

# Combining AHP and genetic algorithms approaches to modify DRASTIC model to assess groundwater vulnerability: a case study from Jiangnan Plain, China

Jing Yang<sup>1</sup> · Zhonghua Tang<sup>1</sup> · Tian Jiao<sup>1</sup> · Akhtar Malik Muhammad<sup>1</sup>

Received: 16 December 2016 / Accepted: 10 June 2017 / Published online: 19 June 2017  
© Springer-Verlag GmbH Germany 2017

**Abstract** Accurate identification of vulnerability areas is critical for groundwater resources protection and management. The present study employed the modified DRASTIC model to assess the groundwater vulnerability of Jiangnan Plain, a major farming area in central China. DRASTIC model was developed by incorporating the land use factor to the original model. The ratings and weightings of the selected parameters were optimized by analytic hierarchy process (AHP) method and genetic algorithms (GAs) method, respectively. A combined AHP–GAs method was proposed to further develop this methodology. The unity-based normalization process was employed to categorize the vulnerability maps into four types, such as very high ( $>0.75$ ), high ( $0.5–0.75$ ), low ( $0.25–0.5$ ), and very low ( $<0.25$ ). The accuracy of vulnerability mapping was validated by Pearson's correlation coefficient between vulnerability index and the nitrate concentration in groundwater and analysis of variance  $F$  statistic. The results revealed that the modified DRASTIC model had a large improvement over the conventional model. The correlation coefficient increased significantly from 41.07 to 75.31% after modification. Sensitivity analysis indicated that the depth to groundwater with 39.28% of mean effective weight was the most critical factor affecting the groundwater vulnerability. The developed vulnerability model proposed in this study could provide important objective information for groundwater and environmental management at local level and innovation for international researchers.

**Keywords** Modified DRASTIC · Vulnerability · AHP · Genetic algorithms · Jiangnan Plain

## Introduction

Groundwater is one of the most precious freshwater resources in many countries, especially in arid and semiarid regions (Neshat et al. 2014). Groundwater quality especially in the unconfined aquifer is sensitive to surface contamination, caused by the agriculture and industrial activities (Al-Hanbali and Kondoh 2008; Javadi et al. 2011a, b). There is a clear urgent need for treating contaminated groundwater, and furthermore, for delineating potential vulnerability areas and applying effective policies to prevent groundwater from being polluted. However, traditional methods such as field investigations at regional scales are often not effective, time-consuming, and costly. Consequently, the assessment of groundwater vulnerability has been subject to intensive research and a variety of methods have been developed (Neukum and Azzam 2009), such as DRASTIC model (Aller et al. 1987).

DRASTIC model is widely used to estimate the groundwater vulnerability (Babiker et al. 2005; Chitsazan and Akhtari 2008; Huan et al. 2012; Sener et al. 2009; Yin et al. 2012). This method is gradually becoming a standardized approach for assessment of groundwater vulnerability due to the property of easy application at regional scales with Geographic Information Systems (GIS). Seven hydrogeological factors, including depth to groundwater, net recharge, aquifer media, soil media, topography, impact of vadose zone, and hydraulic conductivity, are used to evaluate the vulnerability index. As a relative measure, the vulnerability index is dimensionless. The higher values indicate more vulnerable to contamination.

✉ Zhonghua Tang  
zhhtang@cug.edu.cn

<sup>1</sup> School of Environmental Studies, China University of Geosciences, No. 388 Lumo Road, Wuhan 430074, Hubei Province, People's Republic of China

Generally, the ratings and weightings of the seven features used in DRASTIC model are assigned specific values as shown in Table 1 without considering the regional hydrogeological conditions, which then makes DRASTIC model an easy target for criticism. Moreover, this technique disregards the impact of alternative data on human activities, such as land use. Thus, many researchers have modified the traditional model to further develop this method (Al-Hanbali and Kondoh 2008; Hernández-Espriú et al. 2014; Neshat et al. 2013, 2014; Secunda et al. 1998; Sener and Davraz 2012; Thirumalaivasan et al. 2003).

An evolutionary notable modification is applying analytic hierarchy process (AHP) to determine the weightings and/or ratings for parameters used in DRASTIC model (Thirumalaivasan et al. 2003). AHP method is a powerful approach in dealing with multicriteria decision-making technique. A set of pairwise comparisons (PCMs) is used to obtain the weights in regard to the impact of the decision and subdecision criteria (Yalcin et al. 2011). Accordingly, by using AHP technique, the ratings and/or weightings of

geodata layers are optimized by transferring acquisition of knowledge from experts to a model.

While AHP method is practically useful, subjectivity from experts may distort the results of estimation. Genetic algorithms (GAs) is recently the most prominent and widely used adaptive optimization search approach based on a direct analogy to Darwinian natural selection and genetics in biological systems (Holland 1975). This method shows a strong ability in global optimization. Rahimi et al. (2014) used the GAs method to the flood spreading site selection of Gareh Bygone plain, Iran. Giacobbo et al. (2002) presented the feasibility of using GAs for estimating the parameters of groundwater contaminant transport model. Zhang et al. (2014b) employed GAs method to optimize the initial weights determined by artificial neural networks (ANN) to assess the occurrence of earth fissure in Su-Xi-Chang land subsidence area.

Jiangnan Plain, located in the middle reaches of Yangtze River, is a major farming area in central China. Recent

**Table 1** Original DRASTIC weightings and ratings (Aller et al. 1987)

Depth (m)		Recharge (mm/year)		Topography (%)		Conductivity (m/d)	
Weight: 5		Weight: 4		Weight: 1		Weight: 3	
Range	Rating	Range	Rating	Range	Rating	Range	Rating
0–1.5	10	0–50.8	1	0–2	10	0.04–4.1	1
1.5–4.6	9	50.8–101.6	3	2–6	9	4.1–12.3	2
4.6–9.1	7	101.6–177.8	6	6–12	5	12.3–28.7	4
9.1–15.2	5	177.8–254	8	12–18	3	28.7–41	6
15.2–22.8	3	>254	9	>18	1	41–82	8
22.8–30.4	2					>82	10
>30.4	1						
Aquifer media		Impact of vadose zone		Soil media			
Weight: 3		Weight: 5		Weight: 2			
Range	Rating	Range	Rating	Range	Rating		
Massive shale	2	Confining layer	1	Thin or absent	10		
Metamorphic/igneous	3	Silt/clay	3	Gravel	10		
Weathered metamorphic/igneous	4	Shale	3	Sand	9		
Glacial till	5	Limestone	6	Peat	8		
Bedded sandstone, limestone	6	Sandstone	6	Shrinking clay	7		
Massive sandstone	6	Bedded limestone, Sandstone, Shale	6	Sandy loam	6		
Massive limestone	8	Sand and gravel with significant silt and clay	6	Loam	5		
Sand and gravel	8	Metamorphic/igneous	4	Silty loam	4		
Basalt	9	Sand and gravel	8	Clay loam	3		
Karst limestone	10	Basalt	9	Muck	2		
		Karsts limestone	10	No shrinking clay	1		

studies indicate that groundwater is subjected to pollution in many areas due to different anthropogenic activities, such as the extensive fertilizes use and industrialization (Duan et al. 2015; Zhou et al. 2012). Given the above-mentioned considerations, this study is to propose a new methodology to modify DRASTIC model by integrating land use, AHP and GAs methods and then to evaluate the groundwater vulnerability in Jiangnan Plain.

