ORIGINAL PAPER

Contribution of GIS tools and statistical approaches to optimize the DRASTIC model for groundwater vulnerability assessment in arid and semi-arid regions: the case of Sidi Bouzid shallow aquifer

Ikram Jmal[1] · Bachaer Ayed[1] · Moez Bahloul[2] · Emna Boughariou[1] · Salem Bouri[1]

Abstract

The groundwater vulnerability evaluation against pollution becomes an effective tool for sensible resource management and land use planning. These resources are threatened by the emanating from anthropogenic activities and the intensive fertilizer applications in agricultural sector. Over time, groundwater abstraction exceeds aquifer recharge, leading to the groundwater levels decline and constitutes a continuous and serious problem. Protecting groundwater resources from pollution has become necessary especially by the use of the vulnerability models. The objective of this study is to establish a groundwater vulnerability map in the Sidi Bouzid region using geographical information system (GIS) by overlaying the available hydrogeological data. As well, a comparative study between two vulnerability models has been carried out to conduct more realistic management of groundwater quality. In fact, groundwater vulnerability map to pollution was determined by applying the DRASTIC model. This model was almost applied, without any changes, in the Sidi Bouzid aquifer. In this case, the DRASTIC method was modified to estimate vulnerability and pollution risk of shallow aquifers especially for nitrate component. The modified DRASTIC model is based on some of statistical and geostatistical approaches in odor to amend the rates and the weights of DRASTIC parameters via GIS tools. This method has proved to be an effective way to achieve better correlation between nitrate concentration ($[NO_3^-]$) and vulnerability index. The nitrate content was analyzed for thirty-eight samples. The obtained Pearson's coefficients correlation showed that the vulnerability index and the $[NO_3^-]$ relationship was increased from 44 to 82% (before and after modification, respectively). This requires the use of the modified DRASTIC method. Statistical approaches, taking into account parameter classes for the categorical parameters, show a R^2 statistic equal to 75.4% for the total variance, explained by multiple linear regression (MLR) model, for the estimation of NO_3^-.

Keywords Factorial analysis of correspondences · Water management · Vulnerability map · Multiple linear regression · Nitrate concentration · Pearson correlation · Water quality

✉ Ikram Jmal
  ikramjmal@hotmail.fr

1 LR3E, National Engineering School, University of Sfax, Sfax, Tunisia

2 LASED, Preparatory Institute for Engineering Studies of Sfax, Sfax, Tunisia

Introduction

Groundwater is a major source of drinking water available from the natural resources on the earth. Water demand for human, industrial and agricultural activities does not cease to increase in the Sidi Bouzid region, located in the Center of Tunisia, considered an arid region. According to Castellazzi and Schmid (2021), if groundwater abstraction exceeds groundwater recharge for extensive areas, groundwater levels show a significant decline occurring groundwater depletion. Unfortunately, water shortages due to the excessive use and continued mismanagement of water resources have long been a critical constraint to its future development.

Therefore, early and respectable groundwater resource management is essential to preserve the natural resource quality. In fact, these problems have been conducted to increase interest in determining the key factors inducing the vulnerability of groundwater resources to anthropogenic factors. The term "vulnerability" refers to a state of change where there is a risk or possibility of any harmful effect on society (Pal et al. 2019).

The concept of groundwater vulnerability was proposed for decades, and the vulnerability maps are critical references in decision-making for land-use planning, groundwater management, monitoring, and remediation (Wang et al. 2012). The groundwater vulnerability notion was initially introduced by J. Margat in 1960 (Albinet and Margat 1970), who stated that aquifer vulnerability is "the possibility of percolation and diffusion of contaminants from the ground surface into natural water-table reservoirs under natural conditions." However, this definition has gradually changed, e.g., Hrkal (2001) describes groundwater vulnerability as the tendency and the probability that a general contaminant reaches the water table after introduction at the ground surface.

There are two types of groundwater vulnerability: the intrinsic and the specific vulnerability. The first one could be described as the "relative inability" of the soil-rock-groundwater system to protect its water against contamination (Zaporozec 2002). For the second one, in addition to the various components of intrinsic vulnerability, it considers the aquifer vulnerability taking into account the contaminant properties participation in groundwater modifications (Gogu and Dassargues 2000).

Different approaches have been proposed to assess groundwater vulnerability (Javadi et al. 2011a and b; Allouche et al. 2017; Ayed et al. 2017 and 2018; Nasri et al. 2021). These methods are grouped into three categories: process-based simulation, statistical and overlay and index models (Dixon 2004).

The index-overlay methods, such as DRASTIC method, which use a numerical rating system, are based on the selection of several parameters to assess the intrinsic vulnerability using a geographical information system (GIS), which provides efficacious tools to overlay and integrate the different multiple maps (Vu et al. 2019). The DRASTIC method presents the simplest and exhaustively exploited methods which considers seven hydrogeological parameters: depth to groundwater, net recharge, aquifer media, soil media, topography, impact of the vadose zone and hydraulic conductivity. However, the choice of this method depends on the data available to a great extent, the type of aquifer being studied, and takes into account the hydrological, geological, and hydrogeological characteristics of the study area. The plain of Sidi Bouzid, situated in the center of Tunisia, is almost composed of agricultural land.

So, the shallow aquifer was exploited in order to satisfy agricultural needs. For this reason, the DRASTIC method is considered the most suitable for the study region. In fact, this method was mostly used for the modification technique by the nitrate concentration, which can be classified into the standard method and specific method.

The use of process-based models, incorporating appropriate scaling procedures, provides an adequate approach to apply a mathematical method that form the key tool. The application of these methods on a regional scale are more difficult.

Statistical methods represent an effective tool to better determine the role of factors having the highest influence on groundwater vulnerability. Moreover, they allow identifying the likely important sources of contamination and assessing the relative probability of contamination occurrence (i.e., groundwater vulnerability), considering the simultaneous presence, or absence, of these factors (Stevenazzi et al. 2017). Such methods, like factorial analysis of correspondence "FCA" and multiple linear regression "MLR," are capable to classify each parameter that potentially influences the probability of pollution of the aquifer and lead to a score which designates the vulnerability of the groundwater (Focazio et al. 2002; Panagopoulos et al. 2006). Usually, these methods aim to use the fewest predictors to explain the greatest variability in the response variable. Stepwise approaches are used to select the most relevant predictors in regression models (Ouedraogo 2017).

In the current study, we performed the MLR approach to select the predictors. The rationale of this selection method is to combine the measure of fit with a penalty term based on the number of parameters used in the model. If more parameters (i.e., the number of trends or explanatory variables) are used, the model fit can be better, but the penalty for the extra parameters are higher as well (Boy-Roura et al. 2013; Brindha and Elango 2015; Ouedraogo 2017).

Concerning the specific vulnerability assessment to pollution by nitrate, different researchers have made a modification to the original DRASTIC method using nitrate concentrations in order to provide a reliable groundwater vulnerability model (Rupert 2001; McLay et al. 2001; Panagopoulos et al. 2006; Antonakos and Lambrakis 2007; Huan et al. 2012; Javadi et al. 2011a, b; Neshat et al. 2013; Neshat et al. 2014; Sinha et al. 2016; Jmal et al. 2017; Roohollah et al. 2019; Ncibi et al. 2020; Nasri et al. 2021).

The nitrate, originating mainly from agriculture activities, has been considered an indicator of pollution in Sidi Bouzid shallow aquifer. In fact, this anion results from the leaching of nitrogen used to fertilize agriculture land. Therefore, it is necessary to estimate whether certain locations in this aquifer are more susceptible to receive and transmit pollution.

The objective of this study is to assess the most vulnerable areas to nitrate pollution of the Sidi Bouzid region using

a comparative study between the standard and the modified DRASTIC methods.

Study area

The Sidi Bouzid shallow aquifer, the subject of this study, covers around 580 km^2, and it is located in the western part of Central Tunisia (Fig. 1). The study area was classified among arid regions with an average annual temperature and rainfall of about 19.7 °C and 228 mm, respectively, for the period 1975 to 2018 (NIM 2018). This region is almost composed of agricultural land. The Sidi Bouzid shallow aquifer was exploited in order to satisfy the agricultural needs. As knowing, the agricultural activities development is based on a chemical fertilizers intensive use, leading to groundwater contamination. Nitrate is considered the most widespread chemical contaminant in the Sidi Bouzid shallow aquifers, selected to describe urban and agricultural lands. The highest values of nitrate content (reaching 80 mg/l) were recorded in the North and the East of the study area, characterized by the irrigated croplands and the low water table (Jmal et al. 2017). Figure 2 represents the spatial distribution of nitrate content in the study area.
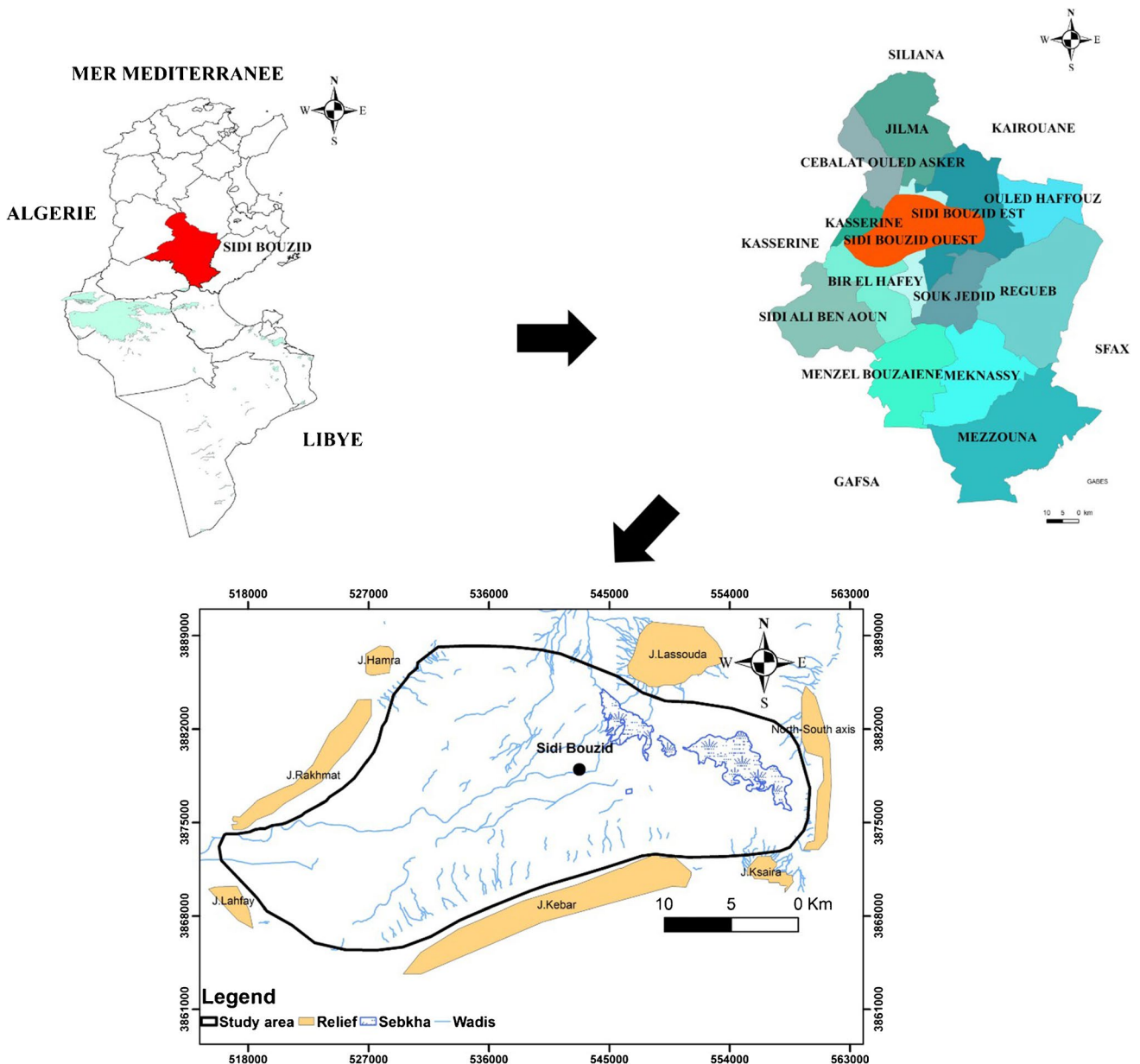

