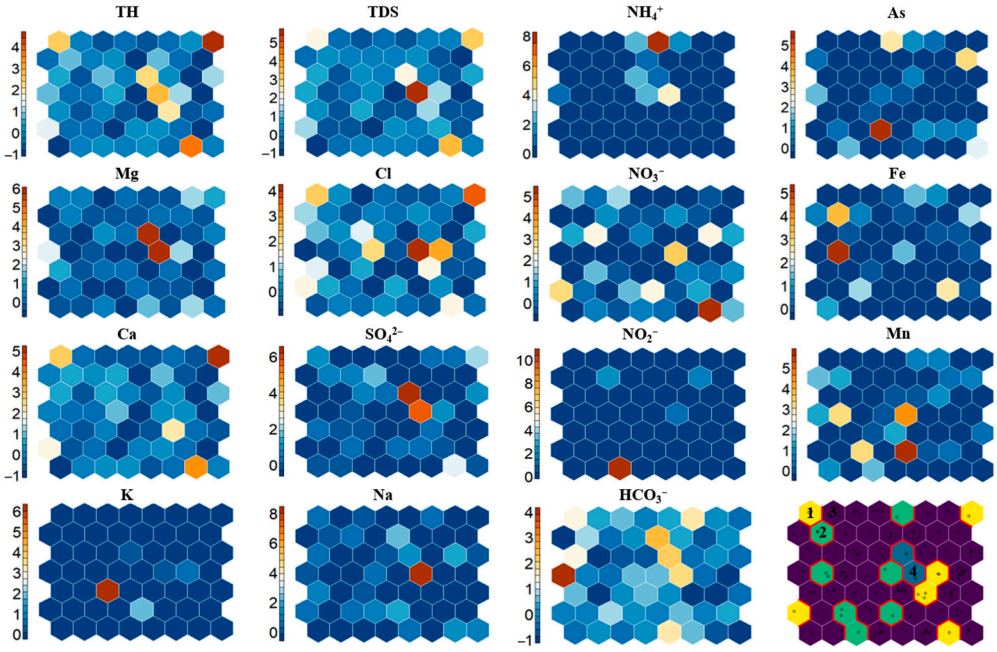


Figure 10. Neuron SOMs for 2014 groundwater quality indicators. (Note: Red and blue colors represent high and low ion concentrations, respectively, in the groundwater samples mapped to the neurons. The yellow, green, purple, and blue colors in the lower-right corner represent Categories 1, 2, 3, and 4, respectively).

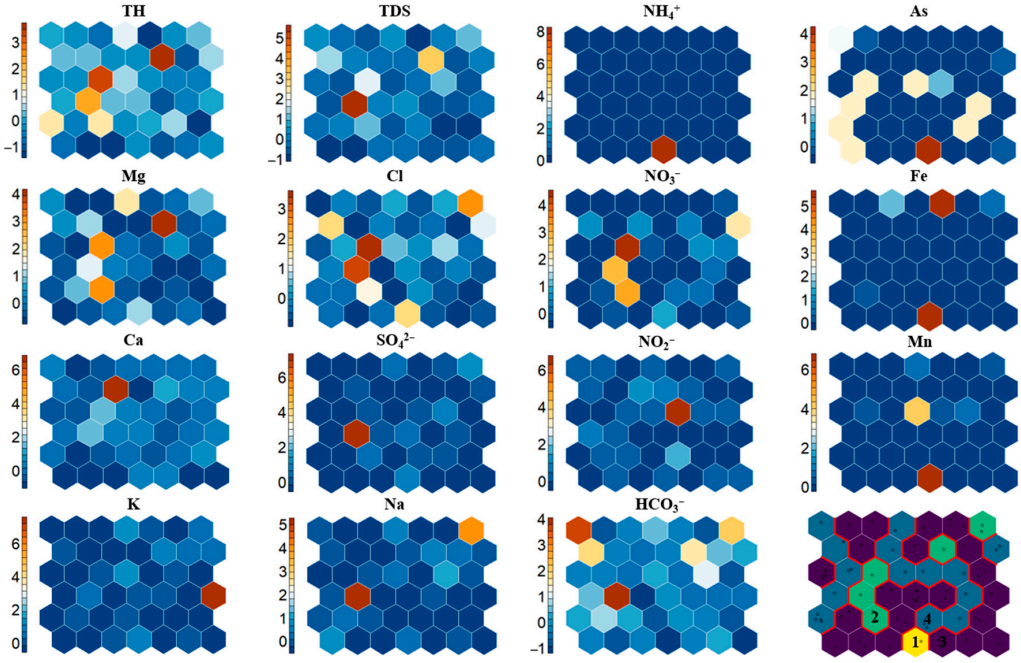


Figure 11. Neuron SOMs for 2022 groundwater quality indicators (Note: Red and blue colors represent high and low ion concentrations, respectively, in the groundwater samples mapped to the neurons. The yellow, green, purple, and blue colors in the lower-right corner represent Categories 1, 2, 3, and 4, respectively).

In 2014, the concentrations of TH, TDS, Ca^{2+}, Na^+, and Cl^- exhibited similar spatial distributions, suggesting a co-clustering mechanism driven by natural groundwater dissolution processes as well as anthropogenic chloride inputs from de-icing salts and chlorinated industrial effluents [44–46]. Notably, the study area lies on the southeastern margin of the Songliao Basin, where Cretaceous sandstone formations contribute Ca^{2+}

through limestone weathering, thereby increasing TH. As Ca^{2+} and Cl^{-} tend to migrate synergistically by forming soluble $CaCl_2$ complexes, enhancing their spatial correlation, their high-concentration zones were predominantly observed along urban industrial belts and major traffic corridors, where chlorinated effluents and de-icing salts are present in winter. The primary sources of Na^{+} include domestic sewage (e.g., sodium from table salt) and industrial discharges (e.g., detergents). The HCO_3^{-}, Mg^{2+}, and SO_4^{2-} indicators exhibited similar distributions. The HCO_3^{-} primarily originated from carbonate dissolution or CO_2–rock interactions, the SO_4^{2-} from gypsum ($CaSO_4$) dissolution or sulfide oxidation, and the Mg^{2+} from dolomite or silicate weathering, collectively reflecting the dominant influence of geochemical weathering. The As, NH_4^{+}, NO_3^{-}, NO_2^{-}, and K^{+} indicators exhibited distinct localized hotspots within different regions of the SOM topology, with low overall concentrations but high point values. These anomalies were driven by various mechanisms: the high As^{3+} concentrations were primarily of geogenic origin; the presence of NO_3^{-} and NH_4^{+} reflected nitrogen leaching from fertilizers and domestic wastewater; the NO_2^{-} anomalies were attributed to incomplete nitrification of fertilizers or sewage [47,48], but could also result from dry and wet deposition following coal combustion [49,50]; the high K^{+} levels were attributed to enrichment from suburban agricultural potassium fertilizer application or sewage leakage; and finally, elevated levels of Fe and Mn were also geogenic and predominantly controlled by the reducing conditions within the primary aquifer system.