### The study area

#### Location and climate

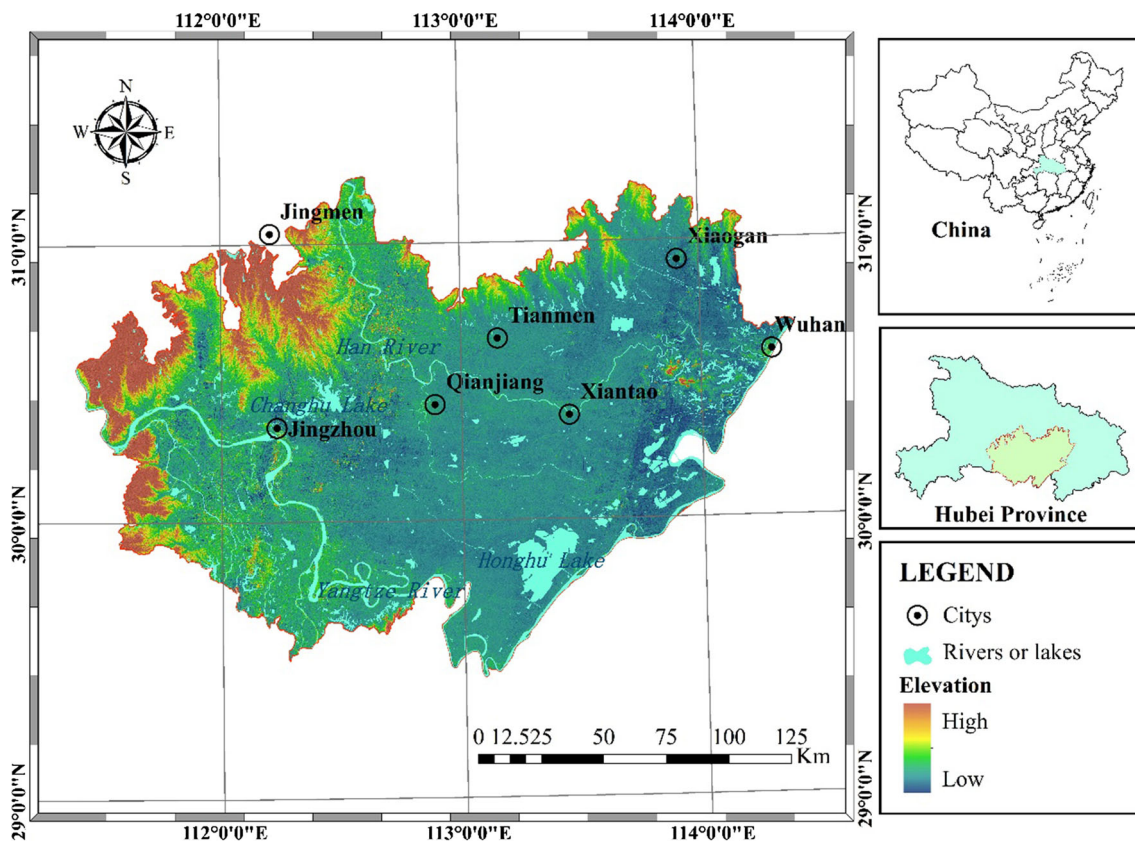
Jiangnan Plain, situated between latitude 29.4°–31.3°N and longitude 111.4°–114.5°E, is an alluvial plain deposited by the Yangtze and Han rivers. It is located in the central and southern regions of Hubei Province (Fig. 1). It is well known as “national base of fish and rice” in China, and agriculture is the main human activity in majority of the study area. Under a north subtropical monsoonal climate, the annual average temperature is 15–17 °C. The annual average precipitation is 1269 mm, increasing from 800 mm in the northwest to 1500 mm in the southeast, 30–50% of which concentrates in summer.

### Topography

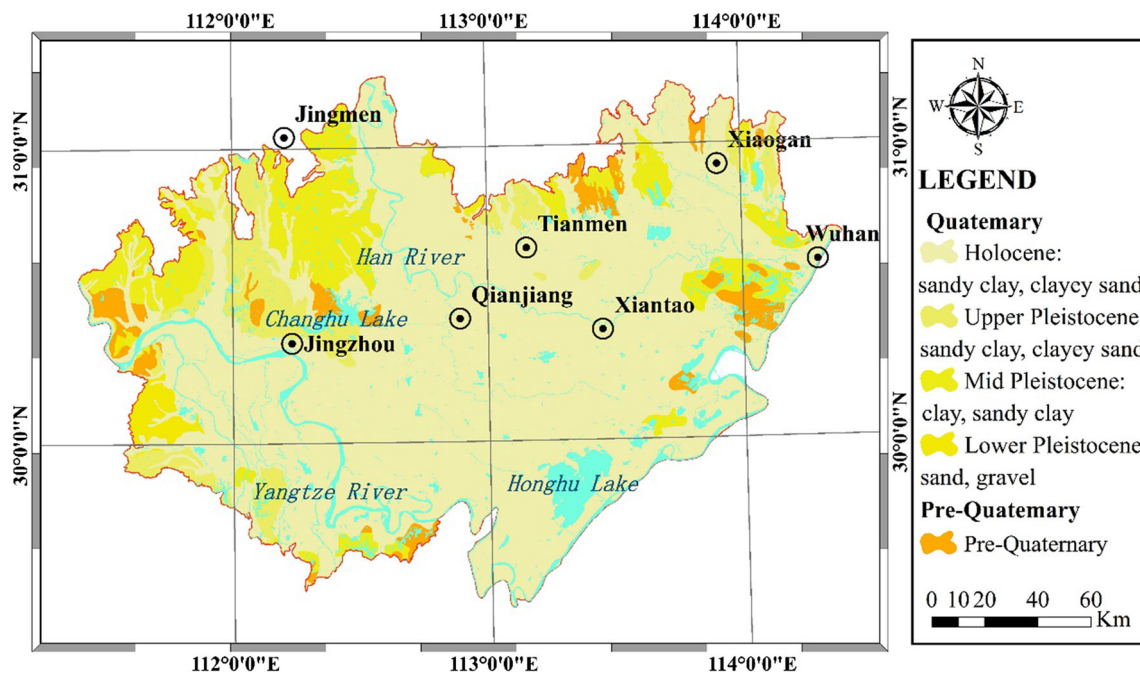
Jiangnan Plain is a semiclosed basin with a higher elevation in the northwest and a lower elevation in the southeast (Fig. 1). The middle region is a low alluvial plain with elevations ranging from 20 m in the southeast to 40 m in the northwest. The outlying areas mostly consist of two-level terraces with elevations of 40–80 m and 80–120 m, respectively. The outer area of the terraces is primarily composed of craggy terrain and low hills above 120 m. The slope ranges between 0° and 22° with an average value of 2°–3° in the central flat plain (Zhou et al. 2012).

### Geology

The geological map of study area was prepared based on lithological units, recent sediments, and landform (Fig. 2). The central plain is composed of Holocene alluvial–lacustrine and upper Pleistocene sediments. The thickness of alluvial–lacustrine sediments varies greatly in different regions with depth of 100–200 m in the center and 20–100 m in the outlying areas. The outlying hills, which have been suffered from erosion, are formed with porous and fissured media of mid Pleistocene, lower Pleistocene, and Pliocene sediments. The lithology of Jiangnan Plain is



**Fig. 1** Location map and digital elevation model of study area



**Fig. 2** Geological map of study area

mainly Quaternary lacustrine sediments, which are mainly composed of sandy clay, clayey sand and clay.

### Hydrogeology

The hydrogeological characteristics are closely related to the topography. The hilly areas primarily consist of aquitard, while the central areas consist of unconsolidated water-bearing sediment layers. These aquitard and unconsolidated layers constitute a complete hydrogeological unit. Based on the geological and hydrogeological settings, aquifer system can be subdivided into unconfined aquifer and confined aquifer (Gan et al. 2014). The unconfined aquifer is mainly composed of clay, clayey silt, and sandy silt. The depth to groundwater is usually 0.5–3.5 m below ground surface. There are two types of confined aquifer: the confined aquifer in Quaternary porous media and the confined aquifer in porous-fissured media of Pliocene and lower Pleistocene sediments. The former is mainly distributed in the lower plain, while the latter occurs in the hilly areas.

### Materials and methods

#### Data collection

Table 2 presents the data used in this study. The features are obtained from different sources to establish thematic layers.

#### Depth to groundwater ( $D$ )

Depth to groundwater represents the depth from the surface to the groundwater table. It determines the thickness of the unsaturated zone through which the infiltrating water must travel before reaching the aquifer. The depth to groundwater was measured in August 2014 from 786 wells in the study area. Besides, available data of 43 monitoring wells, which measured the groundwater level (depth to groundwater) every five days, were collected in August 2014. The minimum depth to groundwater occurs near August, and fertilizers are extensively used in this month (Zhang et al. 2014a). Therefore, this month is selected to consider the worst possible case scenario. A log-normality distribution of the depth to groundwater is confirmed by the Kolmogorov–Smirnov test performed in SPSS 20.0 software. Kriging interpolation is then employed to estimate the distribution of depth to groundwater (Fig. 3a).

#### Net recharge ( $R$ )

Net recharge is the amount of water that penetrates into the aquifer from the surface. It is the principal vehicle that transports the contaminants vertically to the groundwater and horizontally within the aquifer. Net recharge is not only the results of rainfall and irrigation return flow, but also the results of river and lakes recharge (Guan et al. 2016). However, the current study focuses on the vulnerability potential to pollution from the surface, particularly by the human activities. Therefore, the river and lakes

**Table 2** Data and its sources used in this study

Data	Sources
Depth to groundwater	Field measured and monitoring wells
Net recharge	Wuhan Center of China Geological Survey
Aquifer media	Hubei Institute of Hydrogeology and Engineering Geology (Chen et al. 1985)
Soil media	Previous literature (Zhao et al. 2007)
Topography	DEM-based from Chinese Academy of Sciences
Impact of vadose zone	Drilling logs
Hydraulic conductivity	Pumping test results
Land use	TM remote sensing images using supervised classification
Nitrate measurements	Field sampling

recharge is not considered in this study. The net recharge map is obtained based on the rainfall and the irrigation return flow data and calculated by:

$$\text{Net recharge} = \text{rain fall} \times \text{recharge rate} + \text{irrigation return flow} \quad (1)$$

The rain fall map is obtained by interpolating a 20-year mean of annual precipitation (mm/year) from seven representative rainfall stations (Wuhan, Jingzhou, Jingmen, Qianjiang, Tianmen, Xiantao, and Xiaogan). The recharge rate and irrigation return flow maps are obtained from Wuhan Center of China Geological Survey. The net recharge layer is presented in Fig. 3b.

