

Modeling of groundwater productivity in northeastern Wasit Governorate, Iraq using frequency ratio and Shannon's entropy models

Alaa M. Al-Abadi¹

Received: 4 September 2014 / Accepted: 26 March 2015 / Published online: 9 April 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract In recent years, delineation of groundwater productivity zones plays an increasingly important role in sustainable management of groundwater resource throughout the world. In this study, groundwater productivity index of northeastern Wasit Governorate was delineated using probabilistic frequency ratio (FR) and Shannon's entropy models in framework of GIS. Eight factors believed to influence the groundwater occurrence in the study area were selected and used as the input data. These factors were elevation (m), slope angle (degree), geology, soil, aquifer transmissivity (m^2/d), storativity (dimensionless), distance to river (m), and distance to faults (m). In the first step, borehole location inventory map consisting of 68 boreholes with relatively high yield (>8 l/sec) was prepared. 47 boreholes (70 %) were used as training data and the remaining 21 (30 %) were used for validation. The predictive capability of each model was determined using relative operating characteristic technique. The results of the analysis indicate that the FR model with a success rate of 87.4 % and prediction rate 86.9 % performed slightly better than Shannon's entropy model with success rate of 84.4 % and prediction rate of 82.4 %. The resultant groundwater productivity index was classified into five classes using natural break classification scheme: very low, low, moderate, high, and very high. The high–very high classes for FR and Shannon's entropy models occurred within 30 % (217 km²) and 31 % (220 km²), respectively indicating low productivity conditions of the aquifer system. From final results, both of the models were capable to

prospect GWPI with very good results, but FR was better in terms of success and prediction rates. Results of this study could be helpful for better management of groundwater resources in the study area and give planners and decision makers an opportunity to prepare appropriate groundwater investment plans.

Keywords Groundwater · Frequency ration · Shannon's entropy · AUC · Wasit · Iraq

Introduction

Water is a precious natural resource without it life is not possible. The demand for water has rapidly increased over the last few years and this has resulted in water scarcity in many parts of the world. Due to the fact that Iraq is an arid country at least in the central and southern parts, this country is heading towards a water crisis mainly due to the improper management of water resources, water policies in neighboring countries (Turkey, Syria, and Islamic Republic of Iran), and the prevalence of drought conditions due to climatic changes. During the last few decades, groundwater levels in main freshest aquifer in Iraq have been falling due to the increase in extraction rates and very bad management scenarios. The rapid increase of population associated with changing lifestyles, especially after 2003, has also increased the domestic, agricultural, and industrial usages of groundwater in entire Iraq, particularly in central and south Iraq, distant from the centers of the cities. The contamination of these aquifers has also added another dimension for the problem for decision maker and politicians (Jabar Al-Saydi, Expert, Head of Groundwater Commission of Groundwater/Basra Branch, personal communication). In the light of these challenges, there is a truly urgent need for reassessment of groundwater

✉ Alaa M. Al-Abadi
alaaatiaa@gmail.com

¹ Department of Geology, College of Science, University of Basra, Basra, Iraq

resources using modern techniques such as remote sensing, global positioning system (GPS), and geographic information system (GIS). Generally, the conventional approaches for groundwater resources are time consuming, costly, uneconomical and sometimes unsuccessful (Todd and Mays 2005; Jha et al. 2010). With the advent of powerful computers, advance in GPS and GIS, efficient and powerful techniques for groundwater resources have evolved. These techniques have reassigned the ways to manage natural resources in general and groundwater resources in particular.

The term “groundwater productivity (potentiality)” denotes the amount of groundwater available in an area and it is a function of several hydrologic and hydrogeological factors (Jha et al. 2010). From a hydrogeological exploration point of view, this term may be defined as the possibility of groundwater occurrence in an area. The methodology proposed in the literature (Chi and Lee 1994; Krishnamurthy and Srinivas 1995; Kamaraju et al. 1995; Krishnamurthy et al. 1996; Sander et al. 1996; Edet et al. 1998; Saraf and Choudhury 1998, Shahid et al. 2000; Jaiswal et al. 2003; Rao and Jugran 2003; Sikdar et al. 2004; Sener et al. 2005; Ravi Shankar and Mohan 2006; Solomon and Quiel 2006; Madrucci et al. 2008; Ganapuram et al. 2009; Suja Rose and Krishnan 2009; Pradeep Kumar et al. 2010; Chowdhury et al. 2010; Jha et al. 2010; Machiwal et al. 2010; Dar et al. 2010; Manap et al. 2011; Khodaei and Nassery 2011; Sahu and Sikdar 2011; Abdalla 2012; Pandey et al. 2013; and Gumma and Pavelic 2013; Al-Abadi and Al-Shamma’a 2014; Rahmati et al. 2014; Chen et al. 2014) to delineate groundwater potential zones of an area is attained through integrating several thematic layers (maps) from different resources such as conventional, geophysical, and remote sensing data to generate groundwater productivity index (GWPI). Usually, the GWPI is computed using the weighted linear combination technique (Malczewski 1999)

$$GWPI = \sum_{j=1}^m \sum_{i=1}^n (w_j x_i) \quad (1)$$

where x_i is the normalized weight of the i th class/feature of theme, w_j is the normalized weight of the j th theme, m is the total number of themes, and n is the total number of classes in a theme. 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 thematic layers in GIS environment. The AHP provides a flexible, low cost, and easily understood way for analysis complicated problems (Satty 1980). The drawback of AHP is related to its dependency on the expert’s knowledge which is the main source of uncertainty (Chowdary et al. 2013).

In few recent years, several authors have attempted to delineate groundwater productivity and springs potentiality using several knowledge-driven and data-driven models. Most of the used techniques have been applied in other fields of earth and environmental sciences such as mineral prospecting, flood susceptibility, and landslides studies. The used models involve probabilistic frequency ratio (Ozdemir 2011a; Oh et al. 2011; Manap et al. 2011; Moghaddam et al. 2013; Pourtaghi and Pourghasemi 2014; Naghibi et al. 2014; Elmahdy and Mohamed 2014) logistic regression (Ozdemir 2011a, b; Pourtaghi and Pourghasemi 2014), Shannon’s entropy (Naghibi et al. 2014), weights of evidence (Corsini et al. 2009; Ozdemir 2011b; Lee et al. 2012; Pourtaghi and Pourghasemi 2014; Al-Abadi 2015), artificial neural networks (Corsini et al. 2009; Lee et al. 2012), fuzzy logic (Shahid et al. 2014), and more recently evidential belief function (Nampak et al. 2014). The idea behind these techniques is to explore the relationship between groundwater (springs/productive boreholes) locations and influential groundwater occurrence factors. The type and number of factors vary from one study to another and their selection is often arbitrary. Often, personal judgment plays an important role in choosing factors and their class attributes. The factors of geology, soil, land use/land cover (LULC), altitude, slope, aspect, curvature, topographic wetting index (TWI), stream power index (SPI), length steepness factor (LS), distance to roads, distance to faults, faults density, distance to river, drainage density, lineaments and lineaments density are often used in the analysis of groundwater springs and aquifer yields potentiality. The availability of data is the main constrain to use factors from one study to another.

The main objective of this study is to demarcate groundwater productivity at northeastern Wasit Governorate, Iraq through using probabilistic frequency ratio and Shannon’s entropy models in framework of GIS. The objective of this study is achieved by building a geospatial database and investigates the relationship between productive boreholes locations and many groundwater occurrence factors such as elevation (m), slope angle (degree), geology, soil, aquifer transmissivity (m^2/d), specific storage (dimensionless), distance to river (m), and distance to faults (m). The results of this study could help in efficient management of groundwater resources in the study area and help workers in water resources in the country to put suitable plans to manage limited groundwater resources incorporating growing challenges facing water sector.

