Journal of Hydrology 589 (2020) 125343

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Application of a genetic algorithm to groundwater pollution source identification

Kexue Han[a,b], Rui Zuo[a,b,*], Pengcheng Ni[a,b], Zhenkun Xue[a,b], Donghui Xu[a,b], Jinsheng Wang[a,b], Dan Zhang[c]

[a] College of Water Sciences, Beijing Normal University, Beijing 100875, China
[b] Engineering Research Center of Groundwater Pollution Control and Remediation, Ministry of Education, Beijing 100875, China
[c] Beijing Municipal Research Institute of Environmental Protection, National Engineering Research Center of Urban Environmental Pollution Control, Beijing 100037, China

ARTICLE INFO

This manuscript was handled by Huaming Guo, Editor-in-Chief

Keywords:
Advection-Dispersion Equation (ADE)
Genetic algorithm (GA)
Source identification
Sandbox experiment
Site application
Investigation data

ABSTRACT

The accurate identification of groundwater pollution sources is the key to site remediation and management, which directly relates to the cost and effect of remediation during later periods. Studying the identification of some small-area site pollution sources by using a small amount of groundwater monitoring data is difficult and is currently a hot topic in the field of groundwater management. In this study, a method combining the Advection-Dispersion Equation (ADE) of contaminants in groundwater with Genetic Algorithm (GA) was proposed to identify groundwater pollution sources, that is, the location of the pollution source, the time of pollution release, and the intensity of pollution release. A sandbox experiment for the migration and transformation of deterministic petroleum hydrocarbons in groundwater is designed to verify the method. The results show that the method code program has basically the same identification parameters as the deterministic experimental pollution source parameters. The ADE-GA method was applied to an actual site polluted by a gas station in Beijing, China. The parameters of the pollution source obtained from the detailed site investigation are basically consistent with the calculated values. Therefore, the source code established in this work is a reliable tool for identifying the parameters of groundwater pollution sources in small areas.

1. Introduction

Pollution source identification is the key to controlling and remediating groundwater pollution (Xu and Gómez-Hernández, 2016). Therefore, it is very important to identify the source of pollution for previously polluted groundwater, i.e., identify the location, time and amount of the pollution source release (Ayvaz, 2010), through the inverse problem solving method based on data of the position and concentration of the pollutants when monitored downstream of the groundwater (Sun et al., 2006; Yeh et al., 2007; Xia et al., 2019). The inversion methods of pollutant transport (Bagtzoglou and Atmadja, 2005) to solve the inverse problem of pollution source identification mainly include the direct method, stochastic method and optimization method (Sun et al., 2006; Atmadja and Bagtzoglou, 2001).

The geochemical footprint method is commonly used in the direct method. The principle of the geochemical footprint method is to obtain the source parameters by using isotope and geochemical fingerprint methods (Kohl et al., 1971; Liu et al., 2006; Bronders et al., 2012; Aeppli et al., 2013; Chen et al., 2017; Miglietta et al., 2017; Shakya et al., 2019). However, the application of the geochemical footprint method to the analysis of soil and groundwater pollution sources is immature. In addition, when there are multiple pollution sources, uncertainty assessments are lacking, and the quantization accuracy is limited (Stout et al., 1998). In the stochastic statistical method, the parameters for groundwater pollution sources are considered as random variables; this method uses a certain mathematical method to find the probability of each parameter within a certain range and finally generates the parameter as the result (Snodgrass and Kitanidis, 1997; Butera and Tanda, 2003; Michalak and Kitanidis, 2004; Boano et al., 2005; Sun, 2007; Butera et al., 2013; Venkatramanan et al., 2016; Long et al., 2018). Unfortunately, the stochastic statistical method considers less relevant parameters, such as the actual site hydrogeological conditions. In addition, the stochastic statistical method only involves a large amount of data processing, so it cannot truly represent the actual situation. Moreover, there are more unknowns and higher spatial dimensions during the calculation process, which results in a larger calculation load.

The basic principle of the optimization method is to establish the

* Corresponding author at: College of Water Sciences, Beijing Normal University, Beijing 100875, China.
E-mail address: zuo1101@163.com (R. Zuo).

https://doi.org/10.1016/j.jhydrol.2020.125343
Received 14 March 2020; Received in revised form 22 July 2020; Accepted 23 July 2020
Available online 28 July 2020

corresponding solute model. The simulation results were compared with the observed values, and the process was continuously optimized to obtain the pollution source parameters (Singh and Datta, 2006). Optimization methods include classical methods containing gradient calculations and non-classical methods containing non-gradient calculations (Gandhi et al. 2016). Jiang et al. (2018) identified the mass-loading rate of unknown pollution sources by using the classical optimization method of linear programming in MATLAB software. Gandhi et al. (2016) used a non-classical optimization method of direct search in MATLAB software to identify the characteristics of virus sources in groundwater.

In recent years, as a heuristic search algorithm with global convergence (Yeh et al., 2007) and another non-classical optimization method (Gandhi et al. 2016), genetic algorithm has attracted increasing attention (Singh et al., 2008; Jin et al., 2009; Ayvaz, 2016). Genetic algorithms are widely used in the research of optimization problems because of their advantages, such as easy convergence and high accuracy (Ayaz 2017). Moreover, non-gradient search is the characteristic of genetic algorithm, which can find the global optimal solution well. Thus, the disadvantages of gradient-based search method can be overcome (Ayaz 2017). The research shows that this method can still achieve accurate identification with multiple unknown sources of pollution (Singh and Datta, 2006). Mahinthakumar and Sayeed (2005) used the real encoded genetic algorithm (RGA) and the binary encoded genetic algorithm (BGA) in MATLAB software to combine with several different local search methods, so as to realize the identification of the location and release intensity of the assumed two-dimensional pollution site. The genetic algorithm is combined with a groundwater simulation model to form a new method that identifies the groundwater pollution sources. The results show that the initial estimate of the source parameters will not affect the final identification result (Aral et al., 2001). Ayaz (2017) combined the pollutant transport model with the optimization simulation model based on genetic algorithm in MATLAB software to identify the groundwater pollution source of the hypothetical site when the population size of the genetic algorithm was relatively small, and obtained very good results. Sophia and Bhattacharjya (2020) combined MODFLOW and MT3DMS models with genetic algorithm to identify pollution sources in a hypothetical study area with homogeneous and isotropic aquifers. Xia et al. (2019) used the hyper-parameters obtained based on the Taguchi experimental design tuning method to optimize the genetic algorithm, so as to perform more effective pollution source identification on the hypothetical site of irregular geometry and non-uniform media. All the works mentioned above were tested in hypothetical cases (Table 1). Currently, only a small amount of literature on the application of other source identification methods in the laboratory can be found. Cupola et al. (2015) designed a sandbox experiment to verify the performance of the simultaneous release function and source location identification (SRSI) and backward probability model based on adjoint state method (BPM-ASM) in the identification of pollution source location and release time. Chen et al. (2018) verified the ability of restart ensemble Kalman filter (r-EnKF) to identify pollution source and geometry of aquifer by designing a sandbox experiment. Zanini and Woodbury (2016) used the empirical Bayes method combined with Akaike's Bayes Information Standard (ABIC) to deduce the release history of groundwater pollutants using the data obtained through the sandbox experiment. The filling medium and tracer used in the above sandbox experiments are glass beads and sodium fluorescein respectively, and the influence of the natural attenuation of pollutants on the concentration is not considered.

The purpose of this paper is to establish a genetic algorithm calculation program, which is based on the solute transport governing equation in groundwater (ADE-GA), and to improve the genetic algorithm population size in MATLAB software. Simultaneously, a sandbox experiment considering the natural attenuation of pollutants is designed to verify the proposed improved method, so as to use it to identify groundwater pollution sources of actual contaminated site. This is the first time that the improved population-size genetic algorithm in MATLAB software has been verified in a sandbox experiment considering natural attenuation and then applied to a small-scale actual site. We hope that it can solve the problems of method validity verification and actual site application in the previous research, and provide a strong basis for the future application of ADE-GA to the actual site.

The paper is organized as follows: first, the governing equation of solute transport in groundwater, the basic principle of the genetic algorithm and the calculation steps in the identification process of groundwater pollution sources will be reviewed. Then, the migration and transformation of the petroleum contaminants in the sandbox experiment is described. The purposes were i) to obtain the relevant parameters in the process of ADE-GA identification and ii) to analyze the migration and transformation rules of petroleum pollutants entering the groundwater. The above experimental purposes can verify the accuracy of the ADE-GA during the identification of groundwater pollution sources. Subsequently, the validated method is applied to field site with unclear information and relatively small scope. Based on the analysis of the collected data, the reliability of the groundwater pollution source identification results produced by the method is verified. Finally, some conclusions are provided.

2. Methods and materials

2.1. Groundwater flow and transport equation

The solute transport process in sandbox groundwater is simplified into a one-dimensional solute transport model with plane instantaneous point source leakage. It is assumed that there is a uniform flow field in the semi-infinite aquifer. The direction of flow along the groundwater is in the positive direction of x. Then, according to the two-dimensional aquifer flow equation (Pinder and Bredehoeft, 1968; Singh and Datta, 2006; Xia et al., 2019), the equation of groundwater flow in this aquifer is:

$$\frac{\partial}{\partial x}\left(K\frac{\partial h}{\partial x}\right) + W = 0 \tag{1}$$

where K is the hydraulic coefficient of aquifer medium, m d^{-1}; h is hydraulic head, m; W is the amount of water flowing into or out of the aquifer per unit area in vertical direction per unit time (positive for inflow and negative for outflow), m d^{-1}.

Table 1
Overview of references on using genetic algorithms (GA) to identify groundwater pollution sources.