In 2022, the TH, TDS, Ca^{2+}, Mg^{2+}, and Cl^{-} indicators exhibited co-located high-concentration zones, particularly around the urban periphery and near industrial corridors. These patterns can be attributed to long-term overextraction of groundwater, which promotes preferential dissolution of dolomite ($CaMg(CO_3)_2$) from Cretaceous aquifers, replacing earlier calcite-dominated hardness. Additionally, persistent application of chloride-based de-icing agents and leakage from expanded sewage pipelines contributed to elevated Cl^{-} levels, indicating a shift in the mineral weathering sequence of the aquifer. Furthermore, the correlation between TDS and SO_4^{2-} concentrations was significantly stronger in 2022 and exhibited enhanced spatial consistency, primarily owing to the continuous dissolution of the sulfates, such as $CaSO_4$, in Quaternary alluvial sediments, as well as increased replacement of traditional organic fertilizers with sulfur-based fertilizers, which promoted SO_4^{2-} transport via soil infiltration. Other localized high-value indicators, such as As, Fe, and Mn, although largely geogenic, may also share common input sources with NH_4^{+} or be subject to coupled geochemical processes, such as co-release under reducing aquifer conditions.

The results of SOM-based source tracing indicated a transition in groundwater hardness composition from Ca^{2+}-dominated to Ca^{2+}–Mg^{2+}-synergistic control between 2014 and 2022, driven by Cl^{-} input from winter de-icing agents and preferential dissolution of dolomite in aquifers under cold-region conditions [51]. Moreover, the intensification of SO_4^{2-} pollution increased its contribution to TDS levels, whereas concentrations of Fe, Mn, As, and NH_4^{+} remained spatially localized. In particular, the "polarized" migration of point-source As and NH_4^{+} pollution was restricted, exacerbating spatial heterogeneity in regional health risks.

3.3.2. Deterministic Health Risk Assessment

Based on the water quality screening and SOM-based source apportionment results, NH_4^{+}, As, and Cl^{-} were selected as representative contaminants for a deterministic assessment of the cumulative health risks associated with groundwater ingestion. The evaluation results are shown in Figures 12 and 13. This assessment was conducted using the USEPA's FSM to calculate the *HI* and *CR* levels. The spatial distributions of these risks for children

and adults were interpolated using ArcGIS 10.8.2. The results indicated that, under the ingestion exposure pathway, both *HI* and *CR* exhibited similar spatial trends for adults and children: overall, most of the study area presented acceptable risk levels for both populations. In 2014, elevated health risks were observed in the central and eastern regions, where *HI* was higher for children than for adults but *CR* was higher for adults than for children. In 2022, *HI* exceeded the threshold in the central-western region, and *CR* surpassed the limit in the northern region. The cumulative *HI* remained obviously higher in children and *CR* exhibited a slight predominance in children, both with comparable magnitudes. The observed difference between the *HI* levels for adults and children may be attributed to children's lower body weight and underdeveloped immune systems.

