



Optimizing the DRASTIC vulnerability approach to overcome the subjectivity: a case study from Shabestar plain, Iran

Fatemeh Kadkhodaie¹ · Asghar Asghari Moghaddam¹ · Rahim Barzegar^{1,2} · Maryam Gharekhani¹ · Ali Kadkhodaie¹

Received: 16 December 2018 / Accepted: 10 July 2019 / Published online: 12 August 2019
© Saudi Society for Geosciences 2019

Abstract

Due to the seasonality of Shabestar plain rivers, most of the water demands for agricultural and drinking sectors are provided from groundwater. The existence of agricultural activities has expanded the use of chemical and animal fertilizers that are possible to infiltrate and contaminate the groundwater resources. The purpose of this study is to provide an optimized DRASTIC approach for assessing the vulnerability of the Shabestar plain aquifer located in East Azarbaijan Province, North West of Iran. The effective parameters in this method include the depth to groundwater, net recharge, aquifer media, soil media, impact of the vadose zone, topographic slope, and hydraulic conductivity, which are being provided in seven layers in the ArcGIS software environment. After ranking each of the input parameters and applying their equivalent weight coefficients, the DRASTIC vulnerability index for the study area ranged between 53.3 and 118.3. Nitrate concentration values at 66 water wells were used for the validation of the DRASTIC-based vulnerability maps. The r value between the DRASTIC indices and the concentration of nitrate is 0.38 indicating a low correlation. Sensitivity analysis results showed that the impact of the vadose zone and aquifer media is significant on the intrinsic vulnerability of the study area. The Wilcoxon rank-sum test (WRST) was used for rate modification. Also, the weights of the DRASTIC approach were optimized by using the ant colony optimization (ACO) and genetic algorithm (GA). All optimized models increased the r compared to the typical DRASTIC method. The results indicate that WRST and GA methods were most successful for optimization of the rates and weights, respectively, where the WRST-GA-DRASTIC model obtained an r of 0.63.

Keywords DRASTIC · Subjectivity · Optimization · Iran

Introduction

Due to the increasing population growth and water requirements, the use of freshwater sources, such as groundwater resources, has become very important. On the other hand, human activities, such as agriculture, urbanization, and industry, have reduced groundwater quality (Kazakis and Voudouris 2015). The groundwater flow is very slow, and, if contaminated, it will take a lot of time to remove

contamination. Therefore, prevention measures are the most appropriate strategy to combat groundwater contamination. One of the important tools for the management and protection of groundwater is the provision of groundwater vulnerability maps (Patrikaki et al. 2012). So far, various definitions have been presented for the aquifer vulnerability and its concept in hydrogeology. Among them can be pointed to the definition of the National Committee of the United States in 1993 (Stigter et al. 2006). According to the definition of this committee, groundwater vulnerability is the possibility of infiltration and diffusion of pollution to reach the groundwater system from the ground surface (Almasri 2008). The vulnerability of groundwater is divided into two types of intrinsic and specific vulnerability. Intrinsic vulnerability depends on the characteristics of geology, hydrology, and hydrogeology of the region and human activities, while it is independent on the nature of pollution (Hamza et al. 2007). Specific vulnerability indicates groundwater vulnerability to particular pollution or a group of pollution caused by human activities; in fact, it depends on the

Editorial handling: Angela Vallejos

✉ Asghar Asghari Moghaddam
Moghaddam@tabrizu.ac.ir

¹ Department of Earth Sciences, Faculty of Natural Sciences, University of Tabriz, 29 Bahman Boulevard, Tabriz, Iran

² Department of Bioresource Engineering, McGill University, 2111 Lakeshore, Ste Anne de Bellevue, Quebec H9X3V9, Canada

inherent sensitivity, location, and type of the contaminant source (Gogu and Dassargues 2000). There are various approaches to assess intrinsic vulnerability, for example, DRASTIC (Aller et al. 1987), AVI (Van Stempvroot et al. 1993), GOD (Foster 1987), SINTACS (Civita 1994), and Time-Input (Kralik and Keimel 2003). The most common method to assess the groundwater vulnerability is the DRASTIC which was presented by the US Environmental Protection Agency (USEPA) for the first time (Aller et al. 1987). The effective parameters in the DRASTIC method are the depth of groundwater, net recharge, aquifer media, soil media, impact of the vadose zone, topographic slope, and hydraulic conductivity being provided in seven layers in the GIS software environment. Through ranking and assigning weights to the aforementioned input layers, the DRASTIC index is calculated. Different hydrological, hydrogeological, and geological characteristics for different areas can lead to optimizing the DRASTIC method. Also, due to the fact that the rates and weights in the DRASTIC method are somewhat dependent on the expert's decision, therefore, reducing the subjectivity is necessary to provide a reliable vulnerability map. So far, different methods have been used to optimize the DRASTIC method. A number of these methods are presented in Table 1.

In this paper, the optimized methods, such as genetic algorithm (GA), and ant colony optimization (ACO), as a new methodology, were used to optimize the weights. Also, Wilcoxon rank-sum test (WRST) was used to modify the rates of the DRASTIC approach. Accordingly, the various DRASTIC models, including typical DRASTIC, ACO-DRASTIC, GA-DRASTIC, WRST-DRASTIC, WRST-ACO-DRASTIC, and WRST-GA-DRASTIC, were generated. These models were employed to evaluate the groundwater vulnerability of the Shabestar plain, NW Iran, as a case study. The main objective of this study is to compare the performance of the mentioned optimization models in assessing the vulnerability of the study area.

Material and methods

Study areas

Shabestar plain with an area of 350 km² and geographical coordinates, $x = 540500$ to $x = 587500$ eastern longitude and $y = 4215000$ to $y = 4239550$ northern latitude, is situated in East Azarbaijan province in NW Iran. The plain is bordered to the north by the Marand Zilberchay basin, to the east by the Tabriz plain, to the south by the Urmia Lake, and to the west by the Tasuj Basin. The slope of the ground surface in the northern parts is high, and, to the south of the plain, the amount of slope is reduced and the ground becomes almost flat. The DEM (digital elevation model) of the study area is

shown in Fig. 1. The average annual rainfall in the region is 325 mm (Asghari Moghaddam et al. 2017). The highest and lowest rainfall occurred in May and August, respectively. The climate of the area is semi-arid. The most important rivers in the plain are the Daryan-Chay, Payam-Chay, Shanehjan-Chay, and Sis-Chay which flow approximately northeast–southwest (Asghari Moghaddam et al. 2017).

Geologically, the oldest deposit in the area, in the southern part of Mishoo and the northern plain of Shabestar, is the Kahar Formation, which includes micaceous shale with fine sandstone, schist, and a little dolomitic stromatolite with layers of chert and dark limestone. On the southern slopes, there is the Barut Formation that includes sandstone and dolomite. In the Mishoo Mountain, the Dorood and Rute Formations are composed of dark and red limestone recovered by shale. Cretaceous sediments in the area include massive limestone (the eastern part of the region) and sedimentary flesh (the middle section of the region). Cenozoic rocks are dominated by a condensed conglomerate, silt, marl, clay (west and center of the region), and red marl with layers of the conglomerate (northern part of Khamene and northeastern part of Shabestar). The expansion of Quaternary deposits in the region is high, such as conglomerates, long alluvial terraces, old alluvial, alluvial plains, and young alluvial terraces, which are exposed (Asghari Moghaddam et al. 2017).

