Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-021-16242-x

RESEARCH ARTICLE

A hybrid statistical decision-making optimization approach for groundwater vulnerability considering uncertainty

Yalda Norouzi Gharakezloo[1] · Mohammad Reza Nikoo[2] · Ayoub Karimi-Jashni[1] · Mehrdad Ghorbani Mooselu[3]

Received: 6 March 2021 / Accepted: 25 August 2021

Abstract

Recognizing the vulnerable areas for contamination is a feasible way to protect groundwater resources. The main contribution of the paper is developing a hybrid statistical decision-making model for evaluating the vulnerability of Shiraz aquifer, southern Iran, with modified DRASTIC (depth to the water table, net recharge, aquifer media, soil media, topography, impact of the vadose zone, and hydraulic conductivity) by using the genetic algorithm (GA), the analytical hierarchy process (AHP) method, and factorial analysis (FA). First, considering the variation of the uncertain parameters, 32 scenarios were defined to perform factorial analysis. Then using the AHP method and GA, DRASTIC parameters were rated and weighted in all scenarios. To achieve the optimal weights for parameters, the objective function in GA was maximizing the correlation coefficient between the vulnerability index and the nitrate concentration. The single and interactive effects of parameters on groundwater vulnerability were analyzed by factorial analysis. The results revealed that the net recharge had the highest single effect, and the resulted effect between net recharge and hydraulic conductivity was the most significant interactive effect on the objective function. Besides, the variation of aquifer media does not change the objective function. The application of the proposed method leads to a precise groundwater vulnerability map. This research provides valuable knowledge for assessing groundwater vulnerability and enables decision-makers to apply groundwater vulnerability information in future water resources management plans.

Keywords Hybrid statistical model · Decision-making model · Factorial analysis · Groundwater vulnerability · Genetic algorithm · DRASTIC

Introduction

Population growth and increasing water demand have highlighted groundwater resources' significance and necessitate a feasible preservation method. Groundwater vulnerability assessment is an effective method for protecting groundwater resources and evaluating the feasibility of other management approaches such as water, treated wastewater, and waste load allocation (Daneshmand et al. 2014; Nikoo et al. 2016; Mooselu et al. 2019, 2020; Yazdian et al. 2021). In recent years, the DRASTIC approach has been one of the most applicable methods to assess the vulnerability of groundwater. This method was first used by Aller et al. in 1987 in US Environmental Protection Agency (Aller et al. 1987). The DRASTIC method includes seven hydrogeological parameters of depth to water table (D), net recharge (R), aquifer media (A), soil media (S), topography (T), the impact of vadose zone (I), and hydraulic conductivity (C). These parameters are assigned weights and rates commensurate with their effects on the transmission of pollution. Since the

Responsible Editor: Marcus Schulz

✉ Mohammad Reza Nikoo
m.reza@squ.edu.om

Yalda Norouzi Gharakezloo
norouzi.gh_y@yahoo.com

Ayoub Karimi-Jashni
akarimi@shirazu.ac.ir

Mehrdad Ghorbani Mooselu
mehrdad.g.mooselu@uia.no

1 Department of Civil and Environmental Engineering, Shiraz University, Shiraz, Iran

2 Department of Civil and Architectural Engineering, Sultan Qaboos University, Muscat, Oman

3 Department of Engineering Sciences, University of Agder, Kristiansand, Norway

Published online: 07 September 2021

hydrogeological situation of the regions plays a vital role in calculating the vulnerability index, the DRASTIC method should be modified by reliable measures to achieve more accurate and consistent results to the site-specific conditions. The previous researches on developing an accurate model to assess aquifer vulnerability can be divided into two parts. A part of this researches has been trying to modify the structure of the vulnerability model to make it more compatible with the environmental conditions of the aquifers, and the other part has utilized methods to minimize errors due to uncertainty in calculating input parameters.

In recent years, extensive researches have been conducted to improve the compatibility of the DRASTIC method with the study areas. For example, the AHP method has been applied in some investigations to modify the weights and rates of the DRASTIC parameters (Hu et al. 2018; Jesiya and Gopinath 2019; Tomer et al. 2019; Arshad et al. 2020; Mallik et al. 2021; Liu et al. 2021). The relative classes of each parameter are used as modified rates of the DRASTIC parameters. These rates and adjusted weights led to more accurate results by creating a better fit between the actual aquifer conditions and the DRASTIC vulnerability index. In addition, another group of researchers modified the weight and rate of the DRASTIC parameters using the Wilcoxon method in which non-parametric statistical tests evaluate the similarity of two samples related to the rating scale (Jafari and Nikoo 2016; Barzegar et al. 2019; Bordbar et al. 2019; Balaji et al. 2021). Meanwhile, the adequacy of the number and type of DRASTIC parameters was the subject of study for many researchers (Kumar and Pramod Krishna 2019; Liu et al. 2021). Some researchers added parameters to the DRASTIC method based on the study area (Sener and Davraz 2013; Hu et al. 2018; Kumar and Pramod Krishna 2019; Soyaslan 2020; Liu et al. 2021), and others removed some of the DRASTIC parameters (Arezoomand Omidi Langrudi et al. 2016; Nadiri et al. 2019).

The results of deterministic analyses have a certain confidence level due to the uncertainties in measurements and data analysis. This issue necessitates considering the uncertainty analysis in aquifer vulnerability assessment to improve the data reliability and consequently achieving more accurate results. New algorithm-based methods, e.g., genetic algorithm (GA) and fuzzy algorithm, help to develop a model close to the real situation (Ahn et al. 2012; Jafari and Nikoo 2016; Yang et al. 2017; Barzegar et al. 2019; Jafari and Nikoo 2019; Torkashvand et al. 2020; Baalousha et al. 2021; Saranya and Saravanan 2021). Moreover, the effect of every single DRASTIC parameter on aquifer vulnerability has been assessed by sensitivity analysis (Pacheco et al. 2015; Jafari and Nikoo 2016; Jafari and Nikoo 2019; Pourkhosravani et al. 2021).

Different methods such as AHP, GA, and fuzzy have been applied in the literature to obtain an accurate vulnerability map that is compatible with the conditions of the case studies. However, other components such as identifying more effective parameters and defining various vulnerability scenarios are yet to be considered. So far, vulnerability maps were drawn based on the parameters obtained from field studies showing the vulnerability at specific times and conditions, but considering the uncertainty of input parameters and their single or interactive effects on the vulnerability of aquifer is also of significant importance. In this study, in addition to using the GA and AHP models in modifying the weights and rates of parameters in the vulnerability model, a factorial analysis–based optimization model was developed to gain a comprehensive view on assessing the vulnerability of aquifer. Using factorial analysis (FA) in 32 scenarios, the uncertainties of input parameters were considered, and both single and interactive effects of the parameters on the vulnerability were evaluated. Applying the proposed method can show how the different parts of the aquifer will react to the possible changes in aquifer characteristics (input parameters). Also, the results of the FA method in the form of single or interactive effects between parameters can show how small changes in some parameters affect the vulnerability of the whole system. Understanding this sensitivity can prevent making the wrong decision and remove/mitigate irreversible damage. This research measured the impact of every parameter and the interactive effect of parameters on groundwater vulnerability. Hence, the proposed method aims to modify the rate scores and weights of the DRASTIC method considering the uncertainty of input parameters and then investigate the single and interactive effects of DRASTIC parameters on the objective function, which was to maximize the correlation between DRASTIC vulnerability index and nitrate concentration. Nitrate is considered one of the primary contaminants of Shiraz plain's groundwater to evaluate the accuracy and validity of the vulnerability investigations (Baghapour et al. 2016; Jafari and Nikoo 2016, 2019). The high concentration of nitrate in groundwater is mainly due to land use (primarily residential) and the entry of residential, agricultural, and industrial wastewater (containing nitrate as the main contaminant) into the Shiraz plain. The paper contributes in the following ways:

1) Developing a hybrid statistical decision-making method that corrects the rate and weights of parameters
2) Evaluating the uncertainty of input parameters and analyzes their single and interactive effects on the variation of the objective function utilizing the factorial analysis method.

Finally, the vulnerability of the aquifer was mapped by ArcGIS® 10.5 (Bera et al. 2021) for different values of the objective function and provided more precise information for decision-makers in future management plans. The feasibility of the proposed method was evaluated in the Shiraz aquifer, southwestern Iran.

Materials and methods

This paper suggests a feasible method for determining the vulnerability of groundwater which consists of four main steps, including data gathering (step 1), uncertainty analysis by defining the lower and upper bounds of uncertain parameters and then determining vulnerability scenarios (step 2), developing a hybrid AHP-GA-based DRASTIC model in which the rate and the weight of DRASTIC parameters are modified and validation of results in different scenarios (step 3), and analyzing the single and interaction effects of parameters (step 4). A flowchart of the proposed methodology is presented in Fig. 1.

Case study and data collection

Shiraz plain with an area of 300 km^2 is located between longitudes 52° 29′ to 52° 36′ east and latitudes 29° 33′ to 29° 36′ north. The general direction of groundwater flow is parallel to the topographic slope and from northwest of the plain to southeast. Shiraz plain includes calcareous and alluvial aquifers. The alluvial aquifer in the west is coarse-grained and becomes fine-grained as it approaches the east. The alluvial aquifer is layered, and clay layers are located between the water layers. Most of the lands in Shiraz plain are dedicated to urban and agricultural regions, and its non-agricultural areas are located in the south and southeast. The study area includes 30 observation wells. The geographical location of the Shiraz plain and the location of observation wells are presented in Fig. 2.

