A comparison between index of entropy and catastrophe theory methods for mapping groundwater potential in....

2 authors:

Alaa M. Al-Abadi
University of Basrah
20 PUBLICATIONS   76 CITATIONS

SEE PROFILE

Shamsuddin Shahid
Universiti Teknologi Malaysia
158 PUBLICATIONS   1,640 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Study the climate change effect on storm drainage networks by storm water management model [swmm] View project

Project Determination of Lead in Water, Roadside Soil, Plant and Blood in Basrah City – IRAQ View project

CrossMark

A comparison between index of entropy and catastrophe theory methods for mapping groundwater potential in an arid region

Alaa M. Al-Abadi · Shamsuddin Shahid

© Springer International Publishing Switzerland 2015

Abstract In this study, index of entropy and catastrophe theory methods were used for demarcating groundwater potential in an arid region using weighted linear combination techniques in geographical information system (GIS) environment. A case study from Badra area in the eastern part of central of Iraq was analyzed and discussed. Six factors believed to have influence on groundwater occurrence namely elevation, slope, aquifer transmissivity and storativity, soil, and distance to fault were prepared as raster thematic layers to facility integration into GIS environment. The factors were chosen based on the availability of data and local conditions of the study area. Both techniques were used for computing weights and assigning ranks vital for applying weighted linear combination approach. The results of application of both modes indicated that the most influential groundwater occurrence factors were slope and elevation. The other factors have relatively smaller values of weights implying that these factors have a minor role in groundwater occurrence conditions. The groundwater potential index (GPI) values for both models were classified using natural break classification scheme into five categories: very low, low, moderate, high, and very high. For validation of generated GPI, the relative operating characteristic (ROC) curves were used. According to the obtained area under the curve, the catastrophe model with 78 % prediction accuracy was found to perform better than entropy model with 77 % prediction accuracy. The overall results indicated that both models have good capability for predicting groundwater potential zones.

Introduction

The occurrence and distribution of groundwater is a consequence of a finite combination of different factors that together form an integrated dynamic system. These interrelated factors interact in a manner by which the performance of the groundwater system could be predicted. Thus, it is possible to evaluate the general groundwater potential yield of an area by studying these factors and their interrelationship. The term *groundwater potential* denotes the amount of groundwater available in an area, and it is a function of several hydrologic and hydrogeological factors (Jha et al. 2010). There are several methods such as geological, hydrogeological, geophysical, and their combinations which have been employed to delineate groundwater potential zones (Prasad et al. 2008; Pandey and Kazama 2012; Pandey et al. 2013). Geographical information system (GIS) can be used effectively for this purpose to combine different controlling factors of groundwater occurrence

A. M. Al-Abadi (✉)
Department of Geology, College of Sciences, University of Basra, Basra, Iraq
e-mail: alaaatiaa@gmail.com

S. Shahid
Faculty of Civil Engineering, Universiti Teknologi Malaysia, Johor 81310, Malaysia
e-mail: sshahid_ait@yahoo.com

Springer

Table 1 A list of seven catastrophe models (after Kam 1992)

Catastrophe model	Energy function
Fold	$1/3x^3 + ax$
Cusp	$1/4x^4 + 1/2ax^2 + bx$
Swallowtail	$1/5x^5 + 1/3ax^3 + 1/2bx^2 + cx$
Butterfly	$1/6x^6 + 1/4ax^4 + 1/3bx^3 + 1/2cx^2 + dx$
Parabolic umbilici	$x^2y + y^4 + ax^2 + by^2 + cx + dy$
Hyperbolic umbilici	$x^3 + y^3 + axy + bx + cy$
Elliptic umbilici	$1/3x^3 - xy^2 + a(x^2 + y^2) + bx + cy$

objectively and analyze those systemically for delineating groundwater potential zones (Shahid et al. 2000). The common method for integrating several factors in GIS is through weighted linear combination approach, in which the importance of each factor is determined through assigning appropriate weights for factor and their attributes (classes). The multi-criteria decision techniques (MCDM) such as analytical hierarchy process (AHP) or personal judgments based on expert's opinion are often used to assign appropriate weights prior to integrate factor layers in GIS. MCDM approaches provide a systematic procedure to help decision makers to choose the most desirable and satisfactory alternative under uncertain situation (Chen 2000). The AHP and its derivatives give a flexible, low cost, and easily understandable output for complicated decision making problems (Satty 1980). The major drawback of AHP is related with its dependency on the expert's knowledge which is the main source of uncertainty (Chowdary et al. 2013).

In order to avoid subjectivity in assessing weights for groundwater occurrence factors, Ahmed et al. (2014) successfully applied a method based on catastrophe theory. Catastrophe theory is designed to deal with discontinuous dynamic systems governed by a potential energy-like function (Wang et al. 2011). The proposed approach does not involve the decision maker's opinion

in assigning weights for the factors; instead, it calculates the importance of one criterion over others by its inner mechanism and thus greatly reduces the subjectivity (Yang et al. 2012). In the same context, many authors attempted to demarcate groundwater potential yield and groundwater spring potentiality using different data-driven and knowledge-driven techniques such as frequency ratio, weights of evidence, logistic regression, index of entropy, evidential belief functions, artificial neural networks, and fuzzy logic (Corsini et al. 2009; Oh et al. 2011; Ozdemir 2011a, b; Manap et al. 2011; Lee et al. 2012; Moghaddam et al. 2013; Pourtaghi and Pourghasemi 2014; Naghibi et al. 2014; Elmahdy and Mohamed 2014; Shahid et al. 2002; Nampak et al. 2014; Al-Abadi 2015). The idea behind these techniques is to explore the relationship between groundwater springs or productive borehole locations and influential groundwater occurrence factors. The hydrogeologist inspired these techniques from their applications in other fields of earth sciences such as landslide susceptibility, flood vulnerability hazards studies, and ore deposits potentiality. Index of entropy (also called Shannon's entropy) is the average unpredictability in a random variable, which is equivalent to its information content. The entropy of groundwater potential refers to the extent that the various controlling groundwater occurrences influence the groundwater productivity. Several influencing factors give extra entropy into the index system. Therefore, the entropy value can be used to calculate objective weights of the index system (Jaafari et al. 2013). The application of this method for demarcating groundwater potential zones is still limited. The studies by Naghibi et al. (2014) and Al-Abadi (2015) successfully applied this technique for demarcating groundwater qanat potential and groundwater potential yield, respectively.

In this study, a comparison between catastrophe theory and index of entropy methods is made for demarcating groundwater potential yield in Badra area in the eastern part of central of Iraq. The demand for water has

Table 2 Normalization formulas for catastrophe theory (after Ching et al. 1996)