The Shabestar plain aquifer is single-layered and unconfined. However, in the southern parts of the plain, there is a confined aquifer, but the water table falls below the impermeable layer indicating hydrogeologically that the aquifer is not merely a confined aquifer but also acts as an unconfined aquifer. Based on the pumping test data, the amount of transmissivity in the middle of the plain is estimated to be about 1000 to 1500 m²/day, which decreases towards the south of the plain with a value of 100 m²/day. The storage coefficient of the aquifer ranged between 2 and 8%.

DRASTIC method

The DRASTIC method (Aller et al. 1987) is based on the seven hydrological, hydrogeological, and geological parameters, including groundwater depth (D), net recharge (R), aquifer media (A), soil media (S), topography (T), impact of vadose zone (I), and hydraulic conductivity (C). Relative weight is allocated to each of the aforementioned parameters with respect to their importance in the transmission of pollution to the groundwater system. The weight ranged between 1 and 5, which is the least and most important parameters, respectively. The rate of the parameters varied from 1 to 10 based on the characteristics of the region, in which 1 means the lowest and 10 means the most dangerous for groundwater contamination (Panagopoulos et al. 2006). Table 2 shows the rates of DRASTIC index parameters.

Table 1 A number of methods to optimize the DRASTIC method

Researchers	Year	Optimization method/model	Result
Panagopoulos et al.	2006	Statistical methods	The r increased to 33% higher than the original method
Sadat-Noori and Ebrahimi	2016	Statistical methods	The r^2 increased significantly from 0.52 to 0.78
Thirumalaivasan et al.	2003	AHP	The r was equal to 0.84 at 0.01 level
Asefi et al.	2013	AHP	AHP showed a higher r compared to the original model
Yang et al.	2017	AHP-GA	The r increased significantly from 41.07 to 75.31% after modification
Neshat et al.	2014	WRST	Result of the modified DRASTIC is better than that of the original method
Jafari and Nikoo	2016	WRST-GA	The r increased to 0.82
Neshat and Pradhan	2015	SPSA, FR, AHP	FR-AHP with a higher value of r was the optimal model
Dixon	2005	Fuzzy logic	This model had better estimations of groundwater vulnerability
Mahdavi and Zare	2015	Fuzzy logic models	Based on these models, south, west, and northeast showed maximum contamination potential to nitrate.
Barzegar et al.	2016	ANN, SFL, MFL, NF, CMI, SCMAI	SCMAI was the best model for prediction of vulnerability
Ghanbari et al.	2017	ANN, MFL, SFL, NF	The r^2 between NF output and nitrate concentration was higher than the other models
Hamami and Nadiri	2018	MFL, SFL, LFL, SCFL	Combining the all fuzzy logic models (SCFL) used from the advantage of each FL model
Nadiri et al.	2018	WRST, GA, SWM, AIMF	AIMF was the best model and increased r to 0.84
Barzegar et al.	2019	WRST, FR, CC, AHP, GA	WRST-GA was the best model. By combining all optimized models based on bagging, ensemble model r value increased to 0.67

To calculate the aquifer vulnerability index, the parameters used are introduced into a simple linear equation as follows:

$$DI = D_r D_w + R_r R_w + A_r A_w + S_r S_w + T_r T_w + I_r I_w + C_r C_w \tag{1}$$

In the above-mentioned relationship, DI is a vulnerability index. The capital letters represent the seven parameters, r is the rate of value (rank), and w is the weight assigned to each parameter. The larger the index, the greater the potential for contamination.

Sensitivity analysis

Sensitivity analysis provides useful information about ratings and weights assigned to each parameter (Gogu and Dassargues 2000). The effect of each of the used seven DRASTIC parameters can be determined on the vulnerability of the study area. For this purpose, the sensitivity analysis was used as the variation index. The VAR_i is calculated as follows (Sadat-Noori and Ebrahimi 2016):

$$VAR_i = \left(\frac{v_i - v_{xi}}{v_i} \right) \times 100 \tag{2}$$

where VAR_i is the variation index of the removal parameter and v_i and v_{xi} are vulnerability index calculated using Eq. (1) on the i th subarea and vulnerability index of the i th subarea removing one map layer, respectively.

Nitrate concentration data

Nitrate as one of the most common groundwater pollutants has different origins, such as domestic, municipal, and industrial wastewater, chemical fertilizers, and animal wastes. In this study, nitrate was used as an indicator to validate vulnerability assessment, because of the fact that nitrate, produced by human activities, is reached into the groundwater from the ground surface and also is considered as indicator of groundwater quality degradation (US EPA 1996). Therefore, nitrate concentration collected from 66 water wells was used to validate the DRASTIC method. The concentration of nitrate is varied between 0.2 and 70 mg/L with an average of 16.4 mg/L. Figure 2 shows the distribution map of nitrate concentration in groundwater of Shabestar plain.

Ant colony optimization

Ant colony algorithm was first introduced in 1992 by Marco Dorigo in his Ph.D. thesis. This algorithm is based on the actual behavior of the ants, in which the ants can find the shortest route to find food using their simple communication mechanisms. As they move on their path, a chemical material called pheromone remains on the earth which guides the ants to the target. Ants choose their own pathway, which has the greatest effect of pheromone. Due to the fact that pheromone is evaporated with time, the path that more ants cross it has more pheromone and encourages other ants to cross that. As