Figure 3 shows the zoning map of nitrate concentration in the study area in which the nitrate concentration is higher in the southern and southeastern regions. The correlation coefficient between the DRASTIC index and nitrate concentration was chosen to validate the study area's DRASTIC method. According to Fars Regional Water Organization, this aquifer's overall water balance is −8 M m^3/year. Most of the extracted water from 872 discharge wells is applied for agricultural activities.

The hydrogeological information, including DRASTIC parameters, i.e., depth to water table (D), net recharge (R), aquifer media (A), soil media (S), topography (T), the impact of

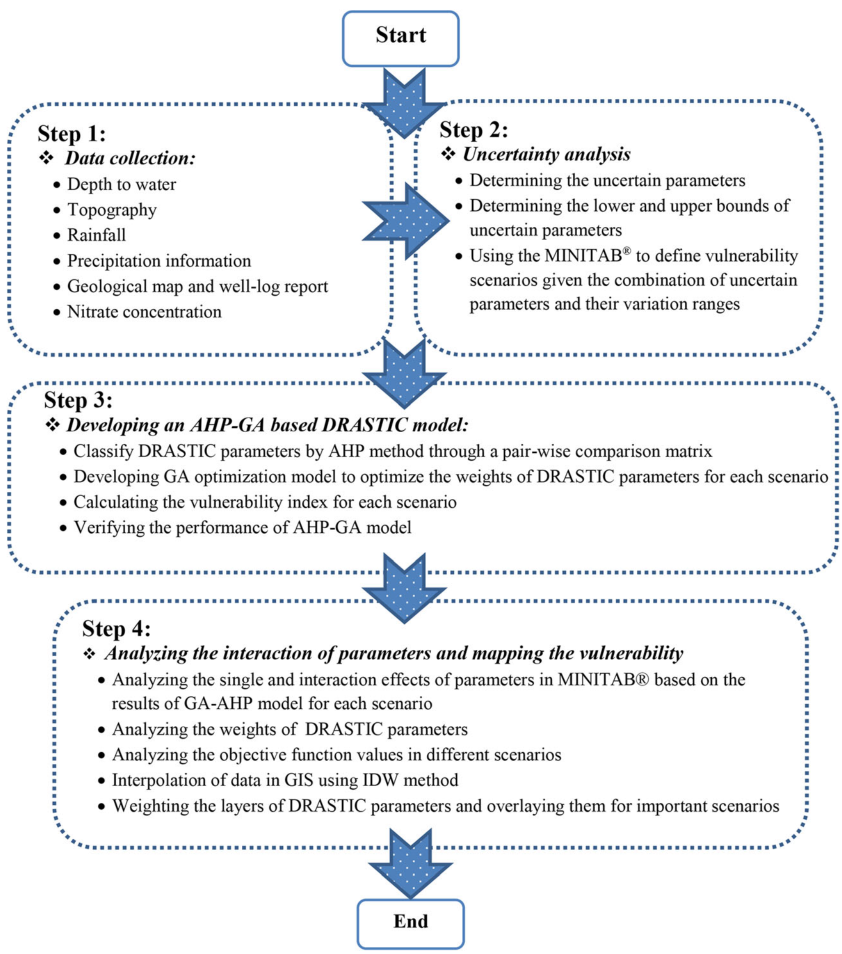


Fig. 1 The structure of the proposed method for the assessment of groundwater vulnerability

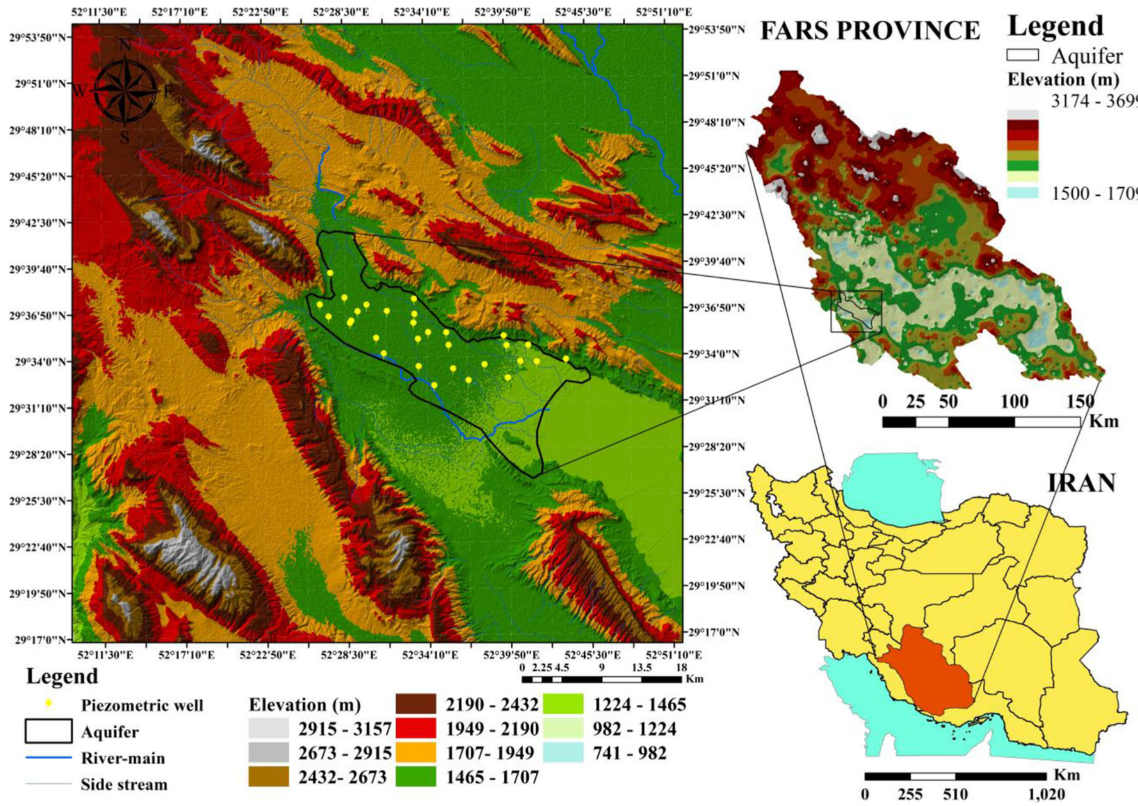
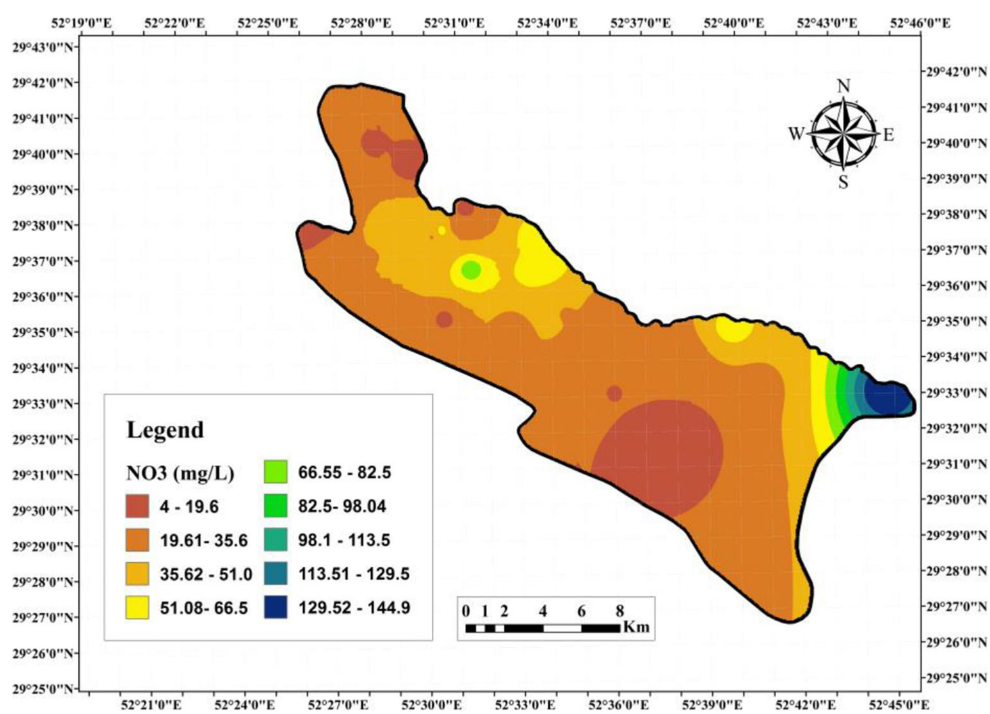


Fig. 2 The location of the Shiraz aquifer and piezometric wells

vadose zone (I), and hydraulic conductivity (C), were obtained through the piezometric wells, geotechnical boreholes, and the spatial distribution map of precipitation, evaporation, and hydraulic conductivity provided by the Fars Regional Water Organization. Data were analyzed for homogeneity and using geostatistical methods, the defects and errors were addressed.

Quantitative data related to 30 observation wells, including depth, net recharge, topography, and conductivity, are shown

Fig. 3 The zoning map of nitrate concentration in the study area



in Fig. 4. The thematic layers of the study area related to qualitative data, including aquifer media (A), soil media (S), and impact of vadose zone (I), can be seen in Fig. 5. For example, as shown in Fig. 4, the average depth is about 17 m, and its standard deviation is about 14 m, and the depth to the water table in observation wells has varied from 0.9 to 59 m. According to the data presented in Fig. 5, it can be seen that the material of the Shiraz aquifer is mainly composed of clay, clay and gypsum, and clay and sand. In addition, the material related to the impact of vadose zone (I) was mainly clay and clay and silt. Soil media (S) in the northern and northwestern parts of loamy sand, while in the southern and southeastern regions, the soil layer was thin or non-existent. In Fig. 4, topography (T) is defined as the natural slope (%) of the plain surface.