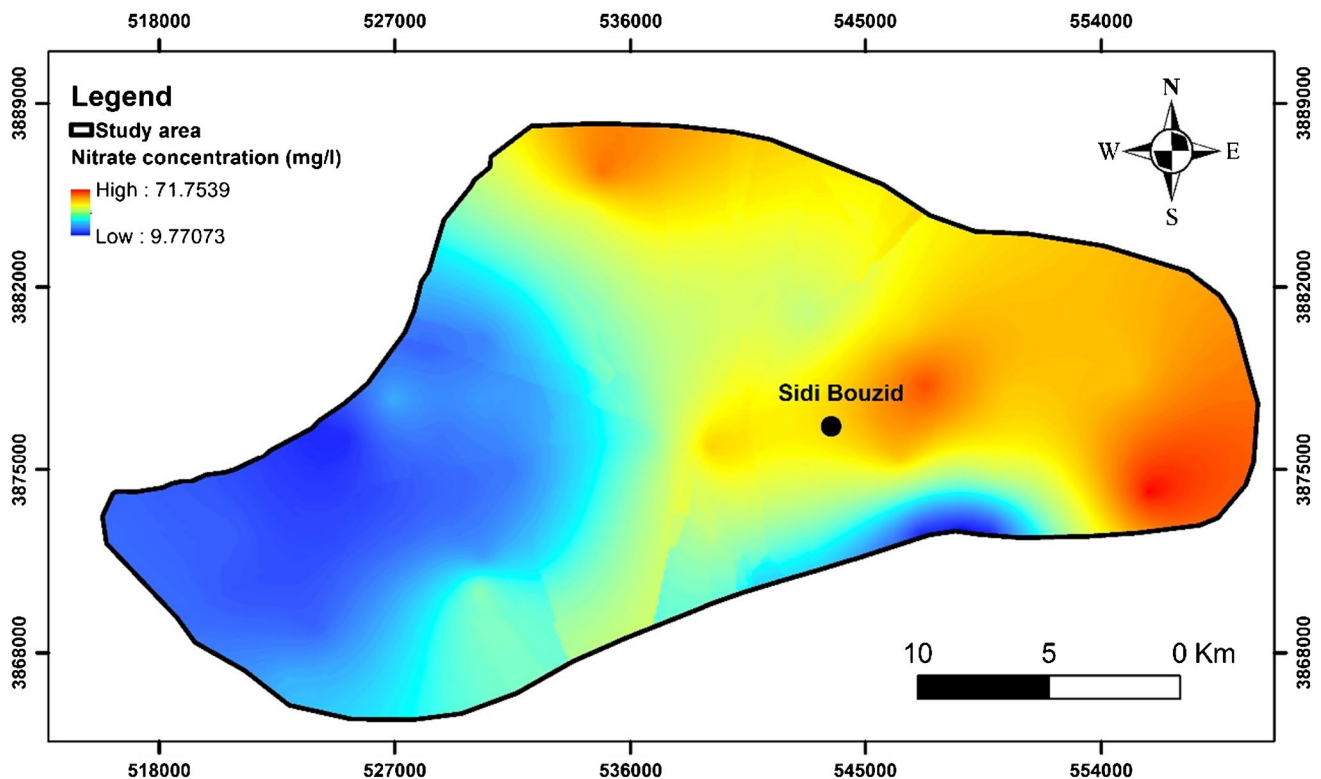**Fig. 1** Localization of the study area

Fig. 2 Spatial distribution of nitrate in Sidi Bouzid shallow aquifer (Jmal et al. 2017)

The plain of Sidi Bouzid is relatively flat and covered essentially by the mixing of sand, clay, and gravel intercalations of the Pleistocene age and recent alluvial deposits. The study area boundaries are covered by the marl of Triassic, the clay and gypsum of Cretaceous, and the clay and sand of Neogene and Paleogene (Fig. 3). The Sidi Bouzid shallow aquifer is formed by detrital deposits of the Mio-Plio-Quaternary layer system (Smida 2008).

The piezometric study was based on 27 static level measures in 2014. The piesometric map of the Sidi Bouzid shallow aquifer shows head values ranging from 290 to 365 m. It shows a general flow direction going from the South to the North. It shows an appearance of hydrodynamic convergence areas located especially in the North and the Southwest parts of the study area (Fig. 3). This can be explained by the excessive exploitation of the aquifer in these irrigated zones.

The study area is constituted of two among 11 hydrogeological basins in the entire region of Sidi Bouzid: the El Fakka Wadi and Serg Edhiba Wadi basins. The extension of these basins coincides with the study area boundaries.

Material and method

Dataset

To evaluate the groundwater vulnerability to pollution and to develop a vulnerability map, a detailed study was conducted in the Sidi Bouzid region by collecting available data (Table 1). The processing of all these data is performed on ArcMap 10.1, which was considered the tool for creating, viewing, querying, editing, composing, and publishing maps.

DRASTIC method

The groundwater vulnerability was almost assessed using the DRASTIC model. It was created by the US Environmental Protection Agency to evaluate the groundwater pollution potential (Rahman 2008). Its acronym is resulting from the initials of its seven parameters: depth to water table, recharge to aquifer, aquifer media, soil media, topography (slope), impact of the vadose zone media (vadose media), hydraulic conductivity of the
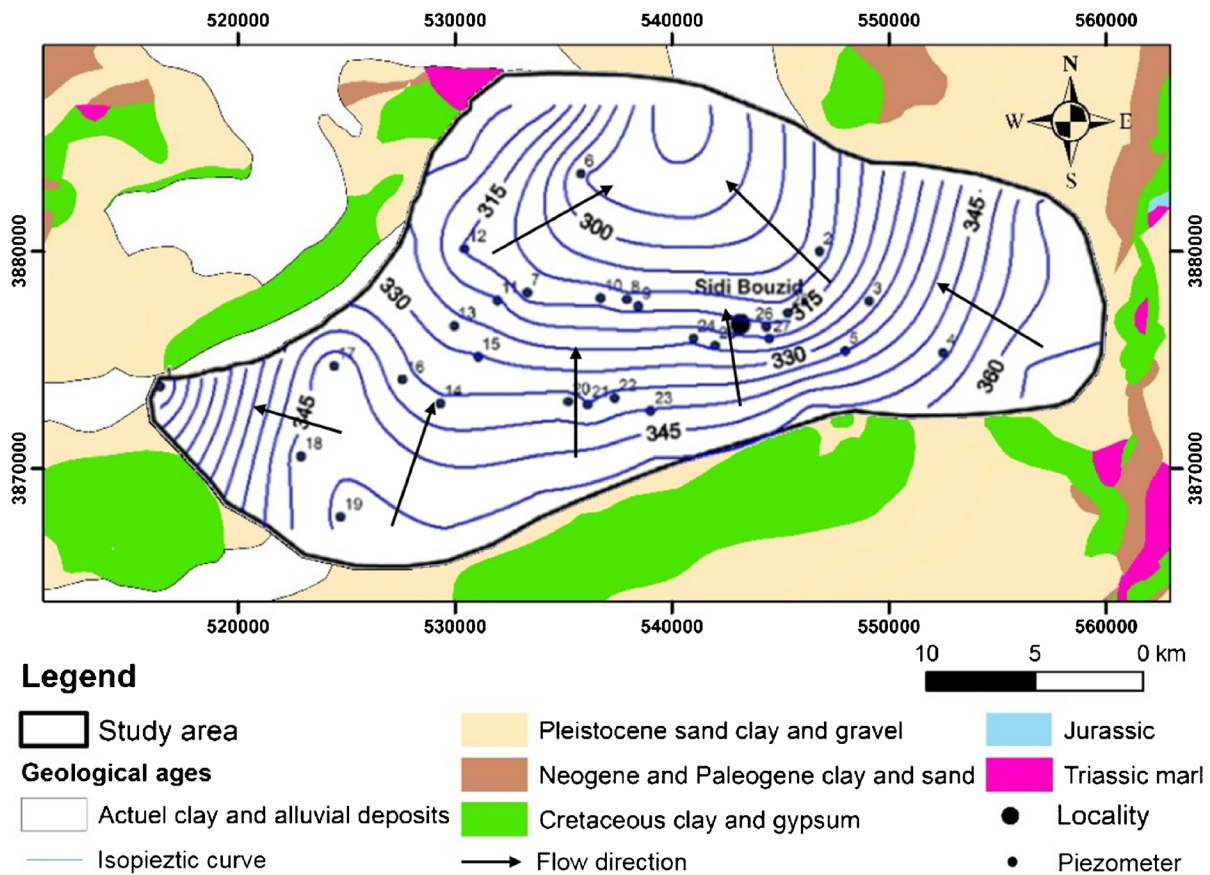
Fig. 3 Geological and piezometric map of the study area (Jmal et al. 2017)

aquifer (Aller et al. 1987). Each of the above seven DRASTIC parameters is mapped and classified either into ranges or into significant media types, and a numerical rating ranging between 1 (least pollution potential) and 10 (highest pollution potential), is assigned to each factor or parameter. The numerical rating represents the degree of the susceptibility to groundwater pollution. Some hydrogeological parameters are more important than others in determining aquifer vulnerability. Each DRASTIC rating is multiplied by a weighting factor ranging from 1 to 5. The DRASTIC Index is then computed applying a linear combination of all factors according to the following equation:

$$V(intrinsic) = D_R*D_W + R_R*R_W + A_R*A_W + S_R*S_W$$
$$+ T_R*T_W + I_R*I_W + C_R*C_W \qquad (1)$$

where $V_{(intrinsic)}$ is the DRASTIC index for intrinsic vulnerability.

"D" is the groundwater depth parameter It is one of the most important factors because it determines the depth of material through which a contaminant must travel before reaching the aquifer. The depth to groundwater was measured in 71 surface wells recorded during March 2014. The

Table 1 Data sources of DRASTIC parameters

Parameters	Data source	Mode of processing
D	Piezometric data (21 groundwater samples) from the piezometric campaign carried out in March 2013 in the region through the Regional Commissary for Agricultural Development (CRDA)	Interpolation
R	Static level statement of the 2001 and 2013 years (CRDA)	Interpolation
A	Lithology classes taken from well logs (CRDA) and geological information (Smida 2008)	Interpolation
S	Soil map (CRDA)	Digitalization
T	Topography maps (1/50,000) (CRDA)	Digitalization
I	Lithology classes taken from well logs (CRDA) and geological maps (Smida 2008)	Interpolation
C	Permeability calculated from the transmissivities taken from well logs (CRDA)	Interpolation
L	Land cover map (CRDA)	Digitalization

depth to water table map is obtained using the Kriging interpolation extension of the Arc Toolbox "spatial analyst tools."

"R" is the net recharge parameter It represents the infiltered portion of rainfall water, irrigation water return flow, and absorption wells to migrate down to the Groundwater. We have used the water table fluctuations method (WTF) to determine the recharge values. This method has been shown to be applied to shallow aquifers that show a large rise and falls in water level (Sophocleous 1991). It links the groundwater storage variation and the water table fluctuations:

$$R_{(tj)} = S_y \times DH_{(tj)} \tag{2}$$

where $R_{(tj)}$ (mm) is the recharge showing between the times t_0 and tj, S_y is the specific yield (storage) or the effective porosity of the unconfined aquifer, and $DH_{(tj)}$ is the peak water level rise associated to the recharge period.

"A" is the aquifer media parameter The aquifer permeability was determined based on the drilling logs of the saturated zone for each well using the horizontal permeability equation (Castany 1982). The aquifer media states the rock types (consolidate and unconsolidated) (Aller et al. 1987).

"S" is the soil media parameter It controls the recharge rate, which can reach the groundwater and its retention pollutant ability into the vadose zone. It was obtained by digitizing the existing soil maps of the study area, with a scale of 1:50,000. These maps were collected from the Regional Agency of Agriculture Laboratory (CRDA).