Control variable	State variable	Catastrophe model	Normalization formula
2	1	Cusp	$x_a = a^{0.5}$ and $x_b = b^{0.33}$
3	1	Swallowtail	$x_a = a^{0.5}$, $x_b = b^{0.33}$, and $x_c = c^{0.25}$
4	1	Butterfly	$x_a = a^{0.5}$, $x_b = b^{0.33}$, $x_c = c^{0.25}$, and $x_d = d^{0.20}$
5	1	Wigwam	$x_a = a^{0.5}$, $x_b = b^{0.33}$, $x_c = c^{0.25}$, $x_d = d^{0.20}$, and $x_e = e^{0.17}$

progressively increased in the study area for various fields such as agriculture, industry, and domestic. The improper management of aquifer in the study area may lead to mining of groundwater reserve and deterioration
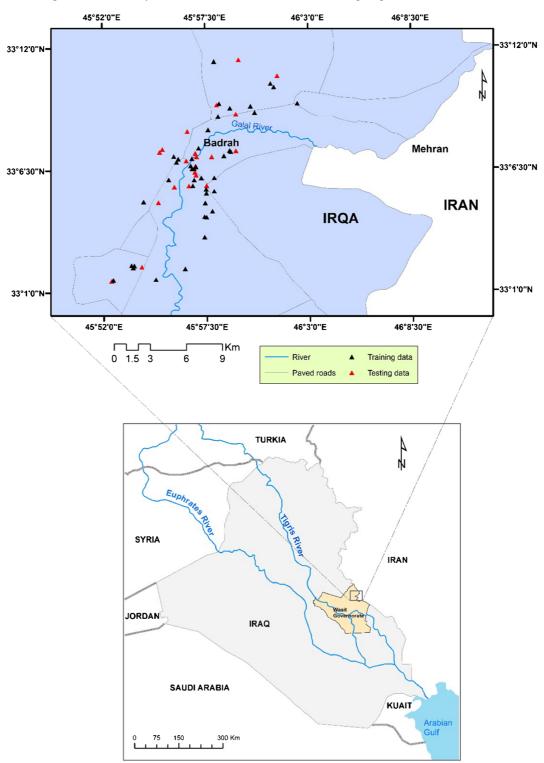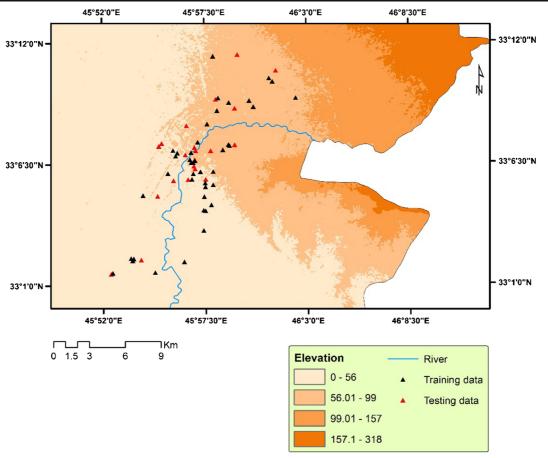

Fig. 1 Location of the study area

Fig. 2 Elevation range in the study area

in groundwater quality. Therefore, it is vital to develop groundwater resource in order to manage this finite resource properly. The first step in this context is to demarcate groundwater potential yield of the aquifer system in the study area. The results of this study could help hydrogeologist to put suitable management scenarios incorporating growing challenges which faced by the water sector in Iraq and especially in Badra. It is also expected that the methodologies presented in this paper can be used as a guiding tool for more accurate assessment of groundwater potential in other regions.

Methods and materials

Modeling techniques

Catastrophe theory

Catastrophe theory is a branch of applied mathematics that was originated by Rene Thom in order to describe certain kinds of biological processes. The theory can be applied to systems where there are sudden, abrupt changes, i.e., changes that are not smooth and continuous (Kam 1992). It is a theory for study of dynamical

Table 3 Description of the lithological units in the study area

Formation	Age	Environment	Description
Injana	Upper Miocene	Sub-marine	Red or gray colored silty marl or clay stones and purple silt stones
Muqdadyia	Pliocene	Continental	Gravely sandstone, sandstone, and red mudstone
Quaternary	Pleistocene-Holocene	Continental	Mixture of gravel, sand, silt, and clay

systems in general, and for the description of discontinuities in particular (Ghorbani et al. 2010). In catastrophe theory, system function variables are divided into *dependent state variable*, which are the internal token variables of system, and *control variables*, which are the external influence factors while system is running (Hui 2008). Therefore, catastrophe theory deals with the discontinuous changes in dependent variables caused only by continuous changes in control variables and not by discontinuous changes in control variables. The dependency of dependent variable is determined by catastrophe fuzzy membership functions, rather than weights assigned by the user (Ahmed et al. 2014). To solve complicated problems using this theory, the dynamic system must be divided into subsystems; each subsystem consists of a number of evaluation indicators. The original data is first normalized in the range from 0 to 1 using catastrophe theory and fuzzy mathematics (Chen et al. 2006). The multidimensional catastrophe fuzzy membership functions assign values ranging from 0 to 1 to resolve incompatibility of various initial data (Wang et al. 2011 in

Ahmed et al. 2014). There are seven catastrophe models, namely fold, cusp, swallowtail, butterfly, hyperbolic umbilical, elliptic umbilical, and parabola umbilical. The list of seven catastrophe models is given in Table 1. The normalization forms of the four catastrophe models which describe all possible discontinuities controlled by no more than four variables are given in Table 2.

Application of this method for delineating groundwater potential zones involves four main steps (Ahmed et al. 2014): (i) selection of groundwater occurrence subsystems (factors); (ii) data standardization; (iii) process of data normalization using catastrophe models; (iv) computation for groundwater potential zone. The following paragraphs illustrate these steps in detail.

(i)   Selection of groundwater occurrence indicators

In general, the type and number of groundwater occurrence indicators (factors) used for assessing groundwater potential vary considerably from one
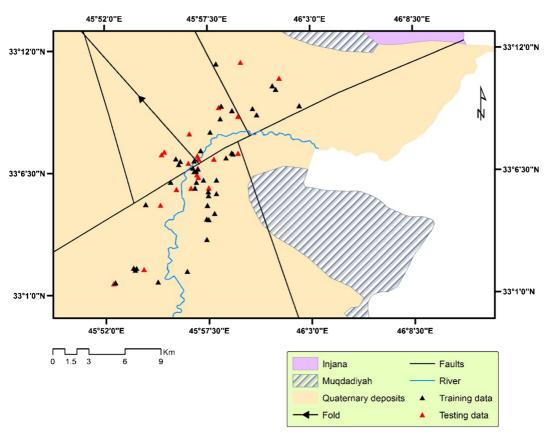

Fig. 3 Geology of the study area

study to another and their selection is arbitrary (Machiwal et al. 2010). The availability of data and the local conditions of an area are the main constrains to use factors in groundwater potential studies. In this study, six factors were selected based on expert opinion (Jabar Al-Saydi, Expert, Head of Groundwater Commission of Groundwater/Basra Branch, personal communication) and literature review. These factors were elevation (m), slope (°), aquifer transmissivity (m²/day), storativity (dimensionless), soil infiltration capacity (mm/h), and distance to faults (m). Elevation is an important factor on groundwater occurrence because weather and climatic conditions vary greatly at different elevation, and this caused differences in soil and vegetation (Aniya 1985). Slope is a rise or fall of land surface. It is an important factor for groundwater potential mapping studies because it controls accumulation of water in an area and, hence, enhances the groundwater recharge (Ozdemir 2011a). Soil has an impact on the amount of recharge, which can infiltrate to the groundwater

and thus groundwater storage and aquifer yield. The transmissivity and storativity are very important factors for modeling groundwater productivity because they control the ability of a specific water bearing layer to transmit and store water, respectively. Therefore, they play very important role in aquifer yield and groundwater potentiality. It is known that structure setting controls the occurrence and movement of groundwater. Most rocks possess fractures and other discontinuities which facilitate storage and movement of fluids through them. On the other hand, some discontinuities, e.g., faults and dykes, may also act as barriers to water (Singhal and Gupta 1999). Therefore, it is consider in this study.