References	Approach	Application
Aral et al., 2001	Progressive genetic algorithm(PGA)	Hypothetical random heterogeneous unconfined aquifer
Mahinthakumar and Sayeed, 2005	Real encoded GA (RGA)/Binary encoded GA (BGA)-Local Search Methods	Hypothetical two-dimensional pollution site
Singh and Datta, 2006	GA-based Linked Simulation Optimization Model	Hypothetical two different study areas
Ayaz, 2017	GA based contaminant transport model	Hypothesized 1D and 2D contaminated sites
Xia et al., 2019	TD(Taguchi design)-based GA	Hypothetical site of irregular geometry and non-uniform media
Sophia and Bhattacharjya, 2020	Combined MODFLOW and MT3DMS simulation models with genetic algorithm	Hypothetical study area with homogeneous and isotropic aquifer

From studies of man-made releases of pollutants into groundwater, pollutants can quickly be transported downstream in groundwater due to advection and dispersion under steady-state flow (Ayvaz, 2016). It is assumed that the emission quality of the pollutant is M. Before the pollutant is released, the initial concentration of the pollutant is zero (initial condition), and the concentration of the pollutant at infinity is zero (boundary condition). Assuming that the transport of pollutants in groundwater is mainly controlled by advection and dispersion, then the solution formulas of the pollutant concentration at different locations are given (Bear, 1979):

$$C(x, t) = \frac{M/S}{2n\sqrt{\pi D_L t}} \exp\left[-\frac{(x-ut)^2}{4D_L t}\right] \quad (2)$$

where $C(x, t)$ is the concentration of contaminant at x for time t, g L^{-1}; D_L is the longitudinal diffusion coefficient, m^2 d^{-1}; u is the Darcy velocity, m d^{-1}; n is the effective porosity; S is cross-sectional area, m^2; and M is the mass of the pollutant, kg.

2.2. Basic principles of genetic algorithm

The genetic algorithm (GA) is a global optimization algorithm (Giacobbo et al., 2002) based on natural selection and genetic laws in the biological world (Goldberg, 1989). It has been widely used in the research of groundwater management and multitarget groundwater pollution (Kalayci et al., 2016; Mitra et al., 1998). The basic principle is to convert the relevant parameters of groundwater pollution source identification into the initial population in the GA. Then, the GA imitates the process of biological evolution, mainly by employing three operators (selection operator, mutation operator and crossover operator) to generate a new generation. This is to reduce the difference between the pollutant concentration of groundwater numerically simulated with the newly generated identification parameters and the concentration detected by the monitoring well downstream of the groundwater, thereby eliminating the results with lower fitness function values. Finally, through the number of predefined iterations or the range of set fitness functions (Bastani et al., 2010), the calculation is terminated to minimize the difference between simulated and monitored pollutant concentrations, so as to obtain the most suitable environment. Therefore, the problem of minimizing the difference between the source identification results and the real situation is solved. This is the optimal solution of the parameters of groundwater pollution source to be solved.

2.3. Steps for identifying groundwater pollution sources by genetic algorithm

The genetic algorithm can be used to identify the problem of groundwater pollution source identification in the following four steps, and the technical roadmap is shown in Fig. 1.

(a) Coding and generating initial populations: The characteristic parameters of the groundwater pollution source, including the location, release intensity and time of the pollution source (Ayvaz, 2010), are encoded by using a binary coding method. The relevant parameters of the identified pollution source are used as decision variables and then converted into genotype structure string data in the genetic algorithm. The genotype structure string form chromosome string. So the initial chromosome population is generated according to the pollution source parameters. The three parameters of the location, release intensity and time of the pollution source determine the number of chromosome genes in the genetic algorithm, and the accuracy of solving the three parameters determines the length of the chromosome. On the basis of fully investigating and understanding the hydrogeological conditions and other relevant information of the site, the range of the relevant parameters of the pollution source is initially given as the range of the initial population in the genetic algorithm. In this way, the algorithm can be quickly converged in the calculation process, which is conducive to the later calculation.

(b) Fitness function evaluation: The selection of the fitness function is the key to the algorithm, and it directly determines the convergence speed and the accuracy of the solution. Through the input of the relevant hydrogeological parameters, the groundwater solute transport model is used to calculate the concentration of pollutants at different times and locations. The fitness function is defined as the difference between the measured value and the simulated value of the concentration at each observation point (Jin et al., 2009) within the study area. Therefore, the fitness of the different individuals is evaluated. The essence of the pollution source inversion identification problem is to find an appropriate set of M, T_0 and X_0 parameters to minimize the difference between the simulated calculated value and the measured value. Therefore, the fitness function is defined as:

$$z = \min \sum_{i=1}^{n} [C(X_0, M, T_0) - C(X, T)]^2 \quad (3)$$

In the above equation, $C(X_0, M, T_0)$ is the simulated value of the pollutant concentration at X_0 for time T_0, which can be calculated according to Eq. (3); $C(X, T)$ is the concentration measured at different times in the monitoring well location. Where X_0 is the distance between the observation point and the pollution source; M is the quality of pollutant leakage; T_0 is the time between the first detection time and the leakage of the pollution source; X is the distance between the monitoring well location and the pollution source; T is the detection time of monitoring well.

(c) Population evolution: In the calculation process, the individuals with the best fitness values are first selected and directly entered into the calculation of the next generation population. For individuals with poor adaptability, crossovers operator are applied to generate individuals with relatively good adaptability before entering the next generation of calculation. Under the condition that neither of the previous two calculation processes can be satisfied, the mutation operator is used to make individuals generate the next generation through mutation (Ayaz 2017). The process of pollution source identification is to use the initially defined source parameter variables as the initial population. The source parameters are used to simulate and generate the simulated concentration of monitoring well position, which is compared with the monitoring concentration to generate the fitness value. When the fitness value is small, the source parameters can be directly selected into the calculation of the next generation. Otherwise, the calculation of crossover and mutation operators will be carried out to generate new individuals that can satisfy relatively good fitness function values and enter the next generation of calculation. According to the set conditions, all calculation processes are completed automatically.

(d) Termination condition judgment: The condition can be terminated by an algorithm, such as setting the maximum number of iterations or setting the range of fitness functions (Bastani et al., 2010). After the algorithm has evolved through multiple iterations, the optimal individual is produced as the result.

2.4. Laboratory sandbox experiment

According to the calculation steps of the genetic algorithm in the above section, it can be seen that the solute transport governing equation in groundwater is necessary to identify groundwater pollution sources by using the genetic algorithm. Simultaneously, obtaining the parameters of the solute transport equation is the key to the whole calculation process. In addition, the study of the solute migration and transformation law in groundwater can verify the accuracy of the

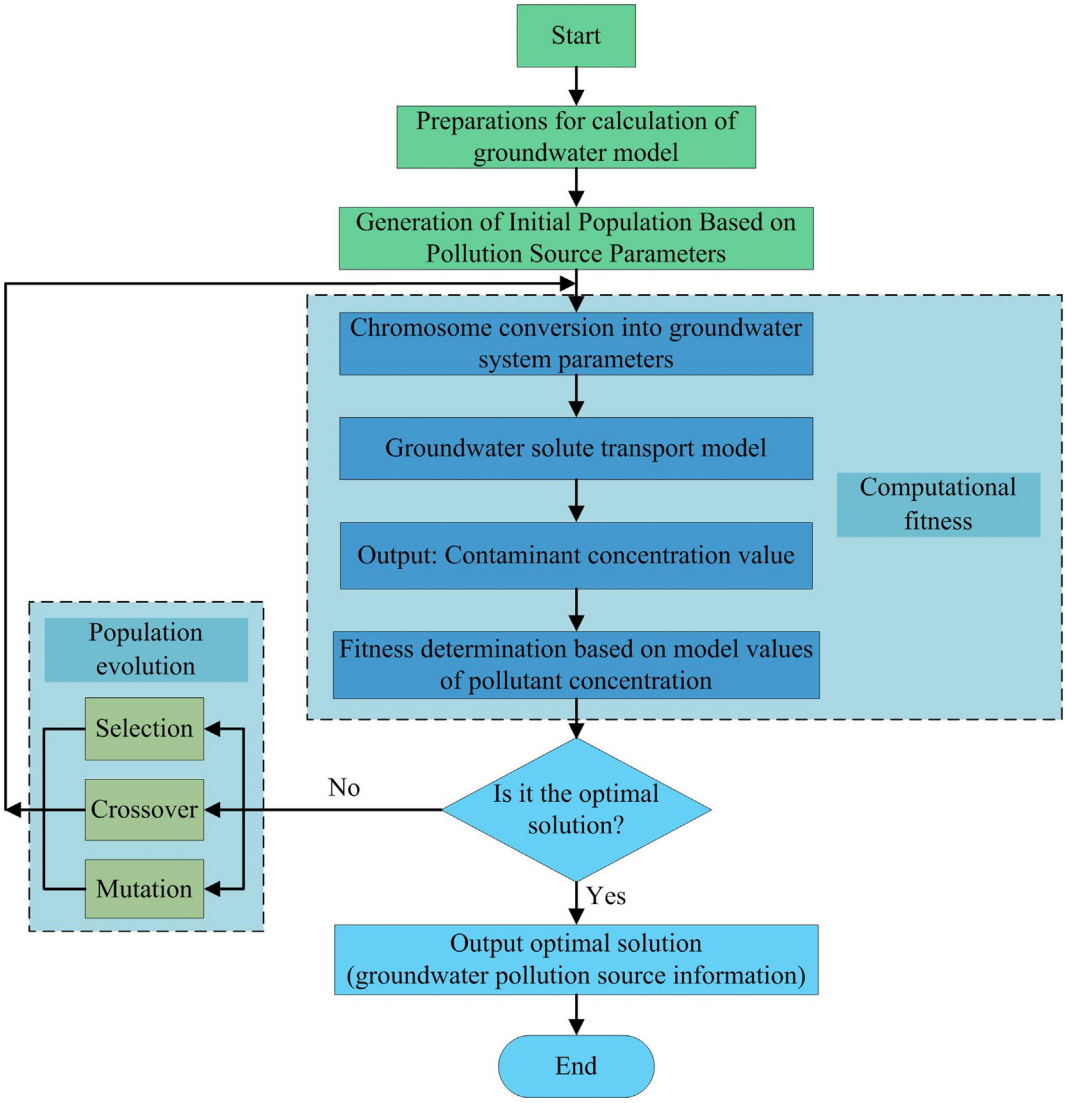

Fig. 1. Genetic algorithm process diagram.

genetic algorithm output results, which is also an important part of the research on employing the genetic algorithm to identify pollution sources. The laboratory sandbox experiment (Chen et al., 2018) can solve the above problems, which is the content described in this section.