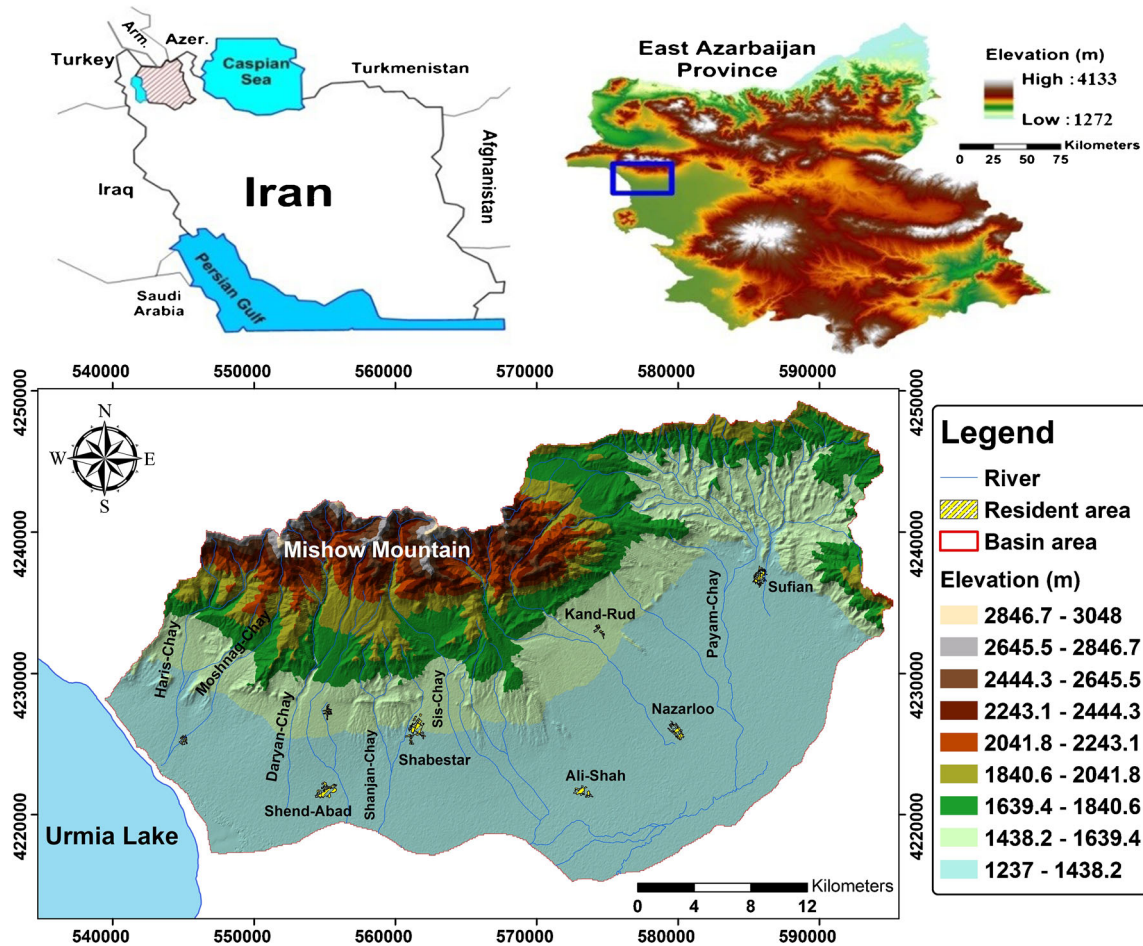


Fig. 1 Digital elevation model (DEM) of the Shabestar basin, NW Iran

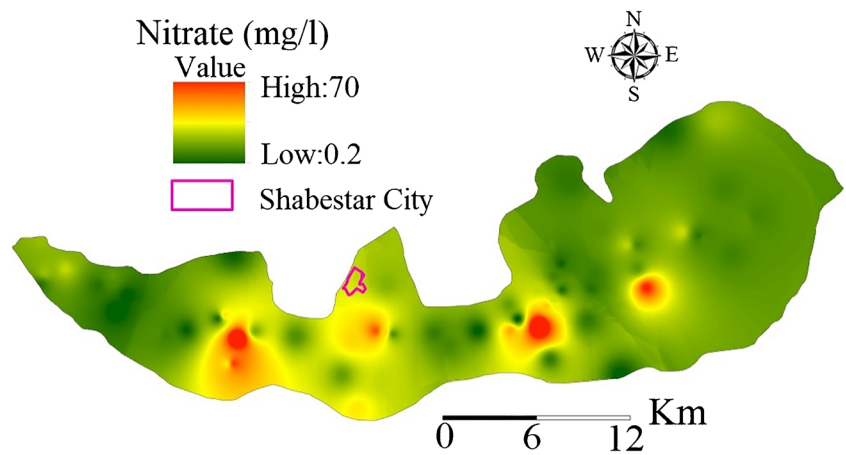
shown in Fig. 3, the mechanism of finding food from the shortest path by ants includes the following three steps:

- 1) The ant by walking a route finds a food source and returns to the nest by placing the effect of the pheromone.

Table 2 The rates of DRASTIC index parameters (Aller et al. 1987)

Range	Rate	Range	Rate	Range	Rate	Range	Rate	Range	Rate
Groundwater depth (m)		Topography (%)		Aquifer media		Impact of vadose zone		Soil media	
0–1.5	10	0–2	10	Massive shale	2	Confining layer	1	Thin or absent	10
1.5–4.6	9	2–6	9	Metamorphic	3	Silt/clay	3	Gravel	10
4.6–9.1	7	6–12	5	Weathered metamorphic	4	Shale	3	Sand	9
9.1–15.2	5	12–18	3	Alluvium	5	Limestone	6	Peat	8
15.2–22.8	3	18 <	1	Sandstone, limestone, shale	6	Sandstone	6	Shrinking clay	7
22.8–30.4	2	Hydraulic conductivity (m d ⁻¹)	0.04–4.1	Massive limestone	6	Sandstone, shale	6	Sandy loam	6
30.4 <	1			Massive sandstone	6	Gravel, sand, clay	6	Loam	5
Net recharge (mm)		4.1–12.3	2	Gravel, sand	8	Metamorphic	4	Silty loam	4
0–50.8	1	12.3–28.7	4	Basalt	9	Gravel, sand	8	Clay loam	3
50.8–101.6	3	28.7–41	6	Karst limestone	10	Basalt	9	Muck	2
101.6–177.8	6	41–82	8			Karst limestone	10	No shrinking clay	1
177.8–254	8	82 <	10						
254 <	9								

Fig. 2 The spatial distribution of the nitrate concentration of Shabestar plain



- 2) Ants follow one of the four possible paths, but the gradual increase in the amount of pheromone in the shortest path increases the tendency of ants to path this route.
- 3) Ants follow the shortest route that has the highest amount of pheromones to reach the food (Dreo 2006).

A simple ant colony optimization composed of two nodes of v_s (representing nest of ant) and v_d (representing a source of food). e_1 and e_2 are between v_s and v_d . Lengths of l_1 and l_2 are allocated to e_1 and e_2 , respectively. l_1 and l_2 indicate short and long paths between v_s and v_d , respectively. An artificial pheromone value t_i is considered for each of the two links $e_i = i = 1, 2$. This shows the strength of pheromone trail on the path. Finally, n_a artificial ants are introduced. The behavior of each ant is as follows (Blum 2005):

$$p_i = \frac{t_i}{t_1 + t_2} \quad i = 1, 2 \tag{3}$$

If $t_1 > t_2$, it means a trail of pheromone in the path e_1 is a strength. Then, the probability of choosing a path e_1 is higher and vice versa. An ant uses the same path as it chose to reach v_d . For returning from v_d to v_s , the ant changes the value of artificial pheromone as follows (Blum 2005):

$$t_i \leftarrow t_i + \frac{Q}{l_i} \tag{4}$$

where Q is a positive constant parameter of the model. The above-mentioned equation shows artificial pheromone value that is added depending on the chosen path. In other words, the shortest path has a high pheromone value. In nature, pheromone deposits evaporate over time. Simulation of this pheromone is as follows (Blum 2005):

$$t_i \leftarrow (1-p) \cdot t_i, \quad i = 1, 2. \tag{5}$$

The parameter $p \in (0, 1]$ regulates the evaporation of pheromone.

Genetic algorithm

A genetic algorithm was introduced by Holland in 1975 (Holland 1975). GA is a randomized optimization method that mimics the same biological concepts. In the GA, a population is initially formed, and, then, the individuals develop from one generation to the next with a repetitive process of evolution. Each evolutionary stage is called a generation. Each person of this generation is judged by the suitability function. The evolutionary routine consists of choice, combination, and mutation. An evolutionary algorithm generally determines the initial population randomly and measures the suitability of each individual based on their fitness value in the environment. In this way, more qualified people will be selected several times to participate in the new population and thus have a greater chance of reproduction.

In this method, each optimization problem has three basic and fundamental components, including the objective function, the decision variables or parameters, and constraints (Jafari and Nikoo 2016). For each optimization problem, the population of the probable solutions chosen becomes the best possible solution. Each solution can be modified based on its compatibility with the selection, crossover, and mutation operators (Jafari and Nikoo 2016) (Fig. 4).

Wilcoxon rank-sum test

The Wilcoxon rank-sum test (WRST) (Wilcoxon 1945) is a non-parametric statistical test that is used to evaluate the similarity of two samples dependent on a rating scale. In addition to considering the differences in symptoms, this test also takes into account the difference between them. So, since it uses more information, it has a more accurate answer (Ahmadi et al. 2013).