"T" is the topography parameter It refers to the slope of the land surface. The surface slope was derived from the 1:50,000 scale topography maps covering the study area.

"I" is the impact of the vadose zone parameter It influences on the potential aquifer pollution. It depends on the permeability and the attenuation characteristics of soil cover. The impact of the vadose zone was determined basing on the drilling logs and using the vertical permeability equation (Castany 1982).

"C" is the hydraulic conductivity parameter The hydraulic conductivity value is controlled by the aquifer properties. It determines the aquifer formation ability to transit water. The hydraulic conductivity values were obtained by using the aquifer thickness and the Transmissivity that is measured from the field pumping test in the Sidi Bouzid region.

The subscripts $_R$ and $_W$ are the corresponding ratings and weights, respectively. Computed maps result from the multiplication between the selected rating and the assigned weight to each parameter.

Once the DRASTIC Index has been computed, it is possible to identify areas which are more likely to be susceptible to groundwater contamination. Values of DRASTIC index for intrinsic vulnerability may range between 23 and 226 (lowest and highest vulnerabilities, respectively).

Nitrate measurements

The DRASTIC model calibration was established using the nitrate concentration. Thirty-eight surface wells were selected for sampling and analysis, with two nitrate samples taken from each well. The first sampling was realized in March 2013 to determine the correlation coefficient between nitrate concentration and groundwater vulnerability, whereas the second nitrate sampling was conducted in March 2014 to calibrate the model.

Modified DRASTIC method

Aiming for a better prediction, a modification of the DRASTIC method has been elaborated using the primary control parameter (Nitrate) according to Panagopoulos et al. (2006). To use nitrate to calibrate the model, the following conditions should be satisfied (Javadi et al. 2011b):

> the source of nitrate has resulted from agricultural activities;
> the area distribution should be homogenous;
> the nitrate leaching was resulted from the surface recharge over a long period of time to establish the correlation between contamination and human activities.

In the modified DRASTIC method, we applied simple statistical procedures to consider the modification of the rating scales and the factor weights of each parameter, and the addition of the land use parameter. This modification is based on their correlation to nitrates concentration.

Statistical approaches

Chemometric approaches, included particularly correspondence analysis (CA) and quality of representation, have been achieved using the ITCF statistical software package (STATIT-CF 1987) so as to evaluate the investigated variables in the samples associations and to recognize the pollution sources.

Correspondence analysis was applied to identify the sources and the processes that may describe the vulnerability

of groundwater analyses. Eigenvector techniques such as CA convert the original variables into a new set of variables (vectors) so that the main portion of data variation is minimized on just two or three of them. Then, this reduced set of vectors can be plotted on a single diagram and grouped in clusters that are formed and interpreted in terms of some physical or chemical process. Fundamentally, vectors are a new set of orthogonal axes geometrically categorized by the eigenvectors and eigenvalues: (1) the elements of the eigenvectors are the angles cosines between the original and transformed axes; (2) the eigenvalues measure the data spread in the vector's direction (Pacheco 1998).

The representation quality of parameters was established by a coefficient giving the maximum variance (F1 × F2 plane), resulting by the sum of the squares of the correlation coefficients between each parameter and the factorial axis of the considered plan (Dutot et al. 1983; Azri et al. 2009). It was defined by the following equation:

$$\text{QLT}(j) = \sum_{j=1}^{n} r_i^2(j)$$

where QLT (j): the representation quality of the parameter j (‰); r_i: the correlation coefficient between the parameter j and the factorial axis i; n: the number of considered factorial axes.

MLR is one of the most flexible statistical tools available to the researcher. It allows the modeling of multiple influences on an outcome and the understanding of the relationship between a dependent variable and one or more independent variables (Mustapha and Abdu, 2012; Pathak 2012; Basu and abd Lokesh, 2014).

The rationale of this selection method is to combine the measure of fit with a penalty term based on the number of parameters used in the model. If more parameters (i.e., the number of trends or explanatory variables) are used, the model fit can be better, but the penalty for the extra parameters is higher as well.

In the present study, the MLR approach will be used in order to well describe the relationship between the nitrate concentrations and the six other variables which could impact the quality of groundwater quality and accordingly assess the groundwater vulnerability to nitrate pollution in arid and semi-arid regions like the Sidi Bouzid shallow aquifer.

According to these researchers, MLR can be expressed using the following equation:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_m X_m + \varepsilon \tag{3}$$

where Y represent the dependent variable; $X_1, \cdots, X_m$ represent the several independent variables (predictor variables); $\beta_0, \cdots, \beta_m$ represent the regression coefficient; ε represent the random error.

From the above model, we could determine the R^2 value, called the coefficient of determination (R-squared) which is considered a statistical metric that is used to measure how much of the variation in outcome can be explained by the variation in the independent variables. R^2 always increases as more predictors are added to the MLR model, even though the predictors may not be related to the outcome values. Moreover, R^2 gives the percentage of variability in the dependent variable accounted for by all the independent variables together.

Additionally, R^2 by itself cannot thus be used to identify which predictors should be included in a model and which should be excluded. R^2 can only be between 0 and 1, where 0 indicates that the outcome cannot be predicted by any of the selected independent variables and 1 indicates that the outcome can be predicted without error from the independent variables. A value of 1.0 indicates a perfect fit and is thus a highly reliable model for future forecasts, while a value of 0.0 would indicate that the calculation fails to accurately model the data at all. But a value of 0.20, for example, suggests that 20% of the dependent variable is predicted by the independent variable, while a value of 0.50 suggests that 50% of the dependent variable is predicted by the independent variable, and so forth.

R^2 measures the explanatory power of the model, which in turn reflects the goodness of fit of the model. It reflects the model adequacy in the sense of how much is the explanatory power of the explanatory variables.

The R^2 value is given as output by the software which will be used.

The most commonly used procedure for selecting the best regression equation is stepwise linear regression analysis (using an F probability of 0.05 for the selected factor), as described by Landau and Everitt (2003) and performed using SPSS.

Results and discussion

Application of the standard DRASTIC model

The major drawback of the DRASTIC method is the particularity of the rating scale and the weighting coefficients determination for the several factors (Table 2). The seven parameters map of the DRASTIC method were, respectively, presented in a raster format from GIS tools (Fig. 4).

The depth to groundwater shows values ranging from 10 to 61.5 m which were divided into six intervals. According to this parameter, the East of the Sidi Bouzid shallow aquifer is considered the most vulnerable zone with a low water depth. The computed depth to groundwater map (Fig. 4.1) corresponds to the first component of DRASTIC model.

Table 2 Assigned weights for DRASTIC features (Aller et al. 1986)

Parameters	Weight
D	5
R	4
A	3
S	2
T	1
I	5
C	3

The resulting map of the net recharge parameter, which was calculated based on the WTF method, reveals three ranges of net recharge: 2.9–5 mm/year, 5–10 mm/year, and 10–17 mm/year. The recharge map shows that the high values of aquifer recharge are recorded in the upstream of the Sidi Bouzid shallow aquifer. This can be explained by the water supply from the El Fakka Wadi during floods. The net recharge map corresponds to the second component of DRASTIC model (Fig. 4.2).

The aquifer rock is composed of Clayey sand, sand, and sand and gravel, which occupies almost all of the study area (Fig. 4.3).

The soil media map consists of mineral soil, rendzina and hydromorphic soil, calcareous brown soil, gypsum soil, soil with little evolution, isohumic soil, halomorphic soil, complex unit soil, and urban zones. The highest rate is attributed to the mineral soil (9), and the lowest rate is attributed to the urban zones (1). The soil media map represents the fourth component of DRASTIC model (Fig. 4.4).

In our case, the topography was divided into six classes (0–3%), (3–5%), (5–10%), (10–15%), (15–25%) and > 25%. Most of the study area is characterized by a slope oscillating from 0 to 3%. The low slope observed in the majority of the study area tends to retain water in the ground surface and favors the infiltration or recharge of water through the vadose zone leading to a great potential for contaminant penetration. The topography map corresponds to the fifth component of the DRASTIC model (Fig. 4.5).

The vadose zone is mainly constituted of sandy clay, clayey sand, fine sand, medium to coarse sand, and gravel with high permeability. The corresponding rates system ranges between 1 and 10. The impact of the vadose zone map characterizes the sixth component of the DRASTIC model (Fig. 4.6).

The hydraulic conductivity values vary between $9.9 \, 10^{-5}$ and $2.6 \, 10^{-5}$ (m/s), divided into 3 ranges. Rates of 6, 7, and 8 were attributed to this parameter. The hydraulic conductivity map corresponds to the seventh component of DRASTIC model (Fig. 4.7).

The groundwater vulnerability map in the study area results from the sum of the different obtained components using GIS tools. It illustrates the presence of three vulnerability classes (Fig. 5 and Table 3):

low vulnerability (74–100) degrees which occupies a very small area (3.5% of the study area). This result can be related to a high groundwater depth (> 33 m) and low permeability of the vadose zone (2.10^{-7}–4.10^{-4} m/s); moderate vulnerability (100–139) degree, spatially scattered, which covers almost all of the study area (71%). Hydrogeological protection seems to be unspecified; high vulnerability (139–163) degree, which covers 25.5% of the study area. It is located mainly in the East and the Northeast and particularly in the center of the study area. It characterizes the shallow areas (10–18 m) and areas with high permeability of the vadose zone and aquifer. Indeed, high vulnerability classes coincide with the Sebkha and urban zone, where industrial activities are developed.

The corresponding vulnerability map represents a decision support tool for future land uses. For more reliable results, it is essential to be validated after their establishment by dependent parameters in the vulnerability equation but does not act directly on the groundwater vulnerability to pollution.

Validation of the DRASTIC model using nitrate concentration

The not validated model creates biased and erroneous results in assessing an environmental system. Thus, the validation process is necessary to avoid wrong conclusions (Kura et al. 2015; Saidi et al. 2011). Nitrate concentration was used to validate the vulnerability DRASTIC model in the Sidi Bouzid shallow aquifer.

Thirty-eight sampled points taken in March 2013 are used and plotted on a DRASTIC map; the corresponding nitrate concentration for each sample was extracted and divided into five categories according to DRASTIC classes (Fig. 6). The highest nitrate concentration values (79 and 73 mg/l) are located especially in the high vulnerability class, located in the agriculture lands near the urban areas, industrial areas and sebkha. However, the low nitrate concentration ones occupy areas with low and mediocre vulnerability class, found in groundwater around the Kebar and Hamra mountains.

Furthermore, nitrate concentration values and vulnerability index were correlated to evaluate the original DRASTIC method using Pearson's correlation factor (r). The results of the correlation are low, in order of 44% (Fig. 6), with a P-value equal to 0.005 (Table 4). The p-value obtained

from this test is less than 0.01, so the data sets are significantly different at the 99% confidence level. This result is statistically significant with 99% of level confidence. Half of the nitrate concentration values increase with the vulnerability index. However, low nitrate concentration values are not equivalent to a low vulnerability index. This means that the intrinsic vulnerability indexes are weakly correlated with the nitrate concentration values. In the same approach, some authors, such as Panagopoulos et al. (2006) and Neshat et al. (2014), have used the correlation between nitrate concentrations and the vulnerability index to validate their vulnerability model. They found correlation coefficients comparable to the coefficient found in our case study.

In order to obtain a better correlation, reflecting actual groundwater vulnerability and a realistic assessment of the potential pollution in the study area, a modification of the original DRASTIC model for groundwater vulnerability using nitrate should be required.