2.4.1. Experimental reagents and instruments

The investigation results of the study area show that 0 # diesel oil is the main pollutant in the contaminated site. In order to keep the sandbox experiment consistent with the pollutants in the field, 0 # diesel oil was selected as the specific pollutant in the experiment. The other reagents required for the experiment include carbon tetrachloride (specifically used for the infrared oil meter), and anhydrous sodium sulfate. The basic properties of the # 0 diesel oil are as follows: melting point is −29.56 °C, boiling point is 282–370 °C, and its relative density is 0.83–0.855 g $(cm^3)^{-1}$. The main instruments required for the experiment are the infrared oil detector, ultrasonic cleaner, peristaltic pump and magnetic stirrer.

2.4.2. Experimental device

The experimental device is a sandbox with a size of 160 cm × 30 cm × 80 cm, and the material is transparent plexiglass (Fig. 2). The left and right sides of the sandbox are the groundwater inlet and outlet zones, respectively. Their lengths are all 10 cm, which was selected to allow for adjustments in the groundwater level differences. A total of 4 × 5 = 20 sampling holes were established in front of the sandbox and constructed via a soil solution sampler (terracotta head). A total of four rows (A, B, C, D) are set in the longitudinal direction of the water intake hole, and there are five water intake holes in each row arranged horizontally. The main sampling holes of this experiment are B1-B5. There are also three layers of thirty-mesh stainless steel mesh on both sides of the migration area. The steel mesh has thirty meshes within a length of 25.4 mm, of which the mesh wire diameter is 0.33 mm and the mesh diameter is 0.52 mm in the GB T 5330-2003 (2003) of China. This ensures that the soil medium does not leak and has a hydraulic connection, so that the hydraulic gradient can remain stable.

2.4.3. Experimental steps

The experimental steps include preparing the pollutants, loading the experimental media, establishing the water level, releasing the pollutants, sampling and testing. The experimental process is shown in Fig. 3.

First, groundwater was added into the 50 L vat, and then an excessive volume of the # 0 diesel oil was added to it until it was saturated. Stirring the groundwater solution continued for 12 h at a speed of 500 r min^{-1} to homogenize and dissolve the solution and to obtain the contaminated liquid. Second, the vadose zone medium was slowly filled into the plexiglass sandbox, and the filling height was fifty-five

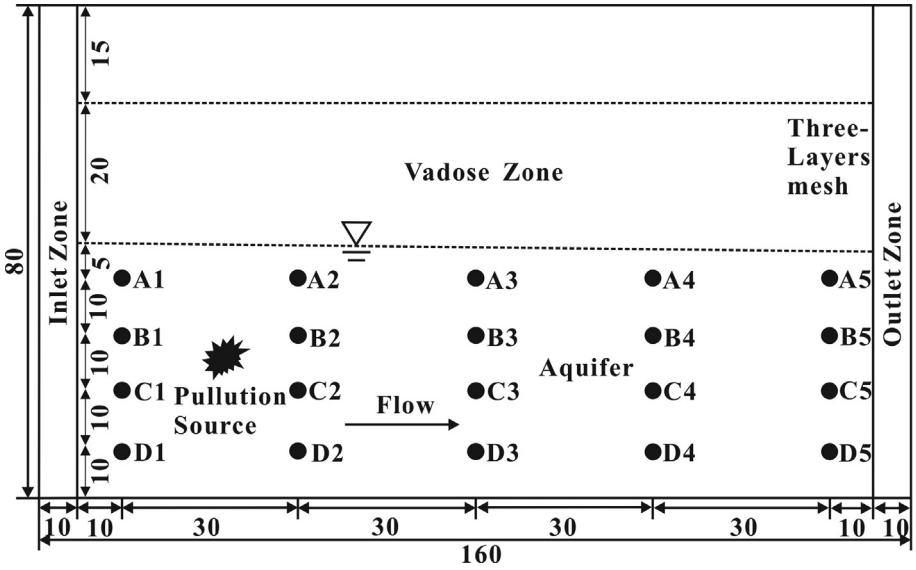

Fig. 2. Experimental device diagram.

centimeters. To ensure the uniformity of the medium in the sandbox during the filling process, the medium should be continuously compacted to avoid the occurrence of voids and cracks that can cause a dominant flow channel. After loading, the two peristaltic pumps are used to allow water to inflow from the left side and outflow from the right side of the sandbox, respectively. The filled medium is filled with water several times through this process to ensure sufficient exhaust in the medium. Subsequently, by repeatedly adjusting the inflow and outflow flow rates of the peristaltic pumps on the left and right sides to ensure that the upstream and downstream areas can evenly enter and exit the water. And make the inlet water level on the left side greater than the right side, which can control the groundwater level on both sides of the sandbox. Finally, a stable hydraulic gradient is formed. After the groundwater flow rate in the plexiglas sandbox is stabilized, the groundwater still keeps flowing in and out from the inlet and outlet, respectively. Then the previously prepared pollutant liquid can be injected into the sandbox (Citarella et al., 2015). Finally, the total petroleum hydrocarbon (TPH) content was measured by an infrared oil meter.

2.5. Study area

The study area is located in the Miyun District of Beijing, China, where the Yanshan Mountains meet the North China Plain and where the interaction of the piedmont alluvial fans in the upper reaches of the Chaobai River occurs. There is a gas station in the study area that was abandoned in the 1970 s. However, due to the unreasonable structure of this old gas station and the oxidation effect over the years, the underground storage tank (including 6 diesel tanks and 4 gasoline tanks) has been corroded. As a result, leakage occurs that contaminates the surrounding soil and groundwater. The contaminated site is in the shape of an inverted trapezoid (Fig. 4) with an area of approximately 8700 m^2. The terrain is flat with an average elevation of 56–57 m and slopes from north to south. The groundwater flow is approximately the same as the topographical change, flowing from north to south. The oil hydrocarbon content of the soil in the study area was within the range of 3.17–16.47 mg kg^{-1}, which showed an overall trend of first increasing and then decreasing. The oil hydrocarbon concentration reached a maximum at 10 m below the surface.

3. Results and discussion

3.1. Sandbox experimental results and discussion

Monitoring wells were used to measure the total petroleum hydrocarbon concentrations, pH values, and concentrations of conventional

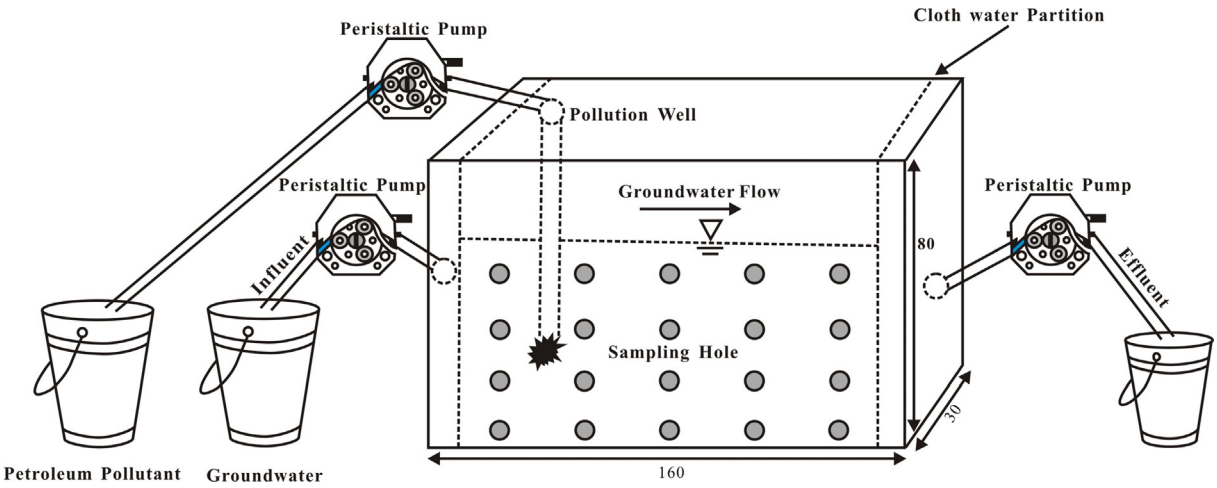

Fig. 3. Experimental flow chart.

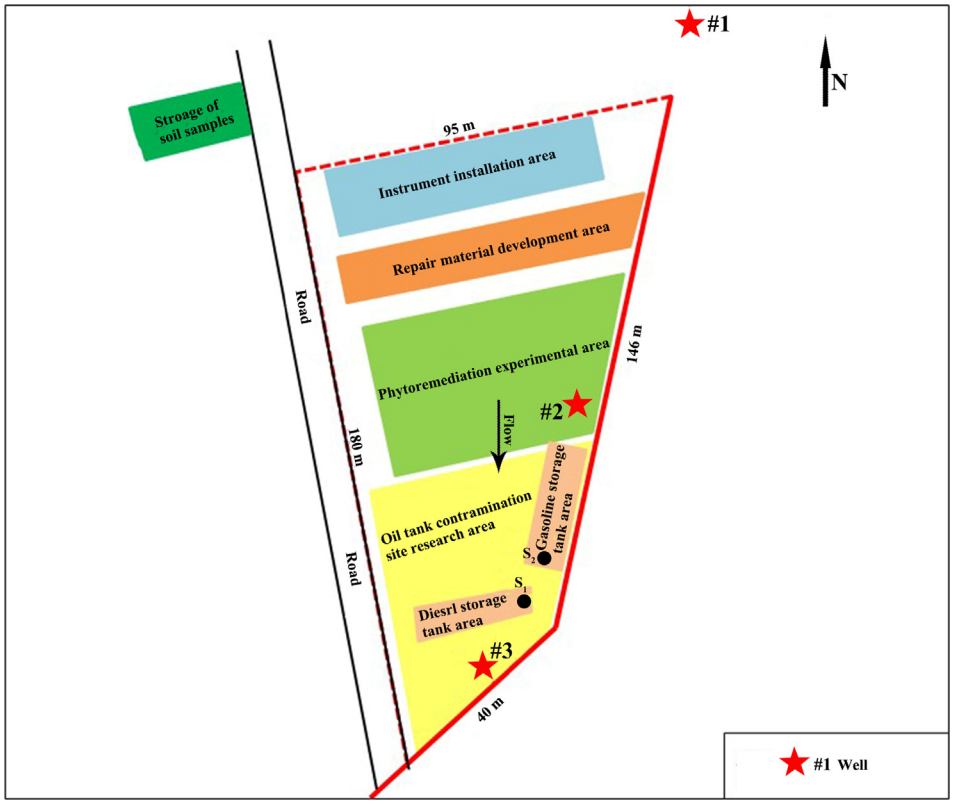

Fig. 4. Location map of each functional block in the study area.

ions such as NO_3^- and SO_4^{2-} in the groundwater. Thus, the migration and transformation rules of petroleum pollutants in the groundwater were analyzed, and the change in the petroleum pollutants in the groundwater were studied more intuitively. This process can be used to verify the accuracy of the identification results of groundwater pollution sources using a genetic algorithm under experimental conditions and to provide a reliable basis for the application of this method at engineering sites.