Uncertainty analysis

The available data were obtained by sampling in the field. Due to the possibility of error in field sampling of these parameters, they can be modeled as uncertain parameters. Performing uncertainty analysis can help better understand the variability of the parameters in the aquifer. In most cases, for uncertainty analysis in groundwater vulnerability assessment, the quality of the available data is not satisfactory to determine the probability distribution functions. In large-scale modeling, even if the probability distribution functions exist due to many components, probability optimization is a time-consuming process and not feasible. In comparison, determining the variations' interval for the uncertain parameters is more applicable (Tavakoli et al. 2014, 2015). Hence, the interval factorial analysis was utilized for uncertainty analysis.

The factorial analysis is a complete sensitivity analysis method that considers each parameter's effect on the output. It can identify and quantify the interactive impact of all uncertain parameters with the help of a purposeful choice of test conditions. In factorial analysis, the lower and upper bounds of each uncertain parameter's variation interval are considered the lower and upper levels. If there are k uncertain parameters, then it will be necessary to perform 2^k different tests. A factorial design of 2^k is a design with k-factors (operator), each one in two levels (i.e., lower and upper), containing k original effects, $\binom{K}{2}$ two-factor interaction effect, $\binom{K}{3}$ three-factor interaction effect, etc., and finally, one k-factor interaction effect. Therefore, for a 2^k factorial design, the complete statistical model has 2^k results. The single and impact of the interaction of parameters can be calculated by the following formulas: (Montgomery 2017)

$$E_x = \frac{(2 \times contrast_x)}{n \times \left(2^{k-1}\right)} \tag{1}$$

$$SS_x = \frac{(contrast_x)^2}{n \times 2^k} \tag{2}$$

kThe number of uncertain parameters

nRun numbers

E_xThe standardized single and interaction effect of parameters

SS_xThe sum of squares of any effect

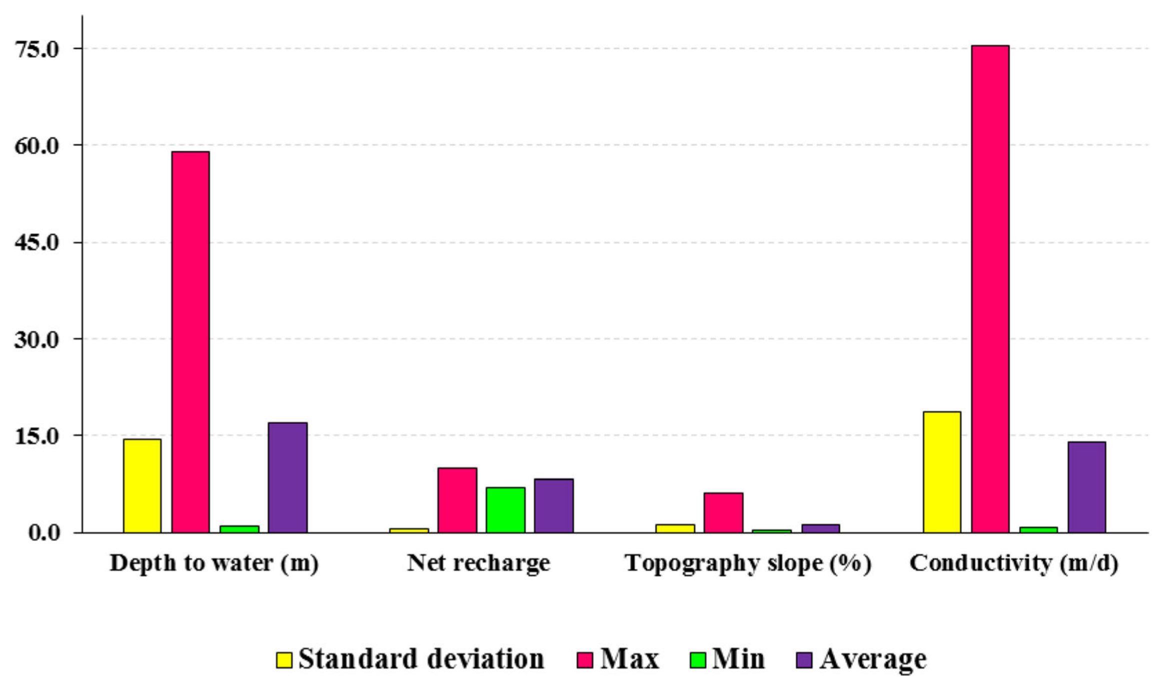


Fig. 4 Statistical data of DRASTIC quantitative parameters

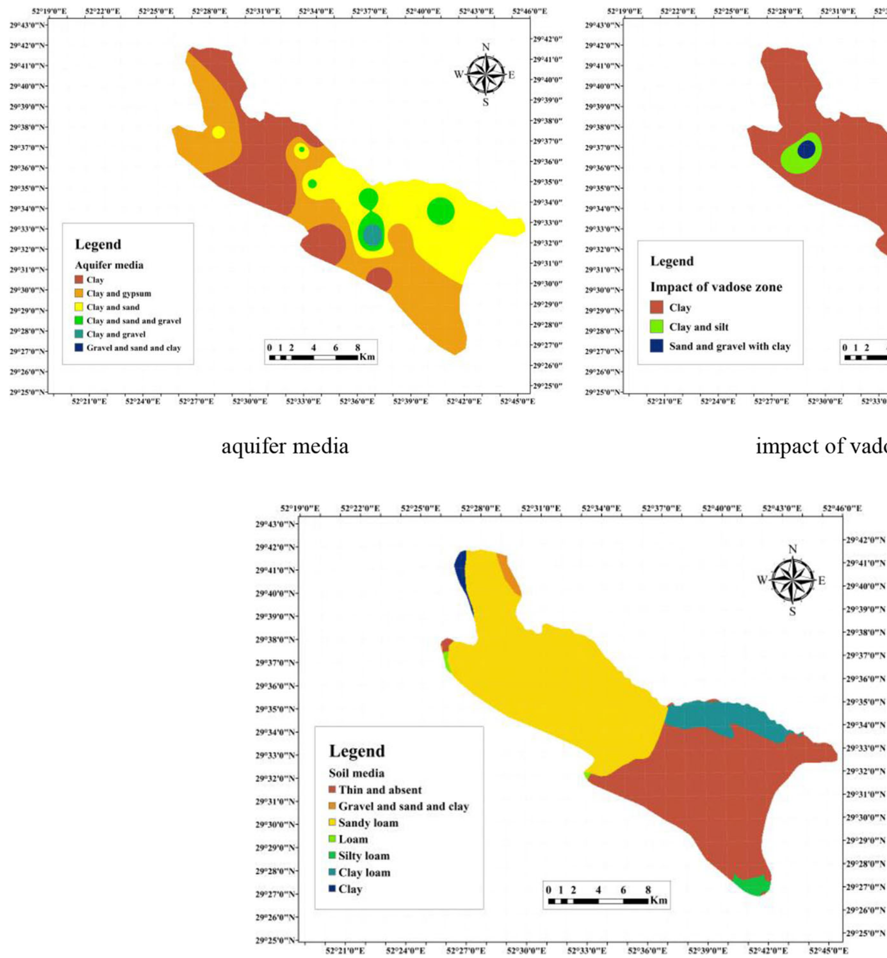


Fig. 5 Spatial distribution of DRASTIC qualitative parameters

$contrast_x$Obtained from plus and minus signs table in (Montgomery 2017)