Application of the modified DRASTIC model in the study area

Modification of the rating scales of each parameter

The modified DRASTIC model becomes an effective tool that was suggested for the study area to obtain significant results. As long as the agriculture land occupies the most of Sidi Bouzid plain, the nitrate concentration was used in order to conduct better results in the vulnerability map. In fact, the rating scale for each parameter in the original model was modified using the median of every class and the corresponding median of nitrate concentration. This modification is based on the Wilcoxon rank-sum nonparametric statistical test (Wilcoxon 1945). This latter was used to study the statistical significance of these different classes according to the nitrate concentrations distribution present in each class. It calculates a P-value which judges the significance of the test, consistent to compare this value to a previously defined threshold (traditionally 5%). For the continuous parameters (depth to groundwater, recharge to aquifer, topography and hydraulic conductivity), the neighboring classes may be combined in such cases when the median of these classes did not differ statistically. Groundwater depth parameter combines forth classes (11–14 m, 14–18 m, 18–25 m, and 25–33 m), which does not present a significant statically results (P-value > 0.05) (Fig. 7). In fact, the concentrations are statistically significant between the "25–33 m" and "33–61.5 m" of the groundwater depth classes. However, the nitrate concentrations in the other neighboring classes are not significant (the classes colored in red in the 1st row) (Fig. 8). When two neighboring classes show nonsignificant results, these two classes are combined, and the combined

class is tested with another neighboring class, and so forth. The final result of the Wilcoxon nonparametric statistical test, applied to the groundwater depth parameter, shows the assembly of three classes "8–11 m," "11–33 m," and "33–61.5 m" (Fig. 8).

For the noncontinuous parameters (aquifer media, vadose zone media, and soil media), all the media present in the zone were kept inconsiderate of statistical diversity (Rupert 2001). The box plot diagrams of the seven DRASTIC parameters are presented in Fig. 9. It shows different spatial distributions of nitrate concentrations applied for the different parameter classes. In fact, the box plot diagrams represent the graphical representation of the statistical analysis applied to the different parameters to assess the importance of each class and determine the new rating through the median of nitrate concentrations calculated for each class. It is designed to display the location, spread, skewness, and tail of the data. The precision of some of this functionality is considered to be more reliable for symmetric data types and thus less appropriate for skewed data such as the extreme data. So, the box plot diagrams explain the repartition of nitrate concentrations in each class of parameter and the location of the highest concentrations values exceeding the threshold, equal to 24.5 mg/l.

The rating system of each parameter was performed according to previous studies (Panagopoulos et al. 2006; Antonakos and Lambrakis 2007; Huan et al. 2012; Javadi et al. 2011b; Neshat et al. 2014; Kazakis and Voudouris, 2015) and the following equation on a scale of 10:

$$r = (C_{(NO3-)}/C_{(NO3-)max}) * 10 \tag{4}$$

where r is the value of the rating, $C_{(NO3^-)}$ is the median nitrate concentration in the classes of corresponding parameter, and $C_{(NO3^-) Max}$ is the highest median nitrate concentration in the corresponding parameter.

The ratings obtained from this equation, corresponding to the ratings from the modified DRASTIC model, are illustrated in Fig. 4 and Table 5.

Modification of the factor weights

A modification of the weighting factors is assigned to all the DRASTIC parameters defined in Eq. (1) using the median nitrate concentration for the thirty-eight sampling points of the study area. The discrete (nominal) parameter was correlated with the nitrate concentrations using their factor scores and not the natural range values due to the fact that the factor scores vary with an interval scale (Panagopoulos et al. 2006). The correlation was calculated using the Spearman's (ρ)and Kendall's (τ)correlation coefficients (Kendall 1975), which are recommended for such parameters type. The new weighting factors were determined based on these coefficients and after their values were reduced to a scale with a
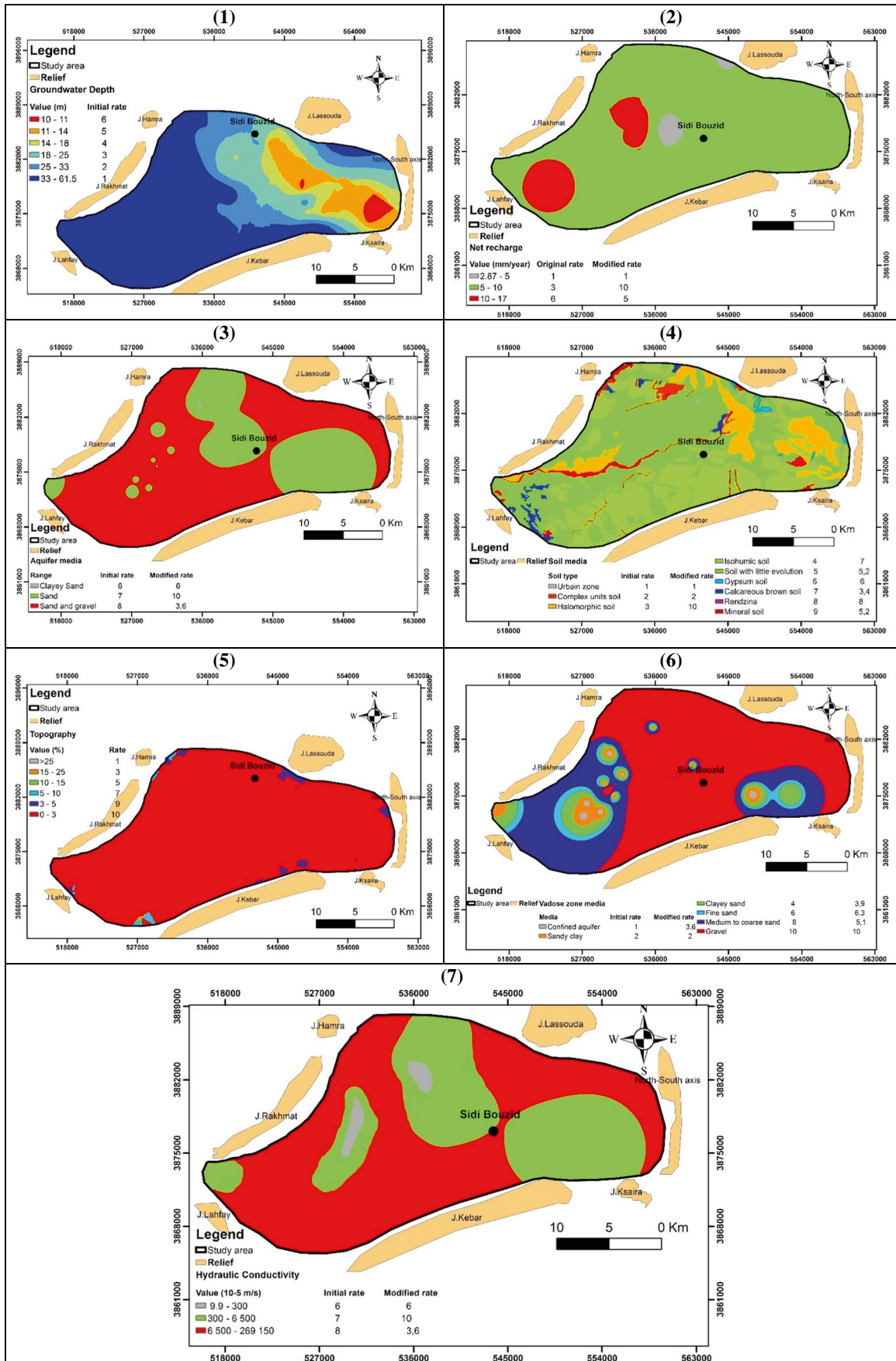
◀**Fig. 4** Seven layers of standard DRASTIC model **1** groundwater depth; **2** recharge; **3** aquifer media; **4** soil media; **5** topography; **6** impact of vadose zone; **7** hydraulic conductivity

maximum value of 5, as described by the DRASTIC model (Panagopoulos et al. 2006). when the parameter represents coefficients not statistically significant, this later will be excluded from the equation of the vulnerability.

The groundwater depth records the highest values of Spearman's and Kendall's correlation coefficients. The result 0.724 and 0.594, respectively (Table 6). The relationship between groundwater depth and nitrate concentrations is very strong. It justifies the importance of this parameter in the groundwater vulnerability to pollution. This relationship is statistically significant at a 99% confidence level.

Recharge and soil media parameters show the lowest values (Table 6), with a critical probability value (P-value) of about 0.12 and 0.15, respectively. So, it will be excluded from the vulnerability equation because these Spearman's and Kendall's coefficients are not statistically significant (statistically comparable). The same conclusion was also

reported by other researchers (Panagopoulos et al. 2006; Antonakos and Lambrakis 2007).

Other parameters are statistically significant to 95% confidence level, exception of topography parameter, which does not represent correlation results. The weighting of each parameter was recalculated using the following equation on a scale of 5:

$$\mathbf{w} = (\mathbf{T}/\mathbf{T_{Max}}) \times 5 \tag{5}$$

where w is the modified weighting of each parameter, T is the correlation value of each parameter with nitrate concentrations, and T_{Max} is the highest correlation value of all parameters with nitrate concentrations. The parameters, the original weighting, the correlation coefficient values, and the modified weighting factors are indicated in Table 6.

The weights obtained from this equation will be used for the evaluation of the modified DRASTIC model. In fact, the weighting factors are comparable to the standard DRASTIC method with the exception of the "impact of the vadose zone" parameter, which has been weakened although it remains relatively high.
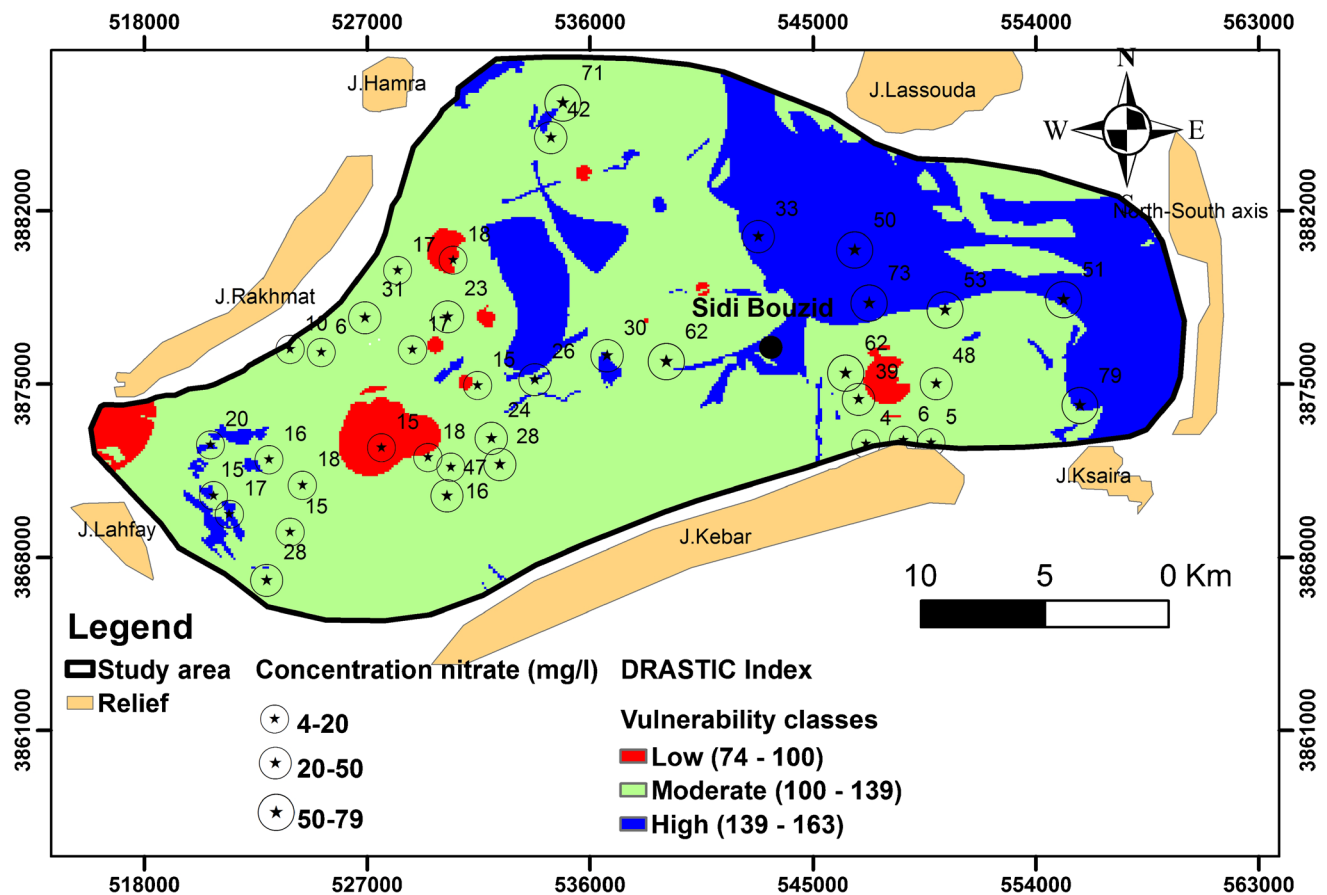