The study area

The study area extends over an area of 707 km² and lies between 33°00′ and 33°14′ latitude and 45°50′ and 46°16′ longitude in the northeastern Wasit Governorate, Iraq

(Fig. 1). It is bounded by Iraqi–Iranian border (Hamrin hills) from the east, wadi Galas from north, and hor Al-Shiwach from east and south. The main city within the question area is Badrah. 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). Three quarters of the study area are plain with a gentle slope and occupy the southwestern parts. The remaining quarter locates in the northeastern part and roughly parallel to the Iranian borders and is characterized by low anticlinal folds with intervening synclinal valleys (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/y with an uneven rainfall distribution between plain and mountain parts. According to the recorded meteorological data in Badra station for the period (1994–2013), the monthly maximum, minimum, and average temperatures are 10.4, 37.8, and 24.56 °C, respectively. Drainage in the question is almost in a southwesterly direction (Parsons 1956). The nature of

Fig. 1 Location map of the study area

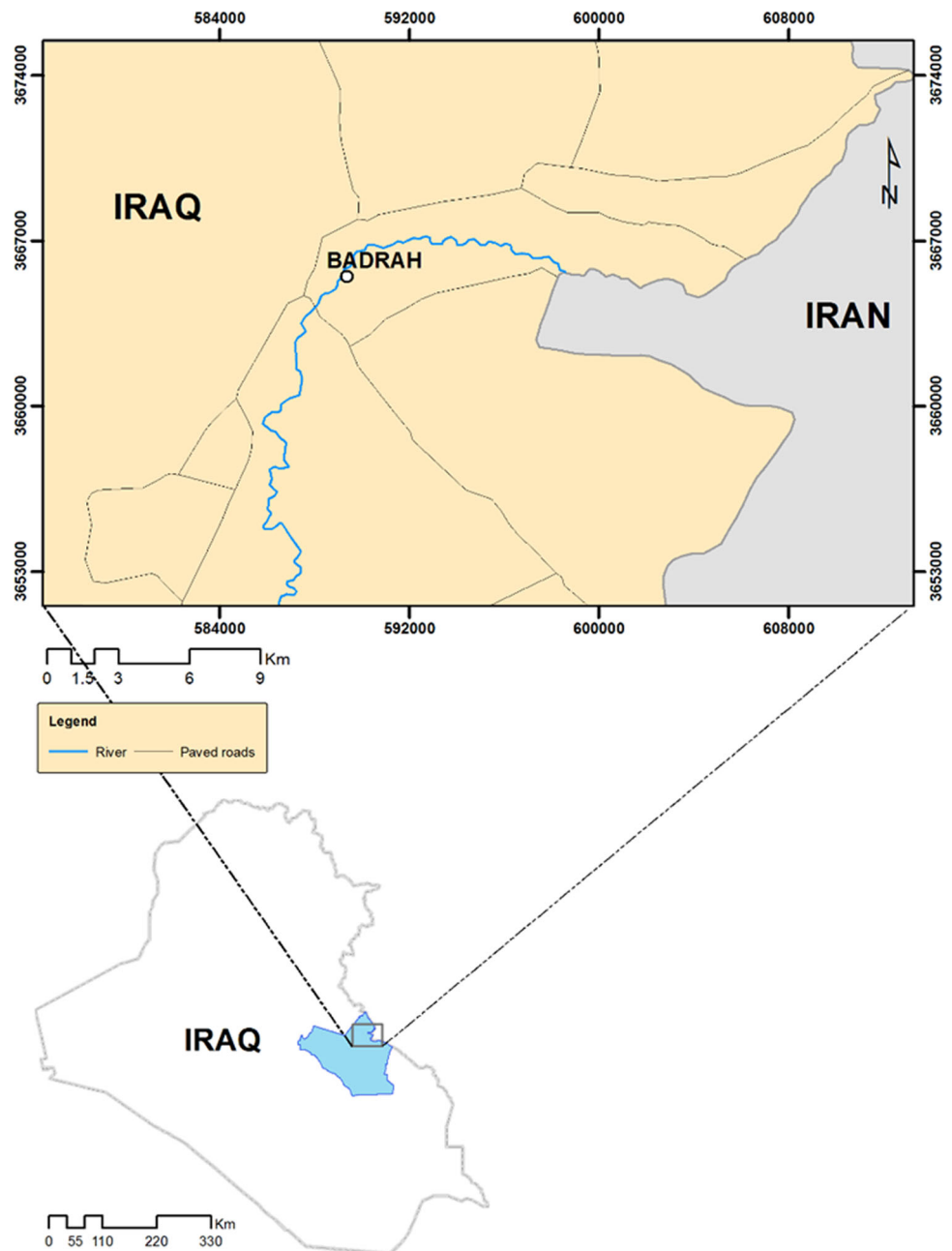
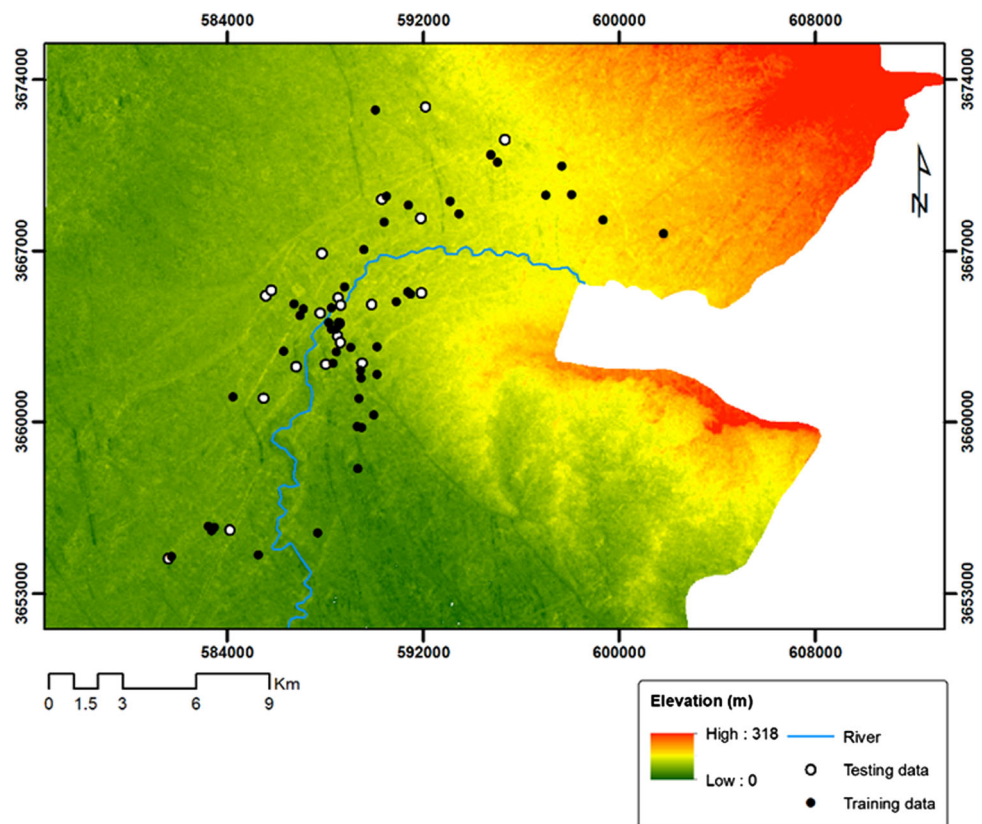


Fig. 2 Ground surface elevation of the study area (extracted from DEM with 30 m resolution)



the galals or streams is intermittent and terminates in the temporary marshes on the delta plain. During heavy rainfall periods, the coming flooding water from the Iranian side submerge the flat plain to the west and causing occasional floods. The major stream in the study area is Galal–Badra River. The mean monthly discharge of this river is 2.5 and 1000 m³/s in drought and flood periods, respectively (Al-Shammary 2006). Due to the prolonged drought conditions and intermittent nature of the streams in the study area, most of the farmers depend on the groundwater for their irrigation needs.