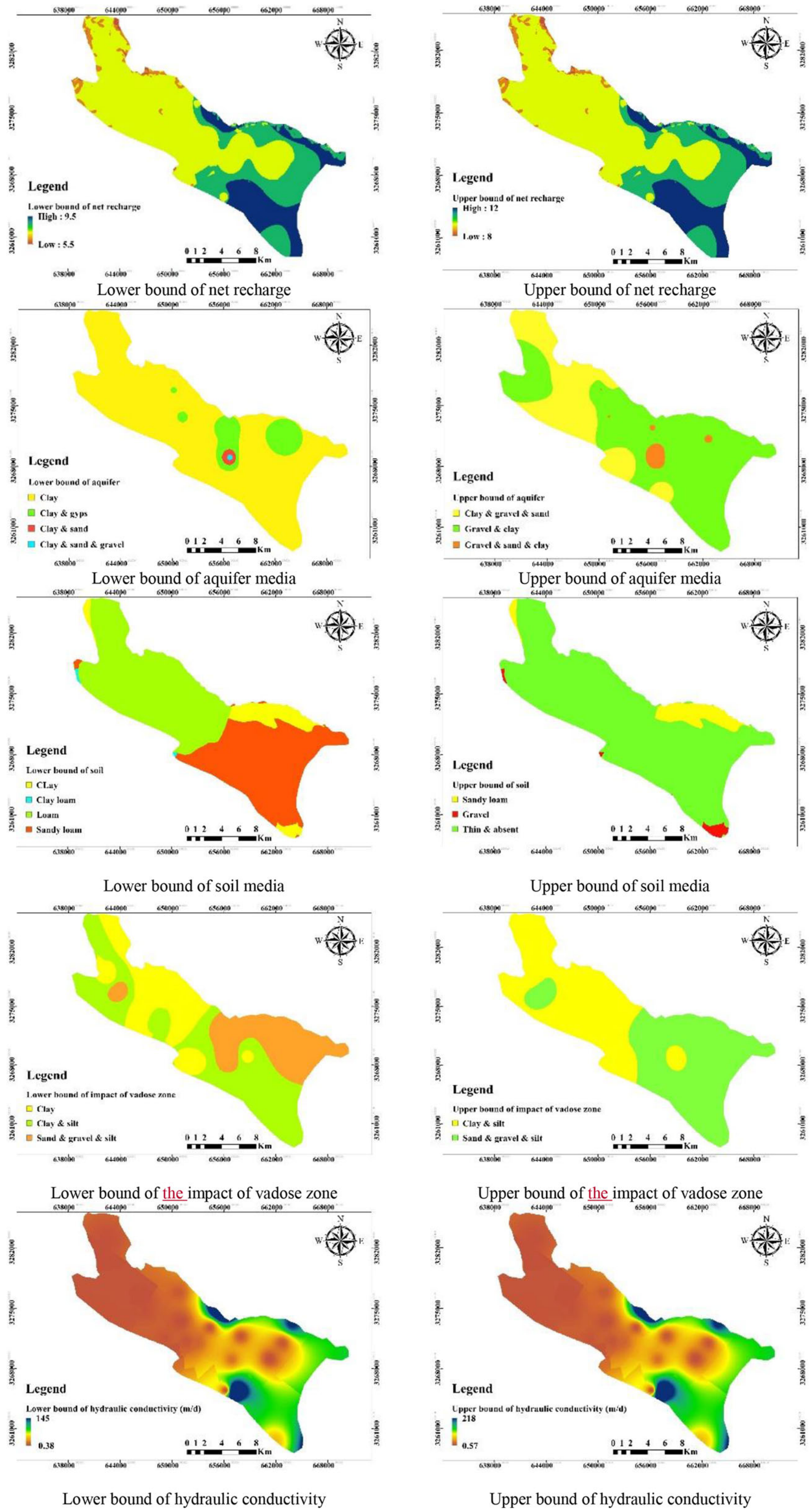
In this study, because of the monotonicity of the objective function's variations in the interval of uncertain parameters, the factorial analysis method effectively considers the uncertainties in DRASTIC parameters. Hence, using factorial design in Minitab® 17.1.0 software, 2^5 different scenarios were defined considering five uncertain parameters of the DRASTIC method, including hydraulic conductivity (C), aquifer media (A), soil media (S), impact of the vadose zone (I), and net recharge (R).

The range of variation in uncertain parameters was determined considering historical time series and engineering judgment. For example, the soil media can vary from sand to clay. Historical time series and engineering judgment also were applied in determining the upper and lower limits of aquifer media (A), and vadose zone (I) impact. The impact of vadose zone (I) is between the saturation zone and the soil environment that is essentially unsaturated or intermittently saturated and controls the passage and dilution of contaminants to the saturation zone. The impact of the vadose zone parameter in the calculation of aquifer pollution is similar to soil media and depends on the permeability of the ingredients and the properties of the unsaturated zone. Information about this parameter and other qualitative parameters were determined by investigating the well-log data collected by the Fars Regional Water Organization. The range of variation in uncertain parameters was determined considering historical time series and engineering judgment. The lower and upper bounds of the uncertain parameters are mapped in Fig. 6. As shown, the aquifer media in the lower bound of analysis is mostly clay, while in the upper bound is predominantly gravel and clay.

DRASTIC method

The DRASTIC method was first developed in 1987 by the US Environmental Protection Agency (EPA) to assess groundwater vulnerability using existing aquifer hydrogeological characteristics (Aller et al. 1987). The DRASTIC method consists of seven hydrogeological parameters. Each of these parameters is given a value between 1 and 10 according to their effect on the entry of pollution into groundwater, with the number 1 corresponding to the lowest pollution potential and the

Fig. 6 The lower and upper bounds of the uncertain parameters over the aquifer area



number 10 indicating the highest pollution potential. In addition, each of these parameters is assigned a weight coefficient based on their degree of importance. These parameters include depth to water table (D), net recharge (R), aquifer media (A), soil media (S), topography (T), the impact of vadose zone (I), and hydraulic conductivity (C). The DRASTIC index is calculated through the following formula (Aller et al. 1987).

$$DRASTIC = D_rD_w + R_rR_w + A_rA_w + S_rS_w + T_rT_w + I_rI_w + C_rC_w \quad (3)$$

r Rate of DRASTIC parameters
w Weight of DRASTIC parameters

AHP method

The basis of AHP is the paired comparisons matrix, which provides a linear relationship between the criteria and sub-criteria. For example, if the preference of element A over B is n, the preference of element B over A would be 1/n, while at different levels of element A, the desirability of B varies. The consistency of the results obtained from the pairwise comparison matrices must be confirmed by calculating the compatibility matrix. The consistency index (CI) is determined based on Eq. 4 (Saaty 1980):

$$CI = \frac{(\lambda_{\max} - n)}{(n-1)} \quad (4)$$

$$CR = \frac{CI}{RI} \quad (5)$$

nThe order of the matrix
$\lambda_{\max}$The eigenvalue of the matrix
CRConsistency ratio
RIThe random index

CR must be calculated to verify the reliability of the pairwise comparison matrices created by personal decisions. As a general rule, the CR value should be less than 0.01 (Soyaslan 2020).

GA method

The genetic algorithm (GA) was first introduced by Holland et al. in 1975. This optimization method is mainly used to optimize very complex and nonlinear problems (Holland 1975). GA is a comprehensive probabilistic search method that follows natural biological evolution. In this method, initial generations (initial answers) are generated randomly or pre-calculated. Several superior chromosomes (superior responses) of the initially developed community whose value of the objective function (fitness function) is more significant in maximization problems and more minor in minimization problems are selected as the parents' chromosomes to generate the next generation (selection operator). These chromosomes are then randomly grouped into binary clusters to produce the next generation. During this process, new generations are driven to the optimal answers, and the calculations continue until, with more repetition, no improvement in the final response (the last generation of superior chromosomes) is observed.

Developing an AHP-GA-based DRASTIC model

The ordinary DRASTIC index only provides a relative evaluation of the aquifer's pollution. To achieve an accurate assessment of the vulnerability, the weights and rates of DRASTIC parameters should be compatible with the nature of the site. DRASTIC parameters were rated using the AHP method, given the lower and upper bounds of uncertain parameters derived from factorial analysis, and then, these modified rates were utilized in the GA model to achieve optimized weights. Each parameter's rates in the upper and lower bounds were adjusted by AHP considering the region's hydrological conditions. Accordingly, 32 matrices of 30×7 were considered for the 32 scenarios defined by factorial analysis, which are the rates of DRASTIC parameters in calculating the aquifer's vulnerability index. In these matrices, the array $A_{i,j}$ shows the value of jth parameter in the ith piezometric well. As a critical step in the proposed methodology, this optimization model considers uncertainties related to net recharge, aquifer media, soil media, vadose zone media, and the aquifer hydraulic conductivity to achieve the optimum weights for DRASTIC parameters. Since the primary purpose of modifying the weight and rate of DRASTIC parameters was to increase the correlation coefficient between DRASTIC vulnerability index and nitrate concentration, the objective function in calculating the optimal weight of DRASTIC parameters by GA was to maximize the correlation between the vulnerability index and nitrate concentration, and the decision variables were weights of seven DRASTIC parameters. The objective function of the optimization model is presented as follow:

$$Max\ F^{\pm} = corr\left(X^{\pm}, Y\right) \quad (6)$$

$$F = -\frac{\sum_{j=1}^{n}\left(\left(X_j^{\pm} - \overline{X}^{\pm}\right)\left(Y_j - \overline{Y}\right)\right)}{\sqrt{\sum_{j=1}^{n}\left(X_j^{\pm} - \overline{X}^{\pm}\right)^2}\sqrt{\sum_{j=1}^{n}\left(Y_j - \overline{Y}\right)^2}} \qquad j = 1, 2, \ldots, n \quad (7)$$

$$X_j^{\pm} = \sum_{i=1}^{7}\left(r_j^{\pm} \times w_i^{\pm}\right) \quad (8)$$

$$r_i^{\pm} = f\left(C_1^{\pm}, C_2^{\pm}, \ldots, C_m^{\pm}\right) \quad (9)$$

$$\overline{X}^{\pm} = \frac{1}{n}\sum_{j=1}^{n} X_j^{\pm} \quad (10)$$

$$Y_j = C_{No_3\ j} \quad (11)$$

$$\overline{Y} = \frac{1}{n}\sum_{j=1}^{n} Y_j \quad (12)$$

$$1 \langle w_i^{\pm} \langle 10 \quad (13)$$

where $F$is the objective function of the optimization model, n is the number of wells, X_j is the vulnerability index of jth well, $\overline{X}$ is the mean of vulnerability indices, $Y_j$is the nitrate concentration (mg/l) in jth well, $\overline{Y}$ is the mean of nitrate concentration, w_i is the weight of DRASTIC model parameters, $r_i$is the rate of DRASTIC model parameters, m is the number of intervals of each parameter, and $f$is the AHP-based submodel used to modify rates of DRASTIC model parameters.

In these equations, input parameters have upper and lower bounds (the positive and negative signs), and the output of the objective function is in the form of intervals.

After modifying the rates, the developed interval-based AHP-GA optimization model was run for different scenarios and provided a set of optimized weights for DRASTIC parameters in each scenario. These optimized weights are utilized to calculate the vulnerability index (VI) of the aquifer using the DRASTIC method in each scenario as follow:

$$VI = \sum_{i=1}^{7} (r \times w) \tag{14}$$

where r and w stand for the rating score and optimum weight assigned to the ith parameter. The hybrid AHP-GA optimization model indicates the aquifer's exact vulnerability proportional to the aquifer's hydrogeological characteristics in different aquifer areas. The greater the index, the higher the risk of contamination. The correlations between the calculated vulnerability index and nitrate concentration in different scenarios were calculated to verify the model's results.