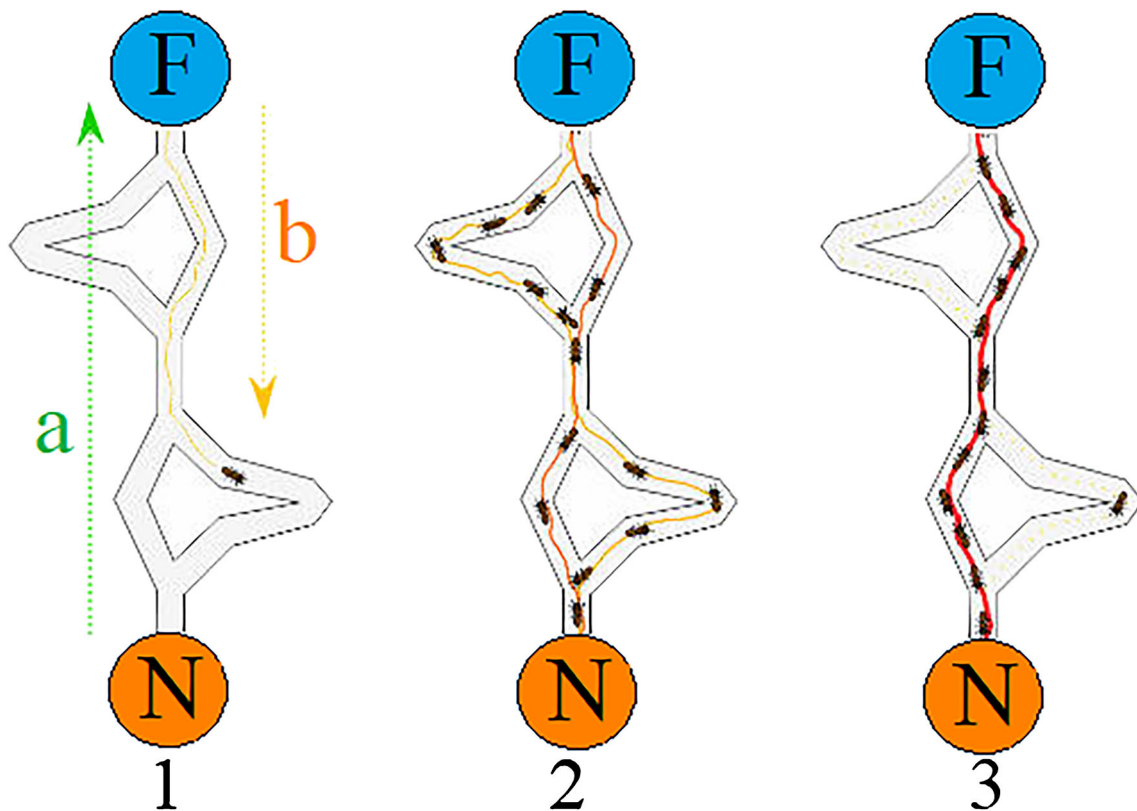


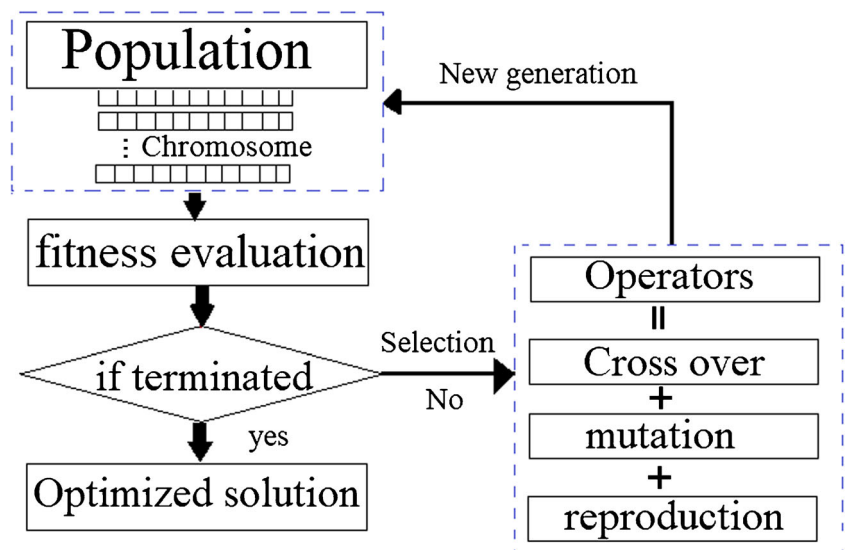
Fig. 3 The mechanism of finding food from the shortest path by ants (Dreo 2006)

Modeling procedure

For arranging seven layers of the DRASTIC method, relevant information was inserted into the GIS environment and the layers were prepared as the raster layers. The groundwater depth parameter represents the depth that the pollutant needs to pass through it to reach the water table. In order to provide the groundwater depth layer, the 24-piezometer water table information in the study area for the year 2016–2017 was

entered into the GIS software environment and interpolated with the IDW method. Afterwards, they were classified using the ratings shown in Table 2. The average depth to groundwater ranged between 2.5 and 86.6 m in the study area. Therefore, rates between 1 and 9 are assigned to this parameter (Fig. 5a). To calculate the net recharge, Piscopo’s (2001) method was employed. For this method, three layers, including precipitation, slope, and soil maps, are needed. To prepare the slope map, a digital elevation model (DEM) of the area

Fig. 4 Flowchart of the genetic algorithm methodology (Chen and Lin 2006)