From a geological point of view, rocks in the investigated 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. The Quaternary deposits mainly consist of a mixture of gravel, sand, silt and conglomerates of post Pliocene deposits. The distribution of these lithological units is shown in Fig. 3. A brief description of these units is provided in Table 1. Approximately 84 % of the study area covers with Quaternary deposits. 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 Iraqi–Iranian border (folded zone). There are many faults in the study area, the bigger and important one is Shbichia–Najaf fault.

The soil of the study area formed from the processes of weathering, erosion and sedimentation during the Quaternary period. 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 is shown in Fig. 4, in which the major portion of the study area (about ~60%) has high infiltration rate (A and B groups).

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

Fig. 3 Geological map of the study area

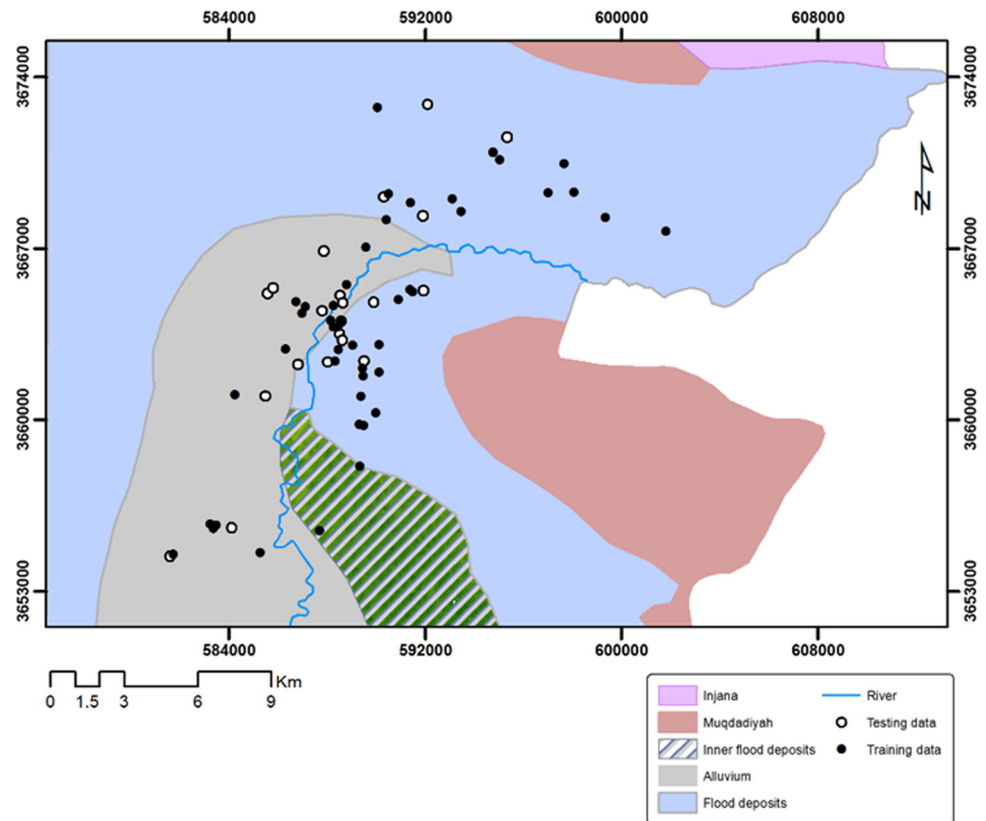


Table 1 Description of the lithological formations in the study area

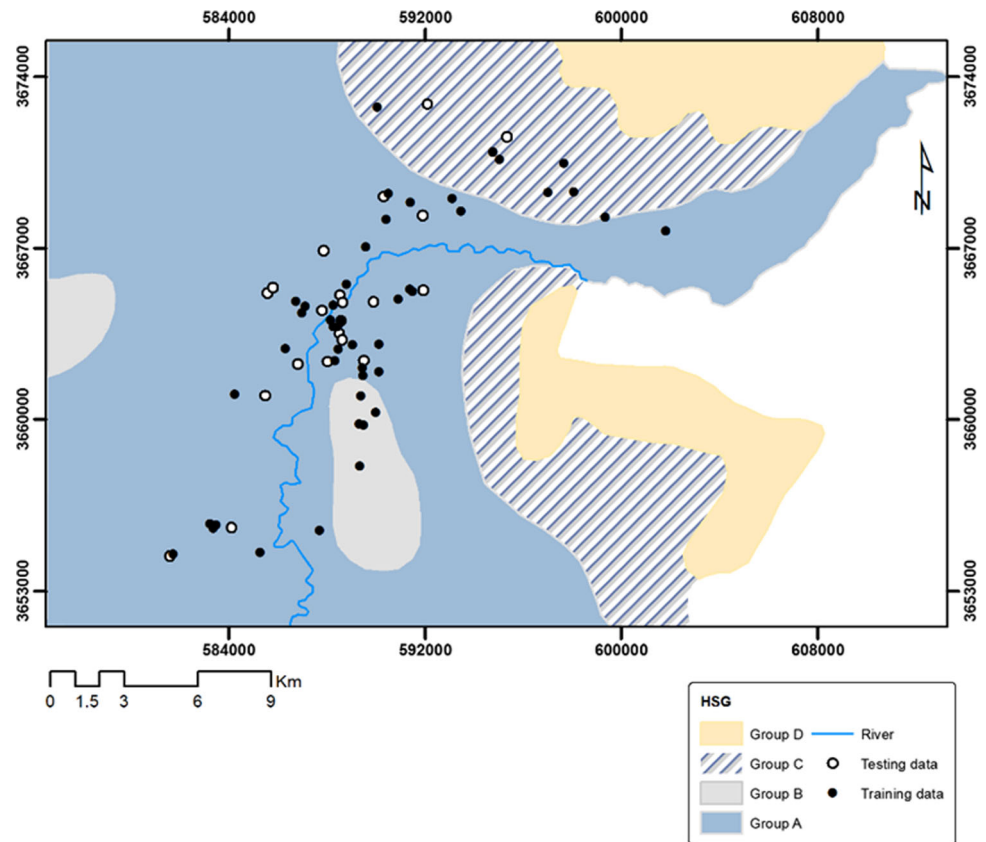
Formation	Age	Environment	Description	Area (km ²)	Area (%)
Injana	Upper Miocene	Sub-marine	Red or gray colored silty marl or clay stones and purple silt stones	8	0.01
Muqdadiah	Pliocene	Continental	Gravelly sandstone, sandstone, and red mudstone	103	0.15
Quaternary	Pleistocene–Holocene	Continental	Mixture of gravel, sand, silt and clay	596	0.84

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. Depths to groundwater range from 26 to 162 m. The spatial distribution of the groundwater depths in the study area is shown in Fig. 5, in which the groundwater depths increase towards eastern and north-eastern parts corresponding with the elevation increase in the same directions. 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, 228.43 m²/d, and 0.012, respectively. For the confined aquifer the values were 3.5, 81.07 m²/d, and 0.0017 for hydraulic conductivity, transmissivity, and storage coefficient, respectively. The spatial distributions of transmissivity and storativity for the whole aquifer system are shown in Figs. 6 and 7. In general, the

hydraulic characteristics of the aquifer system are good in the middle and western side of the study area and become poor in the eastern parts.