(ii)   Standardization

The standardization process makes the data dimensionless and therefore the comprising between variables becomes easier. In the present study, the following standardization formulas were used (Li et al. 2010):
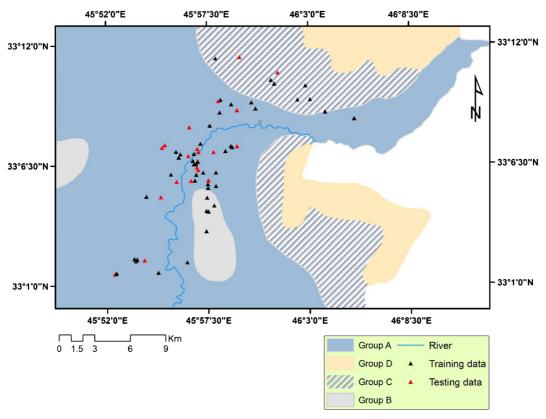

Fig. 4 Hydrological soil group in the study area

For the cost type (larger the better):

$$y_i = \frac{x_i - x_{i(\min)}}{x_{i(\max)} - x_{i(\min)}} \qquad (1)$$

For the efficiency type (smaller the better)

$$y_i = \frac{x_{i(\max)} - x_i}{x_{i(\max)} - x_{i(\min)}} \qquad (2)$$

where i is the index or attribute, x_i is the original value of i, and $x_{i(\max)}$ and $x_{i(\min)}$ are the maximum and minimum values of original data. The Eqs. 1 and 2 normalize the original data into range from 0 to 1.

(iii)   Normalization

The normalization is done using the catastrophe models listed in Table 2. It is the heart of all models developed by catastrophe theory. The catastrophe progression of each control variable is computed from the initial fuzzy subordinate function on normalization formulas (Ahmed et al. 2014). There are two different types of principles:

Table 4 Data used to interpolate groundwater depths and hydraulic characteristics over the study area

Geographic location (UTM)		Borehole	Depth	Transmissivity	Storativity
Easting	Northing				
609382.4	3670038	w1	162	67.32	0.0037
599476.9	3668423	w2	104	20.38	–
590242.3	3666824	w3	–	–	0.0012
587399.2	3663102	w4	61.1	69.03	–
589606.9	3659796	w5	62	263.5	–
587933.2	3655592	w6	39	40.15	–
585244.3	3654552	w7	26	56.75	–
594676.7	3655747	w8	42.4	100.71	–
605111.5	3645504	w9	16	30.2	–
615758.6	3645593	w10	29	30.96	0.00023
586243.6	3685298	w12	81	34.78	0.056
589103	3683722	w13	75.1	16.33	0.039
588287.6	3673797	w14	60	65.88	0.047
597930.5	3668342	w15	108.6	220.32	0.038
597514.2	3667787	w16	101.1	362.88	0.032
599278.4	3667558	w17	114.9	163.01	0.014
597858.7	3665116	w18	112.3	147.74	0.011
588012.3	3661229	w20	57.5	362.22	0.051
586661.4	3667376	w21	50	46.66	0.047
586038.1	3664599	w22	49	1936.06	0.043
587583.1	3662826	w23	60.6	132.7	0.041
589954	3664357	w24	70	98	0.046
588168.8	3664002	w25	71.5	64.46	0.051
590201.7	3662789	w26	65	100.83	0.04
589430.3	3662104	w27	58	117.41	0.044
584925.6	3658399	w28	40.5	23.03	0.038
581443.1	3646758	w29	32.8	7.57	0.073
612992.6	3646731	w32	32	211.9	0.039
		Minimum	16	7.57	0.00023
		Maximum	162	1936.06	0.073
		Average	65.98	177.44	0.036

complementary and non-complementary. In the complementary rule, the control variables of a system such as a, b, c, and d offset each other; thus, each of them tends to reach the average value $x=(x_a+x_b+x_c+x_d)/4$ (Zhang et al. 2009). In contrast, the non-complementary rule means that control variable cannot complement each other; therefore, the value of the state variable is the smallest values of a system ($x=\min\ \{x_a,\ x_b,\ x_c,\ x_d\}$). In this study, the first rule

(complementary rule) was used to compute the catastrophe progression of each control variable.

(iv)   Computation of groundwater potential index

The final step is to compute groundwater potential index (GPI) using linear combination method in GIS environment (Malczewski 1999):
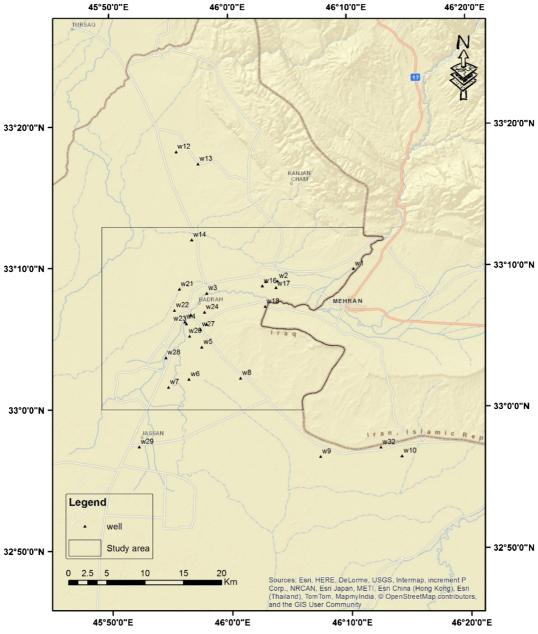

Fig. 5 Geographic locations of wells used for interpolate groundwater depths and hydraulic characteristics

$$\text{GPI} = \sum_{j=1}^{m} \sum_{i=1}^{n} \left(w_j r_i \right) \qquad (3)$$

where w and r are the weight and rank, respectively: weight for each groundwater factor j and rank for each class in groundwater thematic raster layer i. n is the total number of classes in a theme and m is the total number of themes involving in the analysis.

Index of entropy

In information theory, entropy is a measure of uncertainty in a random variable (Ihara 1993). The entropy indicates the extent of the instability, disorder, imbalance, and uncertainty of a system (Shi and Jin 2009). Index of entropy is the average unpredictability in a random variable, which is equivalent to its information content. The following equations are used to calculate the information coefficient w_j (weigh for each factor) (Bednarik et al. 2010, 2012; Constantin et al. 2011; Jaafari et al. 2013):

$$P_{ij} = \text{FR} = \frac{A/B}{C/D} = \frac{b}{a} \qquad (4)$$

$$\left(P_{ij} \right) = \frac{P_{ij}}{\sum\limits_{j=1}^{Sj} P_{ij}} \qquad (5)$$

$$H_j = -\sum_{i=1}^{Sj} \left(P_{ij} \right) \log_2 \left(P_{ij} \right), j = 1, \ldots, n \qquad (6)$$

$$H_{j\max} = \log_2 Sj \qquad (7)$$

$$I_j = \frac{H_{j\max} - H_j}{H_{j\max}}, I = (0, 1), j = 1, \ldots, n \qquad (8)$$

$$w_j = I_j P_{ij} \qquad (9)$$

where FR is the frequency ratio, A is the area of a class for the groundwater factor, B is the total area of the factor, C is the number of pixels in the class area of the factor, D is the number of total pixels in the study area, b is the percentage
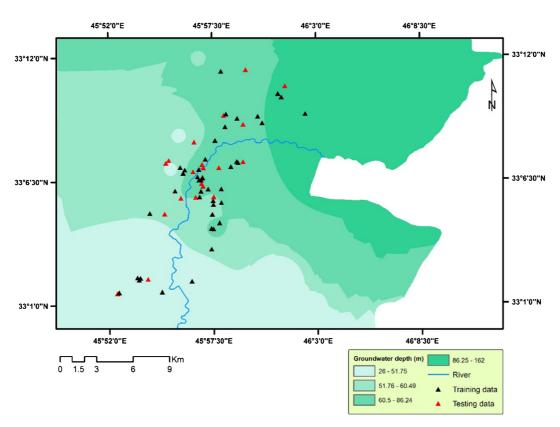

Fig. 6 Spatial distribution of groundwater depths (m) in the study area

for area with respect to a class for the factor and a is the percentage for the entire domain, (P_{ij}) is the probability density, H_j and $H_{j\max}$ refer to entropy values, Sj is the number of classes, I_j is the information coefficient, and w_j is the resultant weight value for the factor as a whole. The range of w_j is between 0 and 1. The final groundwater productivity index is also calculated using Eq. 3.