3.1.1. Variation in petroleum hydrocarbon concentration at different locations

The test results of the total petroleum hydrocarbons in the water samples at the four sampling holes—B2-B5 (0.15, 0.45, 0.75, 1.05 m, respectively, from the pollution sources)—were selected, and the concentration curves were plotted for the different positions (Fig. 5). Once the pollutants enter the aquifer, they will continue to spread under advection and dispersion (Chen et al., 2018), resulting in higher concentrations of pollutants. Then, the pollutants migrate downstream in a concentration gradient. As seen from Fig. 5, the initial concentration of pollutants at B2 is the largest, which is the hole closest to the source of pollution. The maximum initial concentration of pollutants at this point is approximately 11.0 mg L^{-1}, indicating that there is some stagnation and accumulation occurring near the source of pollution after the pollutants enter the aquifer. With the passage of time, the concentration of pollutants shows a gradual decline, which is mainly caused by the instantaneous leakage of pollution sources and the downstream migration of pollutants under the action of groundwater. However, in places far from the pollution sources, such as 1.05 m away, the concentration of pollutants first increases and then decreases. This trend is observed because the pollutants have not yet reached this point during the initial stage. Affected by convection and diffusion (Qian et al. 2018), the concentration of pollutants is only 1.4 mg L^{-1} at this location when the experiment is conducted for 0.2 days. However, at this time, the pollutant concentrations of 0.15 m, 0.45 m and 0.75 m away from the pollution source location are 11.0, 10.0 and 4.0 mg L^{-1}, respectively. It takes one day for the location to reach the peak concentration of 5.0 mg L^{-1}. Subsequently, as with other locations, the concentration gradually decreases due to the effects of natural degradation (Atlas, 1991). From the inlet section to the outlet section, the B2, B3, B4, and B5 monitoring holes sequentially showed the maximum concentration of pollutants, reaching 11.0, 10.0, 6.5, and 5.0 mg L^{-1}, respectively. These concentrations are 91%, 83%, 54%, and 42% of the initial concentrations, respectively. Due to the natural attenuation of petroleum pollutants in aquifers, such as degradation (Atlas, 1991), the peak concentration of pollutants in the sandbox gradually decreases with the extension of the pollutant migration distance. In addition, it can be seen from Fig. 5 that the concentration value at each point after ten days is approximately 2.0 mg L^{-1}.

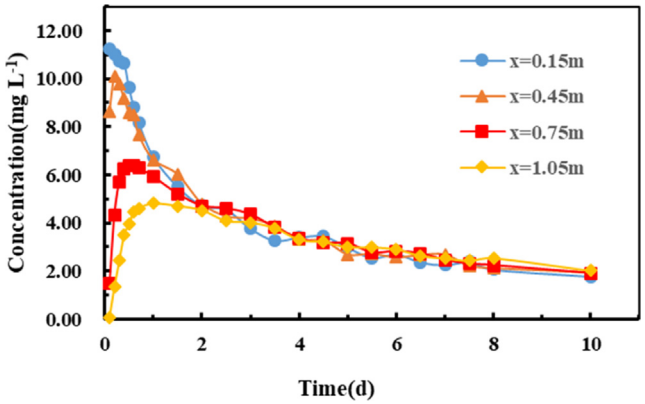

Fig. 5. Curves of the total petroleum hydrocarbon concentration at different locations.

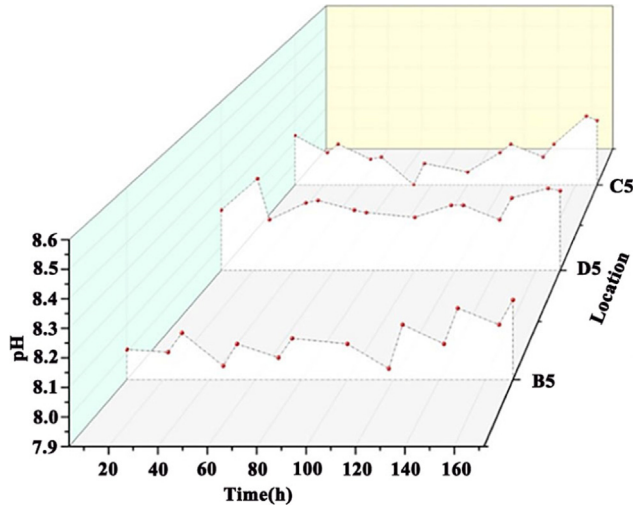

Fig. 6. Changes in pH value with time at the different locations.

3.1.2. Variation in pH value

The diagram showing the variation in pH value with time was drawn (Fig. 6) by testing and analyzing the pH value of the water samples at three locations (B5, C5, D5) near the outlet. As seen from Fig. 6, there is a certain fluctuation in pH value at each location. However, the overall trend in pH value is to first decrease and then increase. The reason for this phenomenon is mainly related to the activities of microorganisms (Atlas, 1991). On the one hand, microorganisms in this process consume oxygen from the water for respiration and produce a certain amount of CO_2, which is slightly dissolved in water and turns into H_2CO_3 (Bellagamba et al., 2016). On the other hand, microorganisms use a variety of organic compounds in petroleum contaminants as carbon sources for their own growth to perform a series of complex redox reactions (Lari et al., 2019). During this process, petroleum hydrocarbons are first oxidized to alcohols. Subsequently, alcohols continue to be oxidized to the corresponding aldehyde under the action of alcohol dehydrogenase. Finally, aldehydes continue to be oxidized to fatty acids in the presence of aldehyde dehydrogenase (Borden et al., 1995; Gieg and Suflita, 2002). These processes can be expressed as follows:

$$R\text{-}CH_2\text{-}CH_3 + O_2 \rightarrow R\text{-}CH_2\text{-}CH_2\text{-}OH \rightarrow R\text{-}CH_2\text{-}CHO\text{-}R \rightarrow R\text{-}CH_2\text{-}COOH$$

The formation of acid in groundwater is an important reason for the decrease in the pH value of the solution (Van Stempvoort et al., 2007).

Once the contaminants are degraded to a certain degree, the degradation of microorganisms will be abated. The medium used in the experiment is mainly aquifer medium obtained from the study area, such as gravel, coarse and medium sand. Affected by the interaction of water and aquifer medium, the environment of groundwater will change, resulting in the increase of pH value. However, HCO_3^- has little change in this process (Qian et al. 2018).

In addition, the magnitude of change at B5 and C5 is significantly stronger than that at D5. The main reason for this phenomenon is that the location of pollutant leakage occurs between the lines of the B and C holes, and microorganisms play a more pronounced role in the flow direction of groundwater. The density of petroleum contaminants is smaller than that of water, and its concentration in the lower layer is low, resulting in weaker microbial effects. Therefore, the change in pH value at point D5 is relatively small. In general, regardless of the pH value, the changes in pH value are very small because of the low concentration of petroleum pollutants in groundwater and relatively weak microbial action.

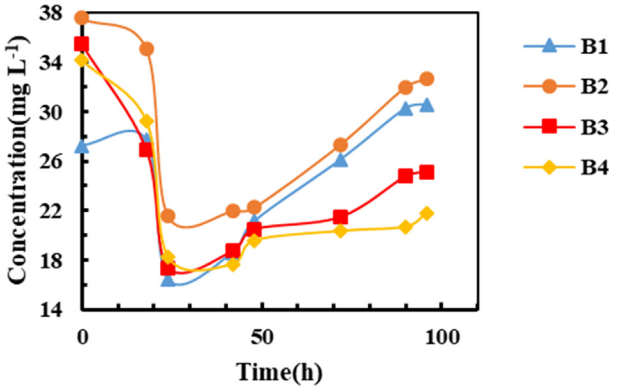

Fig. 7. Concentration curve of NO_3^- at the different locations.

3.1.3. Variation regular ions

3.1.3.1. Changes in NO_3^- content. Fig. 7 shows that the NO_3^- content at the monitoring location first decreases and then increases. Among the locations, the B1, B2 and B3 locations increase considerably in NO_3^-content during the later stage, while that of B4 increases relatively little. Except for the B1 monitoring location, the initial concentration of NO_3^- at the remaining locations was 35.0 mg L^{-1}. As the reaction progressed, the NO_3^-concentration at each location began to decrease, reaching a minimum value at approximately 20 h and presenting a reduction of approximately 50%. The concentration at each monitoring location then increases rapidly. In particular, the NO_3^-content at the B1 and B2 locations were as high as 32.0 mg L^{-1} at 100 h. However, since B3 and B4 are relatively far away from the pollution source, the increase of NO_3^- concentration is small. Moreover, NO_2^- was detected in some samples, which may be caused by the denitrification of microorganisms (Kuhn et al., 1988; Hutchins et al., 1991).