Data preparation

The methodology presented in the literature for modeling aquifer productivity consists of four steps: (1) describing and partitioning the borehole yield data into two sets, training and validation. The training points are solely used in for calibrating the model (relationship between the influencing factors affecting groundwater occurrence and borehole/springs locations), while testing points are used for validation of the results (validation of the calibrated model) (2) data collection and construction of a spatial database for the influencing factors (3) assessing the productivity zones using the relationship between borehole data and influencing factors by means of data-driven and/or

Fig. 4 Hydrological soil groups

data knowledge models (4) validating the results and if more than one methods used, the analysis also involves comparing the performance of the methods and selecting the best one. A flow chart for clarifying this procedure is presented in Fig. 8.

Borehole inventory

The groundwater borehole data were obtained from the General Commission of Groundwater/Ministry of Water Resources, Iraq. The data involved locations of the borehole (UTM), borehole discharge, depth of the borehole, type of aquifer, and chemical analysis of groundwater for major ions. In fact, there are 80 wells in the study area. Only boreholes with high flow rate (>8 l/s) (about 68 boreholes) were used in the rest of the analysis and randomly divided into two sets using MINITAB 16 software. The splitting criteria were 70/30. The training data contained 47 boreholes and testing data contained 21 boreholes.

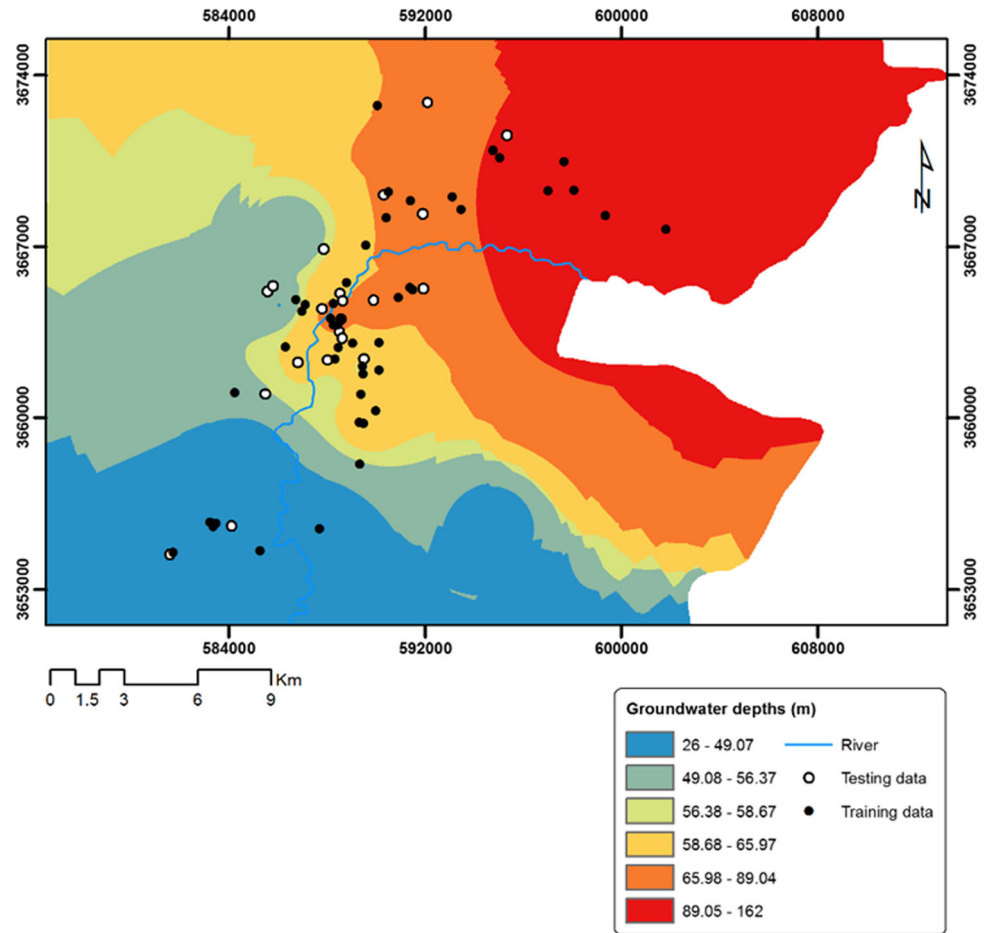
Generating of thematic layers of influential groundwater productivity

Productivity of an aquifer is governed by many surface and subsurface factors such as geology, geomorphology, land

use land cover LULC, soil, topography and related factors, climate, permeability of the water bearing layers, storativity, saturated thickness (Oh et al. 2011). In this study, eight factors were used in the analysis. These factors were elevation (m), slope angle (degree), geology, soil, transmissivity (m^2/d), storativity (dimensionless), distance to river (m), and distance to faults (m). All thematic layers were prepared as a raster format comprising of 30×30 m cell size. The used project coordinate system was (UTM, WGS 1984, 38 N). For classification of continuous values of influential raster layers, natural break classification method was used in this study. The natural break classification scheme, also called the Jenks classification method, is a data clustering method designed to determine the best arrangement of values into different classes. The method seeks to reduce the variance within classes and maximize the variance between classes (Jenks 1967). Selection of this classification scheme is based on literature reviews and author's experience of study area and its condition.

To prepare thematic layers of the topographic factors, i.e. elevation and slope angle, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) (<http://gdem.ersdac.jspacesystems.or.jp/search.jsp>) is used. The ASTER-GDEM was developed by the Ministry of Economy of Japan and the United States National Aeronautics

Fig. 5 Spatial distribution of groundwater depth



and Space Administration (NASA). The spatial resolution of the ASTER-GDEM is approximately 30 m. The raw DEM was reprojected, fill sinks, and clipped for the study area using ArcGIS 10.2 software. Elevation raster was directly created from DEM and was classified into four classes. 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. The slope angle map of study area was prepared from DEM and classified into 4 classes, Fig. 9. It is widely recognized that geology influences the occurrence of groundwater because lithological and structural variations often lead to difference in the strength and permeability of rocks and soils (Ozdemir 2011a). The thematic raster layer of geology was prepared by converting vector layer of geology to raster layer in ArcGIS 10.2. The same converting procedure was made for HSG soil layer vector. 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. The transmissivity and storativity of the aquifer system in the study area were classified into four classes for

both factors, respectively. Maps of distance from faults and river were prepared by applying the distance command in spatial analyst extension of ArcGIS 10.2 and then classified into ten classes for both factors, respectively (Figs. 10, 11).

Modeling techniques

Frequency ratio model

The frequency ratio (FR) is the ratio of the probability of an occurrence to the probability of a non-occurrence for given attributes (Bonham-Carter 1994). The method explores the statistical correlation between boreholes locations and the influencing groundwater occurrence factors. In practical applications, the FR can be calculated as (Ozdemir 2011b).

$$FR = \frac{A/B}{C/D} = \frac{b}{a} \tag{2}$$

where *A* is the area of a class for the influencing 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 for area with

Fig. 6 Spatial distribution of transmissivity (m^2/d)