Interaction of parameters and mapping the vulnerability

Utilizing factorial analysis, the single and interaction effects of parameters on the objective function were analyzed using different scenarios. The values of DRASTIC parameters in other aquifer areas were imported into GIS, where the interpolation was made using the inverse distance weighted (IDW) method. Finally, the DRASTIC parameters' optimized weights were considered in layers and overlaid to plot the final vulnerability map.

Result and discussion

Simple DRASTIC

The vulnerability of the aquifer was first investigated by the conventional DRASTIC method. The vulnerability of the aquifer is mapped in Fig. 7, in which the vulnerability of the aquifer is increased from the northwest to the southeast. The conventional DRASTIC method led to the average correlation coefficient of 59% between the DRASTIC index and the nitrate concentration in 30 piezometric wells, indicating that the simple DRASTIC process cannot reflect the precise estimation of the groundwater vulnerability.

Factorial analysis

The effect of parameters' variation on the objective function should be investigated by data compliance assessment to ensure that factorial analysis can be performed. By changing the values of uncertain input parameters between the upper and lower bounds, the variation of the objective function in Eq. (1) (maximum correlation between vulnerability index and nitrate concentration) was investigated (see Fig. 8). The horizontal axis in Fig. 8 represents the corresponding states in which the values of the parameters change from the lowest possible (state 1) to the highest possible values (state 10). As shown, the variation of the objective function is decreasing uniformly (e.g., R, S, I, and A) or increasing (e.g., C). As a result, it is sufficient to consider only the upper and lower values of the uncertain parameters for uncertainty analysis.

Considering the uncertain parameters (R, A, S, I, and C), 32 scenarios (Table 1) were defined for the hybrid AHP-GA DRASTIC method based on the factorial analysis method. In addition, for each scenario, the modified rates of the DRASTIC parameters (Table 2) were determined by the AHP multi-criteria decision-making method. Having the revised rates, the AHP-GA model was implemented based on 32 defined scenarios, and the modified values of the objective function were determined. These values were used as the results of testing each scenario in the factorial analysis.

Factorial analysis was performed for all 32 scenarios considering the corrected rates of the DRASTIC parameters. The effects of uncertain parameters on the objective function are presented in Fig. 9. The parameters' effects are directly related to the slope of the graphs. The slope of the parameters such as R, A, S, and I is negative, which means that their increase decreases the correlation coefficient value. In contrast, hydraulic conductivity has a positive effect on the objective function. Besides, the higher slope of the parameter, the more significant the impact on the objective function. For example, R is the most affecting parameter on the correlation coefficient.

Figure 10 shows the interactive effect of uncertain parameters on the objective function in which when both parameters change between their lower and upper bounds, parallel lines represent the constant variation rate for the objective function (not significant interactive effect) and non-parallel lines indicate the noticeable interactive effect of the uncertain parameters on the objective function. For example, the analysis of the interactive effect between the parameters of R and A (Fig. 10a) shows that the effect of variation in R on the objective function is independent of variation in A. In other words, there is no interactive effect between R and A on the objective function. The parameters of I and C also do not have an

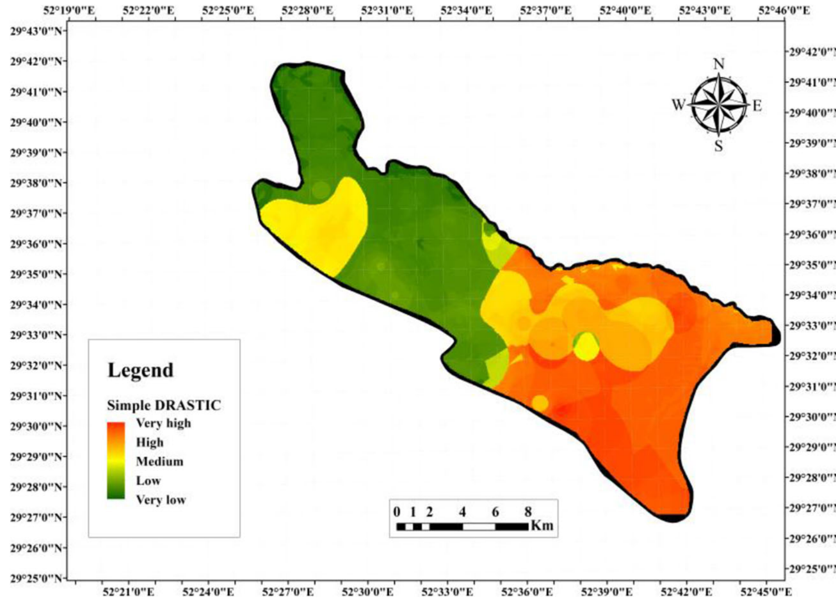


Fig. 7 The vulnerability map of the aquifer derived from simple DRASTIC

interactive effect (see Fig. 10j). In comparison, the effects of R on the objective function are dependent on the value of the parameter I (Fig. 10b), so that when R is changing from lower bound to upper bound, the value of the parameter I can change the rate of variation in the objective function.

The normal probability graph of the effects (Fig. 11a) evaluates the significance of the single and interactive effects of parameters on the objective function. In this diagram, the standard effects of the parameters are assessed based on the distribution fitted line, for which all effects are zero (the green line in Fig. 11a, b).

Accordingly, parameters whose effects are far from the distribution fitted line are essential. The greater the distance, the greater the effect on the objective function. The single effect of the parameters such as R, S, and I and the interactive effect of R and C (RC) are the most critical effects on the objective function. Also, in Fig. 11 a, the position of the points compared to the distribution fitted line is significant. The points on the right side of the line (e.g., the interactive effect of RC and the single effect of C) are consistent with the objective function, which means their changes are directional. In contrast, the parameters on the left side of this line (e.g., R, S, and I) are inconsistent with the objective function. In other words, the variation of such parameters and changes in the objective function are non-directional. The normal probability diagram cannot accurately quantify the effect of each parameter as the absolute value of the effect. Hence, the half-normal plot of the effects (Fig. 11b) shows the magnitude of the effects. As shown, the R has the most significant effect on the objective function, which, compared to that, 97.8% of the data have a less than or equal effect. Similarly, the parameters of S and I and the interactive effect of RC with 94.6, 91.4, and 88.2% are in the following ranks (R > S > I > RC).

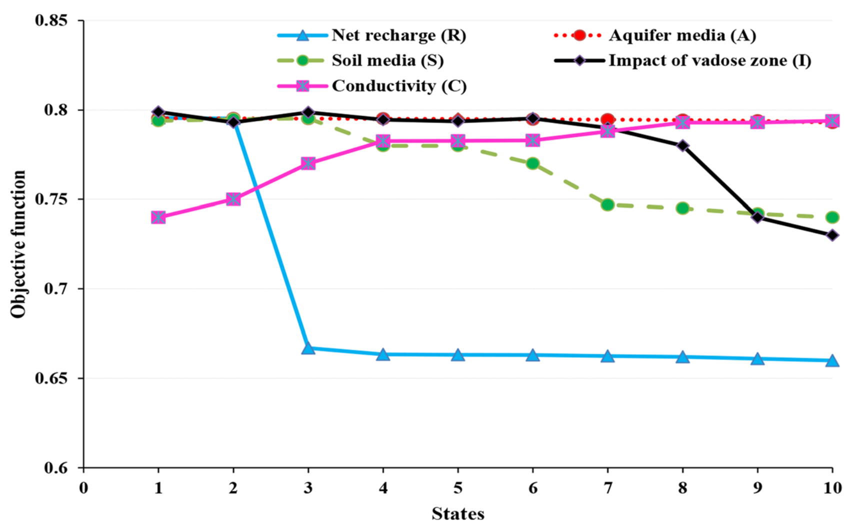


Fig. 8 Variation of the objective function according to different values of uncertain parameters

Table 1 The defined scenarios by the Factorial Analysis given uncertain parameters

Scenario	Uncertain parameter					Scenario	Uncertain parameter				
	R	A	S	I	C		R	A	S	I	C
1	−1*	1**	−1	−1	−1	**17**	1	1	1	−1	−1
2	−1	1	−1	1	1	**18**	1	1	1	1	−1
3	1	−1	1	−1	−1	**19**	1	1	−1	1	−1
4	1	1	−1	−1	−1	**20**	1	−1	−1	1	1
5	−1	−1	−1	−1	−1	**21**	1	1	1	1	1
6	−1	−1	1	1	1	**22**	−1	−1	−1	−1	1
7	−1	1	1	−1	1	**23**	1	−1	−1	−1	1
8	−1	−1	1	1	−1	**24**	1	−1	−1	−1	−1
9	−1	1	−1	1	−1	**25**	−1	1	1	1	−1
10	1	−1	1	1	−1	**26**	1	1	−1	−1	1
11	−1	−1	−1	1	−1	**27**	1	−1	−1	1	−1
12	−1	1	−1	−1	1	**28**	−1	−1	1	−1	−1
13	1	1	1	−1	1	**29**	−1	1	1	1	1
14	1	−1	1	−1	1	**30**	−1	1	1	−1	−1
15	1	1	−1	1	1	**31**	1	−1	1	1	1
16	−1	−1	−1	1	1	**32**	−1	−1	1	−1	1

*−1= lower bound of the parameter, **1= upper bound of parameter

The variation of the objective function in multiple scenarios is shown in Fig. 12. The values of the objective function in various scenarios (Fig. 12) and information of Table 1 show that in scenario no. 1, no. 2, no. 9, and no. 12, where the objective function has the highest value, the parameters of R and S are at their lower bound and A is at the upper bound. Conversely, in scenario no. 10 that the objective function has the lowest value, R and S parameters are at their upper bound, and A is at its lower bound. This point clearly shows the considerable importance of the parameters such as R, S, and I in the objective function variation.