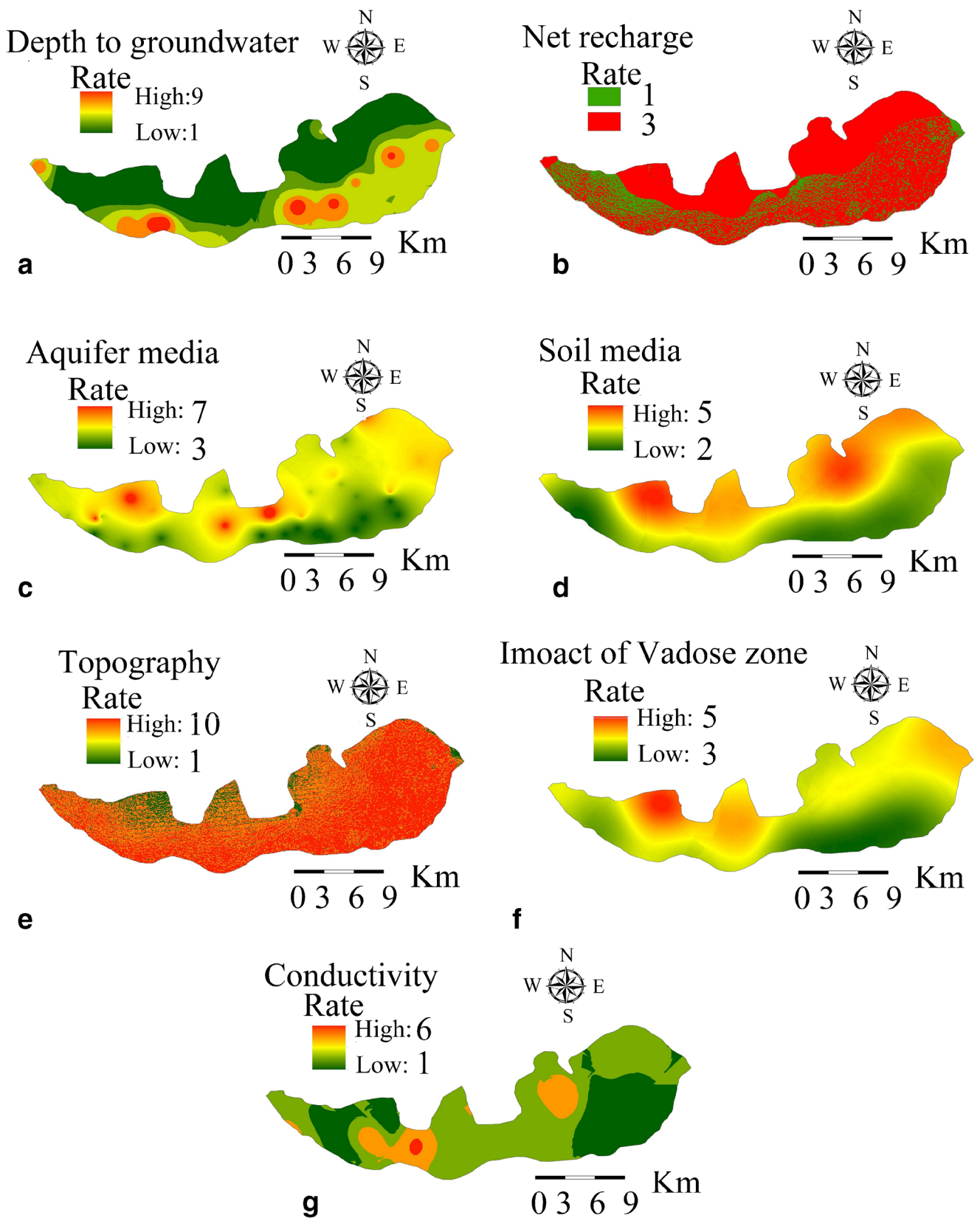


Fig. 5 Input layer maps in the DRASTIC method

was obtained and then the percent slope of the ground surface was extracted and ranked according to the standards in Table 3. In order to prepare the soil map, considering the type of soil layer in 2 m above the logs of the observation wells, the soil map was classified into five classes and interpolated by the Kriging method. Afterward, a rate was assigned to each class based on soil permeability according to Table 3. As regards the fact that the amount of precipitation for the area is less than 500 mm per year, rank 1 was considered for the precipitation layer. Three raster maps were prepared according to Eq. (6) and superimposed. Then, the net recharge layer was reclassified and rated (Fig. 5b). The ratings of net recharge ranged between 1 and 3 in this study.

$$\text{Net recharge} = \text{soil permeability} + \text{rainfall} + \text{gradient} \quad (6)$$

To prepare the aquifer media, soil media, and vadose zone layers, 35 available well logs in the area were considered. Materials of these logs consist of gravel, sand, mud, clay, and silt. To prepare the aquifer media layer, based on the aquifer materials, the ratings between 3 and 7 were given according to Table 2 and interpolated (Fig. 5c). For the soil layer, the upper part of the unsaturated zone or the surface layer (upper 2 m of the well logs in the area) is considered. The ratings between 2 and 5 were assigned according to Table 2 and transformed into a raster layer by the Kriging method (Fig. 5d). The percent slope of the ground surface was calculated in the GIS and classified based on Aller et al. (1987). Then, the rating values were assigned for each class and raster topography map was generated (Fig. 5e). Ratings between 1 and 10 were assigned to topography layer. For the impact of the vadose zone, the material of this zone was considered, and the ratings between 3 and 5 were assigned according to Aller et al. (1987) (Fig. 5f). Hydraulic conductivity refers to the ability of the aquifer-forming material to convey water, depending on the percentage of void spaces associated with the saturated layer. By dividing amounts of transmissivity (obtained from 52 observation wells), saturation thickness amounts of hydraulic conductivity in the study area are ranged between 0.1 and 38.2 m/day. Then, conductivity ratings are ranged between 1 and 6 (Fig. 5g).

Table 3 Ranking of net recharge based on Piscopo (2001)

Soil permeability		Rainfall (mm)		Slope (%)		Net recharge	
Intensity	Rate	Rate	Range	Rate	Range	Rate	Range
High	5	850 <	4	2 >	4	11–13	10
Moderate to high	4	700–850	3	2–10	3	9–11	8
Moderate	3	500–700	2	10–33	2	7–9	5
Low	2	500 >	1	33 <	1	5–7	3
Very low	1					3–5	1

In this study, in order to optimize the weights of the DRASTIC method, ACO and GA methods were used. WRST method was also employed to modify the rates.

In the WRST method, the rates of parameters (e.g., D, R, A, S, T, I, C) were rescaled based on the mean nitrate concentration in each class for each layer, then the highest rate of each parameter was assigned to the class with the highest mean of nitrate concentration, and other rates were modified linearly according to this relation. Table 4 shows the modified rates based on the Wilcoxon’s method.

The ACO is generally expressed as $P = (S, f)$, in which the limited set of objects S and fitness function $f: S \rightarrow R$ are given that attributes a fitness value to every $s \in S$. The aim is to search for an object of minimum fitness value. Knowing that there are n input data x_i and an output data y , the estimation error is calculated as follows:

Table 4 The typical and modified rates based on mean nitrate concentration in the Wilcoxon’s method

Vulnerability parameter	Initial rate	Nitrate concentration	Modified rating
Depth to groundwater	9	0.01	0.001
	7	67.5	9
	5	20.55	2.74
	3	15.79	2.1
	2	16.31	2.17
Net recharge	1	12.49	1.66
	3	14.59	2.13
Aquifer media	1	20.48	3
	3	19.47	7
	4	16.39	5.89
	5	17.26	6.2
Soil media	6	0.01	0.003
	7	0.01	0.003
	5	11.13	3.02
Topography	4	16.72	4.54
	3	18.38	5
	2	12.9	3.5
	10	16.07	9.23
Impact of vadose zone	9	17.41	10
	5	17.26	9.91
	3	3.18	1.82
	1	0.01	0.005
Hydraulic conductivity	5	12.56	3.33
	4	16.27	4.32
Hydraulic conductivity	3	18.81	5
	4	19.68	6
	2	17.88	5.45
	1	11.21	3.41
	6	0.01	0.003

Table 5 Main weighting and modified weighting coefficients using the ant colony algorithm

Modified weighting coefficient (ACO)	Original weighting coefficient	Vulnerability parameter
3.73	5	Groundwater depth
1	4	Net recharge
1.15	3	Aquifer media
1	2	Soil media
1	1	Topography
1.37	5	Impact of vadose zone
1.77	3	Hydraulic conductivity

$$e = y - t(x_1, x_2, \dots, X_n) \tag{7}$$

where t is the output of linear model written as follows:

$$t(x_1, x_2, \dots, X_n) = a_1x_1 + a_2x_2 + \dots + a_nx_n + a_n + 1 \tag{8}$$

where a_1, a_2, \dots, a_n are weighting coefficients of the input parameters (Kadkhodaie 2015). Seven parameters of the DRASTIC were considered as the inputs of the ACO. The output of the ACO is calculated using the following equation:

$$y = \frac{vul_{max}}{(No_3)_{max}} \times (No_3)_i \tag{9}$$

In the above-mentioned equation, y is the adjusted vulnerability index, vul_{max} is the maximum DRASTIC index, $(No_3)_{max}$ is the maximum nitrate concentration, and $(No_3)_i$ is the nitrate concentration. The MSE (mean squared error) of adjusted vulnerability index was employed as the objective function to be minimized by ACO as follows (Kadkhodaie 2015):

$$MSE = \frac{1}{m} \sum_{j=1}^m (y_j - t(x_{ij}))^2 \quad i = 1, 2, \dots, n \tag{10}$$

Table 6 Original and optimized weights of DRASTIC using GA method

Modified weighting coefficient (GA)	Original weighting coefficient	Vulnerability parameter
5	5	Groundwater depth
1	4	Net recharge
5	3	Aquifer media
1.11	2	Soil media
1	1	Topography
5	5	Impact of vadose zone
2.69	3	Hydraulic conductivity

where m and n are the number of estimated points and input parameters for prediction of adjusted vulnerability index, respectively.