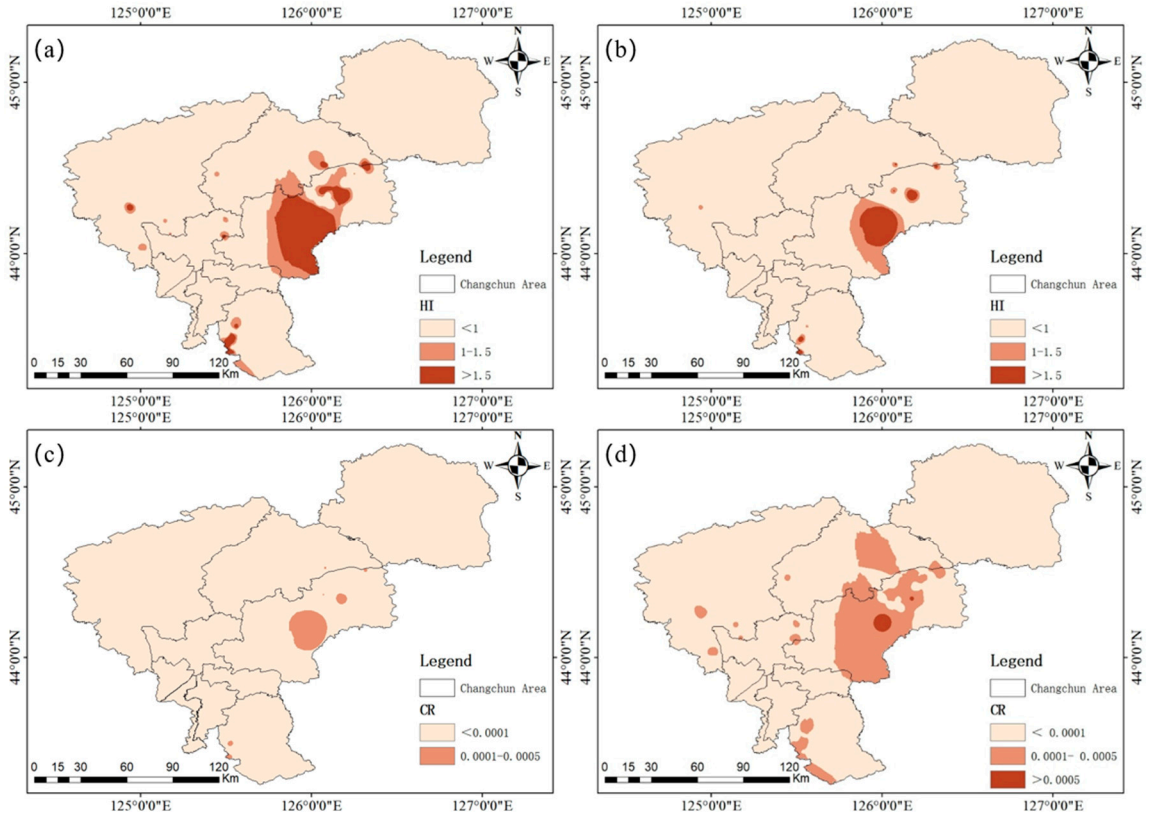


Figure 12. FSM-based deterministic health risk assessment map for 2014: (**a**) *HI* assessment map for children, (**b**) *HI* assessment map for adults, (**c**) *CR* assessment map for children, (**d**) *CR* assessment map for adults.

3.3.3. Probabilistic Health Risk Assessment

(1) TFN and Interval Value Construction Based on KDE

This study employed KDE with a Gaussian kernel function to estimate the probabilistic distributions of NH_4^+, As, and Cl^- concentrations in 2014 and 2022, resulting in the PDFs illustrated in Figure 14. The points with the maximum positive and negative slopes on the rising and falling edges of the KDE curve were designated as "b" and "c," forming the core of the TFN; points where the KDE slope dropped below 10% of the maximum on each side were defined as "a" and "d," forming the support interval. The resulting TFNs are shown in Figure 15. The fuzziness of contaminant concentration decreases as the value of "a" increases and is reflected by the narrowing of the interval between the upper and lower bounds. Given the variability in the groundwater data from the different years, most TFNs were trapezoidal in shape. However, when the most probable interval value (TMPIV) was narrow, some TFNs approximated triangles with extremely short upper bases. Using α-TST,

each TFN was converted into an interval value, with the results presented in Table 4. These interval values were subsequently used as input for the MCSS to generate probabilistic risk assessment outcomes [10].

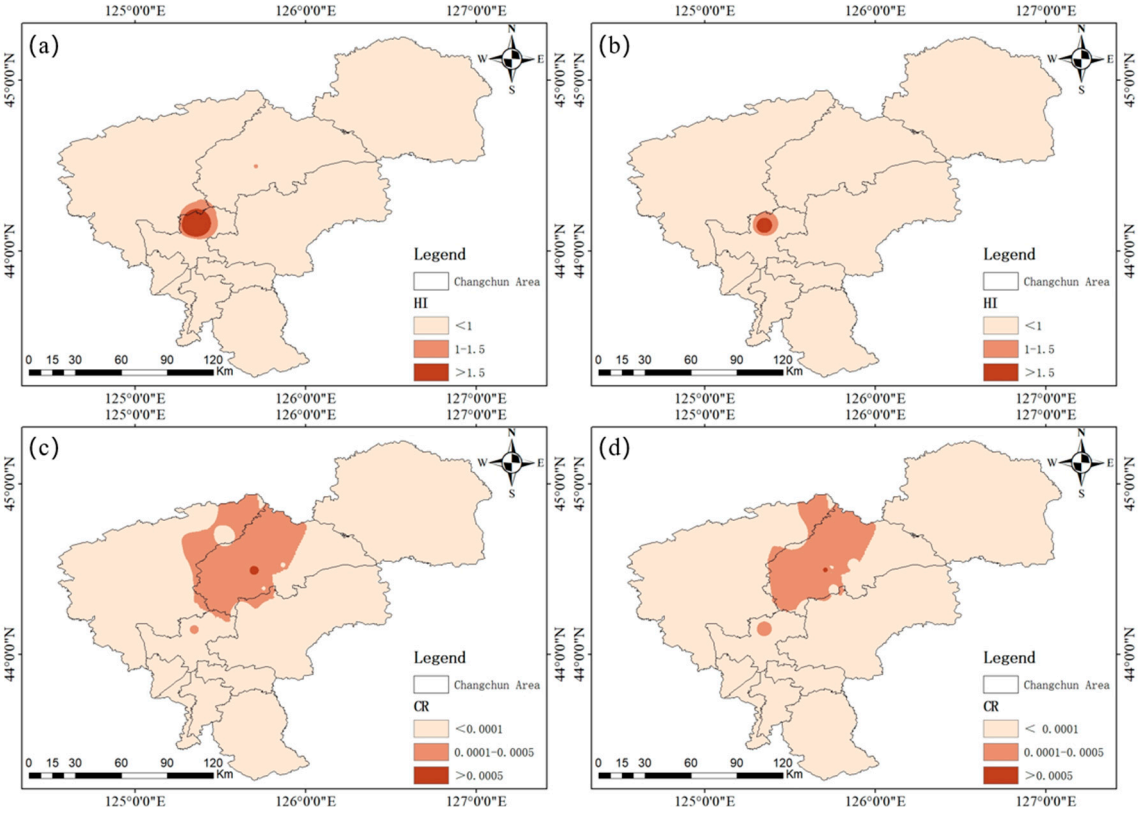
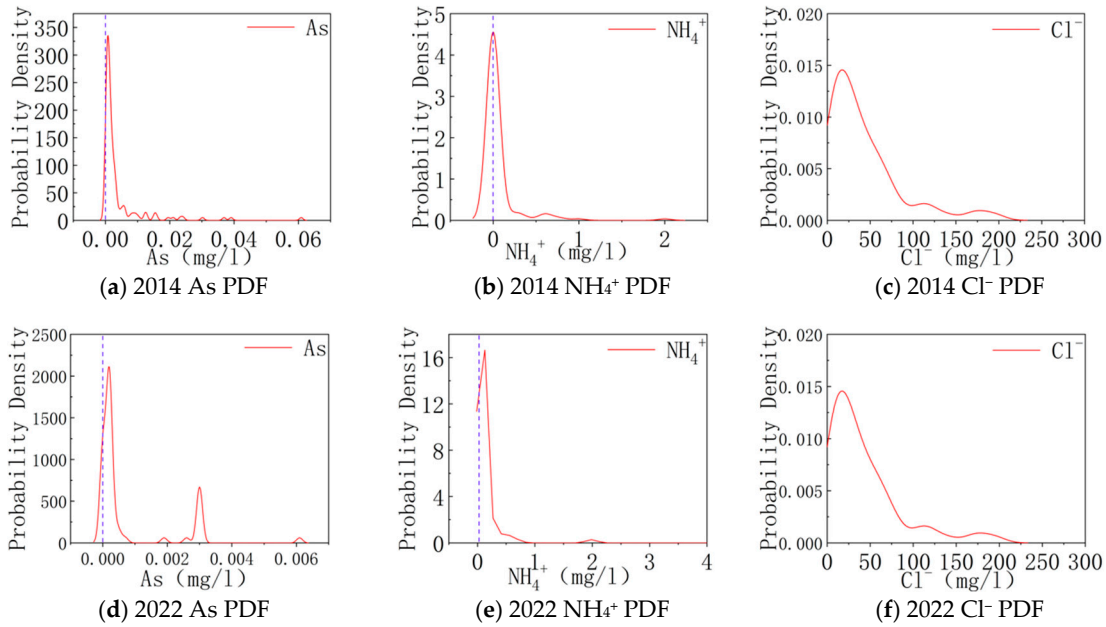