The reason for the decrease in NO_3^- content is that microorganisms require certain carbon sources, energy and electron acceptors during the whole reaction process (Kuhn et al., 1988; Hutchins et al., 1991). The carbon source is organic matter in petroleum pollutants, while the electron acceptor may be dissolved oxygen (DO), nitrate (NO_3^-), ferric iron (Fe_3^+), and sulfate (SO_4^{2-}) in water (Qian et al. 2018). The order in which the above electron acceptors are utilized is DO > NO_3^- > Fe_3^+ > SO_4^{2-} (Borden et al., 1995; Choi and Lee, 2011; Guo et al., 2020). That is, microorganisms prioritize the use of DO as an electron acceptor and only use NO_3^- when the DO is insufficient (Anderson and Lovley, 1997). The biological reaction of NO_3^- as an electron acceptor is as follows:

$$CH_2O + 2NO_3^- \rightarrow H_2O + CO_2 + 2NO_2^-$$

3.1.3.2. Changes in SO_4^{2-} content. From Fig. 8, compared to NO_3^-, the change in SO_4^{2-} concentration over time is relatively small. During the whole reaction process, the content fluctuation range of SO_4^{2-} at the B1-B4 monitoring locations was very small, remaining between 35.0 and 42.0 mg L^{-1}, and the change was only approximately 10%. The reason for this result can be attributed to the order in which SO_4^{2-} is used as an electron acceptor (Guo et al., 2020): as mentioned above, the use of SO_4^{2-} lags behind that of NO_3^-. When NO_3^- ions are still present in the groundwater environment, SO_4^{2-} does not participate in the reaction (Anderson and Lovley, 1997; Guo et al., 2020), so its concentration varies within a small range.

In summary, microbial degradation occurs during groundwater flow (Batlle-Aguilar et al., 2009; Kao et al., 2010). It has been reported that a considerable number of microorganisms, including bacteria and fungi (Anderson and Lovley, 1997), are capable of degrading petroleum hydrocarbons in the aeration zone medium and the groundwater environment (Kwon et al., 2018). Microorganisms use petroleum

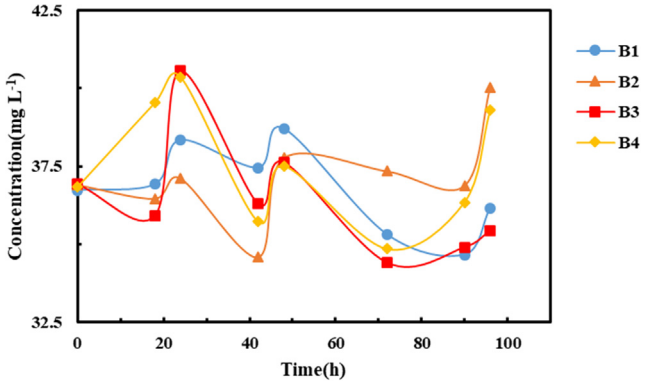

Fig. 8. Concentration curve of SO_4^{2-} at the different locations.

Table 2
Parameter values used by GA in the identification of pollution sources in sandbox experiment and study area.

Parameters	Values in sandbox experiment	Values in study area
Number of chromosomes/ population size	100	80
Number of generations	120	100
Mutation probability	0.05	0.05
Crossover probability	0.8	0.8

pollutants as a carbon source to carry out their own growth and metabolism (Anderson and Lovley, 1997). The essence of this process is that microorganisms participate in the corresponding redox reaction to convert toxic and harmful substances into N_2, H_2O and other nontoxic substances (Guo et al., 2020). After oxygen is consumed rapidly, microorganisms mainly undergo anaerobic degradation (Chiu et al., 2017). According to the thermodynamic principle, the ability of an electron acceptor to compete for electrons is inversely proportional to its free energy; that is, the smaller the free energy is, the easier the electron acceptor is to react, so dissolved oxygen is utilized in preference to nitrates and sulfates (Maric et al., 2019). However, in the actual groundwater environment, especially in mountainous areas with large burial depths, dissolved oxygen is easily consumed but difficult to replenish (Anderson and Lovley, 1997). Therefore, the groundwater system is mostly in a hypoxic environment, and nitrate is first involved in redox reactions as an electron acceptor (Kuhn et al., 1988; Hutchins et al., 1991; Borden et al., 1995). When groundwater is polluted, a

Table 3
The truth values of pollution source parameters in sandbox experiment and genetic algorithm inversion of pollution source parameters results.

Number	Parameters	x = 1.05 m section truth values	x = 1.05 m section simulation values	Relative error(%)
1	M(kg)	0.6500	0.6240	−4.00
2	T_0(d)	0.1000	0.0970	−3.00
3	X_0(m)	1.0500	1.0410	−0.86

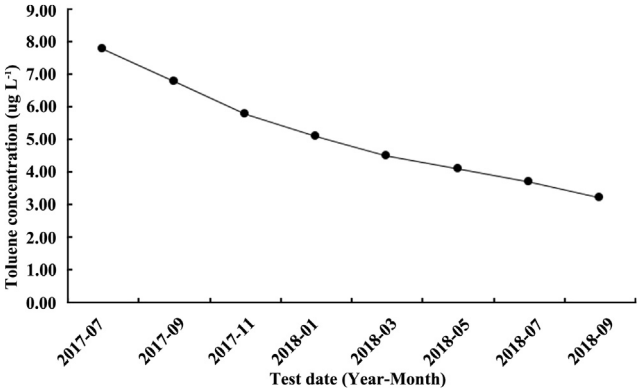

Fig. 10. Changes in toluene content in groundwater.

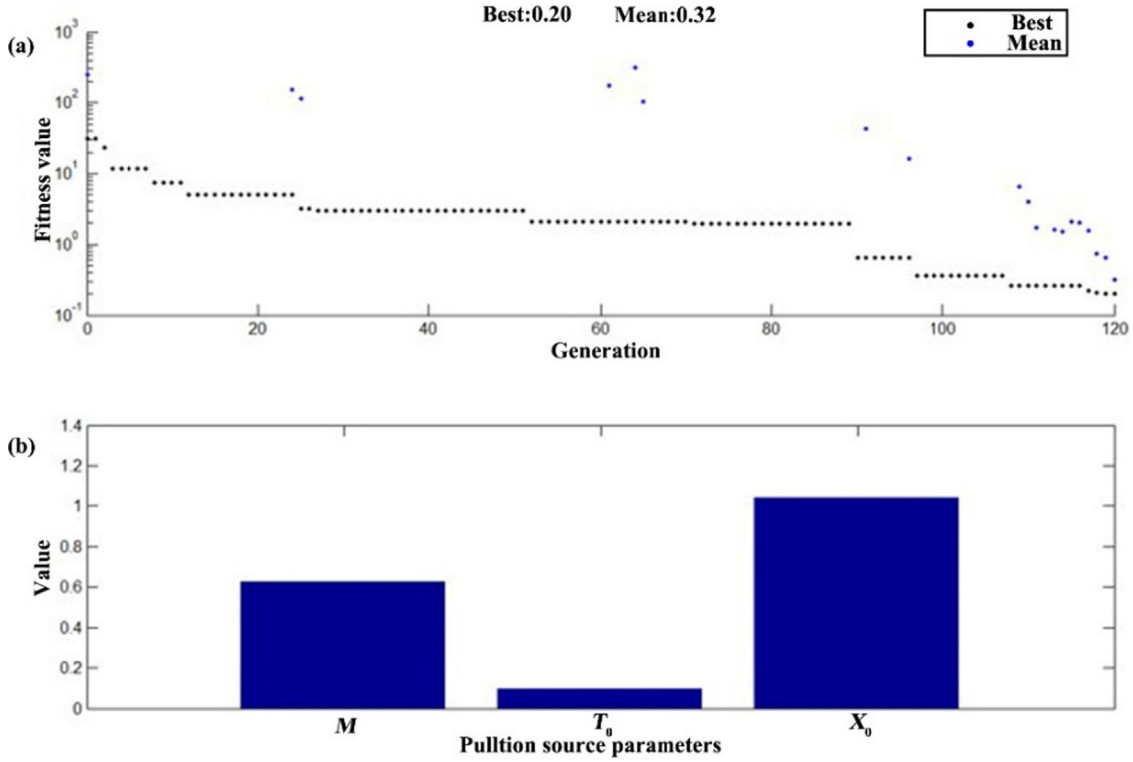

Fig. 9. Genetic algorithm inversion output at the x = 1.05 m section.

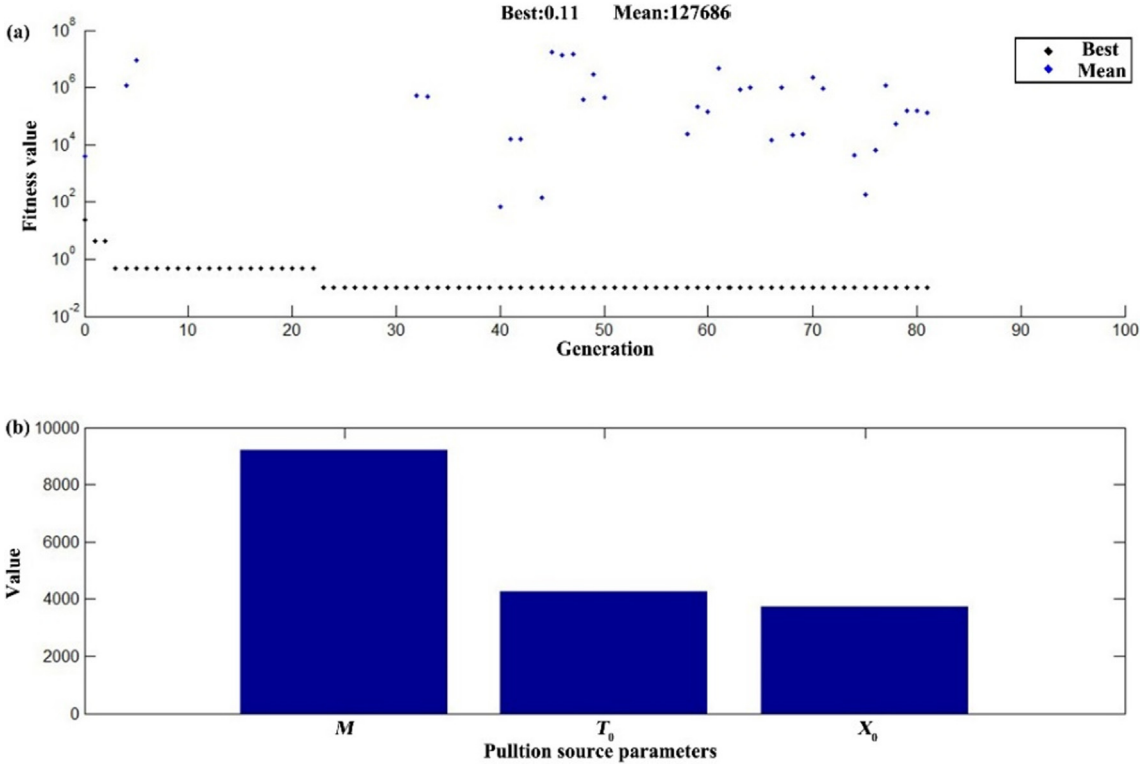

Fig. 11. Genetic algorithm inversion output at the study site.