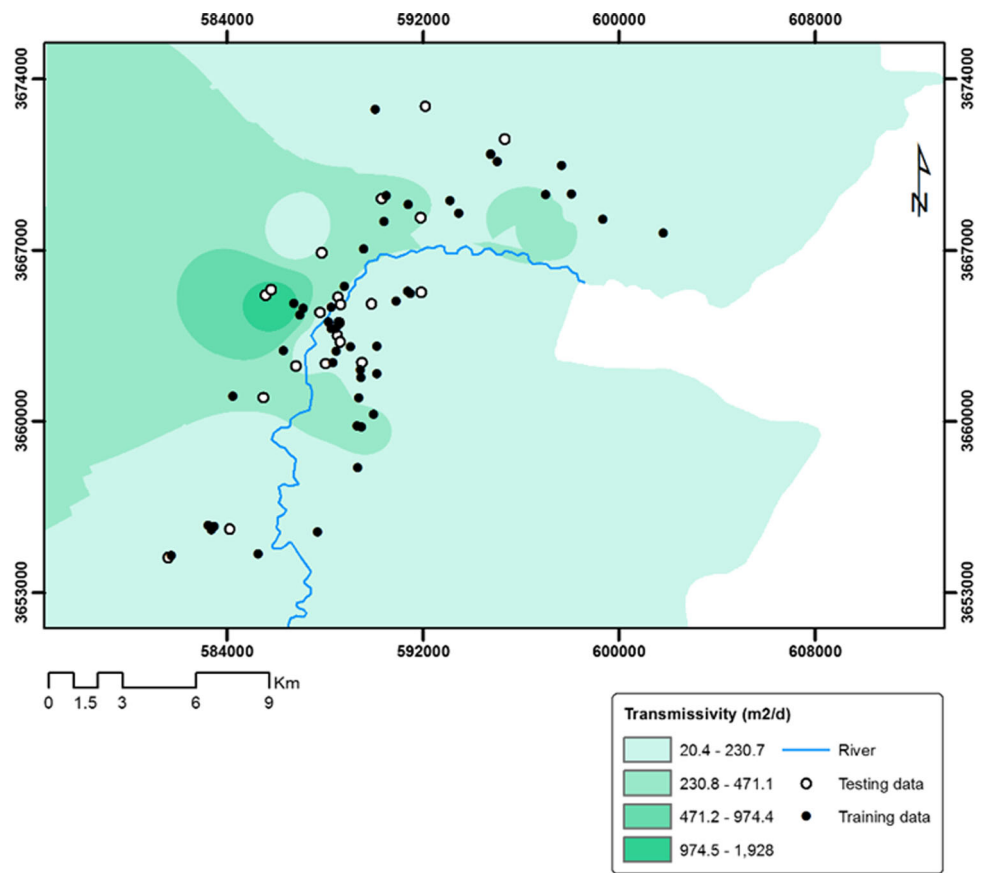
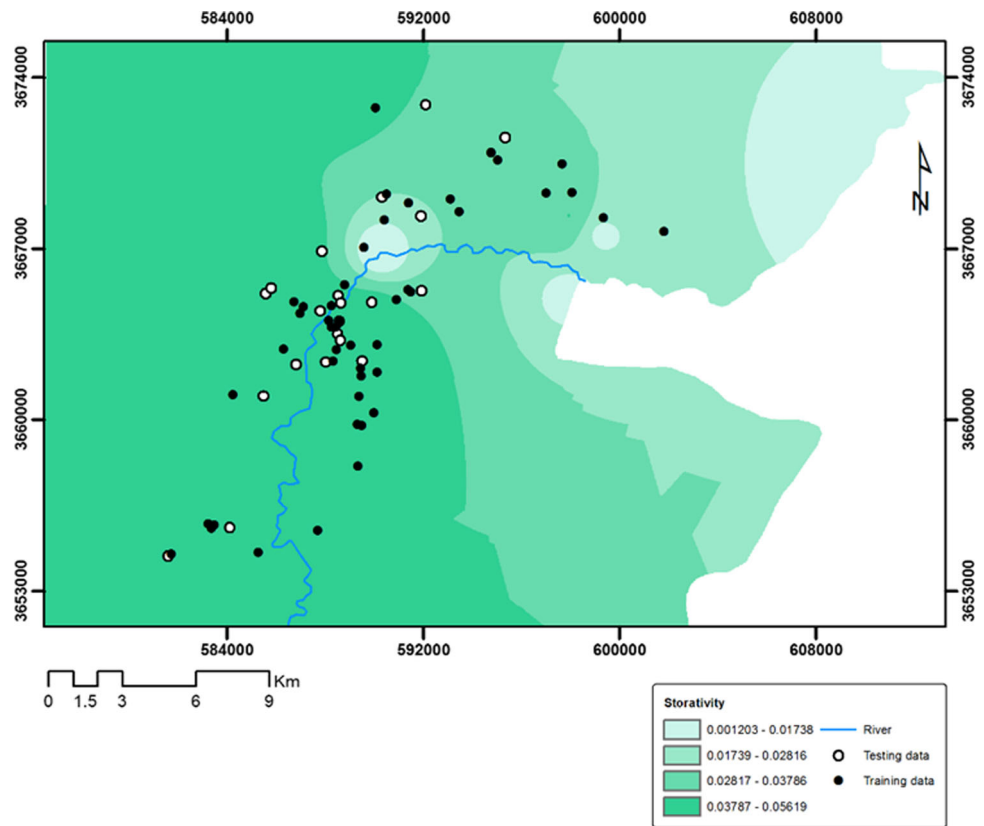


Fig. 7 Spatial distribution of storativity (dimensionless)