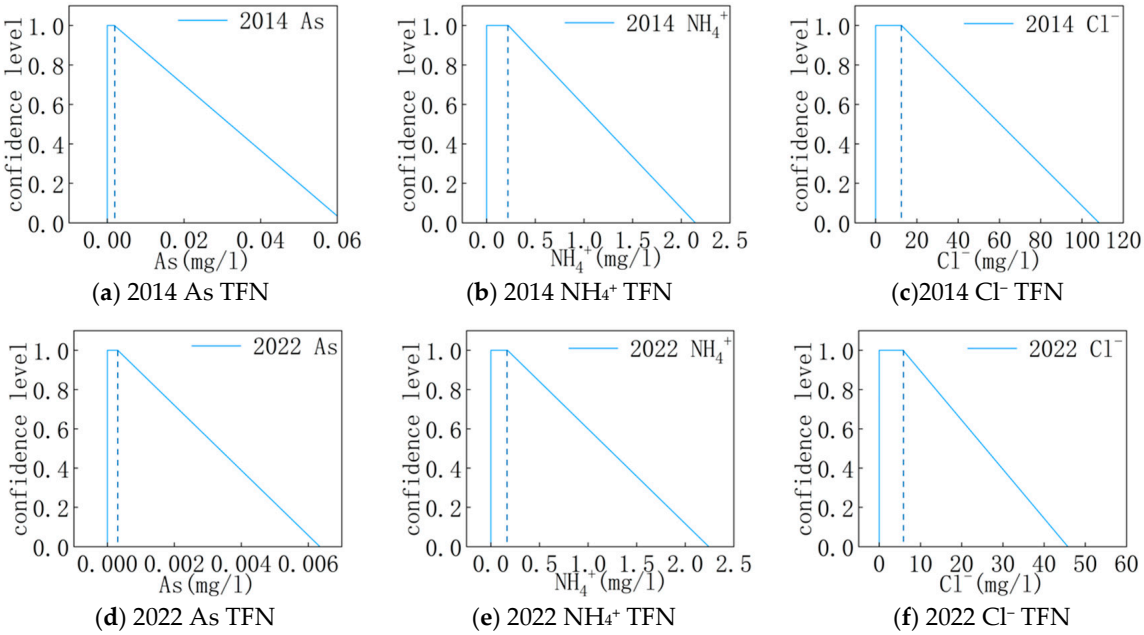
Figure 13. FSM-based deterministic health risk assessment map for 2022: (**a**) *HI* assessment map for children, (**b**) *HI* assessment map for adults, (**c**) *CR* assessment map for children, (**d**) *CR* assessment map for adults.



(**a**) 2014 As PDF (**b**) 2014 NH_4^+ PDF (**c**) 2014 Cl^- PDF

(**d**) 2022 As PDF (**e**) 2022 NH_4^+ PDF (**f**) 2022 Cl^- PDF

Figure 14. KDE PDF plot.



(**a**) 2014 As TFN (**b**) 2014 NH_4^+ TFN (**c**)2014 Cl^- TFN

(**d**) 2022 As TFN (**e**) 2022 NH_4^+ TFN (**f**) 2022 Cl^- TFN

Figure 15. TFN plot.

Table 4. Results of TFN conversion to interval values.

Year	2014		2022	
Interval Value (mg/L)	$a^a{}_L$	$a^a{}_R$	$a^a{}_L$	$a^a{}_R$
As	0	0.007873775	0	0.000914422
NH_4^+	0	0.41469583	0	0.378673057
Cl^-	0	22.1351929	0	9.775605

(2) Health Risk Assessment Results Based on KDE-TFN-MCSS

The probabilistic health risks for children and adults in 2014 and 2022 were estimated using the proposed KDE-TFN-MCSS model; the results are shown in Figure 16. The 95th percentile of the simulated output was adopted as the high-risk threshold. The probability distributions of the total *HI* and *CR* for both age groups were left-skewed, indicating higher probabilities of low to moderate risk levels. Except for the 2022 *CR* in children, the 95th percentile values for all scenarios exceeded the acceptable risk threshold. A comparison between years revealed that overall health risks in 2022 were lower than those in 2014, with a slight increase observed only in the adult *CR*. Within the same year, children were exposed to higher *HI* levels than adults, whereas adults were exposed to significantly higher *CR* levels.