Table 4
The inversion parameters of the genetic algorithm.

Method	Pollution source parameters		
	M(kg)	T_0(d)	X_0(m)
Genetic algorithm	9.2	4269	37.2

transformation process from an oxidizing environment to a reducing environment will take place (Wilson et al., 1986; Anderson and Lovley, 1997).

Through the above analysis, it can be seen that after petroleum hydrocarbon enters groundwater, it has the effects of convection, diffusion and biodegradation (Chiu et al., 2013). Under these effects, the migration range of pollutants in groundwater and the concentration in downstream monitoring wells will change accordingly. These changes will affect the selection of monitoring well locations and the determination of specific contaminants, respectively. These two factors are the key links for identifying the location and release intensity of groundwater pollution sources using GA. In addition, the study on the migration and transformation of petroleum hydrocarbon pollutants in sandbox experiment can help to determine the relevant parameters required by the method.

3.1.4. Parameter acquisition

According to the migration rules of petroleum pollutants in the sandbox groundwater and relevant literature review, the following parameters can be obtained:

Porosity. With reference to empirical values in the literature (Qin et al., 2013; Cao et al., 2017), the porosity of the sand is between 0.25 and 0.5. Combined with the loading medium in the experiment, the porosity n is equal to 0.3.

Hydraulic conductivity and groundwater velocity. During the experiment, when the groundwater level on both sides was stable, the groundwater flow was found to be 18.18 ml min^{-1}. The hydraulic gradient of groundwater is 2.19%. According to $Q = KIA$ (Bear, 1979), it is known that the hydraulic conductivity K is equal to 9.78 m d^{-1}. The above results are combined with the groundwater flow continuity equation (1), so the groundwater velocity u is equal to 0.2139 m d^{-1}.

Longitudinal diffusion coefficient. Diffusion is an important measure of the dispersion characteristics of porous media, and it is usually obtained by field experiments. However, the cost of field experiments is relatively high, and diffusion data is difficult to obtain; thus, diffusion data is generally obtained by the corresponding empirical formula (Gelhar et al., 1992; Neuman and Zhang, 1990; Xu and Eckstein, 1995). In this experiment, the value of a_L is 2.34 m, which is based on the diffusion coefficient $D_L = a_L \times u = 0.5$ m^2 d^{-1} (Chen et al., 2018).

3.1.5. ADE-GA identifies the source of pollution under experimental conditions

The genetic algorithm toolbox based on MATLAB software (Mathworks, 2010) and the parameters used in the genetic algorithm calculation process under experimental conditions were used to invert the pollution source parameters at the $x = 1.05$ m section. The relevant parameters of the genetic algorithm are set as follows (Table 2): The number of chromosomes, that is, the population size is 100. The termination evolutionary number is 120 generations, and the crossover and mutation probabilities use the system default values of 0.8 and 0.05, respectively (Bastani et al., 2010). The result is shown in Fig. 9.

Fig. 9a shows the changes in fitness function values under different biological generations. The ordinate is on a logarithmic scale. During each generation, the black dots represent the best fitness function values, and the blue dots represent the average fitness values. Fig. 9b shows the pollution source-related parameter outputs by the system after multiple optimization iterations. Fig. 9a shows that the optimal fitness function value and average value of the $x = 1.05$ m section are 0.20 and 0.32, respectively. The results of the pollution source parameters obtained by the genetic algorithm are shown in Table 3.

In conclusion, the results of the genetic algorithm are very close to the “true value” of the released pollution source in the experiment. It

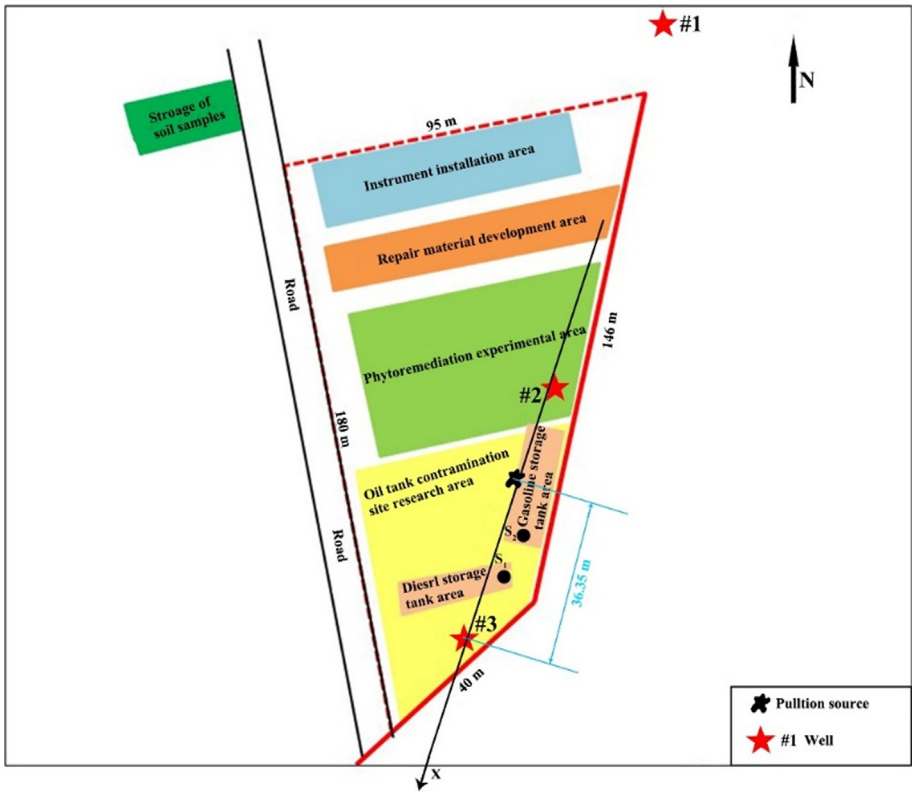

Fig. 12. Location of pollution sources as identified by the inversion of the study site.

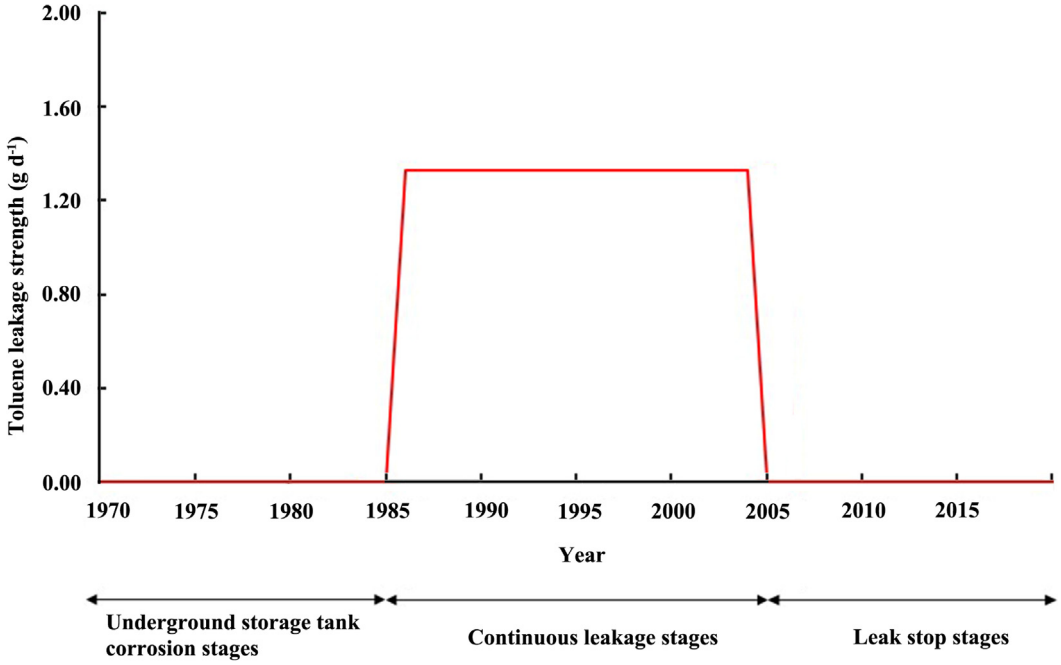

Fig. 13. Pollution history of the inversion of the study site.

has the characteristics of high accuracy and fast convergence. In addition, the accuracy of the genetic algorithm is confirmed by analyzing the migration law of the petroleum pollutants in the sandbox. Therefore, a genetic algorithm can be applied to the inverse problem of groundwater pollution sources.

3.2. Site application results and discussion

3.2.1. Site parameter determination

According to the parameters used in the calculation of the genetic algorithm, the site parameters are determined as follows:

The medium of the aquifer in the site is large gravel with some coarse sand; thus, its hydraulic conductivity is 45 m d^{-1}, and the porosity n is 0.25 based on field tests. The average hydraulic gradient of

the site is 1‰, and the Darcy velocity u is equal to (Bear, 1979) 0.045 m d^{-1}. The dispersion a_L value of the site is 5.3 m, and its dispersion coefficient D_L is 0.2385 m^2 d^{-1}. The average thickness of the aquifer is 3 m, and the width is approximately 40 m; thus, its cross-sectional area S is equal to 120 m^2.