*Aquifer media (A)*

Aquifer media is an important factor controlling groundwater flow path. A weakly permeable aquifer with relatively low recharge rates may result in a low vulnerability, while a highly permeable aquifer with great recharge potential may lead to a high vulnerability. Classification of this parameter is based on a report obtained from Hubei Institute of Hydrogeology and Engineering Geology (Chen et al. 1985) (Fig. 3c).

*Soil media (S)*

Soil media is identified as the uppermost part of the vadose zone, and it influences the infiltrating process of contamination (Baalousha 2010). The soil media map is derived from the previous studies (Zhao et al. 2007) and presented in Fig. 3d.

*Topography (T)*

Topography in DRASTIC model refers to the slope of the land surface and its variation. It controls the runoff and determines the residence time of pollutants after precipitation. Besides, the topography has a significant effect on groundwater flow rate and current direction (Al-Hanbali

and Kondoh 2008). The topography of study area is categorized into five groups according to the digital elevation model (DEM) provided by Chinese Academy of Sciences using ‘‘Slope’’ tool in ArcGIS 10.0 software(Fig. 3e).

*Impact of vadose zone (I)*

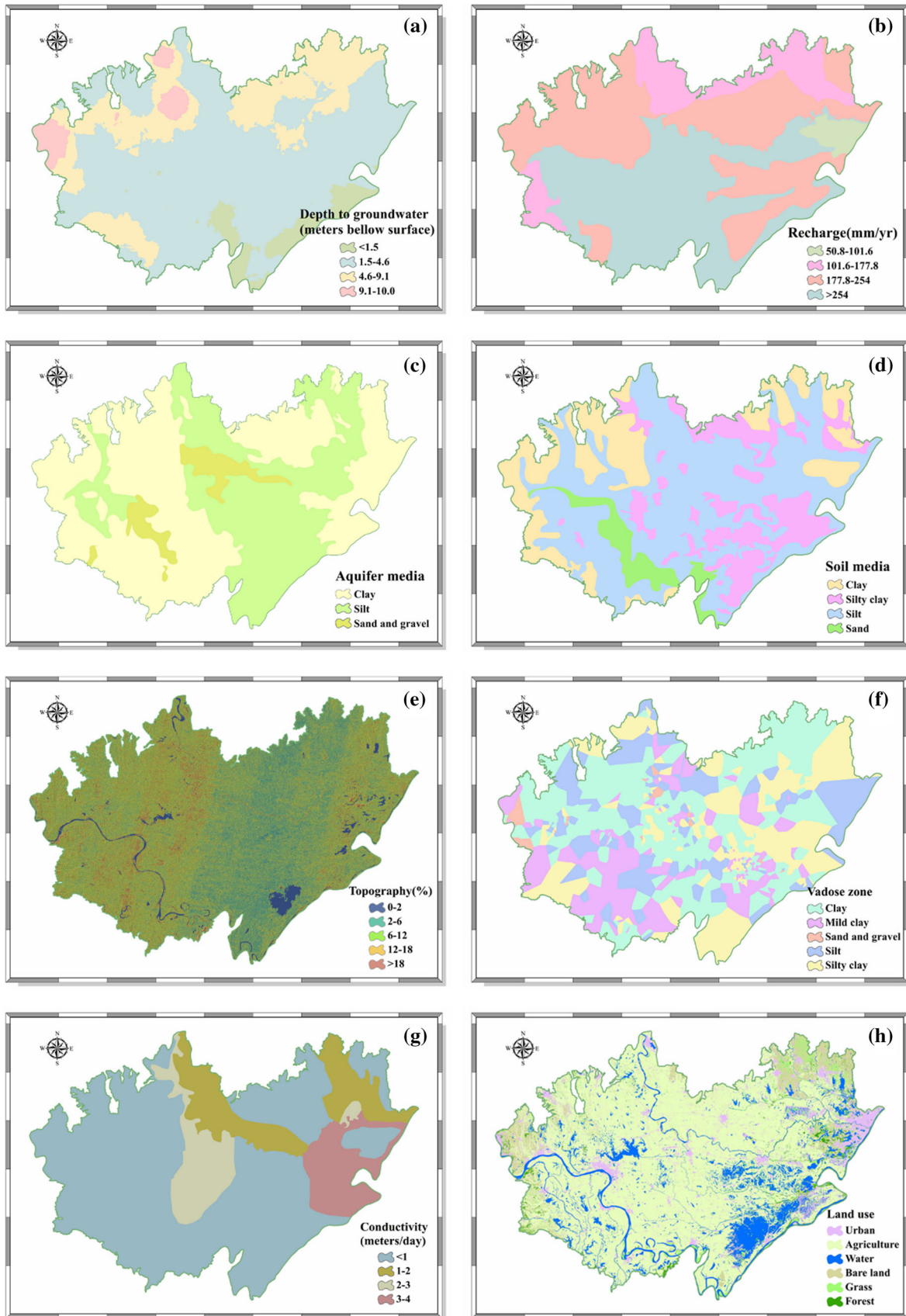
The vadose zone, also refers to the unsaturated zone, is between the land surface and the top of the phreatic zone. It controls the length or duration of hydrologic flow path in controlling the contaminant delivery. Based on 1358 boreholes available information data collected in the study area, the impact of vadose zone layer is generated as shown in Fig. 3f. The drilling logs show a large complexity and variety of the vadose medium which indicated a multilayer vadose system in Jiangnan Plain. The thickest medium from each drilling logs upper the groundwater level is expected to be defined as the vadose material. Thiessen polygons are then constructed so that the centroid of vadose medium from each well is assigned to an area.

*Hydraulic conductivity (C)*

The hydraulic conductivity reflects the ability of aquifer materials to transmit water, which in turn controls the rate at which contaminants will flow with groundwater under a given hydraulic gradient (Neshat et al. 2014). The hydraulic conductivity distribution map is generated using the pumping test results of the study area. Kriging interpolation algorithm is applied to interpolate the hydraulic conductivity and to create the raster layer as shown in Fig. 3g.

*Land use (L)*

Landsat Thematic Mapper (TM) satellite images of August 2014 are used to produce the land use map. TM satellite images are widely applied for land use classification owing to their relatively high spatial resolution (Liu et al. 2014). Supervised classification method with maximum likelihood



**Fig. 3** Geodata layers involving: **a** Depth to groundwater. **b** Net recharge. **c** Aquifer media. **d** Soil media. **e** Topography. **f** Impact of vadose zone. **g** Hydraulic conductivity. **h** land use

clustering technique is employed to classify the land use into six classes: urban, agriculture, water, bare land, grass, and forest, as illustrated in Fig. 3h.

*Nitrate measurements*

The nitrate concentration (expressed as mg/L NO<sub>3</sub>-N) associated with groundwater was measured in August 2014 from 97 wells distributed in the study area (Fig. 4). Outliers are detected using the Z-scores (Haddad et al. 2015). The Z-scores of the observations could be calculated as:

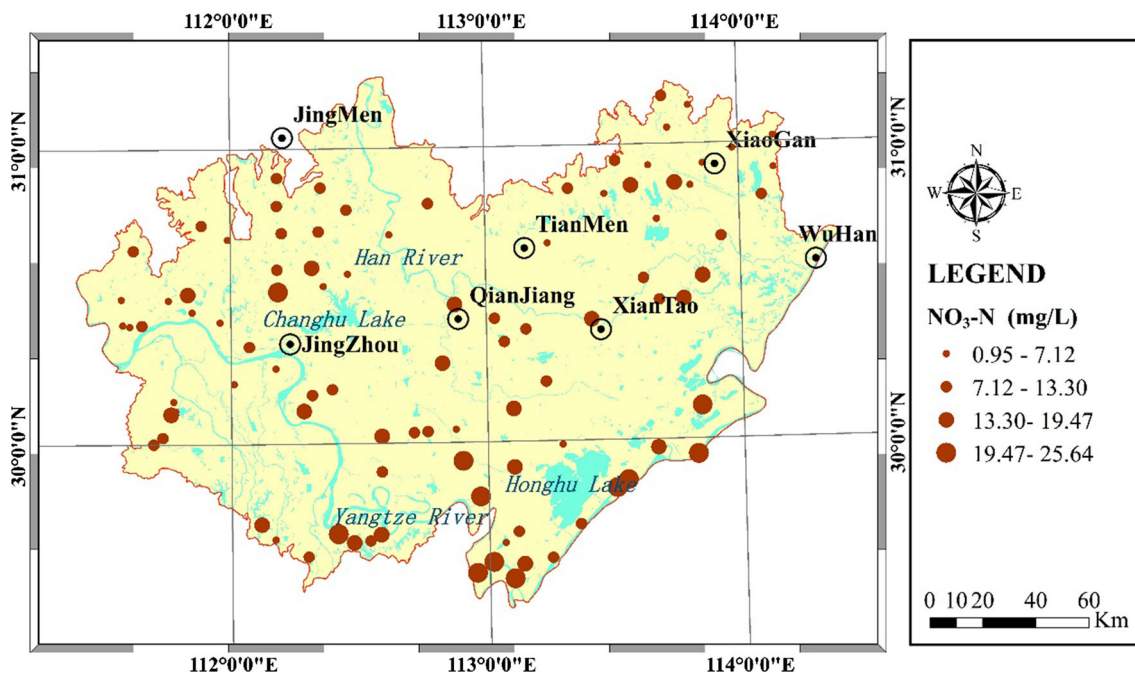
$$Z_i = \frac{x_i - \bar{x}}{s}, \quad \text{with } s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2)$$

where  $x_i$  is the observation value,  $\bar{x}$  is the mean, and  $s$  is the standard deviation. The nitrate concentration values obtained by the laboratory are considered acceptable for Z-scores between -2 and 2 (Chakravarty et al. 2013). The distribution of Z-scores and the corresponding nitrate concentration for the groundwater samples is shown in Fig. 5. As can be seen, two samples are rejected due to the Z-scores felt outside the acceptable limit of ±2.

**Description of the proposed methodology**

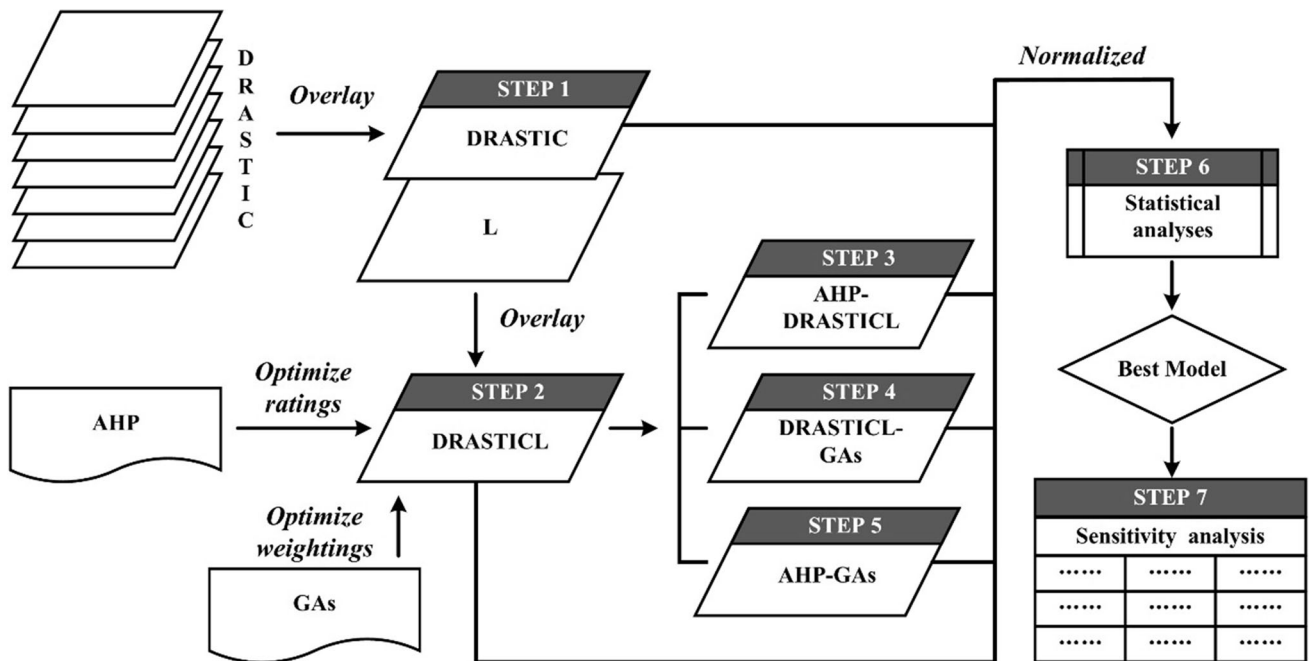
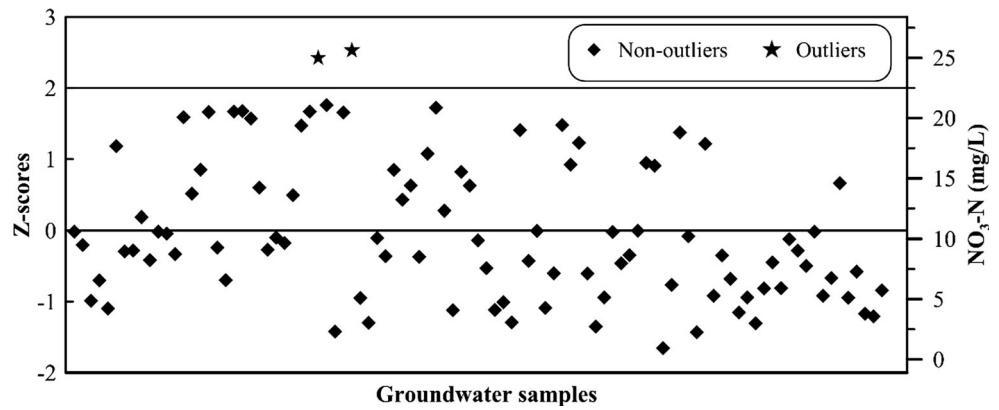
Figure 6 illustrates the methodological flowchart. A combination of land use, AHP and GAs methods is proposed to modify the DRASTIC model to assess the groundwater vulnerability. Detailed steps are described as follows:

1. The conventional DRASTIC method is applied (DRASTIC model).
2. The land use map is overlapped with the resultant DRASTIC model (DRASTICL model).
3. PCMs are used to develop the decision hierarchy structures and determine the priorities of ranges for parameters. New ratings are determined by AHP method (AHP-DRASTICL model).
4. A GAs program is developed to re-ascertained the weightings of the eight factors (DRASTICL-GAS model).
5. AHP and GAs methods are coupled to generate a new technique (AHP-GAS model).
6. Statistical analyses including Pearson’s correlation coefficient and analysis of variance (ANOVA) *F* statistic are adopted to verify the vulnerability mapping by taking nitrate as the typical pollutant.
7. The single-parameter sensitivity analysis is performed to analyze the “effective weights” of the parameters against their theoretical weights.



**Fig. 4** Nitrate sampling locations in the study area

**Fig. 5** Z-scores and NO<sub>3</sub>-N concentration for groundwater samples



**Fig. 6** Flowchart of proposed methodology

**Groundwater vulnerability assessment**

*DRASTIC model*

DRASTIC model involves the following seven hydrogeological features: D (Depth to groundwater), R (Net Recharge), A (Aquifer media), S (Soil media), T (Topography), I (Impact of vadose zone), and C (Hydraulic conductivity). The model calculates an intrinsic vulnerability index with different weighting factors based on the following equation:

$$\text{DRASTIC Index} = \sum_{i=1}^7 (W_i \times R_i) \tag{3}$$

where  $W_i$  is the weighting for parameter  $i$  with an associated rating of  $R_i$ . Each parameter is rated from 1 to 10, a

rating of 10 indicating a highest potential to the pollution, and each parameter is weighted from 1 to 5 using a Delphi (consensus) method to express their relative importance with respect to each other. The ratings and weightings of Aller et al. (1987) are shown in Table 1.

*Development of DRASTIC model*

*Incorporating the land use to DRASTIC model* Land use is the primary factor causing habitat degradation and poor water quality. It can affect the groundwater quality and increase the pollution risk, especially in agriculture areas. For this reason, it should be considered and incorporated into the DRASTIC vulnerability index when evaluating the groundwater vulnerability. The new model can be named DRASTICL by acronyms of the eight factors. Then by



converting the land use map into a raster image to overlay the DRASTIC index, the results of DRASTICL model are expressed as the following equation:

$$\text{DRASTICL Index} = \text{DRASTIC Index} + L_w \times L_r \quad (4)$$

where  $L_w$  and  $L_r$  correspond to the weighting and rating of land use factor, respectively. The land use factor is given a weighting of 5, due to the potential impact of this parameter on groundwater. Detailed ratings are given in Table 3, as illustrated in previous studies (Shirazi et al. 2013).