Both the FSM and KDE-TFN-MCSS models revealed spatial heterogeneity and uncertainty in groundwater health risks across the central Songnen Plain. The deterministic FSM assessment indicated that the *HI* levels in central-eastern regions exceeded the safety threshold (*HI* > 1) in 2014. By 2022, the high-risk zones expanded westward to the central-western region, and the *CR* levels in parts of the north surpassed the threshold (*CR* > 0.0001), which corresponded closely with the increase in NH_4^+ concentrations from 0.06 to 1.26 mg/L and geogenic As enrichment. However, as the FSM is based on fixed-point estimates, it fails to quantify parameter variability, potentially underestimating extreme risks in localized high-concentration areas. By contrast, the proposed probabilistic assessment model used KDE to derive the distributions of NH_4^+ and As, incorporated TFNs to account for parameter fuzziness, and used an MCSS to reveal that, except for the *CR* for children in 2022, the 95th percentile values for *HI* and *CR* exceeded the acceptable thresholds. Furthermore, the left-skewed risk distributions suggested a higher probability of low to moderate risks,

reflecting the spatial heterogeneity of NH_4^+ pollution driven by combined agricultural and domestic sources, along with localized enrichment of As and Cl^-.

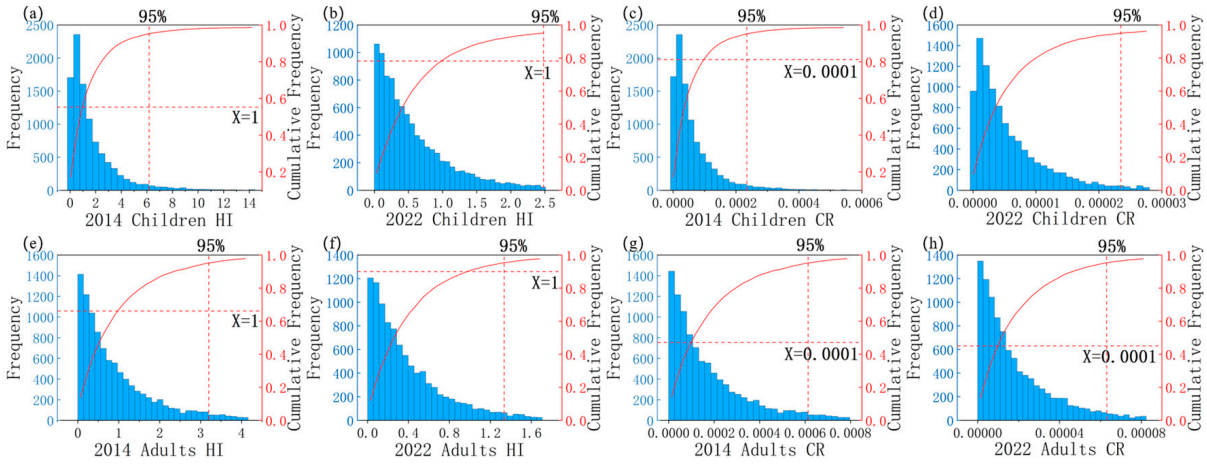


Figure 16. Probabilistic risk assessment plot.

Additionally, both models consistently exhibited higher *HI* levels in children and higher *CR* levels in adults. However, the probabilistic model quantified risk probabilities more precisely. For example, the 95th percentile of the *CR* for adults in 2022 was 0.00012, slightly exceeding the safety threshold and indicating that adults are more susceptible to long-term carcinogenic accumulation, whereas children's lower body weight and immature immune systems render them more vulnerable to non-carcinogenic hazards. The divergence between the results obtained by the two evaluated methods highlights the limitations of deterministic models in cold-region aquifer systems, where freeze–thaw cycles alter aquifer permeability, enhancing pollutant migration and increasing parameter variability. For example, the water quality deterioration observed in 2022 in the central region can be attributed to de-icing salt infiltration and preferential thaw-induced pore development in the vadose zone. In this complex environmental context, the use of a probabilistic assessment method, which integrates both natural and anthropogenic uncertainties, proved more suitable for risk characterization.

3.4. Comprehensive Analysis

In 2022, areas of groundwater quality deterioration and elevated *HI* levels in the central Songnen Plain spatially overlapped with high-vulnerability zones identified by the DRASTICL model, indicating that regions with greater vulnerability were more susceptible to anthropogenic influence. Indeed, shallow aquifers, highly permeable vadose zones, and intensive agricultural land use facilitated the vertical migration of NH_4^+ and Cl^-, directly linking water quality degradation to increased health risks. Furthermore, SOM-based source identification revealed that NH_4^+ anomalies coincided with leakage from decentralized rural sanitation systems, whereas high Cl^-–Mg^{2+} co-concentrations in industrial zones promoted the formation of soluble $CaCl_2$ complexes, enhancing pollutant diffusivity. Further analysis indicated that the mineralization of organic matter and the hydrolysis of urea in agricultural return flow significantly contributed to the NH_4^+ load in the aquifer, especially during the post-fertilization and irrigation seasons. This led to NH_4^+ accumulation and elevated *HI* levels in highly vulnerable areas of the central-western region. Simultaneously, SOM analysis revealed that the spatial distribution of As-related CR hotspots was closely associated with specific hydrogeochemical conditions, notably the prevalence of reducing environments and the presence of aquifers with high natural arsenic background levels. Although mean concentrations of As declined in 2022, its high

carcinogenic slope factor (SF = 1.5) continued to drive *CR* exceedances in localized areas, particularly among adults with prolonged exposure. Sulfate reduction under reducing conditions as well as Cl^- input were found to promote As mobilization, confirming that freeze–thaw cycles enhance hydraulic connectivity in aquifers, thereby facilitating As activation. By contrast, low-vulnerability zones in the eastern region—characterized by deep-seated, compact aquifers—maintained Class I or II water quality with the lowest risk levels, underscoring the natural buffering role of geological barriers [36].

Freeze–thaw cycles in cold plains regions seasonally alter permeability, accelerating Cl^- and NH_4^+ infiltration and enhancing As release under reducing conditions, producing a synergistic amplification of natural and anthropogenic risks. This spatially coupled mechanism of "high vulnerability–water quality degradation–pollutant input" reflects a systemic transformation of groundwater risk in the central Songnen Plain from natural weathering dominance to strong anthropogenic interference [52].