3.2.2. Site-specific pollutant

Toluene is the main component in diesel (Kao et al., 2010; Lin et al., 2012, Chiu et al., 2013). The toluene content in the study area is still very high, so selecting the toluene in aquifer as the site-specific pollutant. From July 2017 to September 2018, the # 3 well in the study area was continuously monitored every two months. The test results are shown in Fig. 10.

3.2.3. Identification of pollution sources

The research site was analyzed according to the ADE-GA and verified under the experimental conditions. The main parameters of the genetic algorithm are set as follows (Table 2): the maximum number is 100 generations, and the population size is 80. The system default values are 0.05 and 0.8 for mutation and crossover probability, respectively (Bastani et al., 2010).

The results of the genetic algorithm are shown in Fig. 11. The research shows that the algorithm converges at the 81st generation, and then the program ends. The optimal fitness function value was 0.1056. According to the results of the genetic algorithm, the pollution source parameters are shown in Table 4.

3.2.4. Verification and discussion

The identification results of pollution sources can be validated and analyzed by using the data collected during the earlier period and from on-site investigations. The calculation result of the source location X_0 is 37.2 m; that is, the source location is 37.2 m upstream of the No. 3 observation well. By using the groundwater flow as the x direction, it can be found that 37.2 m upstream is located in the gasoline irrigation area (Fig. 12), which is consistent with the actual position of the site.

Second, the calculated result of the pollutant leakage time T_0 is 4269 d; that is, the interval between pollutant leakage time and the first sampling test (July 2017) is 4269 d, which is approximately 12 years. The initial estimate is that the leakage approximately occurred in 2005. According to preliminary data, the gas station was abandoned in the 1970 s. The old-fashioned underground storage tanks in China mainly used steel storage tanks with a steel plate thickness of 6 mm. Due to its long-term contact with air and moisture, the corrosion caused the thickness of the steel sheet to become thinner. Eventually, oil contaminants leaked. Under the influence of corrosion, the damage to the storage tank takes 15 years. More than 70% of underground storage tanks leak if they are older than 15 years. Therefore, the site of the study area experienced corrosion damage from 1970 to 1985. The tanks continued to leak for 20 years between 1985 and 2005. After the pollutants in the oil storage tank stopped leaking in 2005, the pollutants were transported downstream under the action of groundwater. The entire pollution history process is shown in Fig. 13.

Finally, the quantity of the identified pollutants is analyzed. The calculated result of M was 9.2 kg; that is, a total of 9.2 kg of toluene entered the aquifer. Then, the leakage strength of toluene can be calculated to be 1.33 g d^{-1}. Therefore, when the steel plate of the storage tank is not completely corroded, the pollutant leakage intensity was zero. After the pollutants in the oil storage tank stopped leaking in 2005, the pollution source intensity was also zero.

In the above process, the whole historical process of pollution is deduced and verified based on the data collected from field exploration and the early stage. The results of the genetic algorithm are in line with the real results in the field. This work shows that the genetic algorithm can be applied to the solution of pollution sources at the field scale.

4. Conclusions

In this paper, a genetic algorithm combined with the advection–dispersion equation are applied to the identification of groundwater pollution sources at the experimental scale and site scale. The monitoring results of the physical experiment obtained by using the sandbox suggest that the identification result is reliable. Analysis of the field data shows that the ADE-GA is also reliable for identifying groundwater pollution sources at the site scale. Among the results, the sandbox test also obtains the regularity of pollutant migration in groundwater, and microbial degradation is accompanied by the migration of pollutants in groundwater, which plays an important role in the rational interpretation of the pollutant source identification at both the experimental scale and site scale. The application of this method to the identification of groundwater pollution sources can not only solve the problem of minimizing the difference between the identification results and the true value, but also provide a reliable reference for the control and efficient remediation of site pollution.

CRediT authorship contribution statement

Kexue Han: Writing - original draft, Methodology, Software, Validation, Formal analysis. **Rui Zuo:** Writing - review & editing, Project administration, Resources, Supervision, Validation, Funding acquisition. **Pengcheng Ni:** Software, Investigation. **Zhenkun Xue:** Data curation. **Donghui Xu:** Investigation. **Jinsheng Wang:** Writing - review & editing, Supervision. **Dan Zhang:** Validation, Formal analysis.

Declaration of Competing Interest

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

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant nos. 41672228 and 41877181) and the National Water Pollution Control and Treatment Science and Technology Major Project (Grant no. 2018ZX07109-003).

References

Aeppli, C., Tysklind, M., Holmstrand, H., Gustafsson, Ö., 2013. Use of Cl and C isotopic fractionation to identify degradation and sources of polychlorinated phenols: mechanistic study and field application. Environ. Sci. Technol. 47, 790–797.
Anderson, R.T., Lovley, D.R., 1997. Ecology and biogeochemistry of in situ groundwater bioremediation. Adv. Microb. Ecol. 15, 289–350.
Aral, M.M., Guan, J., Maslia, M.L., 2001. Identification of contaminant source location and release history in aquifers. J. Hydrol. Eng. 6, 225–234.
Atlas, R.M., 1991. Microbial hydrocarbon degradation bioremediation of oil spill. J. Chem. Tech. Biotechnol. 52, 149–156.
Atmadja, J., Bagtzoglou, A.C., 2001. Pollution source identification in heterogeneous porous media. Water Resour. Res. 37, 2113–2125.
Ayaz, M.d., 2017. Groundwater pollution source identification using genetic algorithm based optimization model. Int. J. Comput. Sci. 5, 65–72.
Ayvaz, M.T., 2010. A linked simulation–optimization model for solving the unknown groundwater pollution source identification problems. J. Contam. Hydrol. 117, 46–59.
Ayvaz, M.T., 2016. A hybrid simulation-optimization approach for solving the areal groundwater pollution source identification problems. J. Hydrol. 538, 161–176.
Bagtzoglou, A.C., Atmadja, J., 2005. Mathematical methods for hydrologic inversion: the case of pollution source identification. Handb. Environ. Chem. 5. https://doi.org/10.1007/b11442.
Bastani, M., Kholghi, M., Rakhshandehroo, G.R., 2010. Inverse modeling of variable-density groundwater flow in a semi-arid area in Iran using a genetic algorithm. Hydrogeol. J. 18, 1191–1203.
Batlle-Aguilar, J., Brouyère, S., Dassargues, A., Morasch, B., Hunkeler, D., Höhener, P., Diels, L., Vanbroekhoven, K., Seuntjens, P., Halen, H., 2009. Benzene dispersion and natural attenuation in an alluvial aquifer with strong interactions with surface water. J. Hydrol. 369, 305–317.
Bear, J., 1979. Dynamics of Fluids in Porous Media. American Elsevier, New York.