For developing the ACO, the number of initial ants of 500, epoch of 20, and $1 < a < 5$ were set. Finally, the optimized weights were obtained when a minimum MSE for the objective function was reached. Table 5 represents the optimized weight values of the DRASTIC method obtained by the ACO method.

For GA-based optimization of the DRASTIC method, seven parameters of the DRASTIC as the variables were considered. The minimum and maximum values between 1 and 5 were defined as constraints for the weights of the variables. The objective function for the GA method was the correlation coefficient between the nitrate concentration and DRASTIC index. The objective function used to optimize the DRASTIC through the GA is presented as follows (Jafari and Nikoo 2016):

$$Max F = Corr(x, y) \tag{11}$$

$$F = \frac{\sum_{j=1}^n ((x_j - \bar{x}) - (y_j - \bar{y}))}{\sqrt{\sum_{j=1}^n (x_j - \bar{x})^2} \sqrt{\sum_{j=1}^n (y_j - \bar{y})^2}} \quad \text{Constrain : } 1 \leq W_i \leq 5 \quad j = 1, 2, \dots, \tag{12}$$

where F represents an objective function, n denotes the number of wells, y_j is contaminant concentration, \bar{y} represents the mean concentration of contaminant, x_j is vulnerability index, \bar{x} is mean vulnerability index, and W_i is the weight of the DRASTIC parameters. Finally, with the purpose of maximizing the function, optimized weights were obtained. The results of the GA-based optimization for DRASTIC are presented in Table 6. Considering the optimized rates using the WRST method, weights of the DRASTIC method were optimized by the GA and ACO. The results of the optimization of the rate and weight of the DRASTIC method are presented in Table 7.

Validation of models

In order to validate the various vulnerability maps, the correlation coefficient (r) was calculated as follows:

$$r = \frac{\sum_{i=1}^n (Vul_i - \overline{Vul})(N_i - \overline{N})}{\sqrt{\sum_{i=1}^n (Vul_i - \overline{Vul})^2} \sqrt{\sum_{i=1}^n (N_i - \overline{N})^2}} \tag{13}$$

where r is the correlation coefficient, Vul_i is the vulnerability index, \overline{Vul} is the mean of a vulnerability index, N_i represents the nitrate concentration, and \overline{N} is the mean of nitrate concentration.

Table 7 Typical and modified weighting coefficients based on WRST-GA and WRST-ACO models

Parameter	Typical DRASTIC weight	WRST-GA Modified weighting coefficient	WRST-ACO Modified weighting coefficient
Depth to groundwater	5	5	2.84
Net recharge	4	3.96	1.80
Aquifer media	3	5	1.44
Soil media	2	2.73	1
Topography	1	1	1
Impact of vadose zone	5	1	1
Conductivity	3	2.30	1.39

Results and discussion

By overlapping the seven effective layers (e.g., D, R, A, S, T, I, C) in the DRASTIC approach and using Eq. (1), intrinsic vulnerability index was calculated from 53.3 to 118.3 in the study area. The vulnerability index was categorized into five categories, according to geometrical interval classification suggested by Huan et al. (2012), including very low, low, moderate, high, and very high vulnerability. The DRASTIC-based vulnerability map of the Shabestar plain is shown in Fig. 6. Based on the typical DRASTIC vulnerability map, very low and very high vulnerable areas are located in the west and south-west of the plain, respectively. Distribution of the class area shows that areas with very low, moderate, high, and very high vulnerability covered 6.79, 39.11, 47.45, 6.02, and 0.6% of the study area, respectively. The amount of r between the DRASTIC index and the concentration of nitrate was calculated as 0.38 indicating an adaptation of the

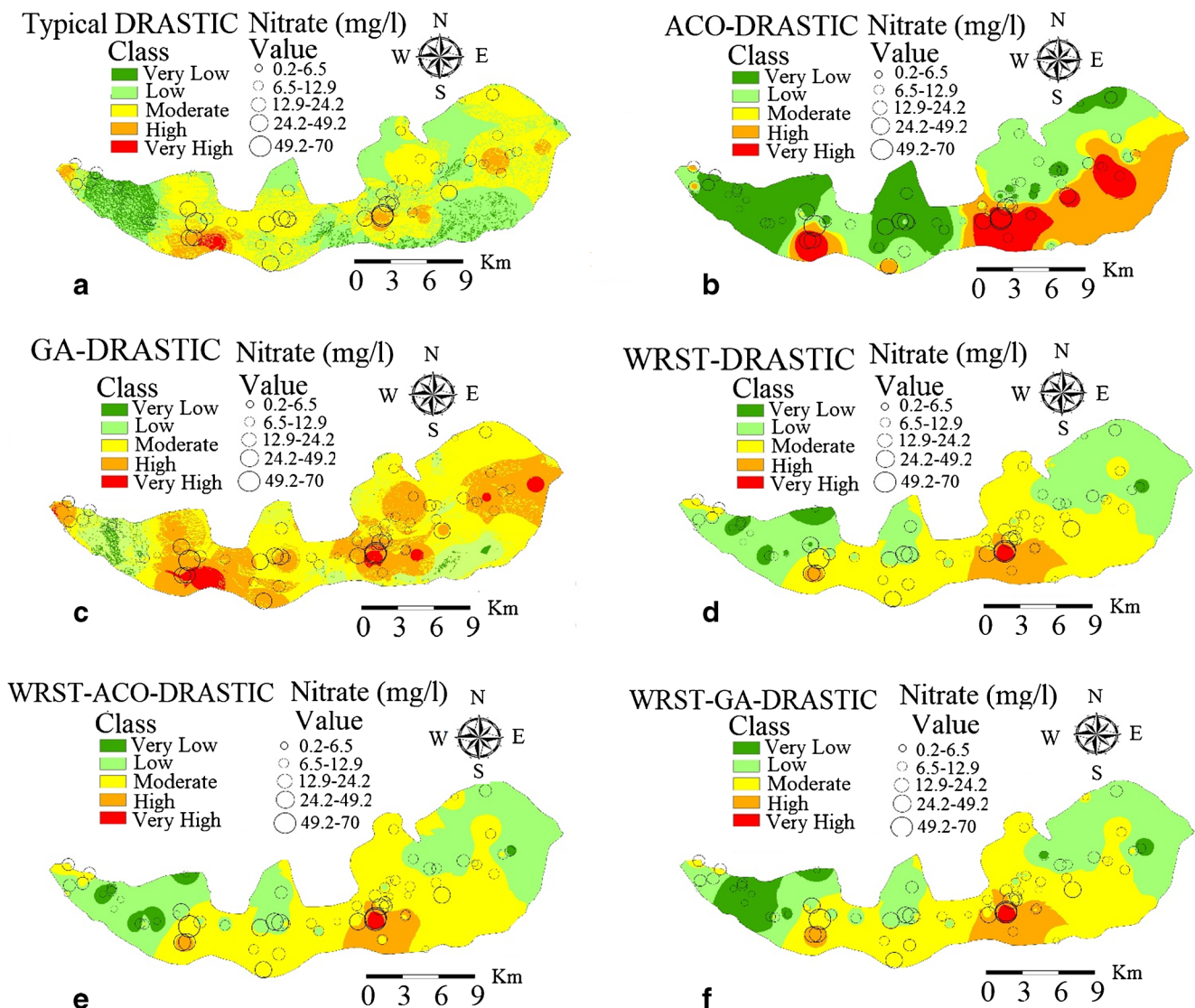
**Fig. 6** Typical and optimized DRASTIC vulnerability maps for the Shabestar aquifer