4. Conclusions

This study used groundwater quality monitoring data collected from the case study area in 2014 and 2022 to derive vulnerability indices and evaluate aquifer vulnerability using a modified DRASTICL model considering regional anthropogenic activities and land use types. The relative importances of key water quality indicators were determined via the RF algorithm to conduct an E-GQI evaluation, and an SOM analysis was undertaken to characterize groundwater pollution patterns. The refined KDE-TFN-MCSS model was proposed to reduce the uncertainty and limitations associated with the TFN-MCSS model for health risk assessment and enhance the accuracy of probabilistic evaluation. Finally, the results of the conventional FSM deterministic evaluation were compared with those of the proposed probabilistic KDE-TFN-MCSS model when assessing the groundwater environment of the central Songnen Plain, providing theoretical support for winter-period pollution control in the northeastern industrial regions of China. The key conclusions of this study are as follows:

(1) The dominant hydrochemical type in the study area was determined to be HCO_3^-–Ca^{2+}, reflecting moderately to weakly alkaline groundwater. No significant shifts in water chemistry were observed between 2014 and 2022, and water hardness remained high. Rock weathering and evaporative crystallization were identified as the primary factors driving hydrochemical evolution. The indicators exceeding their thresholds comprised TH, TDS, pH, Na^+, SO_4^{2-}, NO_3^-, NO_2^-, NH_4^+, F^-, Fe, Mn, and As. Although exceedance rates declined and overall water quality improved from 2014 to 2022, geogenic and nitrogenous pollutants remained significant. Notably, NH_4^+ concentrations rose markedly from 0.06 mg/L in 2014 to 1.26 mg/L in 2022, exceeding the Class III limit (1.0 mg/L).

(2) In 2014, groundwater quality was poor overall, dominated by Class IV and V water with minor distribution of Class III water in the west and northeast and no Class I–II water. In 2022, Class I–IV water predominated, exhibiting a southwest–northeast banded pattern. Although general water quality improved, localized deterioration was observed in central and northern zones.

(3) The study area exhibited a systemic shift from natural weathering dominance to increasing anthropogenic disturbance. The results of the SOM analysis revealed localized high concentrations of As and NH_4^+ indicative of legacy industrial and agricultural pollution, whereas the Cl^- distribution reflected urbanization and de-icing agent inputs.

(4) The NH_4^+, As, and Cl^- indicators were selected for health risk assessments using the conventional FSM and proposed KDE-TFN-MCSS models. Under the ingestion

pathway, most areas in the central Songnen Plain exhibited acceptable risk levels for both children and adults, with a higher probability of low to moderate risks. The *HI* levels were consistently higher in children, whereas the *CR* levels were higher in adults. Although overall groundwater quality improved in 2022, the unique coupling of freeze–thaw cycles and urbanization in the case study cold region led to persistent NH_4^+ and As accumulation in localized areas. Therefore, zonal management of de-icing zones and legacy industrial belts in the case study region is recommended to balance groundwater development with ecological safety.

The findings of this study informed the following general recommendations for groundwater management in cold regions: (1) implement zonal management strategies targeting high-vulnerability areas to restrict intensive agricultural activities and reduce NH_4^+ and Cl^- inputs from de-icing agents and fertilizers; (2) establish dynamic monitoring networks focusing on seasonal freeze–thaw cycles to track pollutant migration and As mobilization in reducing environments; and (3) prioritize remediation in zones where carcinogenic and non-carcinogenic risks exceed thresholds, especially for children and adults with long-term exposure.

The results of this study confirmed that the proposed KDE-TFN-MCSS model provides a robust tool for quantifying uncertainty in health risks, enabling policymakers to design adaptive policies that address both anthropogenic and geogenic pollution. Future research should integrate climate projections to evaluate how warming temperatures and extreme precipitation events alter groundwater quality dynamics in cold regions.

Author Contributions: Conceptualization, J.L. and Y.W.; methodology, J.L. and Y.W.; software, J.L.; validation, J.L., Y.W. and X.F.; formal analysis, data curation, X.S.; writing—original draft preparation, J.L.; writing—review and editing, J.L. and Y.W.; visualization, J.L. and X.F.; supervision, Y.W.; project administration, J.B.; funding acquisition, J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the National Key Research and Development Program of China (Grant No. 2022YFD1500500) and the Projects of Jilin Provincial Department of Science and Technology (Grant No. 20240101059JC). The authors thank all the experts who helped with this article.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Dharmani, B.A.; Verma, M.; Rani, S.; Narang, A.; Singh, M.R.; Saya, L.; Hooda, S. Unravelling groundwater contamination and health-related implications in semi-arid and cold regions of India. *J. Contam. Hydrol.* **2024**, *261*, 104303.
2. Sanjupriya, S.; Poonkothai, M.; Karunanidhi, D.; Rao, N.S.; Subramani, T.; Marghade, D. Evaluating nitrate contamination in groundwater and its health threats from a semi-arid province of southern India using GIS techniques with a special focus on entropy water quality index. *Environ. Geochem. Heal.* **2025**, *47*, 346. [CrossRef]
3. Ahn, S.H.; Jeong, D.H.; Kim, M.; Lee, T.K.; Kim, H.K. Prediction of groundwater quality index to assess suitability for drinking purpose using averaged neural network and geospatial analysis. *Ecotoxicol. Environ. Saf.* **2023**, *265*, 115485. [CrossRef]
4. Jia, X.Q.; Luo, M.M. Groundwater vulnerability assessment in the typical karst watershed of Xiangxi River based on an improved DRASTIC model. *Geol. Sci. Technol. Inf.* **2019**, *38*, 255–261.
5. Chen, S.Y.; Fu, G.T.; Zhou, H.C.; Wang, G.-L. Fuzzy analytical evaluation model and method for aquifer vulnerability. *J. Hydraul. Eng.* **2002**, 23–30.
6. Yu, L.H.; Tao, Z.B.; Hu, S.T.; Gao, S.; Yang, N.; Yu, X.J. Application of the DRASTIC model in groundwater vulnerability assessment. *Yellow River* **2020**, *42*, 45–46+50.