Bellagamba, M., Viggi, C.C., Ademollo, N., Bossetti, S., Aulenta, F., 2016. Electrolysis-driven bioremediation of crude oil-contaminated marine sediments. New Biotechnol. 38, 84–90.
Boano, F., Revelli, R., Ridolfi, L., 2005. Source identification in river pollution problems: a geostatistical approach. Water Resour. Res. 41. https://doi.org/10.1029/2004WR003754.
Borden, R.C., Gomez, C.A., Becker, M.T., 1995. Geochemical indicators of intrinsic bioremediation. Ground Water 33, 180–189.
Bronders, J., Tirez, K., Desmet, N., Widory, D., Petelet-Giraud, E., Bregnot, A., Boeckx, P., 2012. Use of compound-specific nitrogen (d(15)N), oxygen (d(18)O), and bulk boron (d(11)B) isotope ratios to identify sources of nitrate-contaminated waters: a guideline to identify polluters. Environ. Forensics 13, 32–38.
Butera, I., Tanda, M.G., 2003. A geostatistical approach to recover the release history of groundwater pollutants. Water Resour. Res. 39. https://doi.org/10.1029/2003WR002314.
Butera, I., Tanda, M.G., Zanini, A., 2013. Simultaneous identification of the pollutant release history and the source location in groundwater by means of a geostatistical approach. Stoch Environ. Res. Risk Assess 27, 1269–1280.
Cao, G.Z., Qin, R.G., Wu, Y.Q., Wu, J., Xu, Z.G., Zhang, C., 2017. Effects of source size, monitoring distance and aquifer heterogeneity on contaminant mass discharge and plume spread uncertainty. Environ. Fluid Mech. 18, 465–486.
Chen, B., Liu, J., Hu, L.M., Liu, M., Liang, W., Zhang, X.L., Fan, D.J., 2017. Spatio-temporal distribution and sources of Pb identified by stable isotopic ratios in sediments from the Yangtze River Estuary and adjacent areas. Sci. Total Environ. 580, 936–945.
Chen, Z., Gómez-Hernández, J.J., Xu, T., Zanini, A., 2018. Joint identification of contaminant source and aquifer geometry in a sandbox experiment with the restart ensemble Kalman filter. J. Hydrol. 564, 1074–1084.
Chiu, H.Y., Hong, A., Lin, S.L., Surampalli, R.Y., Kao, C.M., 2013. Application of natural attenuation for the control of petroleum hydrocarbon plume: mechanisms and effectiveness evaluation. J. Hydrol. 505, 126–137.
Chiu, H.Y., Verpoort, F., Liu, J.K., Chang, Y.M., Kao, C.M., 2017. Using intrinsic bioremediation for petroleum–hydrocarbon contaminated ground-water cleanup and migration containment: effectiveness and mechanism evaluation. J. Taiwan Inst. Chem. E. 72, 53–61.
Choi, H.M., Lee, J.Y., 2011. Groundwater contamination and natural attenuation capacity at a petroleum spilled facility in Korea. J. Environ. Sci. 23 (10), 1650–1659.
Citarella, D., Cupola, F., Tanda, M.G., Zanini, A., 2015. Evaluation of dispersivity coefficients by means of a laboratory image analysis. J. Contam. Hydrol. 172, 10–23.
Cupola, F., Tanda, Maria, G., Zanini, A., 2015. Laboratory sandbox validation of pollutant source location methods. Stoch. Env. Res. Risk. A. 29 (1), 169–182.
Gandhi, B.G.R., Bhattacharjya, R.K., Satish, M.G., 2016. Simulation–optimization-based virus source identification model for 3D unconfined aquifer considering source locations and number as variable. J. Hazard. Toxic Radioact. Waste. https://doi.org/10.1061/(ASCE)HZ.2153-5515.0000334.
GB T 5330-2003, 2003. National Standard for Industrial Metal Wire Braided Square Screen. People's Republic of China, General Administration of Quality Supervision, Inspection and Quarantine.
Gelhar, L.W., Welty, C., Rehfeldt, K.R., 1992. A critical review of data on field-scale dispersion in aquifers. Water Resour. Res. 28, 1955–1974.
Giacobbo, F., Marseguerra, M., Zio, E., 2002. Solving the inverse problem of parameter estimation by genetic algorithms: the case of a groundwater contaminant transport model. Ann. Nucl. Energy 29, 967–981.
Gieg, L.M., Suflita, J.M., 2002. Detection of anaerobic metabolites of saturated and aromatic hydrocarbons in petroleum-contaminated aquifers. Environ. Sci. Technol. 36, 3755–3762.
Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison Wesley.
Guo, Y.L., Wen, Z., Zhang, C., Jakada, H., 2020. Contamination and natural attenuation characteristics of petroleum hydrocarbons in a fractured karst aquifer, North China. Environ. Sci. Pollut. R. https://doi.org/10.1007/s11356-020-08723-2.
Hutchins, S.R., Downs, W.C., Wilson, J.T., Smith, G.B., Kovacs, D.A., Fine, D.D., Douglass, R.H., Hendrix, D.J., 1991. Effect of nitrate addition on biorestoration of fuel-contaminated aquifer: field demonstration. Ground Water 29, 571–580.
Jiang, S.M., Fan, J.H., Xia, X.M., Li, X.W., Zhang, R.C., 2018. An effective Kalman filter-based method for groundwater pollution source identification and plume morphology characterization. Water. https://doi.org/10.3390/w10081063.
Jin, X., Mahinthakumar, G., Zechman, E.M., Ranjithan, R.S., 2009. A genetic algorithm-based procedure for 3D source identification at the Borden emplacement site. J. Hydroinform. 11, 51–64.
Kalayci, C.B., Polat, O., Gupta, S.M., 2016. A hybrid genetic algorithm for sequence-dependent disassembly line balancing problem. Ann. Oper. Res. 242, 321–354.
Kao, C.M., Chien, H.Y., Surampalli, R., Chien, C.C., Chen, C.Y., 2010. Assessing of natural attenuation and intrinsic bioremediation rates at a petroleum-hydrocarbon spill site: laboratory and field studies. Environ. Eng. Sci. 136, 54–67.
Kohl, D.K., Shearer, G.B., Commoner, B., 1971. Fertilizer nitrogen: contribution to nitrate in surface water in a corn belt watershed. Science 174, 1331–1334.
Kuhn, E.P., Zeyer, J., Eicher, P., Schwarzenbach, R.P., 1988. Anaerobic degradation of alkylated benzenes in denitrifying laboratory aquifer columns. Appl. Environ. Microb. 54, 490–496.
Kwon, M.J., O'Loughlin, E.J., Ham, B., Hwang, Y., Shim, M., Lee, S., 2018. Application of an in-situ soil sampler for assessing subsurface biogeochemical dynamics in a diesel-contaminated coastal site during soil flushing operations. J. Environ. Manage. 206, 938–948.
Lari, K.S., Davis, G.B., Rayner, J.L., Bastow, T.P., Puzon, G.J., 2019. Natural source zone depletion of LNAPL: A critical review supporting modelling approaches. Water Res. 157, 630–646.
Lin, C.W., Wu, C.H., Tang, C.T., Chang, S.H., 2012. Novel oxygen-releasing immobilized cell beads for bioremediation of BTEX-contaminated water. Bioresour. Technol. 124, 45–51.
Liu, C.Q., Li, S.L., Lang, Y.C., Xiao, H.Y., 2006. Using δ15N- δ18O-values to identify nitrate sources in karst ground Water, Guiyang, Southwest China. Environ. Sci. Technol. 40, 6928–6933.
Long, Y.Q., Cui, T.T., Li, W., Gai, Y.W., 2018. A geostatistical approach to groundwater pollution source identification considering first-order reaction. Desalin. Water Treat 123, 35–40.
Mahinthakumar, G.K., Sayeed, M., 2005. Hybrid genetic algorithm-local search methods for solving groundwater source identification inverse problems. J. Water Res. Plan. 131, 45–57.
Maric, N., Strbacki, J., Kurilic, S.M., Beskoski, V.P., Nikic, Z., Ignjatovic, S., Malbasic, J., 2019. Hydrochemistry of groundwater contaminated by petroleum hydrocarbons: the impact of biodegradation (Vitanovac, Serbia). Environ. Geochem. Health. https://doi.org/10.1007/s10653-019-00462-9.
Mathworks., 2010. MATLAB 7.10.0. The MathWorks Inc., Natick, MA. McCarthy, E.L., 1934. Mariotte's bottle. Science 80 (100).
Michalak, A.M., Kitanidis, P.K., 2004. Application of geostatistical inverse modeling to contaminant source identification at Dover AFB. Delaware. J Hydraul. Res. 42, 9–18.
Miglietta, P.P., Toma, P., Fanizzi, F.P., Donno, A.D., Coluccia, B., Migoni, D., Bagordo, F., Serio, F., 2017. A grey water footprint assessment of groundwater chemical pollution: case study in Salento (Southern Italy). Sustainability 9. https://doi.org/10.3390/su9050799.
Mitra, K., Deb, K., Gupta, S.K., 1998. Multiobjective dynamic optimization of an industrial nylon 6 semibatch reactor using genetic algorithm. J. Appl. Polym. Sci. 69, 69–87.
Neuman, S.P., Zhang, Y.K., 1990. A quasi-linear theory of non-Fickian and Fickian subsurface dispersion: 1. Theoretical analysis with application to isotropic media. Water Resour. Res. 26, 887–902.
Pinder, G.F., Bredehoeft, J.D., 1968. Application of the digital computer for aquifer evaluations. Water Resour. Res. 4, 1069–1093.
Qian, H., Zhang, Y.L., Wang, J.L., Si, C.Q., Chen, Z.X., 2018. Characteristics of etroleum-contaminated groundwater during natural attenuation: a case study in northeast China. Environ. Monit. Assess. 190. https://doi.org/10.1007/s10661-017-6449-6.
Qin, R.G., Wu, Y.Q., Xu, Z.G., Xie, D., Zhang, C., 2013. Numerical modeling of contaminant transport in a stratified heterogeneous aquifer with dipping anisotropy. Hydrogeol. J. 21, 1235–1246.
Shakya, B.M., Nakamura, T., Shrestha, S.D., Nishida, K., 2019. Identifying the deep groundwater recharge processes in an intermountain basin using the hydro-geochemical and water isotope characteristics. Hydrol. Res. 50, 1216–1229.
Singh, A., Minsker, B.S., Valocchi, A.J., 2008. An interactive multiobjective optimization framework for groundwater inverse modeling. Adv. Water Resour. 31, 1269–1283.
Singh, R.M., Datta, B., 2006. Identification of groundwater pollution sources using GA-based linked simulation optimization model. J. Hydrol. Eng. 11, 1216–1227.
Snodgrass, M.F., Kitanidis, P.K., 1997. A geostatistical approach to contaminant source identification. Water Resour. Res. 33, 537–546.
Sophia, L., Bhattacharjya, R.K., 2020. A GA based iterative model for identification of unknown groundwater pollution sources considering noisy data. Nature-Inspir. Methods Metaheuristics Optim. https://doi.org/10.1007/978-3-030-26458-1_17.
Stout, S.A., Uhler, A.D., Naymik, T.G., McCarthy, K.J., 1998. Environmental forensics-Unraveling site liability. Environ. Sci. Technol. 32, 260A–264A.
Sun, A.Y., Painter, S.L., Wittmeyer, G.W., 2006. A constrained robust least squares approach for contaminant release history identification. Water Resour. Res. 42. https://doi.org/10.1029/2005WR004312.
Sun, A.Y., 2007. A robust geostatistical approach to contaminant source identification. Water Resour. Res. 43. https://doi.org/10.1029/2006WR005106.
Van Stempvoort, D.R., Armstrong, J., Mayer, B., 2007. Seasonal recharge and replenishment of sulfate associated with biodegradation of a hydrocarbon plume. Ground Water Monit. Remed. 27, 110–121.
Venkatramanan, S., Chung, S.Y., Kim, T.H., Kim, B.-W., Selvam, S., 2016. Geostatistical techniques to evaluate groundwater contamination and its sources in Miryang City. Korea. Environ Earth Sci. 75. https://doi.org/10.1007/s12665-016-5813-0.
Wilson, J.T., Leach, L.E., Henson, M., Jones, J.N., 1986. In situ biorestoration as a ground water remediation technique. Ground Water Monit. R. 6, 56–64.
Xia, X.M., Jiang, S.M., Zhou, N.Q., Li, X.W., Wang, L.C., 2019. Genetic algorithm hyperparameter optimization using Taguchi design for groundwater pollution source identification. Water supply 19, 137–146.
Xu, M., Eckstein, Y., 1995. Use of weighted least-squares method in evaluation of the relationship between dispersivity and field scale. Groundwater 33, 905–908.
Xu, T., Gómez-Hernández, J.J., 2016. Joint identification of contaminant source location, initial release time, and initial solute concentration in an aquifer via ensemble Kalman filtering. Water Resour. Res. 52, 6587–6595.
Yeh, H.D., Chang, T.H., Lin, Y.C., 2007. Groundwater contaminant source identification by a hybrid heuristic approach. Water Resour. Res. 43. https://doi.org/10.1029/2005WR004731.
Zanini, A., Woodbury, A.D., 2016. Contaminant source reconstruction by empirical Bayes and Akaike's Bayesian Information Criterion. J. Contam. Hydrol. 185, 74–86.