*Correcting the ratings using AHP method* AHP, developed by Saaty (1980), is often used to solve strategic decision problems. This method is based on the construction of a series of PCMs. By ranking the importance of the criteria, a set of PCMs is established to calculate the relative weights. The consequence of each PCMs could be summarized in a square matrix, in which each element ranges from 1/9 (absolute unimportance) to 9 (absolute importance) as described in Table 4. Priorities are then computed by normalizing each column of the matrix, to derive the normalized primary right eigenvector, the relative weight vector, by following equation:

$$A \times w = \lambda_{\max} \times w \quad (5)$$

where  $A$  is the PCM;  $w$  is the principal eigenvector;  $\lambda_{\max}$  is the largest eigenvalue of matrix  $A$ . The consistency index (CI) is calculated to determine the quality of the result of the AHP. The formula for the CI is as follows:

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (6)$$

where  $n$  is the dimension of matrix  $A$ . CR could be checked based on CI, by:

$$CR = \frac{CI}{RI} \quad (7)$$

where RI is the ratio index (Lin and Yang 1996). As a general rule, the judgment matrix is acceptable if  $CR \leq 0.1$ . In contrast, if CR is greater than 0.1, the evaluation procedures have to be reviewed and reconsidered.

In this study, AHP method is employed to correct the initial ratings participating in DRASTICL model. To improve the accuracy and objectivity in decision-making,

three hydrogeologists are invited to fill the PCMs to evaluate the ratings of corresponding eight parameters. The experts are supposed to specify their judgments of the relative importance of each class by asking questions like “with respect to parameter  $x$ , how much more important is class  $a$  to class  $b$ ?”. Not only normal comparison values, that is, 1,2,...,9 and 1/2,1/3,...,1/9, are used to constructed the PCMs, but also non-normal comparison values such as 1.1, 2.3,...,9.6 and 1/1.1, 1/2.3,...,1/9.6 are used when experts have conflict opinions. Table 5 reports an example of the inputs of the PCMs used to derive the optimized ratings for the land use parameter. The  $RI = 1.24$  and the  $CR = 0.06 < 0.10$ , which indicates that the judgment matrix passes the consistency check. The weights produced from the AHP method ranges from 0.04 to 0.33, where 0.04 denotes the least susceptible and 0.33 denotes the most susceptible to groundwater vulnerability. The same assessments are applied to compare alternatives with respect to the other parameters. The final alternative ratings derived using AHP method are shown in Table 6.

*Optimizing the weightings using GAs method* GAs is an iterative method based on the process of genetic selection and natural elimination in biological evolution. GAs handles a population of possible solution to optimize problems using techniques inspired by natural evolution, such as selection, crossover, and mutation. It is a useful tool for searching and optimization problems. Candidate solutions are retained and ranked for the each iteration according to their eligibility. A fitness function is used to remove the unqualified solutions. The algorithms are stopped when the termination condition is met. In GAs, the design variable is a string of numbers, usually called chromosome, which represents the solution. The most common way of encoding is a binary string with 1 and 0 s. The length of the string depends on the accuracy. In this study, Python 2.6 software is utilized to establish the GAs program.

The goal of this procedure is to find the optimal weightings of the eight parameters which maximize, as the objective function, the Pearson’s correlation coefficient between the vulnerability index and nitrate concentration. The algorithm is founded on a population size of 100 chromosomes, each one of which consists of 80 bits string (every 10 bits

**Table 3** Description and rating for each land use class

Land use	Rating	Description
Urban	8	All residential, commercial, and industrial areas, including village settlement
Agriculture	8	Land used for cultivation, such as paddy fields and corn fields
Water	4	All areas of open water, including rivers, lakes, fishponds, and reservoirs
Bare land	3	Unused land, including barren land, waste land, and alkaline land
Grass	2	Areas for wild grass ground
Forest	1	Rangeland with trees, shrubs, and vines (natural and artificial forest)

string coding one of the eight weighting values) and 200 generation steps. Table 7 represents the GAs features used in this study. The GAs would be stopped once the maximum generation number has evolved or if there is no change to population's best fitness for 50 iterations.

**Table 4** Scale for PCMs used in AHP (Saaty 1980)

Relative importance	Definition
1	Equally important
2	Equally to slightly important
3	Slightly important
4	Slightly to fundamentally important
5	Fundamentally important
6	Fundamentally to really important
7	Really important
8	Really to absolutely important
9	Absolutely important
Reciprocals of above numbers	If criteria <i>i</i> has one of the above numbers designated to it when compared with criteria <i>j</i> , then <i>j</i> has the reciprocal value when compared with <i>i</i> .

**Combining AHP and GAs methods** AHP method is proposed to correct the ratings of eight parameters, while GAs method is proposed to produce an optimal weighting combination. Additionally, AHP and GAs methods are combined to generate a new technique. The eight factors are reclassified according to the modified ratings optimized by AHP method. Then the GAs program is executed again to find new solutions by using the same features as shown in Table 7.

**Normalization of vulnerability index**

Given the different ranges of the vulnerability index obtained from these models, for example, the DRASTICL index is absolutely higher than the DRASTIC index for a result of direct addition, and the relativity of groundwater vulnerability index, unity-based normalization process is employed for presenting the vulnerability maps based on:

$$\text{Normalized Index} = \frac{V - V_{\min}}{V_{\max} - V_{\min}} \tag{8}$$

where *V* is the vulnerability index, *V<sub>min</sub>* is the minimum vulnerability index, and *V<sub>max</sub>* is the maximum vulnerability

**Table 5** PCM for land use parameter

Land use	Urban	Agriculture	Water	Bare land	Grass	Forest	Criteria weight
Urban	1	1	4	2	5	9	0.33
Agriculture	1	1	2	4	4	5	0.30
Water	1/4	1/2	1	1/2	2	3	0.11
Bare land	1/2	1/4	2	1	2.5	2	0.15
Grass	1/5	1/4	1/2	1/2.5	1	3	0.07
Forest	1/9	1/5	1/3	1/2	1/3	1	0.04

**Table 6** Alternative ratings optimized by AHP method

Depth (m)		Recharge (mm/year)		Topography (%)		Conductivity (m/d)	
Range	Rating	Range	Rating	Range	Rating	Range	Rating
<1.5	0.32	50.8–101.6	0.08	0–2	0.32	<1	0.08
1.5–4.6	0.26	101.6–177.8	0.22	2–6	0.30	1–2	0.12
4.6–9.1	0.24	177.8–254	0.34	6–12	0.20	2–3	0.26
9.1–10.0	0.18	>254	0.36	12–18	0.14	3–4	0.54
				>18	0.04		

Aquifer media		Impact of vadose zone		Soil media		Land use	
Range	Rating	Range	Rating	Range	Rating	Range	Rating
Clay	0.20	Clay	0.12	Clay	0.08	Urban	0.33
Silt	0.22	Mild clay	0.16	Silty clay	0.10	Agriculture	0.30
Sand and gravel	0.58	Silty clay	0.18	Silt	0.12	Water	0.11
		Silt	0.24	Sand	0.70	Bare land	0.15
		Sand and gravel	0.30			Grass	0.07
						Forest	0.04

**Table 7** GAs features used in this study

Number of chromosomes in the population	100
Number of generations	200
Natural selection rate	50%
Selection	Standard roulette
Crossover	Single-point crossover
Crossover probability	0.6
Mutation probability	0.01

index. The normalized vulnerability index ranging from 0 (lowest potential to contamination) to 1 (highest potential to contamination) is then divided into four groups with the equal interval: very low (<0.25), low (0.25–0.50), high (0.5–0.75), very high (0.75–1). It is helpful in comparing the vulnerability maps produced by different models.

**Model validation**

Nitrate, one of the major contaminant from anthropogenic activities, is not usually found in nature groundwater (Baalousha 2010). With regard to the fact that this contaminant is highly risky to human health and it mainly comes from the nitrogen fertilizers, which are frequently used in agricultural areas to enhance the crop production in the study area, the nitrate concentration is selected as an indicator of initial contamination. Statistical analyses including Pearson’s correlation coefficient and ANOVA *F* statistic are performed to validate the vulnerability results of different models.

*Pearson’s correlation*

Pearson’s correlation analysis is investigated to check the degree of association between the vulnerability index and nitrate concentration. The sampling locations map is overlaid on the vulnerability index maps, and the corresponding value for each point is extracted within GIS environment. The attribute file of the sampling locations is updated with the data from vulnerability index map based on the spatial relationship between the features. Then Pearson’s correlation coefficient could be calculated by:

$$\rho = \frac{\text{cov}(\mathbf{X}, \mathbf{Y})}{\sigma_X \sigma_Y} \tag{9}$$

where  $\rho$  is the correlation coefficient, cov is the covariance, and  $\sigma$  is the standard deviation.