7. Xu, C.; Zhou, J.Y.; He, X.J.; Yang, Z.-T.; Zhao, K.-Z.; Jiang, Y.-N. Groundwater vulnerability assessment in Shaanxi Province based on an improved DRASTIC model. *China Rural. Water Hydropower* **2020**, *3*, 44–51.
8. Liao, X.U.; Li, Y.; Miranda-Avilés, R.; Puy-Alquiza, M.J.; Bian, J.; Anguiano, J.H.H.; Muñoz, A.H.S.; Datta, S.; Zha, X.; Liu, J.; et al. Assessments of Pollution Status and Human Health Risk of Potentially Toxic Elements in Primary Crops and Agricultural Soils in Guanajuato, Mexico. *Water Air Soil Pollut.* **2023**, *234*, 670. [CrossRef]
9. Tenebe, I.T.; Emenike, P.C.; Babatunde, E.O.; Neris, J.B.; Fred-Ahmadu, O.H.; Dede-Bamfo, N.; Etu, E.-E.; Ogarekpe, N.M.; Emakhu, J.; Benson, N.U. Assessing the state of rainwater for consumption in a community in dire need of clean water: Human and health risk using HERisk. *Water Pract. Technol.* **2022**, *17*, 2005–2022. [CrossRef]
10. Li, T.; Bian, J.; Ruan, D.; Xu, L.; Zhang, S. Groundwater health risk assessment and its temporal and spatial evolution based on trapezoidal fuzzy number-Monte Carlo stochastic simulation: A case study in western Jilin province. *Ecotoxicol. Environ. Saf.* **2024**, *282*, 116736. [CrossRef]
11. Ruan, D.; Bian, J.; Wang, Y.; Wu, J.; Gu, Z. Identification of groundwater pollution sources and health risk assessment in the Songnen Plain based on PCA-APCS-MLR and trapezoidal fuzzy number-Monte Carlo stochastic simulation model. *J. Hydrol.* **2024**, *632*, 130897. [CrossRef]
12. Yan, J.; Chen, J.; Zhang, W. 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.* **2023**, *243*, 120379. [CrossRef]
13. Mohammadpour, A.; Keshtkar, M.; Samaei, M.R.; Isazadeh, S.; Khaneghah, A.M. Assessing water quality index and health risk using deterministic and probabilistic approaches in Darab County, Iran; A machine learning for fluoride prediction. *Chemosphere* **2024**, *352*, 141284. [CrossRef] [PubMed]
14. *GB/T 7477-1987*; Water Quality—Determination of Calcium and Magnesium—Atomic Absorption Spectrometric Method. Standardization Administration of the People's Republic of China: Beijing, China, 1987.
15. *DZ/T 0064.49-1993*; Methods for Analysis of Groundwater Quality—Part 49: Determination of Carbonate, Bicarbonate and Hydroxide Ions—Titrimetric Method. Ministry of Geology and Mineral Resources of the People's Republic of China: Beijing, China, 1993.
16. *GB/T 5750.4-2006*; Standard Examination Methods for Drinking Water—Organoleptic and Physical Parameters. Standardization Administration of the People's Republic of China: Beijing, China, 2006.
17. *HJ 535-2009*; Water Quality—Determination of Ammonia Nitrogen—Nessler's Reagent Spectrophotometry. Ministry of Environmental Protection of the People's Republic of China: Beijing, China, 2009.
18. *HJ 1147-2020*; Water Quality—Determination of pH Value—Electrode Method. Ministry of Ecology and Environment of the People's Republic of China: Beijing, China, 2020.
19. *GB/T 7493-1987*; Water Quality—Determination of Nitrogen (Nitrite)—Spectrophotometric Method. Standardization Administration of the People's Republic of China: Beijing, China, 1987.
20. *HJ 812-2016*; Water Quality—Determination of Water-Soluble Anions (F^-, Cl^-, $NO_2{}^-$, Br^-, $NO_3{}^-$, $PO_4{}^{3-}$, $SO_3{}^{2-}$, $SO_4{}^{2-}$)—Ion Chromatography Method. Ministry of Ecology and Environment of the People's Republic of China: Beijing, China, 2016.
21. *HJ 84-2016*; Water Quality—Determination of Inorganic Anions (F^-, Cl^-, $NO_2{}^-$, Br^-, $NO_3{}^-$, $PO_4{}^{3-}$, $SO_3{}^{2-}$, $SO_4{}^{2-}$)—Ion Chroma-tography Method. Ministry of Ecology and Environment of the People's Republic of China: Beijing, China, 2016.
22. *HJ 776-2015*; Water Quality—Determination of 32 Elements—Inductively Coupled Plasma Optical Emission Spectrometry. Ministry of Environmental Protection of the People's Republic of China: Beijing, China, 2015.
23. *HJ 694-2014*; Water Quality—Determination of Mercury, Arsenic, Selenium, Bismuth and Antimony—Atomic Fluorescence Spectrometry. Ministry of Ecology and Environment of the People's Republic of China: Beijing, China, 2014.
24. *GB/T 5750.6-2006*; Standard Examination Methods for Drinking Water—Metal Parameters. Standardization Administration of the People's Republic of China: Beijing, China, 2006.
25. *GB/T 7467-1987*; Water Quality—Determination of Chromium (VI)—1,5-Diphenylcarbohydrazide Spectrophotometric Method. Standardization Administration of the People's Republic of China: Beijing, China, 1987.
26. *GB/T 16489-1996*; Water Quality—Determination of Sulfide—Methylene Blue Spectrophotometric Method. Standardization Administration of the People's Republic of China: Beijing, China, 1996.
27. *HJ 586-2010*; Water Quality—Determination of Cyanide Easily Released—Spectrophotometric Method. Ministry of Environmental Protection of the People's Republic of China: Beijing, China, 2010.
28. Saravanan, S.; Pitchaikani, S.; Thambiraja, M.; Sathiyamurthi, S.; Sivakumar, V.; Velusamy, S.; Shanmugamoorthy, M. Comparative assessment of groundwater vulnerability using GIS-based DRASTIC and DRAS-TIC-AHP for Thoothukudi District, Tamil Nadu India. *Environ. Monit. Assess.* **2022**, *195*, 57. [CrossRef] [PubMed]
29. Sarkar, M.; Pal, S.C. Application of DRASTIC and Modified DRASTIC Models for Modeling Groundwater Vulnerability of Malda District in West Bengal. *J. Indian Soc. Remote. Sens.* **2021**, *49*, 1201–1219. [CrossRef]