Fig. 5 Distribution of the standard DRASTIC vulnerability and nitrates concentrations for the study area

Table 3 Classes of groundwater vulnerability in Sidi Bouzid area

DRASTIC Index value	Vulnerability class
74–100	Low
100–139	Moderate
139–163	High

Table 4 Correlation between DRASTIC vulnerability index and nitrate concentrations in groundwater in the study region

		Nitrate	DRASTIC
Nitrate	Corrélation de Pearson	1	0.44[**]
	Sig. (bilatérale)		0.005
	N	38	38
DRASTIC	Corrélation de Pearson	0.44[**]	1
	Sig. (bilatérale)	0.005	
	N	38	38

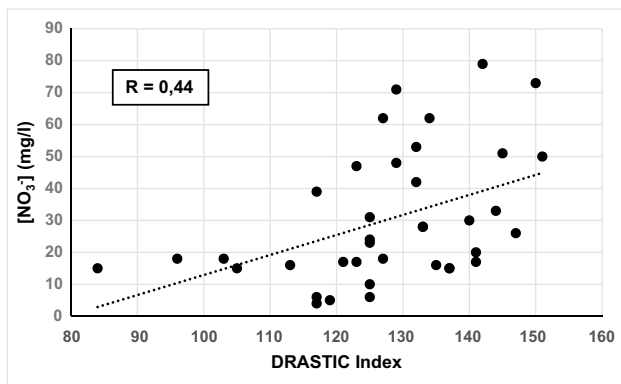[**] La corrélation est significative au niveau 0.01 (bilatéral)

Fig. 6 Correlation between the DRASTIC original vulnerability index and groundwater nitrates concentration for the study area

Addition of the land use parameter

The land use was used as a pollutant loading indicator in the groundwater vulnerability. It represents a major impact upon the nitrate concentrations of the groundwater. In order to achieve more objective vulnerability and evaluate groundwater pollution risk, land use type should be combined with the groundwater vulnerability indexes in certain cases.

The initial land use distribution is rated according to the Secunda et al. (1998), including seven classes of Land use (Fig. 10 and Table 7). An initial weight of five
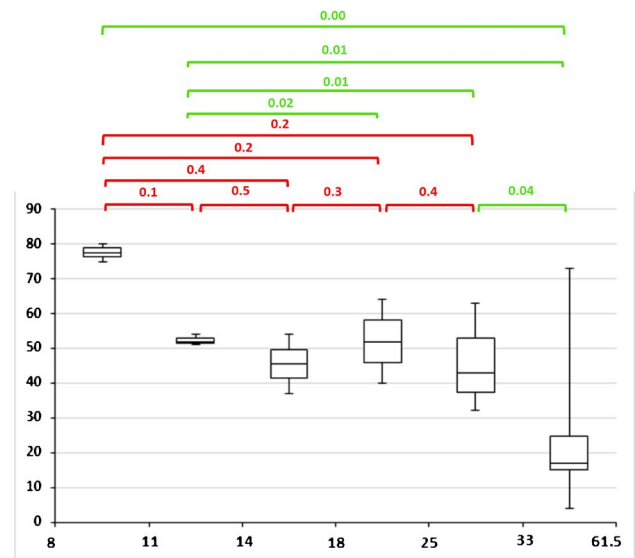
Fig. 7 Application to the Willcoxon nonparametric test for the groundwater depth parameter

has been attributed to the land use parameter due to the potential impact that this factor may have upon the nitrate concentrations determination of the groundwater.

The modification of the rating of each land-use type depended on the revision of the rating scale of the initial land use parameter using the median of nitrate concentrations corresponding to each land-use type, as well as the modification of the DRASTIC parameters' rating scale (Panagopoulos et al. 2006).

The obtained ratings, which correspond to the ratings from the modified land use parameter, are indicated in Table 7. The box plot diagram for each land use category is illustrated in Fig. 11. It shows different spatial distributions for the three land-use categories with dissymmetric boxes. The irrigated field cropland presents the highest nitrate concentration values (median equal to 51) (Fig. 11). This result was due to the intensive use of chemical fertilizer for this crop type. The olives land has low nitrate concentration values (median in order to 15.5).

The land use parameter records highest values of Spearman's and Kendall's correlation coefficients, in the order to 0.766 and 0.620, respectively, statistically significant at 99% level of confidence (Table 6). It shows better relationship between land use and groundwater nitrate concentrations. This parameter is very important in the groundwater vulnerability assessment to pollution.

The land use parameter, the original weighting and the correlation coefficient values are indicated in Table 6.

The modified DRASTIC map was calculated using the new rating and weighting system (Fig. 12). The modified DRASTIC model was classified into three classes of vulnerability:
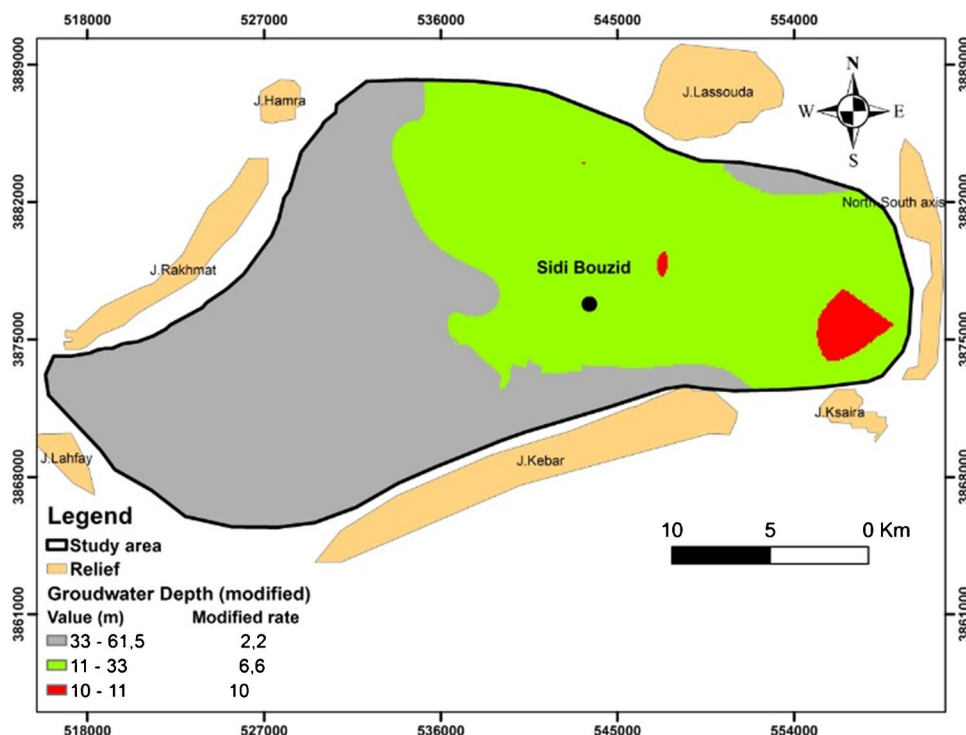
low vulnerability class (48.6–100), which occupies 37%
of the study area. The percentage of the low vulnerabil-
ity degree was brought up than the DRASTIC standard
model (Fig. 13). It can be explained by the presence of
low nitrate concentration values;

moderate vulnerability class (100–139), which covers
23% of the study area. Compared with the DRASTIC
standard model, it was decreased due to the shortage of
average nitrate values;

high vulnerability class (139–193) which represents
40% of the study area, covers mainly the urban area
and the agricultural lands.

Validation of the modified DRASTIC model using nitrate concentration

The application of the modified DRASTIC method produced
a significantly increased correlation coefficient with nitrate
concentration, around 82% (Fig. 14). The correlation factor
was significant statistically at a 99% confidence level.

The calibration results using the nitrate concentrations
recommended that the modified DRASTIC method mean-
ingfully affects the study area. These significant results of
the vulnerability map confirmed that nitrate is crucial in
the calibration method because most of the Sidi Bouzid
region is occupied by agricultural lands. The distribution

of the nitrate concentrations in the study area considers
the maximum permissible limit of 50 mg/l recommended
by the CRDA. The nitrate concentration values exceed-
ing this limit are located in the highest vulnerability areas
(Fig. 12).

The modified DRASTIC map indicates that the high vul-
nerability class represents 40% of the study area. This class
percentage was 25% before the modification. While, the
moderate vulnerability class presents 71% and 23 before and
after modification, respectively. The low-class area increases
from 4% to reach 37 after modification. These improved
results explained the effect of the model modification.

In addition, maps comparison shows the index spatial
distribution before and after modification (Fig. 13). It indi-
cates that 52% of the results had a similar class, and 48%
belonged to other class, thus confirming the success of the
proposed model.

The application of the modified model proved the modi-
fication effect, because the area of higher vulnerability in
the modified DRASTIC map is situated around Sidi Bouzid
center, which is characterized by the presence of the urban
zone, the Sebeka, the major sources of pollution and the
agricultural activities.

Factorial analysis of correspondences

To establish a new vulnerability model based on
the modified DRASTIC method, six parameters
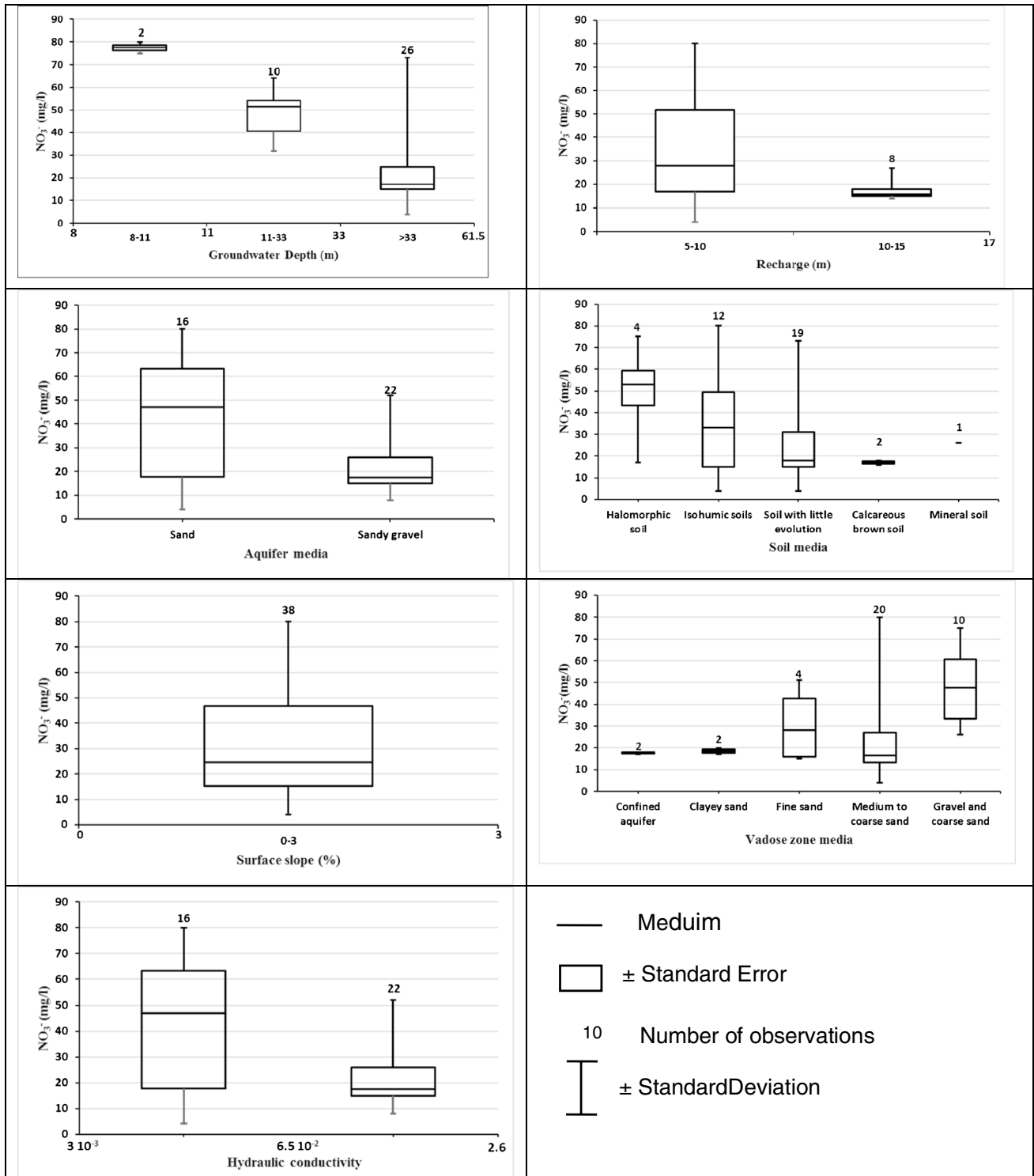
Fig. 9 Box plots showing the distribution of groundwater nitrates concentration for the statistically different classes of all the DRASTIC parameters

(groundwater depth, net recharge, aquifer media, soil media, impact of the vadose zone media, and land use) were chosen for the factorial analysis of correspondences (FACs) study.