*ANOVA F statistic*

*F* statistic is the famous statistic for the ANOVA to compare the means of samples from different levels. Larger values of *F* reject the null hypothesis that the means are

equal (Montgomery 2008). The larger analysis of ANOVA *F* statistic is, the less overlap there is between the nitrate values in different vulnerability classes (Huan et al. 2012). ANOVA *F* statistic can be calculated by:

$$F = \frac{SST/k - 1}{SSE/n - k} \tag{10}$$

where SST and SSE are, respectively, the sum of squares for treatment and the sum of squares for error.  $k - 1$  and  $n - k$  are, respectively, the freedom degree for treatment and freedom degree for error.

**Sensitivity analysis**

Subjectivity is inevitable in the selection of ratings and weightings in all parametric techniques, which can strongly affect the vulnerability assessment. Sensitivity analysis could provide valuable information on the contribution of ratings and weightings of input parameters and could help analyst to judge the significance of subjective elements. There are two types of sensitivity analysis: the map removal sensitivity analysis and the single-parameter sensitivity analysis (Gogu and Dassargues 2000). In this study, single-parameter sensitivity analysis is conducted to evaluate “effective weight” of each parameter.

Single-parameter sensitivity analysis checks the spatial significance of the parameters in the index computation. The effective weight could be defined as follows:

$$W_{pi} = (P_{Ri} \times P_{wi})/V_i \times 100\% \tag{11}$$

where  $W_{pi}$  is the effective weight for each unique condition subarea *i*,  $V_i$  is the vulnerability index,  $P_{Ri}$  and  $P_{wi}$  are the rating values and weighting values of parameter *P* assigned to subarea *i*. A “unique condition subarea” is one or more polygons in the vulnerability map with a unique combination of rating values of the factors used to compute the vulnerability index (Napolitano and Fabbri 1996).

**Results and discussion**

**Application of DRASTIC model**

The assigned layers for the seven DRASTIC parameters are constructed in raster format with a pixel size of 30 m, which are briefly discussed.

The maximum depth to groundwater is found in the northwest of the study area. Depth generally increases from southeast to northwest. According to Aller et al. (1987), ratings of 10, 9, 7, and 5 are assigned to the corresponding four classes, respectively.

The net recharge, calculated by Eq. (1), varies between 92.8 and 326.8 mm/year. Net recharge in 40% of study

area is greater than 254 mm/year. DRASTIC standard ratings of 3, 6, 8, 9 are assigned to the corresponding four classes.

As for the aquifer media of study area, most parts are deposited with clay. Deposits of silt are mainly located in the middle region of the study area. Deposits of sand and gravel contribute to a very low percentage of the study area. The higher permeability of the aquifer, the potential for pollution is greater. In this study, rating of 8 is assigned for the aquifer media type of sand and gravel, 5 for silt, and 3 for clay.

Based on the previous studies, types of soil media mainly consist of clay, silty clay, silt, and sand. Silt and clay could be found in most part of the study area. Sand with high permeability is assigned a rate of 9 following the DRASTIC classification. Rating of 3 is attributed for silt, 2 for silty clay, and 1 for clay.

The topography is divided into five classes according to Aller et al. (1987), and ratings from 1 to 10 are assigned to the different classes.

The vadose zone of the study area is subdivided into five groups: clay, mild clay, silty clay, silt, and sand and gravel, with ratings of 1, 2, 3, 4, and 5, respectively.

The hydraulic conductivity of the study area is smaller than 4.1 m/d. According to Aller et al. (1987), rating of 1 would be assigned to all subareas. However, it is meaningless because the information of this factor would not be completely reflected. A simple modification is proposed relating to the rating values of hydraulic conductivity: 1 for conductivity <1 m/d, 3 for conductivity 1–2 m/d, 5 for conductivity 2–3 m/d, and 7 for conductivity 3–4 m/d.

Figure 7a and Fig. 8a present the normalized DRASTIC vulnerability index and the corresponding percentages, respectively. Obviously, the normalized vulnerability index ranges from 0 to 1 for the whole plain. The DRASTIC results illustrate that 3.98% of the total area has “very high” vulnerability, 45.09% has the “high” vulnerability, and more than half (50.93%) has the “low” and “very low” vulnerability.

## Development of DRASTIC model

### *DRASTICL model*

Figure 7b shows the normalized vulnerability map for DRASTICL model. The difference between the DRASTICL and DRASTIC results is obvious. Figure 8b illustrates that 8.80% of the area is under a “very high” contamination risk, and 57.73% is covered by “high” class of vulnerability to pollution. The “low” and “very low” classes are at 27.66 and 5.81%, respectively.

### *AHP–DRASTICL model*

Figure 7c presents the normalized vulnerability map for AHP–DRASTICL model. Figure 8c indicates that the percentage for “very high” and “high” classes is at 34.87%, while the percentage for “low” and “very low” classes is at 65.13%.

### *DRASTICL–GAs model*

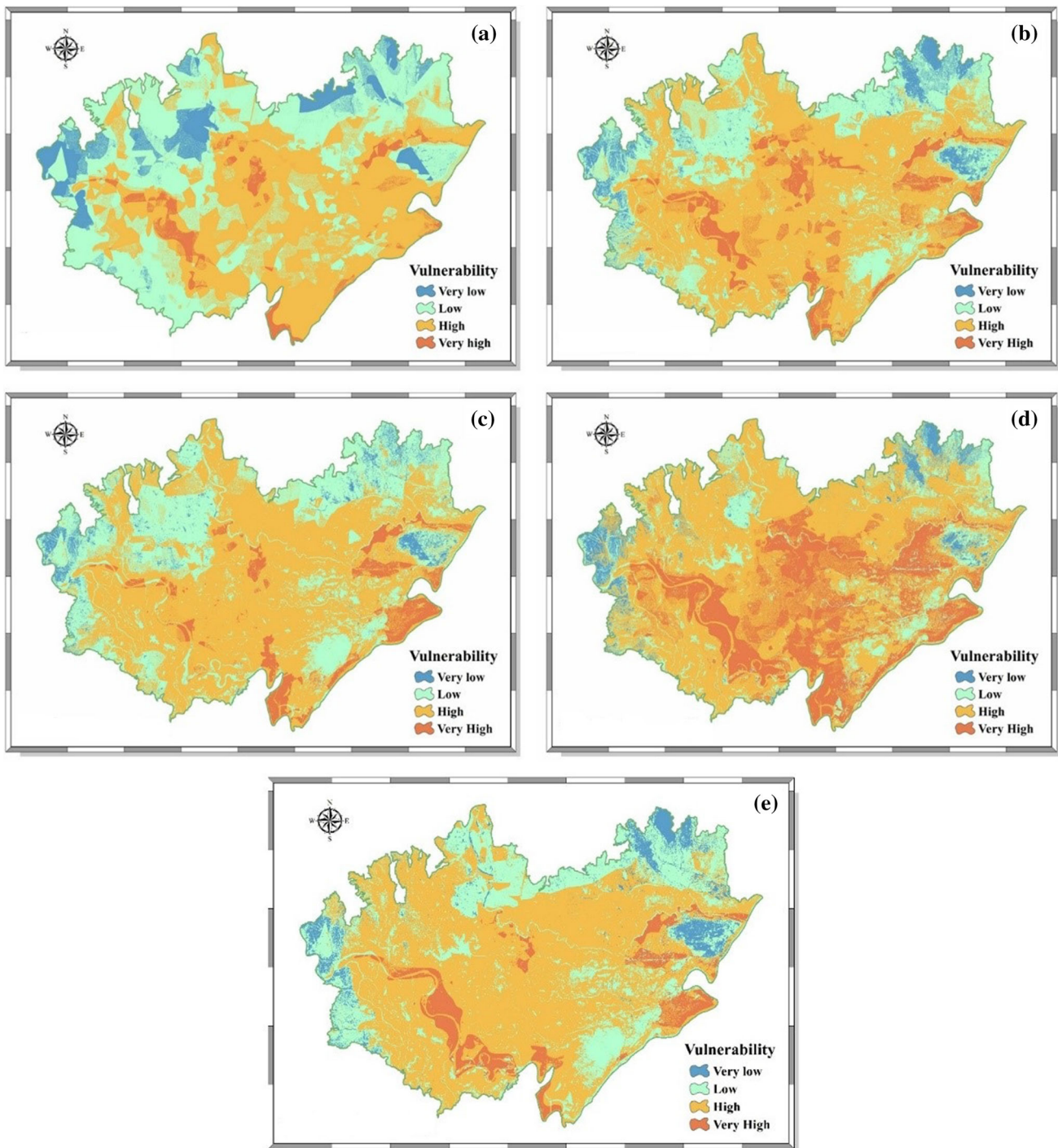
By taking Pearson’s correlation coefficient between the vulnerability index and the nitrate concentration as the objective function and the weightings of the eight parameters as the optimal variables, the GAs stopped when the highest value of fitness function reached 58.17%. After decoding the corresponding weighting values, the vulnerability mapping procedure was carried out. Figure 7d and Fig. 8d illustrate the vulnerability map and percentages of the areas for DRASTICL–GAs model, respectively. As shown in Fig. 8d, the percentages for “very high” and “high” classes are at 24.05 and 55.53%, respectively. 4.76% of the total area is covered with “very low” class of vulnerability to pollution, and 15.66% is covered with “low” class.