The GA method's optimized weights, which lead to the highest correlation coefficient between vulnerability index and nitrate concentration, were selected as the optimal weight of DRASTIC parameters in various scenarios. The optimized weights of uncertain DRASTIC parameters in minimum (0.28) and maximum (0.8) values of the objective function are presented in Fig. 13.

The R parameter adopted a high weight in both cases (maximum and minimum values of the objective function), which indicates the importance of this parameter on the vulnerability of the aquifer. When it comes to S, I, and C parameters, their weights vary in the maximum and minimum values of the objective function, and this variation is visible in calculating

Table 2 The modified rate of DRASTIC parameters by the AHP method

Parameters	Sub-criteria	Rate	Parameters	Sub-criteria	Rate
Depth to water table (m)	0.9–1.75	0.264	**Impact of vadose zone**	Clay	0.103
	1.75–5.51	0.212		Clay and silt	0.174
	5.51–7.84	0.147		Sand and gravel with clay	0.722
	7.84–9.47	0.108	**Hydraulic conductivity (m/day)**	0.47–5.05	0.023
	9.47–10.58	0.094		5.05–9.27	0.030
	10.58–14	0.070		9.27–19.3	0.057
	14–20.27	0.052		19.3–30.83	0.086
	20.27–28.89	0.037		30.83–41.24	0.122
	28.89–45.24	0.023		41.24–56.84	0.207
	45.24<	0.012		>56.84	0.475
Soil media	Thin and absent	0.377	**Topography (%)**	0–0.5	0.320
	Gravel and sand and clay	0.310		0.5–1.5	0.213
	Sandy loam	0.113		1.5–2.9	0.163
	Loam	0.080		2.9–5	0.120
	Silty loam	0.051		5–8.2	0.081
	Clay loam	0.041		8.2–12.5	0.056
	Clay	0.029		12.5–18	0.033
Aquifer media	Clay	0.032		18<	0.015
	Clay and gypsum	0.055	**Net recharge (Piscopo)**	5–7	0.145
	Clay and sand	0.094			
	Clay and sand and gravel	0.118		7–8	0.230
	Clay and gravel	0.253		8–12	0.625
	Gravel and sand and clay	0.448			

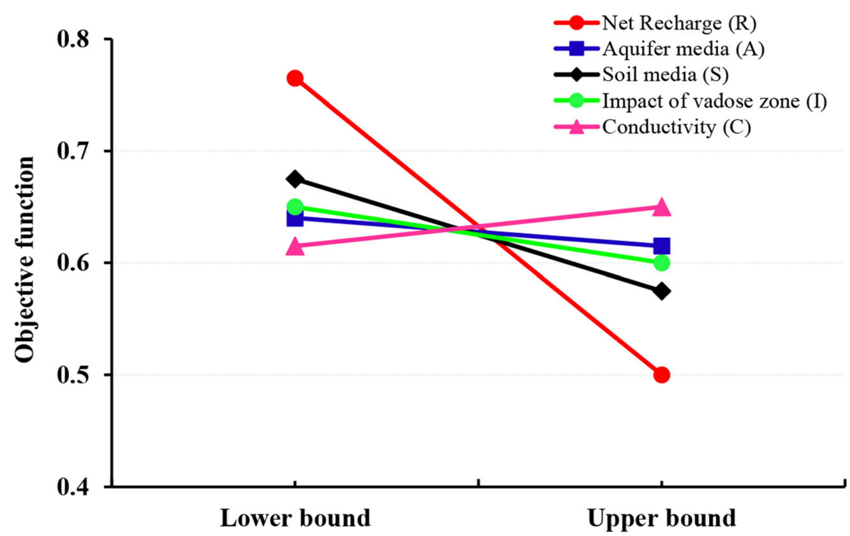


Fig. 9 The effects of uncertain parameters on the objective function

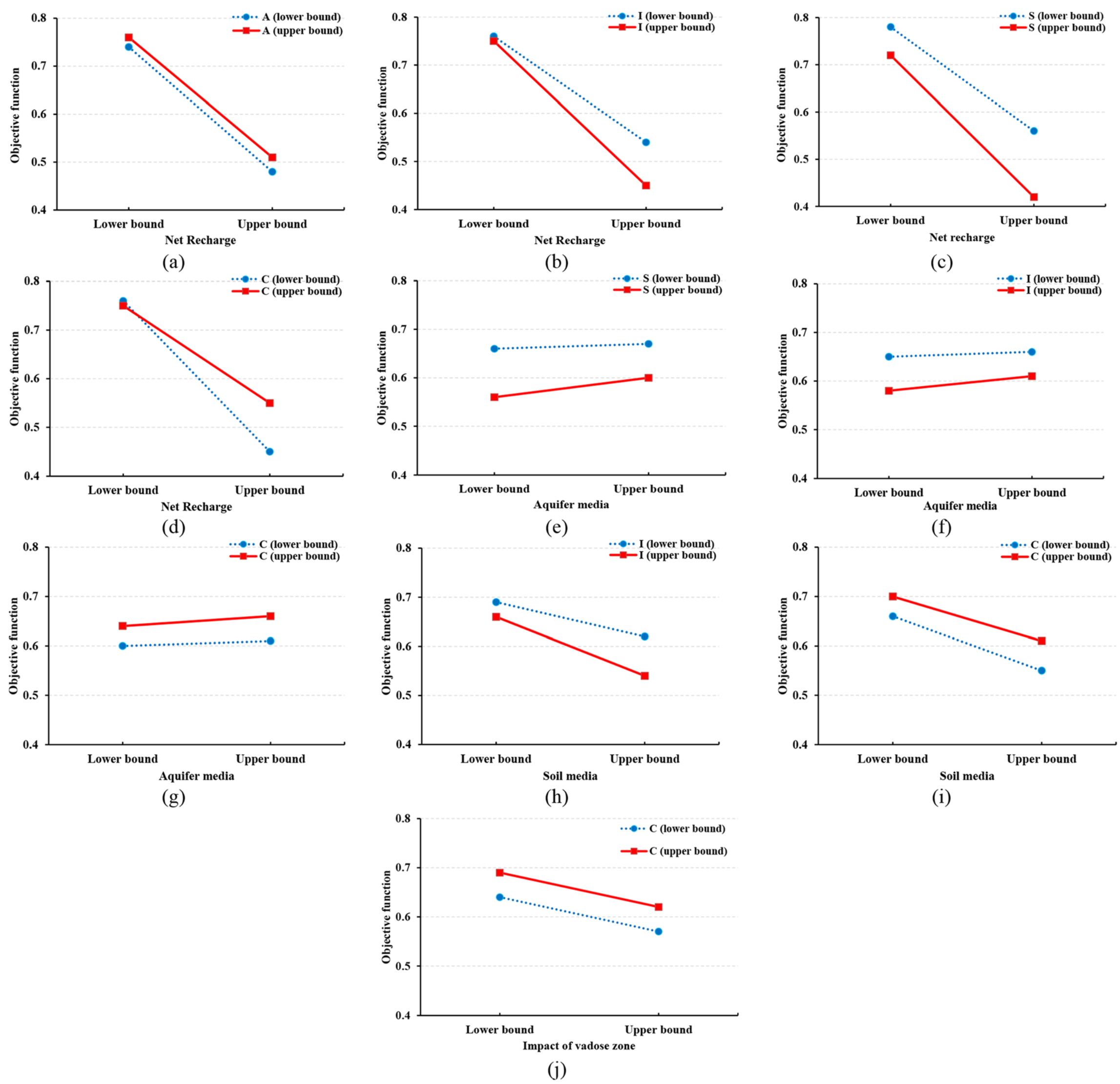


Fig. 10 **a**–**j** The interactive effect of uncertain parameters on the objective function

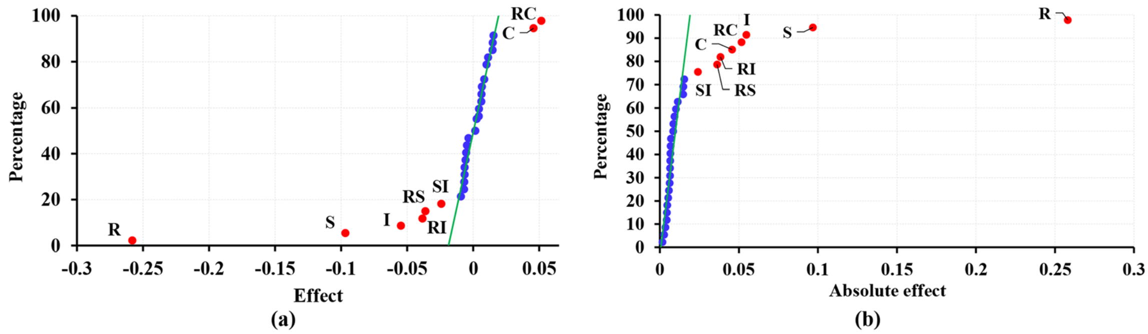


Fig. 11 The probability graph of the parameter effects on the objective function **a** normal graph, and **b** half-normal graph