The relative operating characteristics

Any predictive model (deterministic or stochastic) requires validation before it can be used in prediction purposes. Without validation, the model has no scientific significance (Chung and Fabbri 2003). To examine the accuracy of the catastrophe and index of entropy models, the relative operating characteristic (ROC) was used in this study. The ROC is a common technique to assess the accuracy of a diagnostic test (Egan 1975). It a graphical chart that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is constructed by plotting the trade-off between the false-positive rate (also called sensitivity) on x-axis and true positive rates (also called 1-specificity) on y-axis at various threshold settings. The area under the ROC curve (AUC) characterizes the quality of a forecast system by describing the system's ability to anticipate the correct occurrence or non-occurrence of pre-defined "events" (Pourghasemi et al. 2013). The range of AUC is from 0 to 1. The relationship between AUC and prediction accuracy can be summarized as follows (Yesilnacar and Topal 2005): poor (0.5–0.6); average (0.6–0.7); good (0.7–0.8); very good (0.8–0.9); and excellent (0.9–1).

The study area

General description

The study area covers an area of '707 km² and lies between 33°00′ and 33°14′ latitude and 45°50′ and
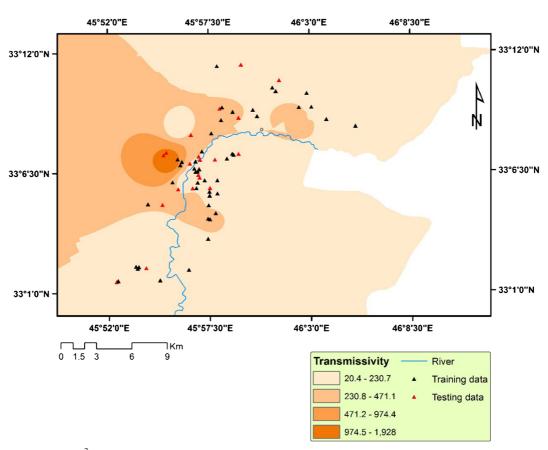

Fig. 7 Transmissivity (m²/day) in the study area

46°16′ longitude in the northeastern Wasit Governorate of Iraq, Fig. 1. The major portion of the study area is flat and featureless. Relief is low with only a few isolated hills rising above the general level of the plain in the east (Parsons 1956). Elevation in the study area ranges from 0 to 318 m with an average of 70 m above sea level, Fig. 2. The study area is generally hot and dry. It is characterized by absence of rainfall in summer (June–September) with rainy season begins from autumn to spring (October–May). The area receives an average annual rainfall of approximately 212 mm/year with an uneven rainfall distribution between plain and mountain parts. Drainage in the area is almost in a southwesterly direction (Parsons 1956). The nature of the streams is intermittent and terminates in the temporary marshes on the delta plain. The major stream in the study area is the Galal-Badra River. The mean monthly discharge of this river is 2.5 and 1000 m^3/s in drought and flood periods, respectively (Al-Shammary 2006). Rocks in the study area range in age from Upper Miocene to Recent. In the western portion, the younger rocks are exposed and increasingly

become old to the east. Most of the area is covered by rocks of alluvial and lacustrine origin, Pliocene or younger in age. The stratigraphic succession composed from Injana, Mukdadiya formations in addition to the Quaternary deposits. A brief description of these lithological units is summarized in Table 3. Tectonically, the platform of the Iraqi territory is divided into two basic units: the stable and unstable shelf (Jassim and Goff 2006). The stable shelf is characterized by reduced thickness of the sedimentary cover and by the lack of folding, while the unstable shelf has a thick and folded sedimentary cover. Folds are arranged in narrow long anticlines and broad flat synclines (Al-Sayab et al. 1982). The greater parts of the study are located in the stable shelf (Mesopotamian plain) and only a small part extends over the unstable shelf close to the Iraq-Iran border (folded zone). There are many faults in the study area; the bigger and important one is Shbichia–Najaf fault. The geological map of the study area is shown in Fig. 3. The soil of the study area formed from the processes of weathering, erosion, and sedimentation during the Quaternary period.
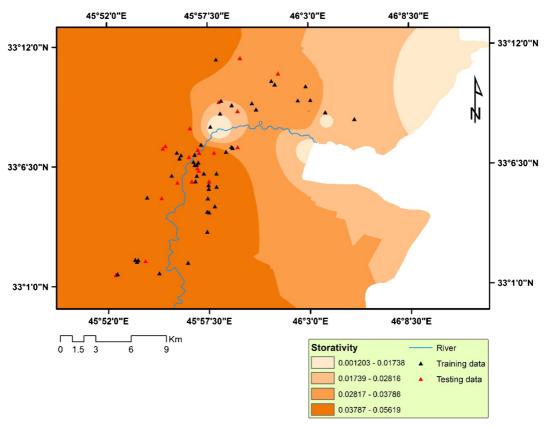

Fig. 8 Storativity (dimensionless) in the study area

Soils are classified into four hydrologic soil groups (HSG's) to indicate the minimum rate of infiltration for bare soil after prolonged wetting (USDA 1986). The four hydrologic soils groups are A, B, C, and D, where A is generally has the greatest infiltration rate (smallest runoff potential) and D is the smallest infiltration rate (greatest runoff potential). The hydrologic soil group map of the study area was shown in Fig. 4, in which the major portion of the study area (about ~60 %) has high infiltration rate (A and B groups).

Aquifer system

The aquifer system in the study area consists of two hydrogeological units. The first one represents the shallow unconfined aquifer consisting mainly from layers of sand, gravel with overlapping clay and silt. This hydrogeological unit is located within the Quaternary lithological layers. The second hydrogeological unit is Mukdadiya water bearing layer. The aquifer condition

of this unit is confined/semi-confined. The regional groundwater flow is from northeast to southwest. The available data used for interpolate groundwater depths and hydraulic characteristics were taken from the General Commission of Groundwater/Wasit Branch, Iraq, and Al-Shammary (2006) work, Table 4. The data include geographic locations of the wells, groundwater depths (m), transmissivity (m/s), and storativity (dimensionless). The well locations covered the interested region and beyond (Fig. 5), and different from the data used for the rest of the analysis. Depths to groundwater ranges from 26 to 162 m with an average of 33.14 m, Fig. 6. The hydraulic characteristics of the two hydrogeological units were estimated by Al-Shammary (2006) by means of pumping test. For the unconfined aquifer, the hydraulic conductivity, transmissivity, and specific yield were 6.3 m/day, 228.43 m^2/day, and 0.012, respectively. For the confined aquifer, the values were 3.5 m/day, 81.07 m^2/day, and 0.0017 for hydraulic conductivity, transmissivity, and storage coefficient,
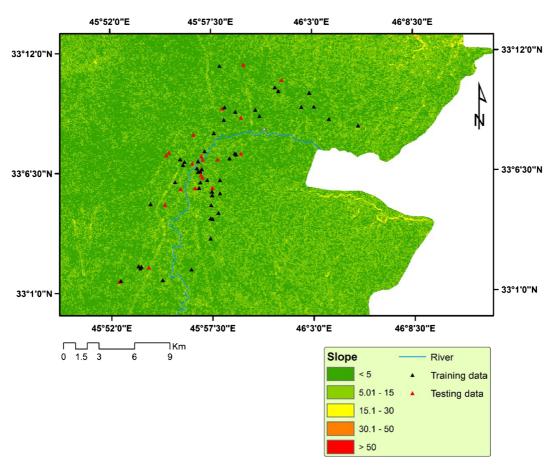

Fig. 9 Slope (in °) in the study area

respectively. The available data of transmissivity and storativity were interpolated using weighted inverse distance method to produce maps (Figs. 7 and 8) of these hydraulic characteristics over the study area. In general, the hydraulic characteristics of the aquifer system are good in the middle and western parts of the study area and become poor in the eastern part.