### *AHP–GAs model*

The GAs stopped when the highest value of fitness function reached 75.31%. The results of AHP–GAs model indicate that 7.11% of the study area belongs to a “very high” class. The percentage for the “high” class is 60.16%. The “low” and “very low” classes are 26.81 and 5.92%, respectively (Fig. 8e). The vulnerability map illustrates the “high” and “very high” vulnerability zones mostly occur in the central Jiangnan Plain, particularly along the Yangtze River (Fig. 7e). There are mainly farmlands where fertilizers are extensive used. In addition, the depth to groundwater is shallower in these areas, therefore, causing a higher groundwater pollution risk.

## Validation of the models

Table 8 reports the validation results of the five models. The mean nitrate concentration values associated with the four classes are calculated for different models. As can be seen, the mean values increase from low vulnerability to high vulnerability. ANOVA *F* statistic shows that there is a markedly significant difference across vulnerability classes for all models ( $P < 0.01$ ). In other words, the null hypothesis of equal means should be clearly rejected and there is a significant difference in means between different classes. High nitrate concentration could be observed in the high groundwater vulnerability areas and vice versa. The



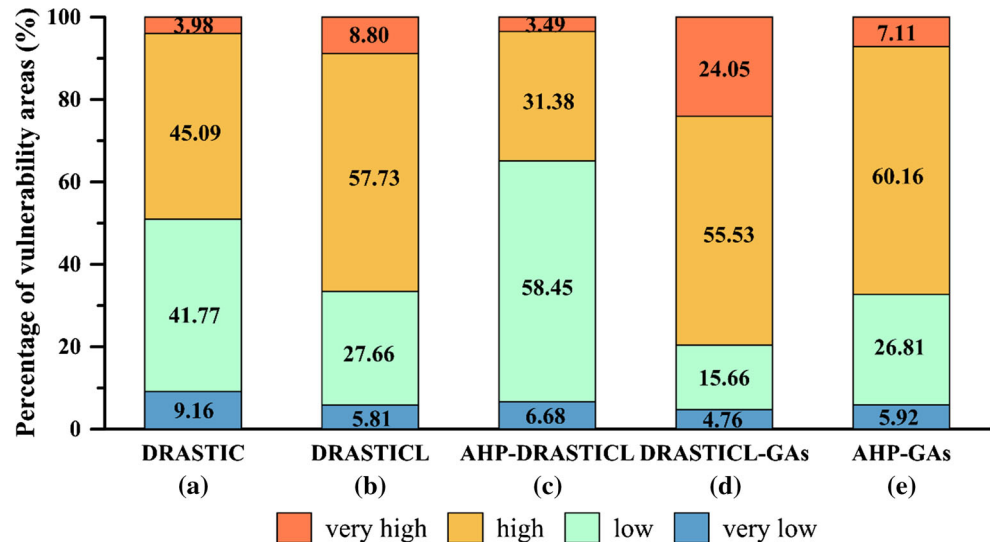
**Fig. 7** Vulnerability maps: **a** DRASTIC. **b** DRASTICL. **c** AHP-DRASTICL. **d** DRASTICL-GAs. **e** AHP-GAs

results demonstrate that the vulnerability assessment is an efficient approach to evaluate the groundwater vulnerability.

In most case, correlation of vulnerability results with actual pollution occurrence is a technique for validating the accuracy of groundwater vulnerability mapping (Huan et al. 2012). It is important to note that a good performance

on only a single contaminant may involve uncertainties. However, considering the extensive agricultural activities and the immoderate use of nitrogen fertilizers in the study area, the nitrate concentration is taken as the primary indicator to validate the accuracy of the vulnerability mapping. Pearson’s correlation value is calculated at 41.07% for the original DRASTIC model. This value is

**Fig. 8** Percentage of vulnerability areas:  
**a** DRASTIC. **b** DRASTICL.  
**c** AHP–DRASTICL.  
**d** DRASTICL–GAs. **e** AHP–GAs



**Table 8** Validation of groundwater vulnerability mapping

Model	Mean NO <sub>3</sub> –N concentration (mg/L) across vulnerability classes (from low to high)	Pearson's correlation coefficient (%)	ANOVA <i>F</i> statistic
DRASTIC	7.84, 9.19, 11.76, 17.30	41.07	4.07*
DRASTICL	4.29, 8.34, 10.92, 18.91	55.60	12.47*
AHP–DRASTICL	4.30, 9.10, 13.14, 19.84	58.05	8.79*
DRASTICL–GAs	4.29, 5.99, 9.83, 14.03	59.83	10.60*
AHP–GAs	4.30, 7.22, 11.66, 18.22	75.31	13.10*

\*  $P < 0.01$

relatively low because the original model does not consider the regional hydrogeological conditions and land use factor. The DRASTIC model needs to be improved to reflect the actual groundwater vulnerability. The correlation coefficient increases to 55.60% when the land use map is introduced. The results indicate that the land use factor has significant effect on the groundwater vulnerability. Fertilizers used in agricultural, industrial activities, septic tanks, and sewer systems within urban areas are important sources of nitrate contamination. These hazards can absolutely increase the groundwater pollution risk. A correlation value of 58.05% is found for the AHP–DRASTICL model, and a correlation value of 59.83% is obtained for the DRASTICL–GAs model. These observations lead to conclusion that the vulnerability map could be more realistic by both optimizing the ratings and weightings. The Pearson's correlation value is calculated at 75.31% for the AHP–GAs model, indicating a strong correlation between the vulnerability index and the measured nitrate concentration. Generally, there is a considerable increase in the correlation coefficient after each modification and the accuracy improves over 30% compared to the original

model. Furthermore, it is noticed that the ANOVA *F* statistic increases to 12.47, which is twice larger than that obtained by the original DRASTIC model, when the land use factor is incorporated. AHP–GAs model shows the largest ANOVA *F* statistic of 13.10, which indicates that the least overlap exists between the nitrate concentration values in different vulnerability classes. Statistically, it can be argued that the modifications proposed in this study are effectiveness and the AHP–GAs model may be the most suitable model for study area.

### Single-parameter sensitivity analysis

A total of 9892 unique condition subareas are used to calculate the “effective weight” that each parameter has according to Eq. (11). Statistical analysis is performed to analyze and display the results as shown in Table 9. In particular, in Jiangnan Plain, the results reveal that the depth to groundwater mostly influences the vulnerability when comparing both theoretical and effective weights. The results confirm that land use with 19.86% of mean effective weight is a parameter that strongly affects the

**Table 9** Statistical analysis of the parameter sensitivity of unique condition subareas for AHP–GAs model

Parameter	Theoretical weight	Theoretical weight (%)	Mean (%)	SD (%)	Median (%)	Minimum value (%)	Maximum value (%)
D	10.00	35.20	39.28	6.75	38.82	23.22	65.36
R	5.00	17.60	23.87	5.60	23.77	5.52	37.96
A	1.36	4.79	5.40	2.42	4.63	2.95	16.56
S	1.54	5.42	3.15	2.57	2.57	1.44	17.95
T	0.05	0.17	0.15	0.08	0.15	0.02	0.40
I	1.10	3.87	3.10	1.01	2.91	1.44	7.38
C	1.86	6.55	5.19	4.19	3.11	1.52	20.66
L	7.50	26.40	19.86	10.33	17.56	4.26	40.89

vulnerability index. The most unexpected result, shown in Table 9, is that the net recharge has a much higher effective weight with an average value of 23.87% against the theoretical weight that is 17.60%. The topography seems to be of low importance in groundwater vulnerability since it has the lowest effective weight. The significance of depth to groundwater and net recharge as well as the land use highlights the importance of obtaining accurate, detailed, and representative information about these factors.