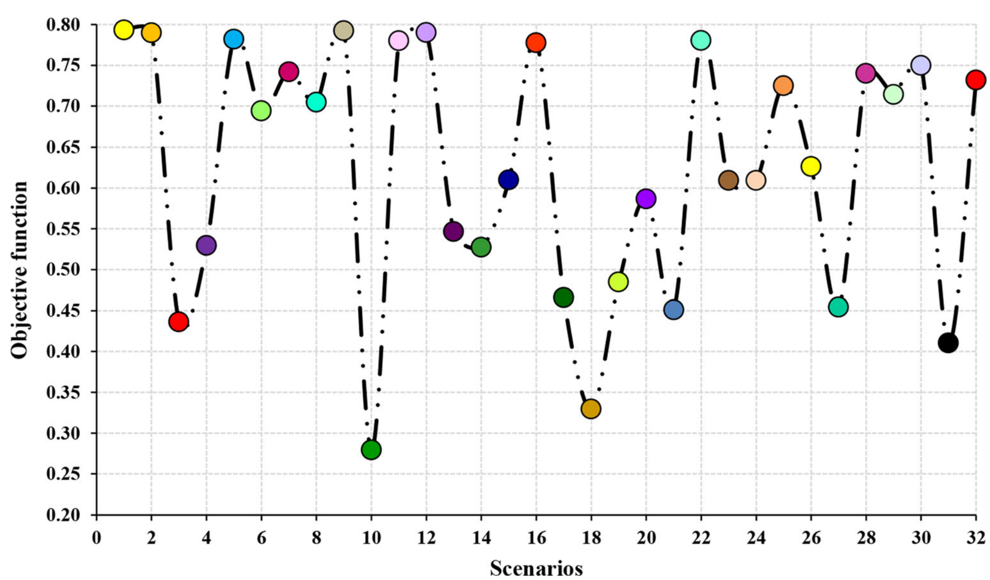
the objective function. For the parameter of A, the weight of the parameter does not change. Hence, due to its low weight and stability, it may conclude that the objective function is independent of the changes in parameter A.

To obtain the vulnerability zoning maps, 32 uncertain states (32 scenarios) were used in the factorial analysis. Each of these different states expresses a situation that may happen to the aquifer due to uncertainties embedded in the input parameters. These maps would indicate the probable vulnerability zoning of the site regardless of the corresponding correlation coefficients. Figure 14 depicts two vulnerability zoning maps for Shiraz plain corresponding to the maximum and minimum values of the correlation coefficient. Selecting these two specific states is because of assessing the aquifer situation in the best and worst conditions. In addition, comparing vulnerability zoning maps in maximum and minimum values of the correlation coefficient revealed that the northern parts of the aquifer have the slightest vulnerability in the most probable scenarios. Conversely, the southern regions will experience the highest vulnerability in any situation.

The results were also compared to researches conducted by Jafari and Nikoo (2019) and Baghapour et al. (2016). The first one has the same data and case study area (Shiraz plain) to use the uncertainty analysis methods for groundwater vulnerability assessment. Meanwhile, there was a difference in methodology to modify the rate of the DRASTIC parameters where the AHP method was employed in this research, but Jafari and Nikoo (2019) used the Wilcoxon method. In addition, in this research, the method used to determine the uncertainty was factorial analysis. However, Jafari and Nikoo (2019) provided a fuzzy model to consider the uncertainties embedded in the input parameters. They presented the results of fuzzy analysis by a membership function related to the correlation coefficient, which shows how the range of correlation coefficient changes at different levels of the parameters' uncertainty (α-cuts) with plotting the vulnerability zoning maps at different α-cuts. By comparing the evaluation process, the following results can be obtained:

1. Factorial analysis is a complete sensitivity analysis method that considers the interactive effects of uncertain parameters in different possible scenarios (e.g., all 32 scenarios presented in Table 1). However, one-

Fig. 12 The values of the objective function in different scenarios



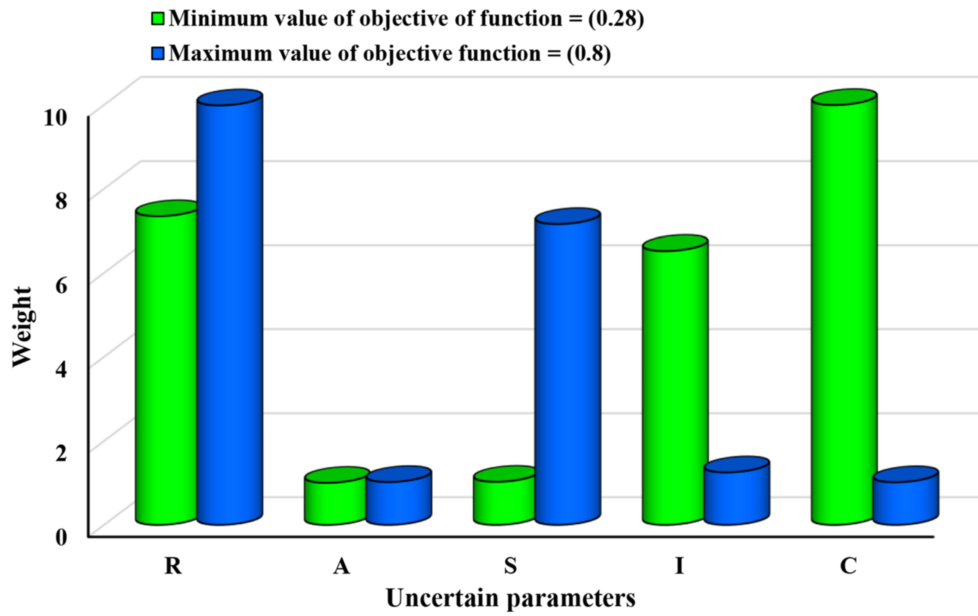
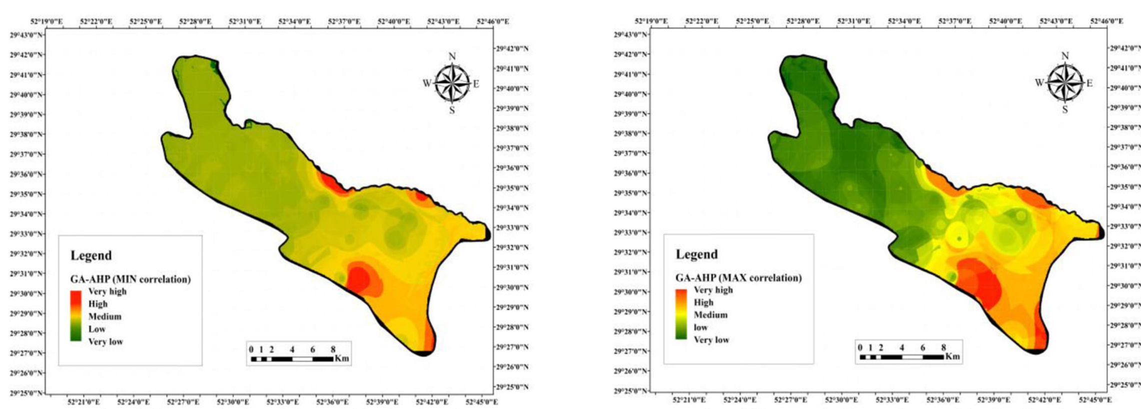


Fig. 13 The weights of uncertain parameters in minimum and maximum values of the objective function

parameter-at-a-time (OAT) sensitivity analysis performed by Jafari and Nikoo (2019) could not consider these parameters' interactive effect. In addition, the results show a clear difference between the DRASTIC parameters' categorization based on their importance degree. Thanks to the factorial analysis method, the authors assert that the factorial analysis method efficiently recognizes the sensitive parameters.

2. Comparing the results of the AHP-GA model used in this research with those of the Wilcoxon-GA model used by Jafari and Nikoo (2019) revealed that both models yielded much better results than the simple DRASTIC method. However, the AHP model provided a higher value of the maximum correlation coefficient than the Wilcoxon model's value (0.8 vs. 0.78).
3. Finally, the vulnerability maps obtained by Jafari and Nikoo (2019) indicated a relatively similar zoning pattern with our results, in which the East and southeastern parts of the Shiraz plain are more susceptible to contamination than the north and northwestern regions. However, a point-to-point comparison of the maps showed differences in the vulnerability zoning, mainly due to the selection of different uncertain parameters with different intervals.

Another case selected for the result's comparison was the research adopted by Baghapour et al. (2016) in which a composite DRASTIC vulnerability model of Shiraz plain was employed. They added land use as a new parameter to the DRASTIC index, optimizing the weight and rate of the DRASTIC parameters using an artificial neural network. Their results showed a similar vulnerability pattern of the Shiraz aquifer so that the southeastern parts were the most vulnerable zones and northwestern areas were less susceptible to contamination.



Vulnerability map in the minimum objective function
Vulnerability map in the maximum objective function

Fig. 14 The vulnerability map of the aquifer in minimum and maximum values of the objective function

Summary and conclusions

In this research, a hybrid AHP-GA-based DRASTIC model was developed, taking into account the uncertainty of the parameters. Considering the variation bounds of uncertain parameters, the factorial analysis defined 32 scenarios. AHP method was utilized to rate DRASTIC parameters, and then the genetic algorithm (GA) optimized the weight of the parameters in all scenarios. The objective function was the correlation coefficient between the vulnerability index and the nitrate concentration. The efficiency of the proposed method was evaluated in the Shiraz aquifer, southwestern Iran, which revealed that the net recharge parameter had the highest single effect on the objective function. Also, the effect between net recharge and hydraulic conductivity was the most significant interactive effect, and the aquifer media was ineffective in the variation of the objective function. Investigating the variation trend of the objective function over the aquifer approved that the proposed method provides a more accurate groundwater vulnerability map. This research provides valuable knowledge for assessing groundwater vulnerability, enabling decision-makers to apply more precise information on groundwater vulnerability in future water resource management plans.