The topography and hydraulic conductivity parameters were excluded since: (i) topography was considered statistically homogeneous because nitrate concentrations do not show a significant variability under the same slope class

Table 5 Modified ratings of the seven DRASTIC factors

Depth to groundwater

Range (m)	Original rating	Median NO$_3^-$ (mg/l)	Modified rating
10–11	6	77.5	10
11–14	5	51.5	6.6
14–18	4		
18–25	3		
25–33	2		
33–61.5	1	17	2.2

Recharge

Range (mm/year)	Original rating	Median NO$_3^-$ (mg/l)	Modified rating
2–5	1	No data	1
5–10	3	34	10
10–17	6	17	5.7

Soil type

Soil type	Original rating	Median NO$_3^-$ (mg/l)	Modified rating
Mineral soil	9	26	5.2
Rendzina	8	No data	8
Calcareous brown soil	7	17	3.4
Gypsum soil	6	No data	6
Soil with little evolution	5	18	5.2
Isohumic soil	4	33	7
Halomorphic soil	3	53	10
Complex units soil	2	No data	2
Urbain zones	1	No data	1

Topography

Range slope (%)	Original rating	Median NO$_3^-$ (mg/l)	Modified rating
0–3	10	24.5	10
3–5	9	No data	9
5–10	8	No data	8
10–15	5	No data	5
15–25	3	No data	3
> 25	1	No data	1

Aquifer Type

Aquifer type	Original rating	Median NO$_3^-$ (mg/l)	Modified rating
Clayey sand	6	No data	6
Sand	7	47	10
Sand and gravel	8	17.5	3.6

Impact of the vadose zone type

Geological formation	Original rating	Median NO$_3^-$ (mg/l)	Modified rating
Confined aquifer	1	17.5	3.5
Sandy clay	2	No data	2
Clayey sand	4	18.5	3.9
Fine sand	6	28	5.9
Medium to coarse sand	8	16.5	3.6
Gravel	10	47.5	10

Hydraulic conductivity

Range (m/s)	Original rating	Median NO$_3^-$ (mg/l)	Modified rating
$9.9*10^{-5}$ – $3*10^{-3}$	6	No data	6
$3*10^{-3}$ – $6.5*10^{-2}$	7	47	10
$6.5*10^{-2}$ – 2.6	8	17.5	3.6

Table 6 Modified weights of the DRASTIC factors and correlation coefficients between DRASTIC factors and nitrates concentration

DRASTIC factors	Original weight	Spearman's rho coefficient	P-value	Kendall's tau coefficient	P-value	Modified factor weight
Depth to groundwater	5	0.724**	0.000	0.594**	0.000	5
Recharge	4	−0.262	0.118	−0.219	0.116	—
Aquifer type	3	−0.384*	0.019	−0.321*	0.021	2.7
Soil type	2	−0.247	0.141	−0.190	0.149	—
Topography	1	No data	No data	No data	No data	1
Impact of the vadose zone	5	0.423**	0.009	0.323**	0.013	2.9
Hydraulic conductivity	3	−0.384*	0.019	−0.321*	0.021	2.7
Land use	5	0.766**	0.000	0.620**	0.000	5

* P-value < 0.05; ** p-value < 0.01

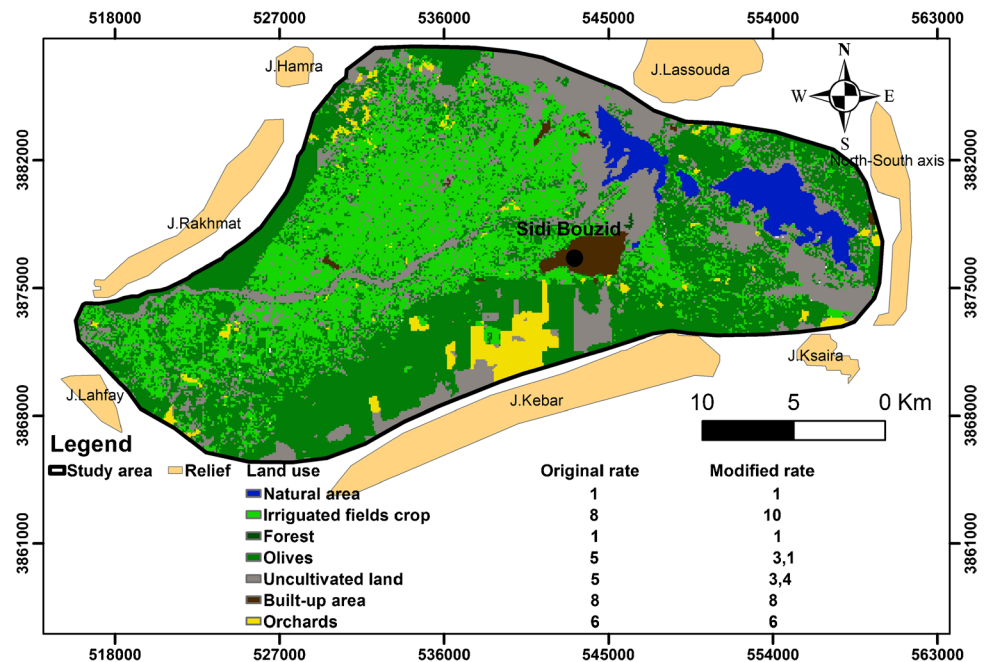**Fig. 10** The land use map with the original rating of the study area

Table 7 Original and modified land use ratings

Land use type	Original rating	Median NO_3^- (mg/l)	Modified rating
Natural area	1	No data	1
Forests	1	No data	1
Uncultivated land	5	17	3.4
Olives	5	15.5	3.1
Orchards	6	No data	6
Irrigated field crop	8	51	10
Built-up area	8	No data	8

(0–5%), and (ii) hydraulic conductivity is comparable to the aquifer media parameter.

In this section, statistical analysis considers the quality of parameter introduced in the FACs model. For this reason, categorical parameters (aquifer media, soil media, impact of the vadose zone media and land use) are specified by the use of the rating classes, extracted from modified DRASTIC model and introduced into used data as binary numbers related to registered classes. While, continuous parameters such as groundwater depth and net recharge have been introduced into the program, as they were (Jmal et al. 2017).

The FACs, applied to all data using the ITCF statistical software package (STATIT-CF, 1987), result basically three
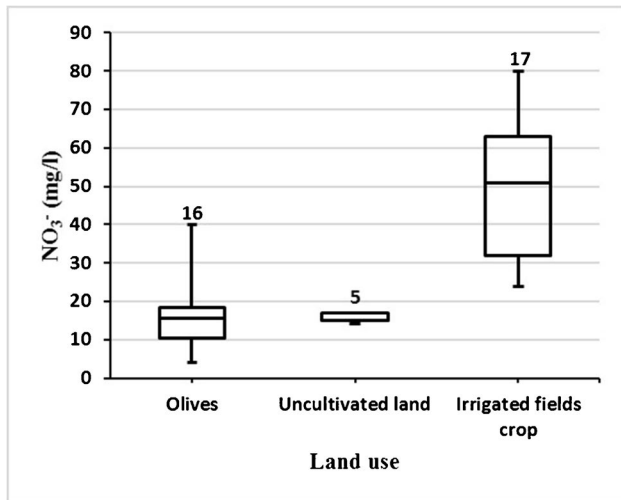
Fig. 11 Box plot showing the distribution of groundwater nitrates concentration for the land use parameter

principal components. A varimax rotation with Kaiser normalization was applied for all data. After a Student's test ($n = 38$), the threshold of significance ($p < 0.05$) is equal to 0.304. The correlations among the selected parameters (variables) and the components are significant and represent about 62.6% of the total variance. The three first FACs represent 44.52, 18.08, and 8.3% of the total variance, for the first, second, and third FACs, respectively.

The projection over the factorial plane for all selected variables (which displays the maximum of inertia; ≈62.60%) shows a distinct data cluster (Fig. 15).

Table 8 represents the correlation matrix between the nitrate concentrations and the different parameters where different correlations have been founded. Results have shown significant correlation between nitrate concentration and the groundwater depth ($R^2 = 0.708$), indicating thus the less groundwater depth corresponding to high potentially nitrate concentration. Moreover, the irrigated fields crop lands (Lu_3) and the olive tree cultivation (Lu_1) present a correlation coefficient equal to 0.765 and 0.633, respectively, due to the large quantities of chemical fertilizers used, such as nitrate in these cultures. And slightly correlated with the net recharge ($R^2 = 0.398$) and the gravel and coarse sand permeability of the vadose zone (I_5), and the medium to coarse sand permeability of the vadose zone (I_4) presented a correlation coefficient equal to 0.497 and 0.392, respectively, resulting of the arid climate in the study area.

Different correspondent parameters are spatially arranged on the factorial plan (F1 × F2) (Fig. 15), showing two principal groups:

the first group (G_1) regroups the third class of land use (Lu_3), which characterizes the irrigated agriculture areas, and the fifth class of vadose zone (I_5), which represents high permeability that has a great effect in the presence of nitrate in groundwater;

the second group (G_2) includes the D, R, I_4, and Lu_1 classes, which are negatively correlated with nitrate concentrations. Primary, the deeper the water table, the less the presence of nitrate in groundwater. This explains the antagonistic effect of the groundwater depth parameter (D). Secondly, olive tree cultivation (Lu_1) does not