## Conclusion

DRASTIC model has been widely applied to evaluate the contamination risk of an aquifer. However, it may be not appropriate for accurate assessment of groundwater vulnerability due to the obvious drawbacks of ignoring the effect of local hydrogeological conditions. The present study resulted in a new DRASTIC-based model, new techniques to optimize the ratings and weightings of selected parameters. And then it was used to assess the groundwater vulnerability in Jiangnan Plain, China. The results showed that the modified model had a large improvement over the conventional DRASTIC method. 7.11% of the study area has “very high” pollution potential, 60.16% has “high” pollution potential, 26.81% has “low” pollution potential, and the remainder of the study area (5.92%) has “very low” risk for groundwater pollution. The depth to groundwater was detected as the most significant factor affecting the groundwater vulnerability with 39.28% of mean effective weight, followed by net recharge (23.81%) and land use (19.86%). It highlights the importance of obtaining accurate, detailed, and representative information for a more efficient interpretation of the vulnerability index about these factors. Detailed and frequent monitoring should be carried out for observing the changing level of pollutants in groundwater. Further investigations are also required in order to understand the relationship between the groundwater vulnerability and some other contaminants.

In conclusion, the proposed method could be considered as a promising tool in coordination with the original DRASTIC method for groundwater vulnerability assessment in other study areas globally.

**Acknowledgements** This work was supported by China Geological Survey (Nos. 1212011121142 and 1212011120084). The authors would like to thank Dr. Quanrong Wang and the three anonymous reviewers for their valuable suggestions which greatly improve the quality of the manuscript.

## References

- Al-Hanbali A, Kondoh A (2008) Groundwater vulnerability assessment and evaluation of human activity impact (HAI) within the Dead Sea groundwater basin, Jordan. *Hydrogeol J* 16:499–510
- Aller L, Bennett T, Lehr J, Petty R, Hackett G (1987) DRASTIC: a standardized system for evaluating groundwater pollution potential using hydrogeologic settings. EPA, Washington, EPA-600/2-87-035
- Baalousha H (2010) Assessment of a groundwater quality monitoring network using vulnerability mapping and geostatistics: A case study from Heretaunga Plains, New Zealand. *Agric Water Manag* 97:240–246
- Babiker IS, Mohamed MA, Hiyama T, Kato K (2005) A GIS-based DRASTIC model for assessing aquifer vulnerability in Kakamigahara Heights, Gifu Prefecture, central Japan. *Sci Total Environ* 345:127–140
- Chakravarty S, Mohanty A, Ghosh B, Tarafdar M, Aggarwal SG, Gupta PK (2013) Proficiency testing in chemical analysis of iron ore: comparison of statistical methods for outlier rejection. *Mapan* 29:87–95
- Chen J, Sun X, Ma T, Zhang W (1985) A preliminary study on environmental geology of cold spring paddy soil in Jinan Plain, Hubei Province (in Chinese). Hubei Institute of Hydrogeology and Engineering Geology, Wuhan
- Chitsazan M, Akhtari Y (2008) A GIS-based DRASTIC model for assessing aquifer vulnerability in Kherran Plain, Khuzestan, Iran. *Water Resour Manag* 23:1137–1155
- Duan Y, Gan Y, Wang Y, Deng Y, Guo X, Dong C (2015) Temporal variation of groundwater level and arsenic concentration at Jiangnan Plain, central China. *J Geochem Explor* 149:106–119
- Gan Y, Wang Y, Duan Y, Deng Y, Guo X, Ding X (2014) Hydrogeochemistry and arsenic contamination of groundwater in the Jiangnan Plain, central China. *J Geochem Explor* 138:81–93
- Giacobbo F, Marseguerra M, Zio E (2002) Solving the inverse problem of parameter estimation by genetic algorithms: the case

- of a groundwater contaminant transport model. *Ann Nucl Energy* 29:967–981
- Gogu RC, Dassargues A (2000) Sensitivity analysis for the EPIK method of vulnerability assessment in a small karstic aquifer, southern Belgium. *Hydrogeol J* 8:337–345
- Guan Y, Zhang L, Hu X, Xiang J, Wang F (2016) National survey and evaluation groundwater and environment in Jiangnan–Dongting Plain (in Chinese). Hubei Institute of Hydrogeology and Engineering Geology, Wuhan
- Haddad M, Hachemi H, Taibi H (2015) Assessment of Gravity Anomaly Surfaces (DTU10, EGM2008 and ITG-Goce02) in Western Mediterranean Sea. *Mediterr J Model Simul* 3:87–99
- Hernández-Espriú A et al (2014) The DRASTIC-Sg model: an extension to the DRASTIC approach for mapping groundwater vulnerability in aquifers subject to differential land subsidence, with application to Mexico City. *Hydrogeol J* 22:1469–1485
- Holland JH (1975) *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. University of Michigan Press, Michigan
- Huan H, Wang J, Teng Y (2012) Assessment and validation of groundwater vulnerability to nitrate based on a modified DRASTIC model: a case study in Jilin City of northeast China. *Sci Total Environ* 440:14–23
- Javadi S, Kavehkar N, Mohammadi K, Khodadadi A, Kahawita R (2011a) Calibrating DRASTIC using field measurements, sensitivity analysis and statistical methods to assess groundwater vulnerability. *Water Int* 36:719–732
- Javadi S, Kavehkar N, Mousavizadeh MH, Mohammadi K (2011b) Modification of DRASTIC model to map groundwater vulnerability to pollution using nitrate measurements in agricultural areas. *J Agric Sci Technol* 13:239–249
- Lin Z, Yang C (1996) Evaluation of machine selection by the AHP method. *J Mater Process Technol* 57:253–258
- Liu J, S-Y Wang, D-M Li (2014) The analysis of the impact of land-use changes on flood exposure of Wuhan in Yangtze River Basin, China. *Water Resour Manage* 28:2507–2522
- Montgomery DC (2008) *Design and analysis of experiments*, 7th edn. Wiley, New York
- Napolitano P, Fabbri AG (1996) Single-parameter sensitivity analysis for aquifer vulnerability assessment using DRASTIC and SINTACS. In: Paper presented at the HydroGIS 96: application of geographical information systems in hydrology and water resources management, Proceedings of Vienna conference, April 1996
- Neshat A, Pradhan B, Pirasteh S, Shafri HZM (2013) Estimating groundwater vulnerability to pollution using a modified DRASTIC model in the Kerman agricultural area, Iran. *Environ Earth Sci* 71:3119–3131
- Neshat A, Pradhan B, Dadras M (2014) Groundwater vulnerability assessment using an improved DRASTIC method in GIS. *Resour Conserv Recycl* 86:74–86
- Neukum C, Azzam R (2009) Quantitative assessment of intrinsic groundwater vulnerability to contamination using numerical simulations. *Sci Total Environ* 408:245–254
- Rahimi S, Shadman Roodposhti M, Ali Abbaspour R (2014) Using combined AHP–genetic algorithm in artificial groundwater recharge site selection of Gareh Bygone Plain, Iran. *Environ Earth Sci* 72:1979–1992
- Saaty TL (1980) *The analytic hierarchy process*. McGraw-Hill, New York
- Secunda S, Collin ML, Melloul AJ (1998) Groundwater vulnerability assessment using a composite model combining DRASTIC with extensive agricultural land use in Israel's Sharon region. *J Environ Manage* 54:39–57
- Sener E, Davraz A (2012) 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
- Sener E, Sener S, Davraz A (2009) Assessment of aquifer vulnerability based on GIS and DRASTIC methods: a case study of the Senirkent-Uluborlu Basin (Isparta, Turkey). *Hydrogeol J* 17:2023–2035
- Shirazi SM, Imran HM, Akib S, Yusop Z, Harun ZB (2013) Groundwater vulnerability assessment in the Melaka State of Malaysia using DRASTIC and GIS techniques. *Environ Earth Sci* 70:2293–2304
- Thirumalaivasan D, Karmegam M, Venugopal K (2003) AHP–DRASTIC: software for specific aquifer vulnerability assessment using DRASTIC model and GIS. *Environ Model Softw* 18:645–656
- Yalcin A, Reis S, Aydinoglu AC, Yomralioglu T (2011) A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. *Catena* 85:274–287
- Yin L, Zhang E, Wang X, Wenninger J, Dong J, Guo L, Huang J (2012) A GIS-based DRASTIC model for assessing groundwater vulnerability in the Ordos Plateau, China. *Environ Earth Sci* 69:171–185
- Zhang T, Chen S, Fu J (2014a) Analysis of three-nitrogen concentration and spatial-temporal distribution of groundwater in Sihui Region (in Chinese). *Resour Environ Yangtze Basin* 23:1295–1330
- Zhang W, Gao L, Jiao X, Yu J, Su X, Du S (2014b) Occurrence assessment of earth fissure based on genetic algorithms and artificial neural networks in Su-Xi-Chang land subsidence area, China. *Geosci J* 18:485–493
- Zhao D, Liu Z, Xiong Q (2007) The vulnerability evaluation for groundwater pollution in Jiangnan Plain (in Chinese). *Resour Environ Eng* 21:62–67
- Zhou Y, Wang Y, Li Y, Zwahlen F, Boillat J (2012) Hydrogeochemical characteristics of central Jiangnan Plain, China. *Environ Earth Sci* 68:765–778