Table 8 Statistical results of the map removal sensitivity analysis for typical DRASTIC model

Variation index (%)	Removal parameter						
	D	R	A	S	T	I	C
Maximum	37.96	20.87	23.7	16.55	17.3	36.16	16.63
Minimum	1.28	4.07	11.92	5.66	4.88	19.43	2.63
Mean	14.14	11.8	17.93	9.55	11.92	27.55	8.09
Standard deviation	8.75	4.81	2.79	2.25	2.08	4.03	3.47

vulnerability map to the distribution of nitrate concentration data does not show much correlation. In other words, zones with a high vulnerability index do not show high nitrate concentration. Accordingly, the typical DRASTIC index cannot be precise enough to evaluate the vulnerability of Shabestar plain aquifer. On the other hand, the typical DRASTIC map is a general model and the weight assigned to each parameter should be modified based on the hydrogeological characteristics of each specific area. Sensitivity analysis by variation index method was performed to determine the effect of ratings and weights assigned to each parameter. By removing one parameter of the typical DRASTIC method and using Eq. (2), the statistical results of the sensitivity analysis were calculated (Table 8). The results show that the effect of the vadose zone and aquifer media has the greatest impact on the vulnerability of the area, while hydraulic conductivity and soil media have the least effect. On the other hand, the sensitivity of each of the parameters in addition to the weight is influenced by the rates. Therefore, it can be stated that both weights and rates of the DRASTIC method need to be optimized. Accordingly, in this study, the GA and ACO were used to optimize weights, and WRST was employed to modify the rates. Afterward, various optimized models were developed to assess the vulnerability of Shabestar plain aquifer (Fig. 6). Validation of the models also was calculated by *r* between the nitrate concentration and optimized DRASTIC index. The validation results of optimized models are shown in Table 9. It can be concluded that all optimized models elevated the *r* compared to the typical DRASTIC model. Using ACO and GA methods to optimize weights, *r* increased to 0.47 and 0.52 compared to the typical DRASTIC model, respectively. The WRST method was

Table 9 Validation results of the various DRASTIC models

Model	<i>r</i>
Typical DRASTIC	0.38
ACO-DRASTIC	0.47
GA-DRASTIC	0.52
WRST-DRASTIC	0.60
WRST-ACO-DRASTIC	0.62
WRST-GA-DRASTIC	0.63

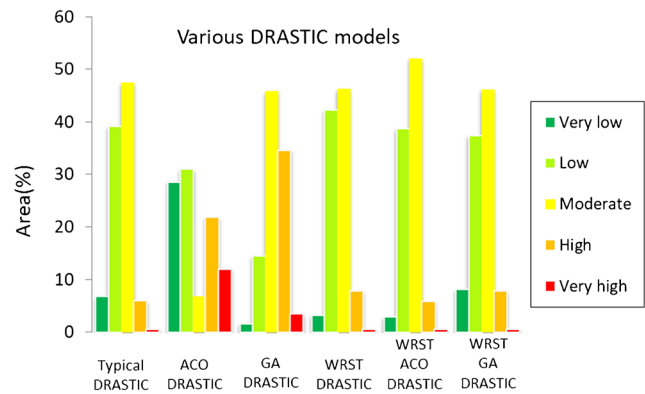


Fig. 7 Distribution of the class areas (%) according to each applied method

used to optimize the ratings of increased *r* to 0.6 compared to the typical DRASTIC model. By applying ACO and GA based on the WRST-DRASTIC model, *r* increased to 0.62 and 0.63, respectively. These results indicate that the WRST-GA-DRASTIC model with higher *r* outperforms the other optimized models in assessing the vulnerability of the Shabestar plain aquifer. Based on this model, zones with very low, low, moderate, high, and very high vulnerability cover 8.08, 37.32, 46.17, 7.82, and 0.58% of the total area, respectively. Various optimized models showed different percentages for different vulnerability categories (Fig. 7). Based on ACO-DRASTIC model, zones with very low vulnerability index are located in north and northwest (28.57% of total area) while zones with very high vulnerability index are situated in south and southeast (11.91% of total area). According to the GA-DRASTIC model, zones with very low and very high vulnerability indices are located in the west (1.6% of total area) and parts of the east and southeast (3.5% of total area), respectively. Based on the WRST-DRASTIC model, zones with very low and very high vulnerability indices are located in western and southern parts, respectively, which cover 3.72 and 5.64% of the total area, respectively. Based on the WRST-ACO-DRASTIC model, zones with very low and very high vulnerability indices are located in the western parts (2.84% of total area) and a small part of the southeast (0.52% of total area).

Conclusion

DRASTIC vulnerability index was calculated from 53.3 to 118.3 in the Shabestar plain. In this model, the areas with very low, moderate, high, and very high vulnerability cover 6.79, 39.11, 47.45, 6.02, and 0.6% of the study area, respectively, while the zones with very low and very high vulnerability indices are located in the west and south-west of the study area, respectively. Nitrate concentration data were used for validation of the DRASTIC model. The *r* value of 0.38 between the DRASTIC indices and nitrate concentrations was calculated, which

indicates a low correlation. Results of sensitivity analysis show that the vadose zone and aquifer media with mean variation index of 27.55 and 17.93%, respectively, have the greatest impact on the vulnerability of the area while hydraulic conductivity and soil media with mean variation index of 8.09 and 9.55%, respectively, are less effective on vulnerability. WRST was used to modify the rates, while ACO and GA were used to optimize the weights. Afterward, various optimized frameworks were developed to assess the vulnerability of the plain. By applying ACO and GA methods to optimize weights, the r value was increased to 0.47 and 0.52 compared to the typical DRASTIC approach, respectively. The WRST method used to optimize the rates increased the r value to 0.6 compared to the typical DRASTIC method. By applying ACO and GA based on the WRST-DRASTIC framework, the r increased to 0.62 and 0.63, respectively. These results indicate that the WRST-GA-DRASTIC framework with higher r value outperforms the other optimized models in evaluating the vulnerability of the Shabestar plain aquifer. Based on the proposed model, zones with very low, low, moderate, high, and very high vulnerability index cover 8.08, 37.32, 46.17, 7.82, and 0.58% of the total area, respectively. Based on the model, new industrial units should be constructed in the western part of the study area with very low vulnerability.