Fig. 12 Distribution of the modified DRASTIC vulnerability and nitrates concentrations for the study area
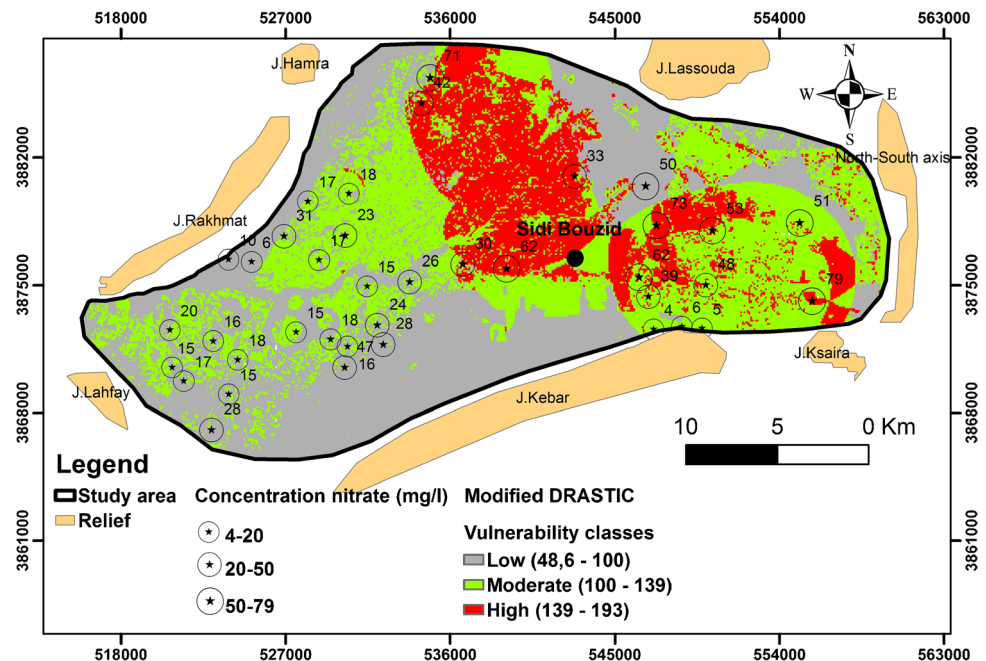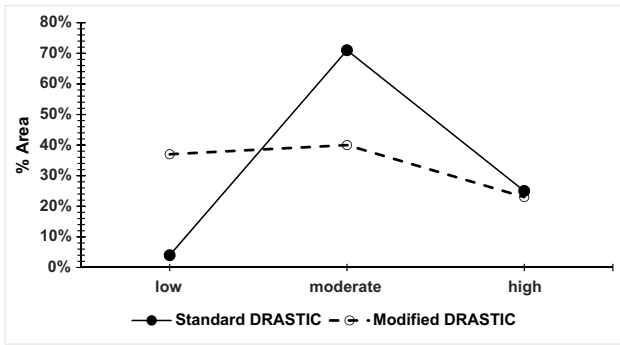
Fig. 13 The percentage of priority between standard DRASTIC and modified DRASTIC
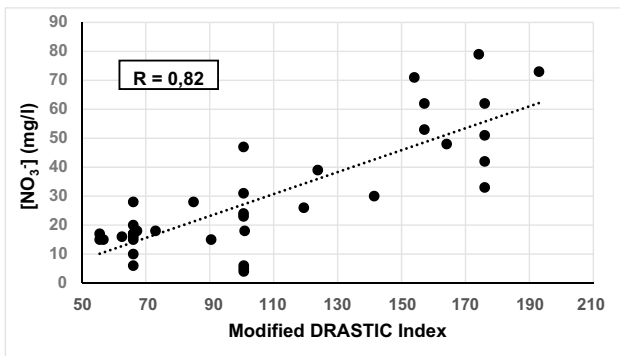

Fig. 14 Correlation between the modified DRASTIC vulnerability index and groundwater nitrates concentration for the study area

require chemical fertilizers such as nitrate. Thirdly, the net recharge (R) and the medium to a coarse sand class of vadose zone media (I_4) are negatively correlated with the nitrate concentration due to the low precipitation and soil permeability.

One notices that the other parameters are statistically insignificant. These results explain the major effect of the groundwater depth and the land use parameters in the vulnerability assessment of the Sidi Bouzid shallow aquifer.

To a greater precise analysis, six parameters (groundwater depth, net recharge, aquifer media, soil media, impact of the vadose zone and land use) were introduced in a representation quality (QLT) (Table 9) in order to better testify the most significant parameter in terms of affecting NO_3^- concentrations in different areas and under different conditions.

The representation quality of parameters showed that the "groundwater depth" parameter has the best quality of representation (QLT equal to 879.213). As well as the, "land use" parameter has a meaningful quality of representation (QLT equal to 708 and 664.191 for Lu3 and Lu1, respectively). In

fact, the role of groundwater depth and land use is reflected not only in their relationship with nitrate concentrations in the groundwater but also in the vulnerability assessment of the Sidi Bouzid shallow aquifer.

Multiple linear regression

In the present study, MLR analysis was applied to the six most significant parameters in the vulnerability hypothesis (caused by nitrate concentration in analyzed water) for all sampling sites. The MLR applied to all selected data has been expressed as follows:

$$NO_3^- = 0.380*(OS3) + 0.021*(R) - 0.130*(I4) - 0.050*(I5)$$
$$- 0.030*(OS1) - 0.007*(D) + 1.434$$

$$R^2 = 0.754 \qquad (6)$$

The above model shows the R^2 value, which gives the percentage of variability in the dependent variable accounted for by all the independent variables together. According to the R^2 statistic, 75.4% of the total variance for the estimation of NO_3^- is explained by the MLR model.

The positive sign of the beta coefficients in the case of OS3 and R indicates that there is a positive relationship between NO_3^- concentrations and these variables. However, the MLR model shows a negative sign of the beta coefficients in the case of I4, I5, OS1, and D, indicating a negative relationship between them and NO_3^- concentrations. Such results show to be in good accordance with previous ones.
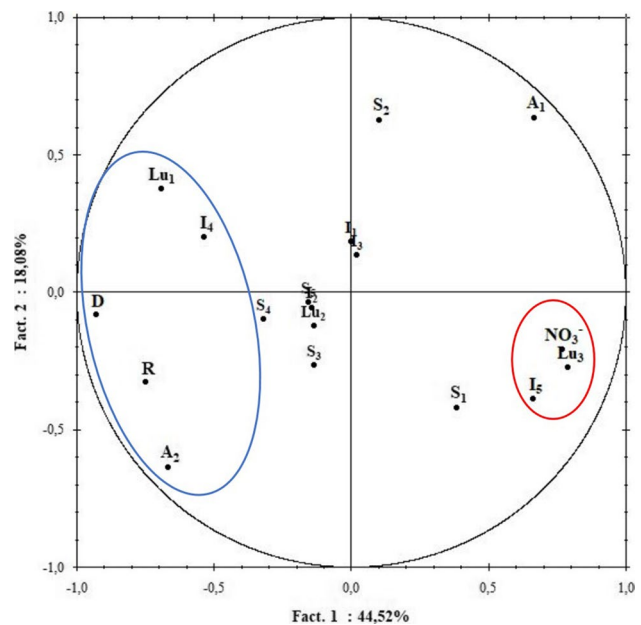

Fig. 15 Distribution of correspondent parameters in the correlation circle (for $p < 0.05$ and $n = 38$)

Table 8 Correlation matrix between the different parameters and the nitrate concentrations

	NO_3^-	A_1	A_2	S_1	S_2	S_3	S_4	S_5	I_1	I_2	I_3	I_4	I_5	Lu_1	Lu_2	Lu_3	D	R
NO_3^-	1																	
A_1	0.258	1																
A_2	−0.258	−1.000	1															
S_1	0.267	0.055	−0.055	1														
S_2	0.039	0.223	−0.223	−0.233	1													
S_3	−0.161	−0.107	0.107	−0.343	−0.679	1												
S_4	−0.102	−0.201	0.201	−0.081	−0.160	−0.236	1											
S_5	0.020	−0.140	0.140	−0.056	−0.112	−0.164	−0.039	1										
I_1	−0.093	0.276	−0.276	−0.081	−0.160	0.236	−0.056	−0.039	1									
I_2	−0.077	−0.201	0.201	−0.081	0.093	0.000	−0.056	−0.039	−0.056	1								
I_3	0.049	0.055	−0.055	−0.118	0.136	0.000	−0.081	−0.056	−0.081	−0.081	1							
I_4	**−0.392**	−0.258	0.258	−0.190	0.078	−0.105	0.224	0.156	−0.248	−0.248	−0.081	1						
I_5	**0.497**	0.217	−0.217	0.379	−0.149	0.000	−0.141	0.193	−0.141	−0.141	−0.205	−0.362	1					
Lu_1	**−0.633**	−0.188	0.188	−0.293	0.109	−0.107	0.276	−0.064	0.038	0.038	0.055	0.275	−0.389	1				
Lu_2	−0.200	−0.174	0.174	0.120	−0.097	0.078	−0.092	−0.212	0.257	0.257	0.120	−0.098	−0.233	−0.332	1			
Lu_3	**0.765**	0.305	−0.305	0.209	−0.042	0.053	−0.212	−0.148	−0.212	−0.212	−0.136	−0.206	0.544	−0.767	−0.350	1		
D	**−0.708**	−0.646	0.646	−0.417	−0.161	0.232	0.279	0.152	0.082	0.098	−0.111	0.436	−0.508	0.563	0.212	−0.703	1	
R	**−0.398**	−0.580	0.580	−0.182	−0.340	0.239	0.366	0.079	0.106	−0.012	0.024	0.265	−0.364	0.417	0.137	−0.507	0.703	1

Table 9 Representation qualities (QLT) of selected parameters as a function of the two first axes of the (1×2) factorial plane (‰) (the threshold of significance = 24.5 mg/l for $p < 0.05$ and $n = 38$ (Jmal et al. 2017))

Parameters	Factor 1	Factor 2	F1 × F2 factorial plane QLT
D	0.873	0.007	**879,213**
Lu_3	0.634	0.074	708,001
Lu_1	0.441	0.224	664,191
R	0.578	0.078	655,589
I_5	0.428	0.133	561,272
I_4	0.263	0.081	344,008

After all, the obtained specific vulnerability model, characterizes the Sidi Bouzid shallow aquifer using nitrate concentrations as pollutant factor. Several authors have used other specific methods in the study region, such as DRASTICU (Safa et al. 2020), studying the "assessment of urban groundwater vulnerability in arid areas: case of Sidi Bouzid aquifer (central Tunisia)." SVI method was used by Ncibi et al. (2021) for determining the "spatial variation of groundwater vulnerability to nitrate pollution under excessive fertilization using index overlay method in Central Tunisia (Sidi Bouzid basin)." The contribution of the modified DRASTIC method consists of the modification of the rates and the weights of each parameter using the nitrate concentrations. As well as the elimination of hydrogeological parameters which does not affect directly the water table, leading primarily to five parameters by the application of this method and conducting to two parameters by the application of the FACs model. The groundwater depth and land use have an important role in the groundwater vulnerability to pollution in this area. finally, the considered specific vulnerability model characterizes the study region using a little number of hydrogeological variables.

Conclusion

In the study area, the DRASTIC model application provided a relative satisfactory intrinsic vulnerability evaluation of groundwater to contamination. Given the pollution caused by the use of the chemical fertilizers and presented by the nitrate leaching from the surface to the groundwater. In fact, the intrinsic vulnerability map, using the standard DRASTIC method, shows that the study area is subdivided into three vulnerability classes: low class (3.5% of the study area), medium (71% of study area), and high vulnerability class located (25.5% of the total area). This latter characterizes especially the east part of the study area. The original DRASTIC model required modification to obtain more accurate results. It was applied to the DRASTIC method by the

revision of the rating and weighting scale of each parameter and the addition of land use parameter using nitrate concentration. The modified DRASTIC method was established to estimate the specific vulnerability and the pollution risk of nitrates in Sidi Bouzid shallow aquifer. This method subdivides the study region into three different vulnerability zones: a zone of low vulnerability covering 37% of the study area, a zone of medium vulnerability distributed to the north and the northeast of the region, which occupies 40% of the total surface, and the high vulnerability class, with 23% of the study area.

The Pearson's correlation of nitrate concentrations improved by 82% compared to the original DRASTIC method (44%), resulting from the use of nitrate concentration in the model calibration. Therefore, the nitrate concentration could be used as a modifying parameter to improve the vulnerability index. This modification could lead to more realistic management of groundwater quality.

Statistical analysis applied to determine the relationship between NO_3^- concentrations and selected parameters proves significant results principally for groundwater depth and land use parameters.

Computing data showed that nitrate concentration was successfully used as a modifying parameter with substantial improvement in the resulting index. This method is suggested for agricultural areas characterized by extensive nitrate use.

Declarations

References

Albinet M, Margat J (1970) Cartographie de la vulnérabilité à la pollution des nappes d'eau souterraine. Bull. B.R.G.M, 2, 3, 4, 13–22, Paris.