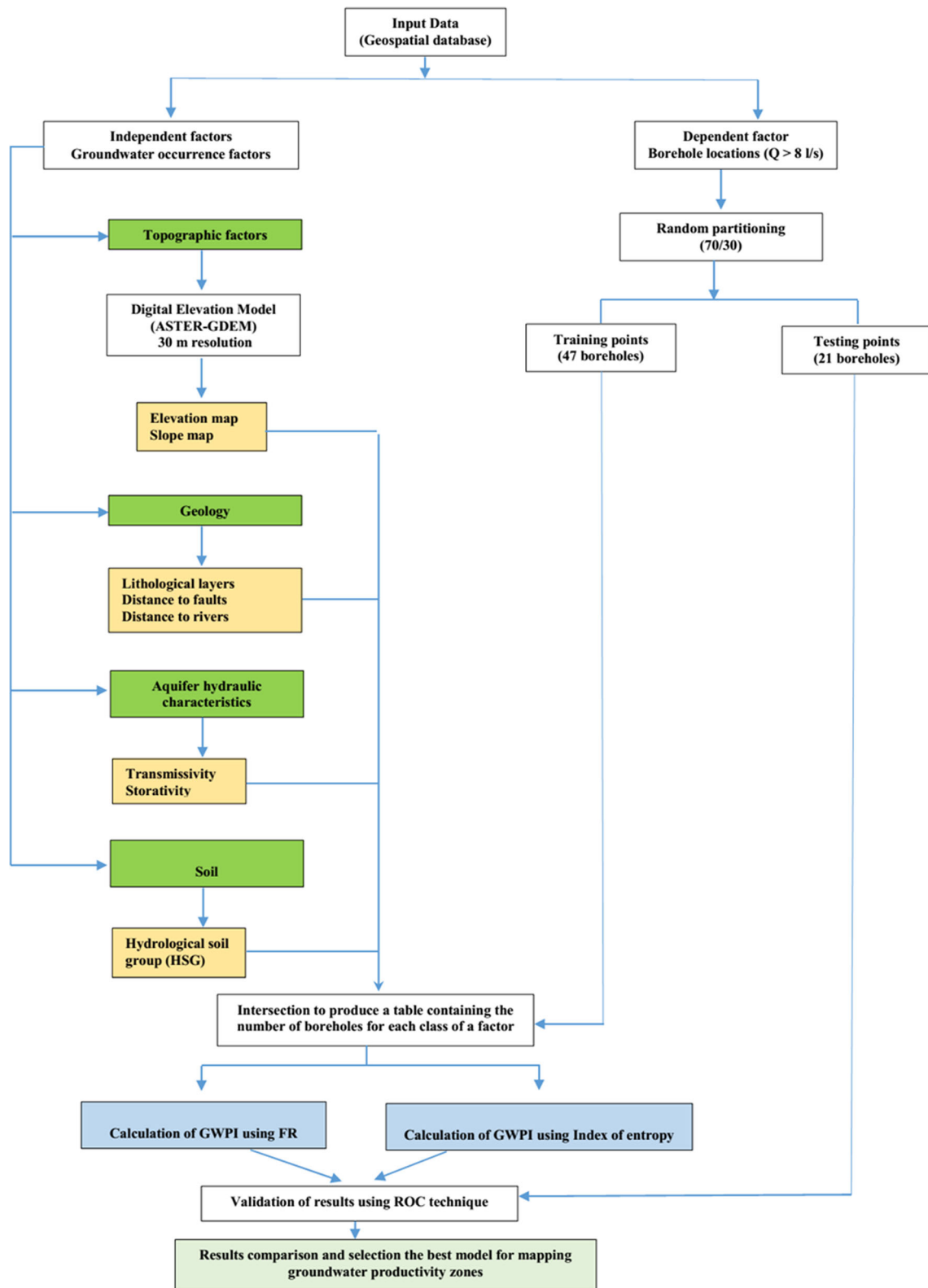
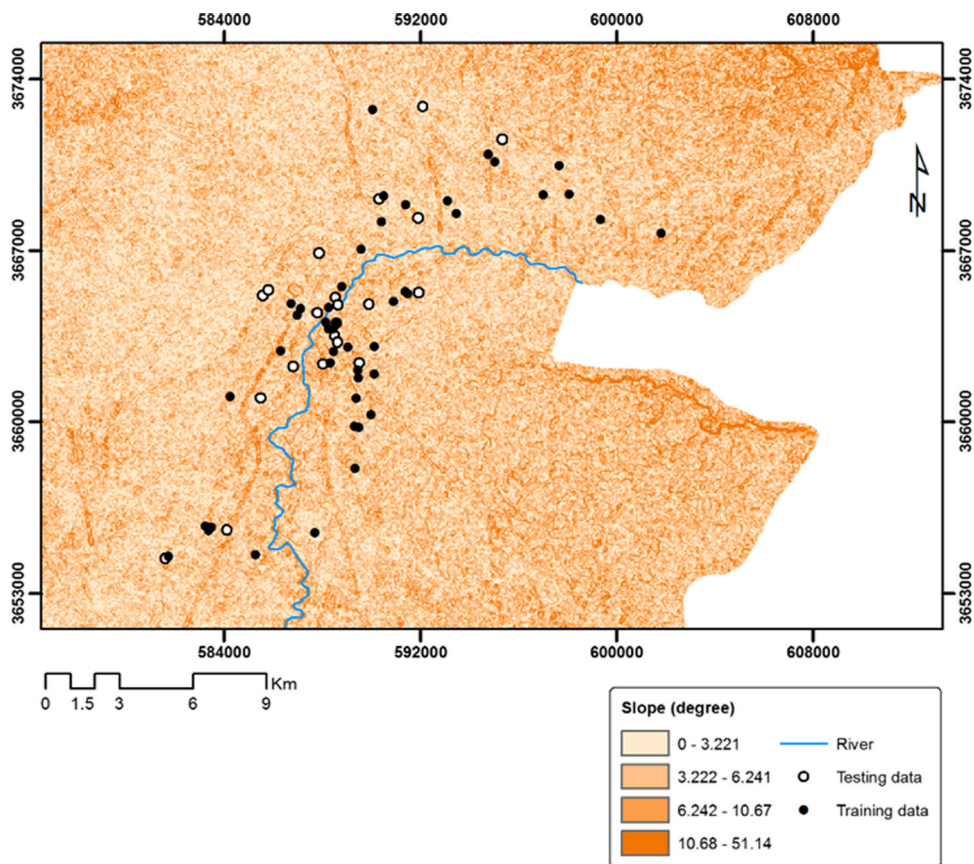


Fig. 8 Flow chart for mapping groundwater productivity index in this study

respect to a class for the factor and a is the percentage for the entire domain. The larger the FR, the stronger the relationship between groundwater production and the

given factor's attribute. The groundwater productivity index based on this technique is calculated as: (Ozdemir 2011b; Jaffari et al. 2013; Naghibi et al. 2014)

Fig. 9 Slope (°) map



$$GWPI = \sum_{i=1}^n FR_i \tag{3}$$

where FR_i is the frequency ratio for a factor and n is the total number of used factors. A detailed mathematical background of this method can be found in Lee et al. (2006).

Shannon’s entropy model

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 (Yufeng and Fengxiang 2009). Shannon entropy is the average unpredictability in a random variable, which is equivalent to its information content. The entropy of groundwater reservoir yield 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 following equations are used to calculate the information coefficient W_j (weigh value for each influencing factor): (Bednarik et al. 2010, 2012; Constantin et al. 2011; Jaafari et al. 2013)

$$P_{ij} = FR = \frac{b}{a} \tag{4}$$

$$(P_{ij}) = \frac{P_{ij}}{\sum_{j=1}^{Sj} P_{ij}} \tag{5}$$

$$H_j = - \sum_{i=1}^{Sj} (P_{ij}) \log_2(P_{ij}), \quad j = 1, \dots, n \tag{6}$$

$$H_{jmax} = \log_2 Sj \tag{7}$$

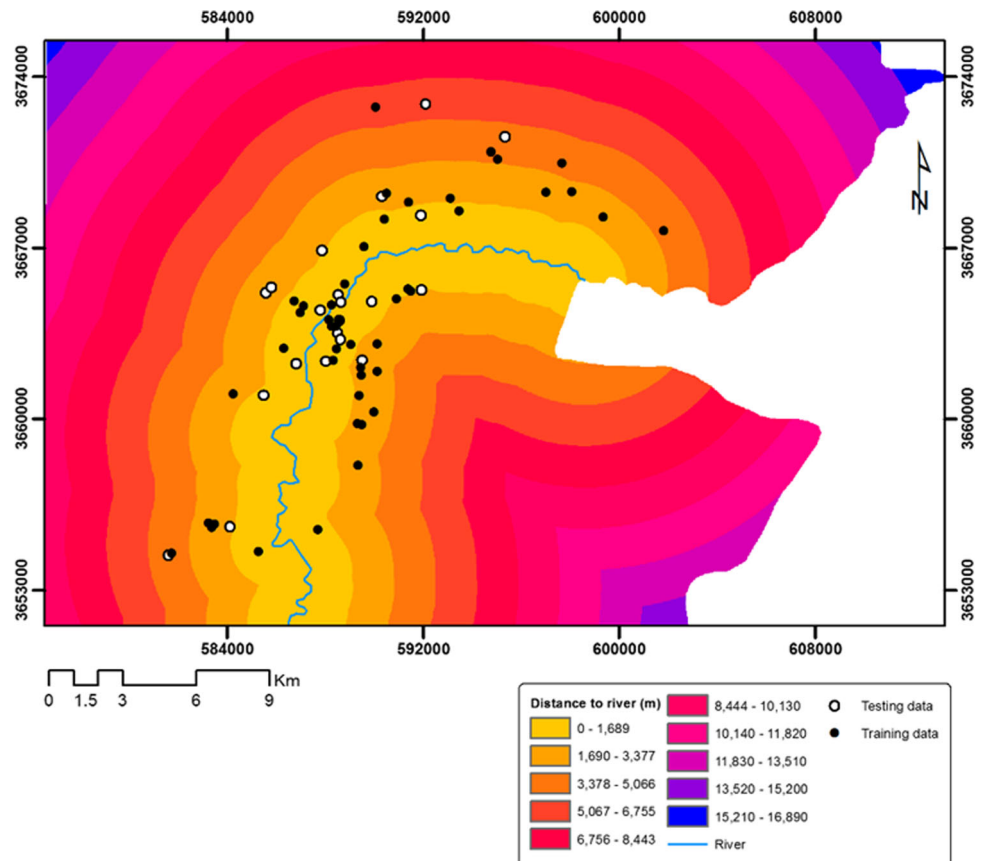
$$I_j = \frac{H_{jmax} - H_j}{H_{jmax}}, \quad I = (0, 1), \quad j = 1, \dots, n \tag{8}$$

$$w_j = I_j P_{ij} \tag{9}$$

where FR is the frequency ratio, (P_{ij}) is the probability density, H_j and H_{jmax} 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 calculated as: (Devkota et al. 2013; Jaafari et al. 2013)

$$y = \sum_{i=1}^n \frac{z}{m_i} \times C \times W_j \tag{10}$$

where y is the sum of all the classes; i is the number of particular factor map; z is the number of classes within