Generation of thematic raster layers

Thematic layer of topographic factors, i.e., elevation and slope angle, were prepared from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) data. The ASTER data is available from web site (http://gdem.ersdac.jspacesystems.or.jp/search.jsp) with 30 m resolution. Four tiles were downloaded and then merged to create a new elevation raster. The new elevation raster was then reprojected, fill sinks, and clipped for the study area polygon. Elevation

thematic layer was directly created from DEM and was classified into four classes, Fig. 2. The natural break classification scheme in ArcGIS 10.2 was used in this study to reclassify all continuous values of used factors. Selection of this classification scheme was based on literature reviews and author's experience of study area and its condition. The thematic raster layer of slope of the study area (Fig. 9) was prepared from ASTER DEM data and classified into five classes; flat-gentle slope (0–5°), fair slope (5–15°), moderate slope (15–30°), steep slope (30–50°), and very steep slope >50° (Pourghasemi et al. 2013). The slope angles in most parts of the study area (99 %) belong to flat-gentle and fair slopes classes.

The soil map was prepared from hydrologic soil group map of the study area, (Fig. 4) by converting polygons to raster thematic layer. The transmissivity and storativity of the aquifer system in the study area were classified into four classes (Figs 7 and 8). Map of distance from faults (Fig. 10) was prepared by applying
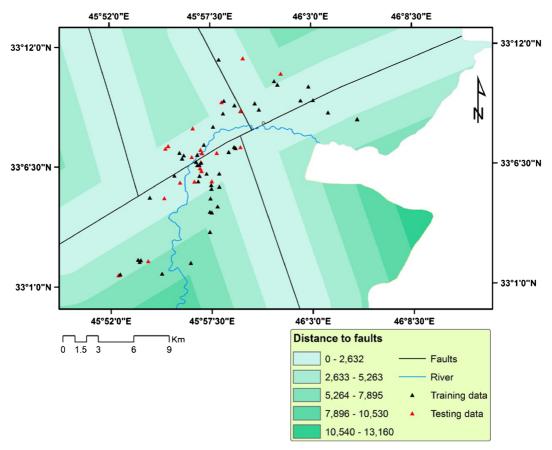

Fig. 10 Distance to faults (m)

the distance command in spatial analyst extension of ArcGIS 10.2 and then classified into five classes.

All previous six layers were prepared using ArcGIS 10.2 with (30×30) grid size. The used projected coordinate system was Universal Transverse Mercator (UTM, WGS 1984, 38N). The total number of cells was 784,910. For application of index of entropy approach, an inventory of existing borehole data is required. The productive borehole data in the study area were obtained from General Commission of Groundwater/Ministry of Water Resource/Iraq. Only boreholes with high flow rate (>8 l/s) were used in the rest of the analysis (Al-Abadi 2015). These data were randomly divided into two sets: 47 (70 %) boreholes for training and 21 (30 %) for testing using MINITAB 16 commercial software.

Results and discussion

The application of catastrophe theory for demarcating groundwater potential yield in the study area is presented in Table 5. The steps of calculation are summarized below. The subsystems (groundwater occurrence) and their classes (indicators) are given in columns 1 and 2, respectively. The elevation, soil, transmissivity, and storativity meet the butterfly model, while the slope and distance to faults meet the Wigwam model. The index values (column 3) are the average values for different class range for each subsystem. The procedure followed to compute these average values involving multi-steps. First, the raster map of the groundwater factor (for example slope) was reclassified and then converted to polygon. After

Table 5 Results of application of catastrophe model

Subsystem	Range		Index value	Standardized value	Catastrophe model	Calculated weight (normalization value)	Rank
Elevation (B1)	0–56	C1	36.48	1.00	Butterfly	0.68	5
	56.01–99	C2	76.31	0.91			
	99.01–157	C3	123.39	0.81			
	157.1–318	C4	191.27	0.00			
Slope (B2)	0–5	C5	2.86	1.00	Wigwam	0.74	6
	5.001–15	C6	7.34	0.97			
	15.01–30	C7	18.07	0.91			
	30.01–50	C8	34.01	0.81			
	50.01–70	C9	51.34	0.00			
HSC (B3)	A	C10	10	1.00	Butterfly	0.65	3
	B	C11	8	0.92			
	C	C12	2.5	0.68			
	D	C13	0.5	0.00			
Transmissivity (B4)	20.4–230.7	C14	149.84	0.00	Butterfly	0.58	2
	230.8–471.1	C15	314.23	0.52			
	471.2–974.4	C16	622.38	0.80			
	974.5–1,928	C17	1320.30	1.00			
Storativity (B5)	0.001203–0.01738	C18	0.011	0.00	Butterfly	0.66	4
	0.01739–0.02816	C19	0.023	0.73			
	0.02817–0.03786	C20	0.032	0.91			
	0.03787–0.05619	C21	0.043	1.00			
Distance to faults (B6)	0–2632	C22	1251.31	1.00	Wigwam	0.57	1
	2633–5263	C23	3759.47	0.25			
	5264–7895	C24	6347.48	0.84			
	7896–10,530	C25	8891.09	0.76			
	10,540–13,160	C26	11,516.32	0.00			

that, the *Zonal Statistics as Table* in ArcToolbox was applied to produce a table containing the summary of statistical parameters of each class. The index values were standardized using Eqs. 1 or 2 (in column 4 of Table 5). The slope subsystem is described here to illustrate the standardization process. In this subsystem, there were five index values; the minimum and maximum values were 2.86 and 51.34, respectively. The slope was considered as "smaller the better," i.e., as the slope increases the groundwater potential becomes poor, and therefore, Eq. 2 was used for standardization. The produced value for the first indicator was 1 ($(51.34-2.86/51.34-2.86)=1$), for the second indicator was 0.97 ($(51.34-7.34/51.34-2.86)=0.97$), and so on. After standardization process, the normalized values (column 5) were obtained by applying catastrophe models given in Table 2. For example, the slope subsystem follows the Wigwam catastrophe model, and therefore, the normalized value were calculated as, $XC5=1^{0.5}=1$, $XC6=0.91^{0.33}=$

0.97, $XC7=0.69^{0.25}=0.91$, $XC8=0.36^{0.2}=0.81$, and $XC9=0^{0.17}=0$. The average of the indicators of slope was 0.74 ($B2 = \sum_{i=1}^{5} XC_i = 0.74$). By applying the same procedure for other factor, the normalized values for the remaining subsystem were obtained and tabulated in Table 5. The normalized values represent weights in Eq. 3. The subsystems were ranked depending on the normalized values. The highest normalized value in Table 5, i.e., slope was assigned the highest rank (=6); elevation was ranked as 5, and so on, Table 5. Finally, Eq. 3 was used to compute GPI, Fig. 11. The resultant GPI has a range of 0.57 to 11.52. The GPI was classified based on natural break classification scheme into five classes: very low, low, moderate, high, and very high. The areas covered by these classes were presented in Table 7. The very low and low classes extend over an area of about 14 % (97 km^2), the moderate class occupies 13 % (91 km^2), and the high and very high
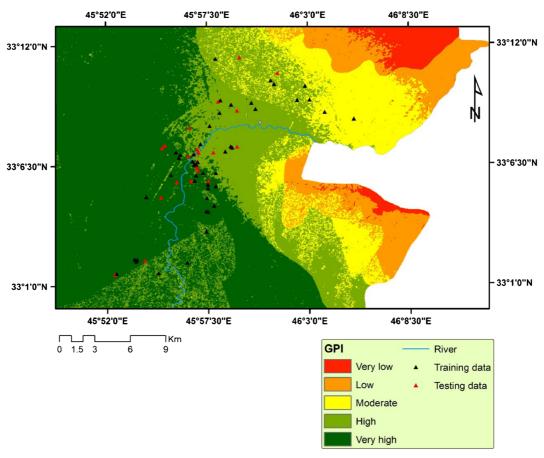

Fig. 11 Groundwater potential index obtained using catastrophe model

classes occupy 73 % (518 km^2). The large extent of high-very high classes indicates high potential condition of the aquifer system in the study area.