Aller L, Bennett T, Lehr JH, Petty RJ, Hackett G (1987) DRASTIC: a standardized system for evaluating ground water pollution potential using hydrogeologic settings. EPA-600/2–87–035, EPA, Washington, DC

Allouche N, Maanan M, Gontara M, Rollo N, Jmal I, Bouri S (2017) A global risk approach to assessing groundwater vulnerability. Environ Model Softw 88:168–182. https://doi.org/10.1016/j.envsoft.2016.11.023

Antonakos AK, Lambrakis NL (2007) Development and testing of three hybrid methods for assessment of aquifer vulnerability to nitrates, based on the DRASTIC model, an example from NE Korinthia. Greece J Hydrol 333(2–4):288–304. https://doi.org/10.1016/j.jhydrol.2006.08.014

Aydi W (2012) Evaluation of the groundwater vulnerability to pollution using an intrinsic and a specific method in a GIS Environment: application to the plain of Sidi Bouzid (Central Tunisia). Arab J Sci Eng 38:1815–1831. https://doi.org/10.1007/s13369-012-0417-9

Ayed B, Jmal I, Sahal S, Bouri S (2017) Assessment of groundwater vulnerability using a specific vulnerability method: case of Maritime Djeffara shallow aquifer (Southeastern Tunisia). Arab J Geosci 10(12):1–14. https://doi.org/10.1007/s12517-017-3035-8

Ayed B, Jmal I, Sahal S, Bouri S (2018) The seawater intrusion assessment in coastal aquifers using GALDIT method and groundwater quality index: the Djeffara of Medenine coastal aquifer (Southeastern Tunisia). Arab J Geosci 11(20):1–19

Azri C, Chaabane M, Medhioub K (2009) Diurnal evolutions of nitrogen oxides (NOx), ozone (O3) and PM10 particles at a busy traffic cross-road in the city of Tunis. Environ Prog Sustainable Energy 28(1):143–154. https://doi.org/10.1002/ep.10315

Boy-Roura M, Bernard TN, Anna M, Josep Mas-Pla (2013) Regression model for aquifer vulnerability assessment of nitrate pollution in the Osona region (NE Spain). J Hydrol (2013) https://doi.org/10.1016/j.jhydrol.2013.09.048

Brindha K, Elango L (2015) Cross comparison of five popular groundwater pollution vulnerability index approaches. J Hydrol 524:597–613. https://doi.org/10.1016/j.jhydrol.2015.03.003

Castany G (1982) Principes et méthodes de l'hydrogéologie. Dunod Université. 236p.

Castellazzi P, Schmid W (2021) Interpreting C-band InSAR ground deformation data for large-scale groundwater management in Australia. J Hydrol: Reg Stud 34:100774. https://doi.org/10.1016/j.ejrh.2021.100774

Dixon B (2004) Prediction of groundwater vulnerability using an integrated GIS-based neuro-fuzzy techniques. J Hydrol 4(309):17–38. http://spatialhydrology.net/index.php/JOSH/article/view/34.

Dutot AL, Elichegaray C, Vie Le Sage R (1983) Application de l'analyse des correspondances a l'étude de la composition physico-chimique de l'aérosol urbain. Atmos Environ 17(1):73–78. https://doi.org/10.1016/0004-6981(83)90009-4

Focazio MJ, Reilly TE, Rupert MG, Helsel DR (2002) Assessing ground-water vulnerability to contamination: providing scientifically defensible information for decision makers: U.S. Geological Survey Circular 1224, 33 p. https://pubs.usgs.gov/circ/2002/circ1224/pdf/circ1224_ver1.pdf.

Gogu RC, Dassargues A (2000) Current trends and future challenges in groundwater vulnerability assessment using overlay and index methods. Environ Geol 39 :549–59 https://doi.org/10.1007/s002540050466

Hrkal Z (2001) Vulnerability of groundwater to acid deposition, Jizerske Mountains, northern Czech Republic: construction and reliability of a GIS-based vulnerability map. Hydrogeol J 9:348. https://doi.org/10.1007/s100400100141

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. https://doi.org/10.1016/j.scitotenv.2012.08.037

Ouedraogo I (2017) Mapping groundwater vulnerability at the pan-African scale (Doctoral dissertation, Doctoral dissertation. UCL-Université Catholique de Louvain)

Javadi S, Kavehkar N, Mohammadi K, Khodadi A, Kahawita K (2011) Calibration DRASTIC using field measurements, sensitivity analysis and statistical method to assess groundwater vulnerability. Water Int 36(6):719–732. https://doi.org/10.1080/02508060.2011.610921

Javadi S, Kavehkar N, Mousavizadeh MH, Mohammadi K (2011a) Modification of DRASTIC model to map groundwater vulnerability to pollution using nitrate measurements in agricultural areas. J Agr Sci Tech 13(2):239–249 http://jast.modares.ac.ir/article-23-10411-en.html.

Jmal I, Ayed B, Boughariou E, Allouche N, Saidi S, Hamdi M, Bouri S (2017) Assessing groundwater vulnerability to nitrate pollution using statistical approaches: a case study of Sidi Bouzid shallow aquifer. Central Tunisia Arab J Geosci 10:364. https://doi.org/10.1007/s12517-017-3143-5

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. https://doi.org/10.1016/j.jhydrol.2015.03.035

Kendall MG (1975) Rank correlation methods, 4th edn. Griffin, London

Kura NU, Ramli MF, Ibrahim S, Sulaiman WNA, Aris AZ, Tanko AI, Zaudi MA (2015) Assessment of groundwater vulnerability to anthropogenic pollution and seawater intrusion in a small tropical island using index-based methods. Environ Sci Pollut Res 22(2):1512–1533. https://doi.org/10.1007/s11356-014-444-0

Landau S, Everitt BS (2003) A handbook of statistical analyses using SPSS. CRC Press Company, London, UK

McLay CDA, Dragten R, Sparling G, Selvarajah N (2001) Predicting groundwater nitrate concentrations in a region of mixed agricultural land use: a comparison of three approaches. Environ Pollut 115(2):191–204. https://www.ncbi.nlm.nih.gov/pubmed/11706792.

Mustapha A, Abdu A (2012) Application of principal component analysis & multiple regression models in surface water quality assessment. J Environ Earth Sci 2:16–23 https://www.iiste.org/Journals/index.php/JEES/article/view/1516.

Basu S and abd Lokesh KS (2014) Application of multiple linear regression and manova to evaluate health impacts due to changing river water quality. Appl Math 5:799–807. https://doi.org/10.4236/am.2014.55076

Nasri G, Hajji S, Aydi W, Boughariou E, Allouche N, Bouri S (2021) Water vulnerability of coastal aquifers using AHP and parametric models: methodological overview and a case study assessment. Arab J Geosci 14(1):1–19. https://doi.org/10.1007/s12517-020-06390-8

National Institute of Meteorology NIM (2018) Monthly climatological table (1975–2018) for Sidi Bouzid station, Tunisia.

National Research Council (1993) Groundwater vulnerability assessment: predicting relative contamination potential under conditions of uncertainty. National Academy Press, Washington, DC, Committee for assessing ground water vulnerability

Ncibi K, Chaar H, Hadji R et al (2020) A GIS-based statistical model for assessing groundwater susceptibility index in shallow aquifer in Central Tunisia (Sidi Bouzid basin). Arab J Geosci 13:98. https://doi.org/10.1007/s12517-020-5112-7

Ncibi K, Hadji R, Hajji S et al (2021) Spatial variation of groundwater vulnerability to nitrate pollution under excessive fertilization using index overlay method in central Tunisia (Sidi Bouzid basin). Irrig Drain 2021:1–18. https://doi.org/10.1002/ird.2599

Neshat A, Pradhan B, Dadras M (2014) Groundwater vulnerability assessment using an improved DRASTIC method in GIS. Resour Conserv Recycl 86:74–86. https://doi.org/10.1016/j.resconrec.2014.02.008

Neshat AR, 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 1–13https://doi.org/10.1007/s12665-013-2690-7

Pacheco FAL (1998) Application of correspondence analysis in the assessment of groundwater chemistry. Math Geol 30 2 https://doi.org/10.1023/A:1021718929576

Pal SC, Das B, Malik S (2019) Potential landslide vulnerability zonation using integrated analytic hierarchy process and GIS technique of Upper Rangit catchment area, West Sikkim, India. J Indian Soc Remote Sens https://doi.org/10.1007/s12524-019-01009-2.

Panagopoulos GP, Antonakos AK, Lambrakis NJ (2006) Optimization of the DRASTIC method for groundwater vulnerability assessment via the use of simple statistical methods and GIS. Hydrol J 14:894–911. https://doi.org/10.1007/s10040-005-0008-x

Pathak H (2012) Evaluation of ground water quality using multiple linear regression and mathematical equation modeling. Annals Univ Oradea—Geography Series, 2, 304–307. http://geografie-uoradea.ro/Reviste/Anale/Art/2012-2/AUOG_595_Hemant.pdf.

Rahman A (2008) A GIS based DRASTIC model for assessing groundwater vulnerability in shallow aquifer in Aligarh, India. Appl Geogr 28:32–53. https://doi.org/10.1016/j.apgeog.2007.07.008

Roohollah N, Hooman G, Bjørn K, Jan FA, Akbar B (2019) Modified-DRASTIC, modified-SINTACS and SI methods for groundwater vulnerability assessment in the southern Tehran aquifer. J Environ Sci Health A 54(1):89–100. https://doi.org/10.1080/10934529.2018.1537728

Rupert MG (2001) Calibration of the DRASTIC groundwater vulnerability mapping method. Groundwater 39(4):625–630. https://doi.org/10.1111/j.1745-6584.2001.tb02350.x

Safa G, Najiba C, El Houda BN, Monji H, Soumaya A, Kamel Z (2020) Assessment of urban groundwater vulnerability in arid areas: case of Sidi Bouzid aquifer (central Tunisia). J Afr Earth Sci https://doi.org/10.1016/j.jafrearsci.2020.103849

Saidi S, Bouri S, Dhia HB, Anselme B (2011) Assessment of groundwater risk using intrinsic vulnerability and hazard mapping: application to Souassi aquifer, Tunisian Sahel. Agr Water Manage 98:1671–1682. https://doi.org/10.1016/j.agwat.2011.06.005

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 Manag 54:39–57. https://doi.org/10.1006/jema.1998.0221

Sinha MK, Verma MK, Ahmad I et al (2016) Assessment of groundwater vulnerability using modified DRASTIC model in Kharun Basin, Chhattisgarh. India Arab J Geosci 9:98. https://doi.org/10.1007/s12517-015-2180-1

Smida H (2008) Apports des Systèmes d'Informations Géographiques (SIG) pour une approche intégrée dans l'étude et la gestion des ressources en eau des systèmes aquifères de la région de Sidi Bouzid (Tunisie centrale). Université de Sfax, Thèse de la Faculté des Sciences de Sfax, p 283p

Sophocleous MA (1991) Combining the soil water balance and 475 water level fluctuation methods to estimate natural groundwater recharge: practical aspects. J Hydrol 124:229–241. https://doi.org/10.1016/0022-1694(91)90016-B

STATIT-CF (1987) Services des études statistiques de l'Institut Technique des Céréales et Fourrages (I.T.C.F.), Boigneville

Stevenazzi S, Masetti M, Pietro Beretta G (2017) Groundwater vulnerability assessment: from overlay methods to statistical methods in the Lombardy Plain area. Italian J Groundw - AS21- 276:17–27.

Vu TD, Ni CF, Li WC, Truong MH (2019) Modified index-overlay method to assess spatial–temporal variations of groundwater vulnerability and groundwater contamination risk in areas with variable activities of agriculture developments. J Water 11:2492. https://doi.org/10.3390/w11122492

Wang J, He J, Chen H (2012) Assessment of groundwater contamination risk using hazard quantification, a modified DRASTIC model and groundwater value, Beijing Plain. China Sci Total Environ 432:216–226. https://doi.org/10.1016/j.scitotenv.2012.06.005

Wilcoxon F (1945) Individual comparisons by ranking methods. Biometr Bull 1:80–83

Zaporozec A (2002) Groundwater contamination inventory: a methodological guide. IHP-VI, Series on groundwater No. 2. Paris, France: UNESCO-160