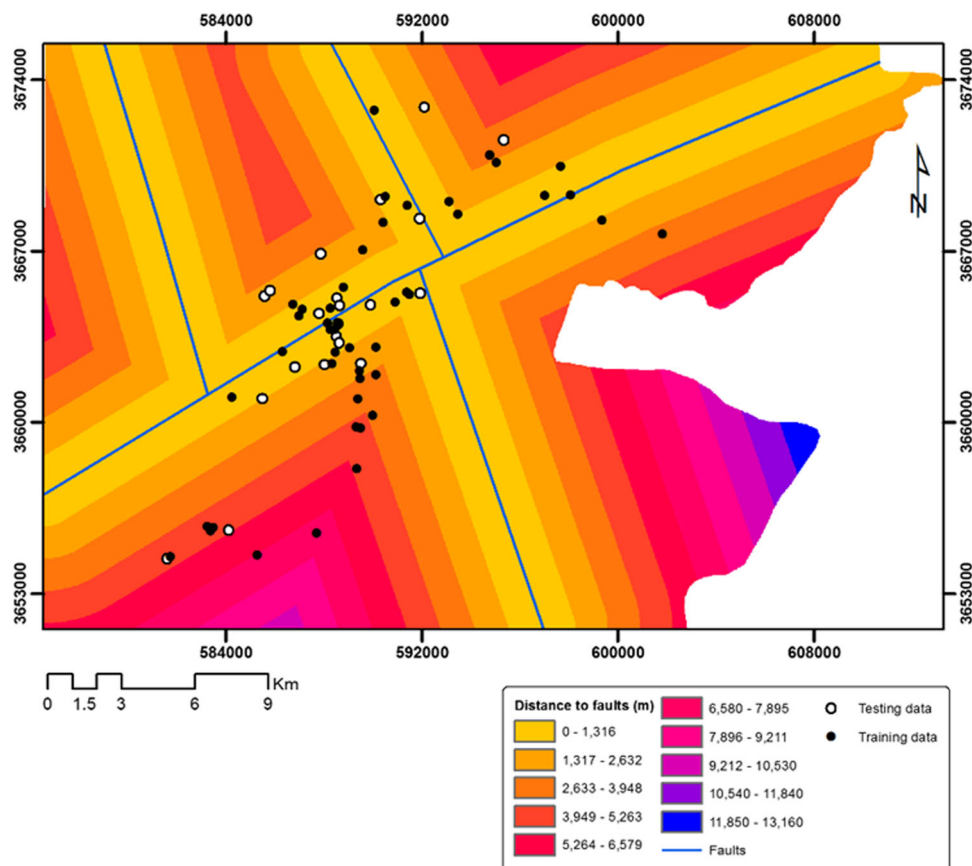
Fig. 10 Distance to river map

factor map with the greatest number of classes; m_i is the number of classes within particular factor map; C is the value of the class after secondary classification; and W_j is the weight of a factor (Bednarik et al. 2010)

Results and discussion

The results of application the two methods were summarized in Table 2. With respect to the FR results, the FR ratios for first elevation ranges (0–56 m) and (56–99 m) were 1.039 and 1.624, respectively, imply high groundwater productivity for these class ranges. The FR ratio for the other classes was low indicating low probability of groundwater productivity. In the literature, it is accepted that groundwater occurrence decreases as the elevation increases. In case of slope, the FR ratio is >1 for the first slope range (0–3.22°) indicating a high correlation between this slope range and groundwater productivity. It is accepted that as the slope increases, then the runoff increases as well leading to less infiltration (Jaiswal et al. 2003). With respect to the study results, the FR decreases as the slope increases, but with the third slope range (6.24–10.67°) it suddenly increases with slope increase and then decreases. To interpret this, it is important to relate

this range with other used factors such geology. The aerial extension of this range is mainly associated with the extension of flood deposits. These deposits consist mainly of sand and gravel and having higher values of hydraulic conductivity. The higher values of FR for flood deposits (1.087) support this conclusion. In case of geology, the Quaternary lithological layers have relatively higher values of FR (1.087, 1.662, and 0.741) for flood deposits, alluvium, and inner flood deposits, respectively. The FRs for the rest of the lithological layers were zero indicating the low probability of groundwater occurrence. If we consider the relationship between groundwater potential and soil factor, it can be seen that FRs are high for the A and B soil groups and low for other groups. The higher infiltration rates of these groups support the resultant higher FR values. As the infiltration rate increases the groundwater recharge increases as well leading to more productivity conditions. In the case of transmissivity and storativity factors, the FR values increase as hydraulic characteristics increase indicating high aquifer productivity conditions in the higher values of these factors. For distance to river factor, the highest FR values of 3.103 and 3.258 concentrate on the first two classes (0–1688 m) and (1688–3377 m), respectively. As the distance to river increase, the FR value decreases until it has no effect on groundwater productivity as

Fig. 11 Distance to faults map

FR becomes zero up to ≈ 6 km. For distance to faults, the highest values of FRs occur on the first fifth classes. Up to 7 km, the FR ratios become zero. This implies the importance of the structural setting on the groundwater occurrence in the study area.

The final groundwater productivity index for the study area was calculated using the Eq. 3 and demonstrated in a map in Fig. 12. The obtained GWPI was classified based on natural break classification scheme into very low, low, moderate, high, and very high classes. The areas covered by each of these classes are summarized in Table 3 in which the high to very high classes extend over an area of 30 % (217 km²). The very low–moderate classes occurred within ≈ 70 % (490 km²) of the study area indicating low productivity conditions of the aquifer system.

Results of applying Shannon's entropy model in the study area, Table 2, revealed that elevation, soil, geology, and slope were the most important factors influencing groundwater productivity conditions in the study area. The weights for these factors were 0.085, 0.073, 0.070, and 0.060, respectively. On the other hand, the other factors (distance to river, transmissivity, distance to faults, and storativity) had a minor effect on groundwater productivity. The calculated weights for these factors were 0.054, 0.035,