30. Wang, Z.; Wang, Y.; Shi, M.; Ji, W.; Li, R.; Wang, X. Coordinated analysis of groundwater spatiotemporal chemical characteristics, water quality, and potential human health risks with sustainable development in semi-arid regions. *Environ. Geochem. Heal.* **2024**, *46*, 370. [CrossRef]
31. Gu, Z.Q.; Bian, J.M.; Wang, Y.; Sun, X.; Ruan, D. Groundwater quality assessment and index optimization of water quality monitoring in the water source area of Changbai Mountain. *China Environ. Sci.* **2023**, *43*, 5257–5264.
32. Alasiri, A.; Kumar, P.M. Investigating the influence of solar distillation on improving the groundwater quality index in the southern region of Riyadh, Saudi Arabia. *Groundw. Sustain. Dev.* **2024**, *25*, 101155. [CrossRef]
33. *GB/T 14848-2017*; Standard for Groundwater Quality. Standardization Administration of the People's Republic of China: Beijing, China, 2017.
34. Yuan, W.; Wang, Z.; Zhang, T.; Liu, Z.; Ma, Y.; Xiong, Y.; An, F. Assessment and Prediction of Groundwater Vulnerability Based on Land Use Change—A Case Study of the Central Urban Area of Zhengzhou. *Water* **2024**, *16*, 3716. [CrossRef]
35. Balamurali, M.; Silversides, K.L.; Melkumyan, A. A comparison of t-SNE, SOM and SPADE for identifying material type domains in geological data. *Comput. Geosci.* **2019**, *125*, 78–89. [CrossRef]
36. Yan, Y.; Zhang, Y.; Yang, S.; Wei, D.; Zhang, J.I.; Li, Q.; Yao, R.; Wu, X.; Wang, Y. Optimized groundwater quality evaluation using unsupervised machine learning, game theory and Monte-Carlo simulation. *J. Environ. Manag.* **2024**, *371*, 122902. [CrossRef] [PubMed]
37. Zhang, Y.Y.; Zhou, Y.Z.; Zhou, J.L.; Zeng, Y.; Ding, Q.; Jiang, F. Hydrochemical characteristics and health risk assessment of groundwater in the Barkol Lake Basin Plain, Xinjiang. *J. Irrig. Drain.* **2024**, *43*, 95–104.
38. Liang, Z.Q.; Yang, J. Health risk and ecological risk assessment of ammonia nitrogen in surface water of Tianjin. *Environ. Sci. Surv.* **2024**, *43*, 68–73.
39. Xu, L.W.; Bian, J.M.; Sun, X.Q.; Lou, Y.Q.; Sun, G.J. Impact of irrigation return flow on regional groundwater quality and health risk assessment. *China Environ. Sci.* **2023**, *43*, 1688–1695.
40. Debnath, K.; Debnath, P.; Choudhury, S.; Saha, A.K.; Majumdar, A. A framework of trapezoidal fuzzy best-worst method in location selection for surface water treatment plant. *Pollution* **2024**, *10*, 1–15.
41. Chen, W.; Yu, S.W.; Liao, J.; Wang, S.C.; Liu, G.N. Source apportionment and health risk assessment of heavy metals in groundwater in northeastern Nanchang. *Environ. Sci.* **2025**, *46*, 854–863.
42. Azlaoui, M.; Karef, S.; Foufou, A.; Haied, N.; Zeddouri, A.; Bengusmia, D. Integration of land use/land cover factors with machine learning in groundwater vulnerability assessment models for semi-arid regions Algeria. *Desalin. Water Treat.* **2025**, *323*, 101256. [CrossRef]
43. Xiong, Y.; Zhang, T.; Sun, X.I.; Yuan, W.; Gao, M.; Wu, J.; Han, Z. Groundwater Quality Assessment Based on the Random Forest Water Quality Index—Taking Karamay City as an Example. *Sustainability* **2023**, *15*, 14477. [CrossRef]
44. Jiang, Q.; Liu, Q.; Liu, Y.U.; Zhu, J.; Chai, H.; Chen, K. Chemical composition of groundwater and its controlling factors in the Liuzhuang coal mine, Northern Anhui Province, China. *Water Supply* **2023**, *23*, 4937–4956. [CrossRef]
45. Rudolph, D.L.; Melchin, J.; Stone, M.; Sarwar, G.; Hodgins, E. Efficacy of urban road salt reduction strategies on public supply well quality. *Sci. Total. Environ.* **2023**, *900*, 166466. [CrossRef] [PubMed]
46. Ming, R.T. Research on Migration and Accumulation of Deicing Agents in Ski Resorts Based on SWAT Model. Master's Thesis, Beijing University of Civil Engineering and Architecture, Beijing, China, 2021.
47. Li, Y. Research on River Health Assessment in Changchun City Based on PSR and FAHP-CRITIC Models. Master's Thesis, Dalian University of Technology, Dalian, China, 2022.
48. Liu, X.Y. Study on Distribution and Transport Patterns of Nitrate-Nitrogen Pollutants in Groundwater of Typical Watersheds in Changchun. Master's Thesis, Jilin University, Changchun, China, 2023.
49. Yang, Z.P.; Lu, W.X.; Liu, X.R.; Xin, X. Source apportionment of heavy metals in topsoil of Changchun urban area. *Urban Environment Urban Ecol.* **2009**, *5*.
50. Wang, X.; Xiao, C.; Liang, X.; Li, M. Groundwater Quality Assessment in the Northern Part of Changchun City, Northeast China, Using PIG and Two Improved PIG Methods. *Int. J. Environ. Res. Public Health* **2022**, *19*, 9603. [CrossRef]
51. Baraza, T.; Cassidy, K.J.; Hasenmueller, E.A. Road salt applications mobilize trace elements from roadside soil to shallow groundwater. *Sci. Total. Environ.* **2024**, *942*, 173435. [CrossRef] [PubMed]
52. Xie, X.; Ji, H.; Yang, Q.; Bao, S.; Han, C.; Qu, W.U. A hybrid approach for identifying the seasonal variation of groundwater quality, source apportionment and health risk in a coastal area driven by natural and anthropogenic factors. *J. Hazard. Mater.* **2025**, *496*, 139283. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.