The result of application of index of entropy approach is presented in Table 6. Equations from 4 to 9 were used for calculating weight for each subsystem. Results in Table 6 demonstrate that the most important factor for groundwater potential yield in the study area was slope (w=0.12) followed by elevation (w=0.09), and then distance to faults factor (w=0.08). On the other hand, the calculated weights for the rest of the factors were 0.065, 0.035, and 0.020 for soil, transmissivity, and storativity factors, respectively. This indicates that these factors had a minor effect on groundwater potential yield in the study area. Ranks for this model were also obtained based on the order of getting weights: the subsystem with highest weight was assigned the highest rank. The final GPI for this model was also computed using Eq. 3. The obtained GPI was also classified according to natural break classification scheme into five classes: very low, low, moderate, high, and very high. The GPI map obtained using index of entropy method is shown in Fig. 12. The area covered by very low and low classes occupy 21 % (148 km^2), the moderate class extend over 18 % (162 km^2), and the high and very high classes occupy 61 % (434 km^2) (Table 7). It can be seen from Table 7 that results of catastrophe model are approximately consistent with results of the index of entropy model.

Table 6 Results of application of entropy model

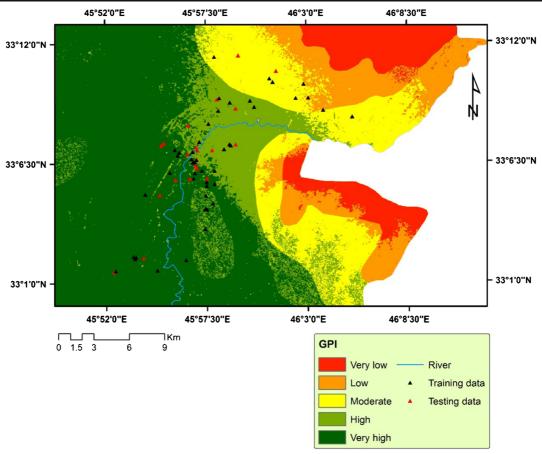| Factor	Range	Area pixels	Area % (a)	Boreholes no.	Boreholes% (b)	FR (b/a)	P_{ij}	H_j	$H_{j\max}$	I_j	W_j	Rank
Elevation	0–56	418,091	0.533	26	0.553	1.039	0.354	1.323	2.00	0.339	0.085	5
	56.01–99	195,357	0.249	19	0.404	1.624	0.553					
	99.01–157	121,533	0.155	2	0.043	0.275	0.094					
	157.1–318	49,929	0.064	0	0.000	0.000	0.000					
Slope	0–5	498,607	0.635	27	0.574	0.9043	0.432	0.987	2.32	0.575	0.115	6
	5.001–15	280,677	0.358	20	0.426	1.1900	0.568					
	15.01–30	5497	0.007	0	0.000	0.0000	0.000					
	30.01–50	128	0.000	0	0.000	0.0000	0.000					
	50.01–70	1	0.000	0	0.000	0.0000	0.000					
Soil	A	488,213	0.622	34	0.723	1.163	0.385	1.481	2	0.260	0.065	3
	B	98,648	0.126	8	0.170	1.354	0.448					
	C	164,896	0.210	5	0.106	0.506	0.167					
	D	33,154	0.042	0	0.000	0.000	0.000					
Transmissivity	20.4–230.7	574,041	0.731	33	0.702	0.960	0.118	1.716	2	0.142	0.035	2
	230.8–471.1	190,037	0.242	11	0.234	0.967	0.119					
	471.2–974.4	16,878	0.022	2	0.043	1.979	0.243					
	974.5–1928	3956	0.005	1	0.021	4.222	0.519					
Storativity	0.001203–0.01738	51,860	0.066	1	0.021	0.322	0.104	1.840	2	0.080	0.020	1
	0.01739–0.02816	144,842	0.185	5	0.106	0.576	0.186					
	0.02817–0.03786	180,085	0.229	10	0.213	0.927	0.300					
	0.03787–0.05619	408,122	0.520	31	0.660	1.269	0.410					
Distance to faults	0–2632	412,881	0.526	31	0.660	1.254	0.486	1.457	2.32	0.373	0.075	4
	2633–5263	253,664	0.323	14	0.298	0.922	0.357					
	5264–7895	82,129	0.105	2	0.043	0.407	0.157					
	7896–10,530	29,596	0.038	0	0.000	0.000	0.000					
	10,540–13,160	6640	0.008	0	0.000	0.000	0.000					

Fig. 12 Groundwater potential index obtained using entropy model
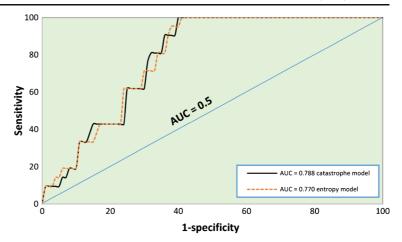
Models validation

The ROC module in IDRISI Selva software was used for validation of results. For constructing ROC for catastrophe model, the total number of boreholes with high yield (68 boreholes) was used because these boreholes are not used in constructing the model itself. On the other hand, for constructing prediction ROC curve for the entropy model, only the testing data (21 boreholes) were used because the rest of the boreholes (47 boreholes) were already used for building the model. The AUCs for catastrophe and entropy models were shown in Fig. 13, in which the prediction accuracy of catastrophe model (0.788) was slightly better than

Table 7 Areas covered by different groundwater potential zones

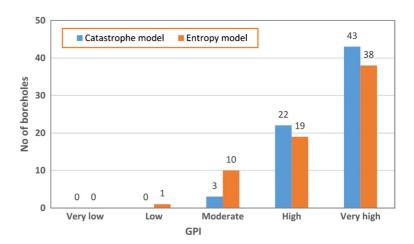| GWPI category	Catastrophe model			Index of entropy model		
	Range	Area%	Area (km^2)	Range	Area%	Area (km^2)
Very low	0.57–5.45	0.04	28	0.0076–0.92	0.09	67
Low	5.44–7.19	0.10	69	0.92–1.18	0.11	81
Moderate	7.19–8.62	0.13	91	1.18–1.36	0.18	126
High	8.62–9.74	0.22	156	1.36–1.48	0.14	98
Very high	9.74–11.52	0.51	363	1.48–1.70	0.48	336

Fig. 13 Results validation using
ROC technique

entropy model (0.770). In general, both models have good capability for prediction groundwater potential yield in the study area.