References

- Ahmadi J, Akhondi L, Abbasi H, Khashei-Siuki A, Alimadadi M (2013) Determination of aquifer vulnerability using DRASTIC model and a single parameter sensitivity analysis and acts and omissions case study: Salafchegan-Neyzar Plai. *J Water Soil Conserv* 20(3):1–25 (In Persian)
- Aller L, Bennet T, Leher J. H, Petty R. J, Hackett G (1987) DRASTIC: a standardized system for evaluating groundwater pollution potential using hydro-geological settings. Kerr Environmental Research Laboratory, U.S. Environmental Protection Agency Report (EPA/600/2-87/035)
- Almasri MN (2008) Assessment of intrinsic vulnerability to contamination for Gaza coastal aquifer. *J Environ Manag* 88(4):577–593
- Asefi M, Radmanesh F, Zarei H (2013) Optimization of DRASTIC model for vulnerability assessment of groundwater resources using analytical hierarchy process (case study: Andimeshk plain). *J Irrig Sci Eng* 37(1):55–67 (In Persian)
- Asghari Moghaddam A, Barzegar R, Fijani E, Saberi Mehr S, Nadiri AA (2017) Study of quantitative and qualitative situation of Shabestar plain drinking wells and qualitative modeling for identification risky wells and provide corrective solutions. Final report of the research project, East Azarbaijan water and wastewater company in Iran, 271 pages (In Persian)
- Barzegar R, Asghari Moghaddam A, Baghban H (2016) A supervised committee machine artificial intelligent for improving DRASTIC method to assess groundwater contamination risk a case study from Tabriz plain aquifer, Iran. *Stoch Env Res Risk* A 30(3):883–899
- Barzegar R, Asghari Moghaddam A, Adamowski J, Nazemi A. H (2019) Delimitation of groundwater zones under contamination risk using a bagged ensemble of optimized DRASTIC frameworks. *Environ Sci Pollut Res* 1–15
- Blum C (2005) Ant colony optimization: introduction and recent trends. *Phys Life Rev* 2:353–373
- Chen CH, Lin ZS (2006) A committee machine with empirical formulas for permeability prediction. *Comput Geosci* 32:385–496
- Civita M (1994) Le carte della vulnerabilita degli acquiferi all' inquinamento Teoria and practica (Aquifer vulnerability maps to pollution) (In Italian). Pitagora Ed, Bologna, p. 325
- Dixon B (2005) Groundwater vulnerability mapping: a GIS and fuzzy rules based integrated tool. *J Appl Geogr* 25:327–347
- Dorigo M (1992) Optimization, learning and natural algorithms, PhD dissertation, Politecnico di Milano, Italy
- Dreo J (2006) Shortest path find by an ant colony. http://en.wikipedia.org/wiki/File:Aco_branches.svg filehistory, 27.5
- Foster S (1987) Fundamental concepts in aquifer vulnerability, pollution risk and protection strategy. International Conference Noordwijk a Zee, Netherlands, pp 1–30
- Ghanbari N, Rangzan K, Kabolizade M, Moradi P (2017) Improve the results of the DRASTIC model using artificial intelligence methods to assess groundwater vulnerability in Ramhormoz alluvial aquifer plain. *J Water Soil Conserv* 24(2):45–65 (In Persian)
- Gogu RC, Dassargues A (2000) Current trends and future challenges in groundwater vulnerability assessment using overlay and index methods. *Environ Geol* 39(6):549–559
- Hamamin D. F, Nadiri A. A (2018) Supervised committee fuzzy logic model to assess groundwater intrinsic vulnerability in multiple aquifer systems. *Arab J Geosci* 1–14
- Hamza MH, Added A, Rodriguez R, Abdeljaoued S, Ben Mammou A (2007) GIS-based DRASTIC vulnerability and net recharge reassessment in an aquifer of a semi-arid region (Metline-Ras Jebel-Raf Raf aquifer, Northern Tunisia). *J Environ Manag* 84:12–19
- Holland JH (1975) Adaptation in natural and artificial systems. University of Michigan, Ann Arbor
- Huan H, Wang J, Teng Y (2012) Assessment and validation of ground water vulnerability to nitrate based on a modified DRASTIC model: a case study in Jilin city of northeast China. *Sci Total Environ* 440:14–23
- Jafari SM, Nikoo MR (2016) Groundwater risk assessment based on optimization framework using DRASTIC method. *Arab J Geosci* 9(742):7–14
- Kadkhodaie A (2015) A systematic approach for estimation of reservoir rock properties using ant colony optimization. *Geopersia* 5(1):1–11
- Kazakis N, Voudouris KS (2015) Groundwater vulnerability and pollution risk assessment of porous aquifers to nitrate: modifying the DRASTIC method using quantitative parameters. *J Hydrol* 525:13–25
- Kralik M, Keimel T (2003) Time-input, an innovative groundwater-vulnerability assessment scheme: application to an alpine test site. *Environ Geol* 44(6):679–686
- Mahdavi A, Zare Abyane H (2015) Determination of aquifer vulnerability potential based on DRASTIC and fuzzy logic models (case study: Hamedan-Bahar plain). *Water and Soil* 25(1):1-17 (In Persian)
- Nadiri AA, Gharekhani M, Khatibi R (2018) Mapping aquifer vulnerability indices using artificial intelligence-running multiple frameworks (AIMF) with supervised and unsupervised learning. *Water Resour Manag* 32:3023–3040
- Neshat A, Pradhan B (2015) An integrated DRASTIC model using frequency ratio and two new hybrid methods for groundwater vulnerability assessment. *Nat Hazards* 76:543–563
- Neshat AR, Pradhan B, Pirasteh S, Shafri HZM (2014) Estimating groundwater vulnerability to pollution using a modified DRASTIC model in the Kerman agricultural area, Iran. *Environ Earth Sci* 1–13
- Panagopoulos G, Antonakos A, Lambrakis N (2006) Optimization of DRASTIC model for groundwater vulnerability assessment, by the use of simple statistical methods and GIS. *Hydrogeol J* 14:894–911

- Patrikaki O, Kazakis N, Voudouris K (2012) Vulnerability map: a useful tool for groundwater protection: an example from Mouriki Basin, North Greece. *Fresenius Environ Bull* 21(8c):2516–2521
- Piscopo G (2001) Groundwater vulnerability map, explanatory notes, Castlereagh Catchment, NSW, Department of Land and Water Conservation, Australia
- Sadat-Noori M, Ebrahimi K (2016) Groundwater vulnerability assessment in agricultural areas using a modified DRASTIC model. *Environ Monit Assess* 188(19):1–18
- Stigter T, Riberio L, Carvalho Dill A (2006) Evaluation of an intrinsic and a specific vulnerability assessment method in comparison with groundwater salinization and nitrate contamination level in two agriculture regions in the south of Portugal. *Hydrogeol J* 14:79–99
- Thirumalaivasan D, Karmegan M, Venugopal K (2003) AHP-DRASTIC: software for specific aquifer vulnerability assessment using DRASTIC model and GIS. *J Environ Model Softw* 18:645–656
- Van Stempvoort D, Evert L, Wassenaar L (1993) Aquifer vulnerability index: a GIS compatible method for groundwater vulnerability mapping. *Can Water Resour J* 18:25–37
- Wilcoxon F (1945) Individual comparisons by ranking methods. *Biomet Bull* 1:80–83
- Yang J, Zhonghua T, Jiao T, Malik Muhammad A (2017) Combining AHP and genetic algorithms approaches to assess groundwater vulnerability: a case study from Jiangnan plain, China. *J Environ Earth Sci* 76(426)