Author contributions Yalda Norouzi Gharakezloo: methodology; software; investigation; writing—original draft, review, and editing; visualization; resources. Mohammad Reza Nikoo: project administration; conceptualization; investigations; writing—review and editing; resources; supervision. Ayoub Karimi Jashni: supervision, review and editing. Mehrdad Ghorbani Mooselu: methodology; software; investigations; writing—original draft, review, and editing; visualization; resources; supervision.

Funding Not applicable

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication All the authors approved the final manuscript and agreed to its publication in the Environmental Science and Pollution Research.

Competing interests The authors declare no competing interests.

References

Ahn JJ, Kim YM, Yoo K, Park J, Oh KJ (2012) Using GA-Ridge regression to select hydrogeological parameters influencing groundwater pollution vulnerability. Environ Monit Assess 184:6637–6645. https://doi.org/10.1007/s10661-011-2448-1

Aller L, Bennett T, Lehr JH, Petty RJ, Hackett G (1987) DRASTIC: a standardized system for evaluating ground water pollution potential using hydrogeologic settings, vol 455. US Environmental Protection Agency, Washington, DC

Arezoomand Omidi Langrudi M, Khashei Siuki A, Javadi S, Hashemi SR (2016) Evaluation of vulnerability of aquifers by improved fuzzy drastic method: case study: Aastane Kochesfahan plain in Iran. Ain Shams Eng J 7:11–20. https://doi.org/10.1016/j.asej.2015.11.013

Arshad A, Zhang Z, Zhang W, Dilawar A (2020) Mapping favorable groundwater potential recharge zones using a GIS-based analytical hierarchical process and probability frequency ratio model: a case study from an agro-urban region of Pakistan. Geosci Front 11:1805–1819. https://doi.org/10.1016/j.gsf.2019.12.013

Baalousha HM, Tawabini B, Seers TD (2021) Fuzzy or Non-Fuzzy? A comparison between fuzzy logic-based vulnerability mapping and DRASTIC approach using a numerical model. A Case Study from Qatar. Water 13:1288

Baghapour MA, Nobandegani AF, Talebbeydokhti N, Bagherzadeh S, Nadiri AA, Gharekhani M, Chitsazan N (2016) Optimization of DRASTIC method by artificial neural network, nitrate vulnerability index, and composite DRASTIC models to assess groundwater vulnerability for unconfined aquifer of Shiraz Plain. Iran J Environ Heal Sci Eng 14:1–16

Balaji L, Saravanan R, Saravanan K, Sreemanthrarupini NA (2021) Groundwater vulnerability mapping using the modified DRASTIC model: the metaheuristic algorithm approach. Environ Monit Assess 193:1–19

Barzegar R, Asghari Moghaddam A, Adamowski J, Nazemi AH (2019) Delimitation of groundwater zones under contamination risk using a bagged ensemble of optimized DRASTIC frameworks. Environ Sci Pollut Res 26:8325–8339. https://doi.org/10.1007/s11356-019-04252-9

Bera A, Mukhopadhyay BP, Chowdhury P, Ghosh A, Biswas S (2021) Groundwater vulnerability assessment using GIS-based DRASTIC model in Nangasai River Basin, India with special emphasis on agricultural contamination. Ecotoxicol Environ Saf 214:112085

Bordbar M, Neshat A, Javadi S (2019) A new hybrid framework for optimization and modification of groundwater vulnerability in coastal aquifer. Environ Sci Pollut Res 26:21808–21827

Daneshmand F, Karimi A, Nikoo MR, Bazargan-Lari MR, Adamowski J (2014) Mitigating socio-economic-environmental impacts during drought periods by optimizing the conjunctive management of water resources. Water Resour Manage 28(6):1517–1529. https://doi.org/10.1007/s11269-014-0549-7

Holland JH (1975) Adaptation in natural and artificial systems, University of Michigan press. Ann arbor, MI 1:5

Hu X, Ma C, Qi H, Guo X (2018) Groundwater vulnerability assessment using the GALDIT model and the improved DRASTIC model: a case in Weibei Plain. China Environ Sci Pollut Res 25:32524–32539

Jafari SM, Nikoo MR (2016) Groundwater risk assessment based on optimization framework using DRASTIC method. Arab J Geosci 9. https://doi.org/10.1007/s12517-016-2756-4

Jafari SM, Nikoo MR (2019) Developing a fuzzy optimization model for groundwater risk assessment based on improved DRASTIC method. Environ Earth Sci 78:109

Jesiya NP, Gopinath G (2019) A customized FuzzyAHP - GIS based DRASTIC-L model for intrinsic groundwater vulnerability assessment of urban and peri urban phreatic aquifer clusters. Groundw Sustain Dev 8:654–666. https://doi.org/10.1016/j.gsd.2019.03.005

Kumar A, Pramod Krishna A (2019) Groundwater vulnerability and contamination risk assessment using GIS-based modified DRASTIC-LU model in hard rock aquifer system in India. Geocarto Int 35:1149–1178. https://doi.org/10.1080/10106049.2018.1557259

Liu M, Xiao C, Liang X, 2021. Assessment of groundwater vulnerability based on the modified DRASTIC model: a case study in Baicheng City, China

Mallik S, Bhowmik T, Mishra U, Paul N (2021) Local scale groundwater vulnerability assessment with an improved DRASTIC model. Nat Resour Res 30:2145–2160

Montgomery DC (2017) Design and analysis of experiments. John wiley & sons

Mooselu MG, Nikoo MR, Sadegh M (2019) A fuzzy multi-stakeholder socio-optimal model for water and waste load allocation. Environ Monit Assess 191(6):1–16

Mooselu MG, Nikoo MR, Latifi M, Sadegh M, Al-Wardy M, Al-Rawas GA (2020) A multi-objective optimal allocation of treated wastewater in urban areas using leader-follower game. J Clean Prod 267: 122189

Nadiri AA, Norouzi H, Khatibi R, Gharekhani M (2019) Groundwater DRASTIC vulnerability mapping by unsupervised and supervised techniques using a modelling strategy in two levels. J Hydrol 574: 744–759. https://doi.org/10.1016/j.jhydrol.2019.04.039

Nikoo MR, Beiglou PHB, Mahjouri N (2016) Optimizing multiple-pollutant waste load allocation in rivers: an interval parameter game theoretic model. *Water Resour Manag 30*(12):4201–4220

Pacheco FAL, Pires LMGR, Santos RMB, Sanches Fernandes LF (2015) Factor weighting in DRASTIC modeling. Sci Total Environ 505: 474–486. https://doi.org/10.1016/j.scitotenv.2014.09.092

Pourkhosravani M, Jamshidi F, Sayari N (2021) Evaluation of groundwater vulnerability to pollution using DRASTIC, composite DRASTIC, and nitrate vulnerability models. Environ Heal Eng Manag J 8:0

Saaty TL (1980) The analytic hierarchy process Mcgraw Hill, New York. Agric Econ Rev 70

Saranya T, Saravanan S (2021) Evolution of a hybrid approach for groundwater vulnerability assessment using hierarchical fuzzy-DRASTIC models in the Cuddalore Region. India Environ Earth Sci 80:1–25

Sener E, Davraz A (2013) Assessment of groundwater vulnerability based on a modified DRASTIC model, GIS and an analytic hierarchy process (AHP) method: The case of Egirdir Lake basin (Isparta, Turkey). Hydrogeol J 21:701–714. https://doi.org/10.1007/s10040-012-0947-y

Soyaslan İİ (2020) Assessment of groundwater vulnerability using modified DRASTIC-Analytical Hierarchy Process model in Bucak Basin. Turkey Arab J Geosci 13:1–12

Tavakoli A, Kerachian R, Nikoo MR, Soltani M, Estalaki SM (2014) Water and waste load allocation in rivers with emphasis on agricultural return flows: application of fractional factorial analysis. Environ Monit Assess 186:5935–5949. https://doi.org/10.1007/s10661-014-3830-6

Tavakoli A, Nikoo MR, Kerachian R, Soltani M (2015) River water quality management considering agricultural return flows: application of a nonlinear two-stage stochastic fuzzy programming. Environ Monit Assess 187(4). https://doi.org/10.1007/s10661-015-4263-6

Tomer T, Katyal D, Joshi V (2019) Sensitivity analysis of groundwater vulnerability using DRASTIC method: a case study of National Capital Territory, Delhi. India Groundw Sustain Dev 9:100271. https://doi.org/10.1016/j.gsd.2019.100271

Torkashvand M, Neshat A, Javadi S, Yousefi H (2020) DRASTIC framework improvement using Stepwise Weight Assessment Ratio Analysis (SWARA) and combination of Genetic Algorithm and Entropy. Environ Sci Pollut Res 1–21

Yang J, Tang Z, Jiao T, Malik Muhammad A (2017) Combining AHP and genetic algorithms approaches to modify DRASTIC model to assess groundwater vulnerability: a case study from Jianghan Plain. China Environ Earth Sci 76. https://doi.org/10.1007/s12665-017-6759-6

Yazdian M, Rakhshandehroo G, Nikoo MR, Mooselu MG, Gandomi AH, Honar T (2021) Groundwater sustainability: developing a non-cooperative optimal management scenario in shared groundwater resources under water bankruptcy conditions. J Environ Manag 292:112807