To further check the results, the total number of boreholes location was compared with GPI for both models, Fig. 14. For catastrophe model, the test revealed that 96 % (65 boreholes) falls into high and very high potential zones, 4 % (3 boreholes) falls into moderate class, and no boreholes in very low and low zones. With respect to entropy model, this test revealed that 84 % (57 boreholes) falls into high-very high zones, 15 % (10 boreholes) falls into moderate zone, 1 % (1 boreholes) falls into low zone, and no boreholes fall into very low zone. It can be seen from these results that both models have good prediction efficacy because most of the borehole locations are correctly predicted by GPI map produced by models, but again the catastrophe model is more effectiveness than entropy model.

Sensitivity analysis

Sensitivity analysis serves to acknowledge uncertainty in variable estimation by observing the changes in results using different set of variable input (Majandang and Sarapirome 2012). Sensitivity analysis can be used to indicate the most important variables to be used for delineating groundwater zones. The results of sensitivity analysis can guide hydrogeologist to select the groundwater influencing factors which require more detail information for accurate estimation of groundwater potential. In general, there are two types of sensitivity analysis, namely map removal and single variable analysis (Babiker et al. 2005; Napolitano and Fabbri 1996). In this study, the map removal sensitivity analysis was used. The single variable analysis is unnecessary because factor's weight and rating are estimated objectively using used models. The map removal analysis identifies the sensitivity of GPI by removing one or more

Fig. 14 Number of existing
boreholes for each GPI zone

Table 8 Statistical summary of sensitivity analysis by the map removal

Variation index (%)	Variable removed					
	Elevation	Slope	Soil	Transmissivity	Storativity	Distance to faults
Minimum	0.03	0.01	0.07	1.20	0.32	0.12
Maximum	6.89	12.02	10.66	14.55	6.40	13.83
Mean	1.93	2.96	6.57	11.41	5.23	10.01
SD	0.91	3.17	2.12	3.36	0.87	3.45

SD standard deviation

layer maps and is computed using the following equation (Lodwik et al. 1990):

$$S = \frac{\left| \frac{V}{N} - \frac{V'}{n} \right|}{V} \times 100 \qquad (10)$$

where S is the variation index, V and V' are the unperturbed and the perturbed GWP indices, respectively, and N and n are the number of data layers used to compute V and V', respectively. The unperturbed GPI is the actual index obtained by using all six factors and the perturbed GPI was computed using a lower number of factors.

Table 8 presents the variation index as a result of removing only one factor at a time. It can be seen that variation indices are highest upon the removal of transmissivity, distance to faults, and soil, respectively. The mean variation indices for these factors were 11.41, 10.1, and 6.57, respectively. This indicates the importance of these factors to delineate groundwater potential zones in the study area. The least sensitive factors were storativity, slope, and elevation with 5.23, 2.56, and 1.93 variation indices, respectively. In general, all used factors have relatively higher values of variation index implying the importance of all factors for accurate demarcation of groundwater potential zones.

Conclusions

The main conclusions drawn from this study were as follows: (1) The two used methods for delineating groundwater potential in the study have very good capability for predicting groundwater potential zones with a model accuracy of 79 and 0.77 % for catastrophe and index of entropy models, respectively. (2) The catastrophe method is a more efficient technique to demarcate groundwater potential zone than index of entropy in arid region. (3) The slope and elevation play a major role in controlling groundwater availability in the study area; the calculated weights for these factors were the highest among the other factors for both used models. (4) The obtained GPI values are classified into five classes: very low, low, moderate, high, and very high. The large extent of high-very high classes (73 % for catastrophe model and 61 % for entropy model) implies that aquifer system in the study area has highly groundwater potentiality. (5) The groundwater potential zone maps produced by this study can provide valuable information for hydrogeologist, planners, and decision makers to put suitable plans for managing groundwater in the study area.

References

Ahmed, K., Shahid, S., bin Harun, S., Ismail, T., Nawaz, N., & Shamsudin, S. (2014). Assessment of groundwater potential zones in an arid region based on catastrophe theory. *Earth Sciences Information*. doi:10.1007/s12145-014-0173-3.

Al-Abadi, A. M. (2015). Modeling of groundwater productivity in northeastern Wasit Governorate, Iraq by using frequency ratio and Shannon's entropy models. *Applied Water Science*. doi:10.1007/s13201-015-0283-1.

Al-Sayab, A., Al-Ansari, N., Al-Rawi, D., Al-Jassim, J., Al-Omari, F., & Al-Shaikh, Z. (1982). *Geology of Iraq*. Iraq: Mosul University (In Arabic).

Al-Shammary, SH. (2006). *Hydrogeology of Galal Basin-Wasit east, Iraq*. Unpublished Ph D thesis, Baghdad.

Aniya, M. (1985). Landslide-susceptibility mapping in the Amahata river basin, Japan. *Annual Association American Geographers, 75*, 102–114.

Babiker, I., Mohamed, M., Hiyama, T., & Kato, K. (2005). A GIS-based DRASTIC model for assessing aquifer vulnerability in Kakamigahara Heights, Gifu Prefecture, central Japan. *The Science of the Total Environment, 345*, 127–140.

Bednarik, M., Magulova, B., Matys, M., & Marschalko, M. (2010). Landslide susceptibility assessment of the Kral Ovany–Liptovsky' Mikula's˘ railway case study. *Physics Chemistry Earth Parts A/B/C, 35*, 162–171.

Bednarik, M., Yilmaz, I., & Marschalko, M. (2012). Landslide hazard and risk assessment: a case study from the Hlohovec–Sered'landslide area in south-west Slovakia. *Natural Hazards, 64*, 547–575.

Chen, C. T. (2000). Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy Sets and System, 114*, 1–9.

Chen, Y. F., Yin, C. F., & Lu, G. F. (2006). The catastrophic model of water bloom: a case study on Lake Chaohu. *Acta Ecologica Sinica, 26*, 878–883.

Ching, H. S., Ying-Hua, C. L., Yin, L. (1996). Evaluating a weapon system using catastrophe series based on fuzzy scales. In: *Fuzzy systems symposium, soft computing in intelligent systems and information processing, proceedings of the 1996 Asian, 11–14.* pp 212–217. doi:10.1109/afss.1996.583593.

Chowdary, V., Chakraborthy, D., Jeyaram, A., Murthy, Y. K., Sharma, J., & Dadhwal, V. (2013). Multi-criteria decision making approach for watershed prioritization using analytic hierarchy process technique and GIS. *Water Resources Management, 27*, 1–17.

Chung, C. F., & Fabbri, A. G. (2003). Validation of spatial prediction models for landslide hazard mapping. *Natural Hazards, 30*, 451–472.

Constantin, M., Bednarik, M., Jurchescu, M. C., & Vlaicu, M. (2011). Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). *Environmental Earth Science, 63*, 397–406.

Corsini, A., Cervi, F., & Ronchetti, F. (2009). Weight of evidence and artificial neural networks for potential groundwater mapping: an application to the Mt. Modino area (Northern Apennines, Italy). *Geomorphology, 111*, 79–87. doi:10.1016/j.geomorph.2008.03.015.

Egan, J. P. (1975). *Signal detection theory and ROC analysis*. New York: Academic Press.

Elmahdy, S. I., & Mohamed, M. M. (2014). Probabilistic frequency ratio model for groundwater potential mapping in Al Jaww plain, UAE. *Arabian Journal of Geosciences*. doi:10.1007/s12517-014-1327-9.

Ghorbani, M. A., Khatibi, R., Sivakumar, B., & Cobb, L. (2010). Study of discontinuities in hydrological data using catastrophe theory. *Hydrological Sciences Journal, 55*, 1137–1151.

Hui, Q. (2008). *Niche, factor interaction and business evolution—the enterprise niche research of the growth business*. Hangzhou: Zhejiang University Press.

Ihara, S. (1993). *Information theory for continuous systems*. USA: World Scientific Pub Co Inc.

Jaafari, A., Najafi, A., Pourghasemi, H. R., Rezaeian, J., & Sattarian, A. (2013). GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran. *International Journal of Environmental Science and Technology, 11*, 909–926. doi:10.1007/s13762-013-0464-0.

Jassim, S. Z., & Goff, J. C. (2006). *Geology of Iraq*. Dolin: Prague and Moravian Museum. 431p.

Jha, M. K., Chowdary, V. M., & Chowdhury, A. (2010). Groundwater assessment in Salboni Block, West Bengal (India) using remote sensing, geographical information system and multi-criteria decision analysis techniques. *Hydrogeology Journal, 18*, 1713–1728. doi:10.1007/s10040-010-0631-z.

Kam, J. K. (1992). Are chaos and catastrophe theories relevant to environmental sciences? *Journal of Environmental Sciences, 4*, 39–42.

Lee, S., Kim, Y. S., & Oh, H. J. (2012). Application of a weight-of-evidence method and GIS to regional groundwater productivity potential mapping. *Journal of Environmental Management, 96*, 91–105. doi:10.1016/j.jenvman.2011.09.016.

Li, P.-Y., Hui, Q., & Jian-Hua, W. U. (2010). Groundwater quality assessment based on improved quality index in Pengyang County, Ningxia, northeast China. *Journal of Chemistry, 7*, 209–216.

Lodwik, W. A., Monson, W., & Svoboda, L. (1990). Attribute error and sensitivity analysis of maps operation in geographical information systems–suitability analysis. *International Journal of Geographical Information System, 4*, 413–428.

Machiwal, D., Madan, K. J., & Bimal, C. M. (2010). Assessment of groundwater potential in a semi-arid region of India using remote sensing, GIS and MCDM techniques. *Water Resources Management, 25*, 1359–1386.

Majandang, J., & Sarapirome, S. (2012). Groundwater vulnerability assessment and sensitivity analysis in Nong Rua, Khon Kaen, Thailand, using a GIS-based SINTACS model. *Environmental Earth Science, 68*, 2025–2039. doi:10.1007/s12665-012-1890-x.

Malczewski, J. (1999). *GIS and multicriteria decision analysis*. New York: Wiley.

Manap, M. A., Sulaiman, W. N., Ramli, M. F., Pradhan, B., & Surip, N. (2011). A knowledge-driven GIS modeling technique for groundwater potential mapping at the Upper Langat Basin, Malaysia. *Arabian Journal of Geosciences, 6*, 1621–1637. doi:10.1007/s12517-011-0469-2.

Moghaddam, D. D., Rezaei, M., Pourghasemi, H. R., Pourtaghie, Z. S., & Pradhan, B. (2013). Groundwater spring potential mapping using bivariate statistical model and GIS in the Taleghan Watershed, Iraq. *Arabian Journal of Geosciences*. doi:10.1007/s12517-013-1161-5.

Naghibi, S. A., Pourghasemi, H. R., Pourtaghi, Z. S., & Rezaei, A. (2014). Groundwater qanat potential mapping using frequency ratio and Shannon's entropy models in the Moghan watershed, Iraq. *Earth Sciences Information*. doi:10.1007/s12145-014-0145-7.

Nampak, H., Pradhan, B., & Manap, M. A. (2014). Application of GIS based data driven evidential belief function model to predict groundwater potential zonation. *Journal of Hydrology, 513*, 283–300.

Napolitano, P., Fabbri, AG. (1996). Single parameter sensitivity analysis for aquifer vulnerability assessment using DRASTIC and SINTACS. In: *Proceedings of the Vienna*

conference on HydroGIS 96: application of geographic information system in hydrology and water resources management*. IAHS Pub. No. 235. pp 559–566.

Oh, H. J., Kim, Y. S., Choi, J. K., Park, E., & Lee, S. (2011). GIS mapping of regional probabilistic groundwater potential in the area of Pohang City, Korea. *Journal of Hydrology, 399*, 158–172.

Ozdemir, A. (2011a). Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey). *Journal of Hydrology, 405*, 123–136. doi:10.1016/j.jhydrol.2011.05.015.

Ozdemir, A. (2011b). GIS-based groundwater spring potential mapping in the Sultan Mountains (Konya, Turkey) using frequency ratio, weights of evidence and logistic regression methods and their comparison. *Journal of Hydrology, 411*, 290–308.

Pandey, VP., Kazama F. (2012). Groundwater storage potential in the Kathmandu Valley's shallow and deep aquifers. In: Shrestha, S., Pradhanand, D,. Pandey, VP. (Eds.) *Kathmandu valley groundwater outlook*. Publisher: AIT/SEN/CREEW/ICRE-UY. pp 31–38.

Pandey, V. P., Shrestha, S., & Kazama, F. (2013). A GIS-based methodology to delineate potential areas for groundwater development: a case study from Kathmandu Valley, Nepal. *Applied Water Science, 3*(2), 453–465.

Parsons, R. M. (1956). *Ground-water resources of Iraq, Khanaqin-Jassan area, vol. 1. Development board*. Baghdad: Ministry of Development Government of Iraq.

Pourghasemi, H., Pradhan, B., Gokceoglu, C., & Moezzi, K. D. (2013). A comparative assessment of prediction capabilities of Dempster-Shafer and weights-of-evidence models in landslide susceptibility mapping using GIS. *Geomatics, Natural Hazards and Risk, 4*, 93–118.

Pourtaghi, Z. S., & Pourghasemi, H. R. (2014). GIS-based groundwater spring potential assessment and mapping in the Birjand Township, southern Khorasan Province, Iran. *Hydrogeology Journal, 22*, 643–662. doi:10.1007/s10040-013-1089-6.

Prasad, R. K., Mondal, N. C., Banerjee, P., Nandakumar, M. V., & Singh, V. S. (2008). Deciphering potential groundwater zone in hard rock through the application of GIS. *Environmental Geology, 55*, 467–475.

Satty, T. L. (1980). *The analytic hierarchy process*. New York: McGraw-Hill.

Shahid, S., Nath, S. K., & Roy, J. (2000). Groundwater potential modeling in a GIS. *International Journal of Remote Sensing, 21*, 1919–1924.

Shahid, S., Nath, S. K., & Kamal, A. S. (2002). GIS integration of remote sensing and topographic data using fuzzy logic for ground water assessment in Midnapur District, India. *Geocarto International, 17*, 69–74. doi:10.1080/10106040208.

Shi, Y., & Jin, F. (2009). Landslide stability analysis based on generalized information entropy. *International Conference on Environmental Science Information Applied Technology, 2*, 83–85.

Singhal, B. B., & Gupta, R. P. (1999). *Applied hydrogeology of fractured rocks*. Netherlands: Springer. 429p.

USDA—United State Department of Agriculture (1986) *Urban hydrology for small watersheds*, technical release No. 55, Washington DC.

Wang, W., Liu, S., Zhang, S., & Chen, J. (2011). Assessment of a model of pollution disaster in near-shore coastal waters based on catastrophe theory. *Ecological Modelling, 222*, 307–312.

Yang, F., Shao, D., Xiao, C., & Tan, X. (2012). Assessment of urban water security based on catastrophe theory. *Water Science and Technology, 66*, 487–493.

Yesilnacar, E., & Topal, T. (2005). Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Engineering Geology, 79*, 251–266.

Zhang, T. J., Ren, S. X., Li, S. G., Zhang, T. C., & Xu, H. J. (2009) Application of the catastrophe progression method in predicting coal and gas outburst. *Mining Science and Technology, 19*(4), 430–434.