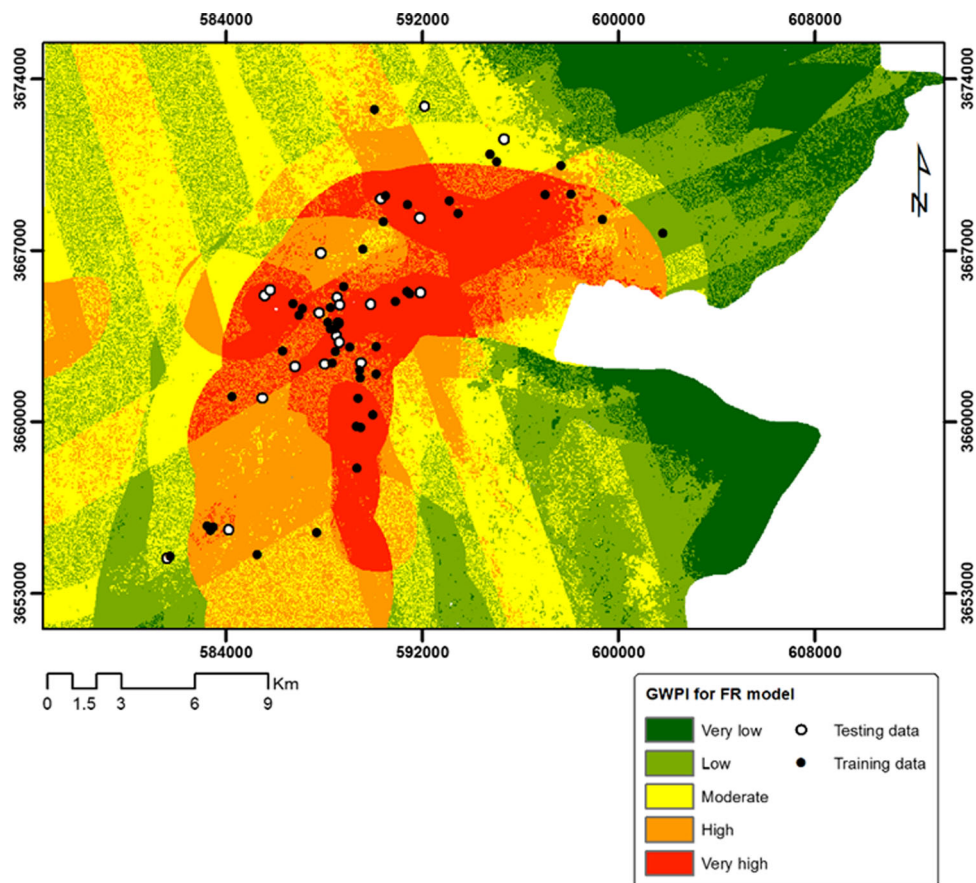
0.033, and 0.020 for distance to river, transmissivity, distance to faults, and storativity factors, respectively. The final GWQI map for this model was developed using Eq. 10. The obtained GWPI was also classified into five classes based on natural break classification scheme, Fig. 13. The area covered by high–very high classes distributed over an area of 31 % (217 km²) consistent with the results of the FR model, Table 3.

Validation of the results

Any predictive model (deterministic or stochastic) requires validation before it can be used in prediction purposes. Without validation, the model will have no scientific significant (Chung and Fabbri 2003). In this context, the Receiver Operating Characteristic (ROC) curve is usually used for examining the quality of deterministic and probabilistic detection and forecast system (Swets 1988). In the ROC curve, the sensitivity of the model (the percentage of boreholes pixels correctly predicted by the model) is plotted against 1-specificity (the percentage of predicted boreholes pixel over the total). The area under the curve (AUC) describes the quality of a forecast system through the system's ability to correctly predict the occurrence or non-occurrence of predefined events (Devkota et al. 2013).

Table 2 Frequency ratio and information coefficient values for the considered factors

Thematic layer	Class	Area (pixels)	Area % (a)	No. of boreholes	Boreholes % (b)	FR (b/a)	(P_{ij})	H_j	H_{jmax}	I_j	W_j
Elevation	0–56	418,091	0.533	26	0.553	1.039	0.354	1.323	2.00	0.339	0.085
	56–99	195,357	0.249	19	0.404	1.624	0.553				
	99–157	121,533	0.155	2	0.043	0.275	0.094				
	157–318	49,929	0.064	0	0	0	0				
Slope	0–3.22	295,915	0.377	21	0.447	1.185	0.354	1.523	2.00	0.238	0.060
	3.22–6.24	311,563	0.397	13	0.277	0.697	0.208				
	6.24–10.67	147,753	0.188	13	0.277	1.469	0.438				
	10.67–51.13	29,679	0.038	0	0	0	0				
Geology	Injana	8407	0.011	0	0	0	0	1.508	2.32	0.350	0.070
	Muqdadiya	114,435	0.146	0	0	0	0				
	Flood deposits	476,358	0.607	31	0.660	1.087	0.311				
	Inner flood deposits	45,070	0.057	2	0.043	0.741	0.212				
Soil	Alluvium	140,640	0.179	14	0.298	1.662	0.476				
	A	488,224	0.622	34	0.723	1.163	0.259	1.419	2.00	0.291	0.073
	B	33,171	0.042	5	0.106	2.517	0.561				
	C	164,860	0.210	8	0.170	0.810	0.180				
Transmissivity	D	98,655	0.126	0	0	0	0				
	20.4–230.7	574,036	0.731	33	0.702	0.960	0.118	1.717	2.00	0.142	0.035
	230.8–471.1	190,032	0.242	11	0.234	0.967	0.119				
	471.2–974.4	16,881	0.022	2	0.043	1.979	0.244				
Specific storage	974.5–1928	3961	0.005	1	0.021	4.216	0.519				
	0.0012–0.0173	51,859	0.066	1	0.021	0.322	0.104	1.840	2.00	0.080	0.020
	0.0173–0.0281	144,846	0.185	5	0.106	0.576	0.186				
	0.0281–0.0378	180,084	0.229	10	0.213	0.927	0.300				
Distance to river	0.0378–0.0561	408,121	0.520	31	0.660	1.269	0.410				
	0–1688	107,641	0.137	20	0.426	3.103	0.420	1.521	3.32	0.542	0.054
	1688–3377	102,523	0.131	20	0.426	3.258	0.441				
	3377–5065	112,311	0.143	6	0.128	0.892	0.121				
	5065–6754	122,070	0.156	1	0.021	0.137	0.019				
	6754–8443	125,165	0.159	0	0	0	0				
	8443–10,131	96,386	0.123	0	0	0	0				
	10,131–11,820	64,562	0.082	0	0	0	0				
	11,820–13,509	37,093	0.047	0	0	0	0				
	13,509–15,197	15,059	0.019	0	0	0	0				
Distance to faults	15,197–16,886	2100	0.003	0	0	0	0				
	0–1315	221,280	0.282	23	0.489	1.736	0.346	2.210	3.30	0.330	0.033
	1315–2631	191,780	0.244	8	0.170	0.697	0.139				
	2631–3947	153,762	0.196	7	0.149	0.760	0.152				
	3947–5263	99,549	0.127	7	0.149	1.174	0.234				
	5263–6579	51,718	0.066	2	0.043	0.646	0.129				
	6579–7895	30,425	0.039	0	0	0	0				
	7895–9211	20,394	0.026	0	0	0	0				
	9211–10,526	9282	0.012	0	0	0	0				
	10,526–11,842	4556	0.006	0	0	0	0				
11,842–13,158	2164	0.003	0	0	0	0					

Fig. 12 Groundwater potential index map (FR model)**Table 3** Distribution of the GWPI classes and areas covered by each of these classes

GWPI class	FR model			Shannon's entropy model		
	Range	Area (%)	Area (km)	Range	Area (%)	Area (km)
Very low	2.277–5.595	0.14	97	0.097–0.283	0.12	83
Low	5.596–7.195	0.25	175	0.284–0.377	0.24	172
Moderate	7.196–8.593	0.31	217	0.378–0.457	0.33	231
High	8.594–10.30	0.18	129	0.458–0.557	0.19	134
Very high	10.31–14.58	0.12	88	0.558–0.746	0.12	86

The predictive capability of the model is excellent if $AUC = 1-0.9$; very good $0.8-0.9$; good $0.8-0.7$; $0.7-0.6$ average; and poor $0.6-0.5$ (Yesilnacar 2005). The AUC was obtained for both the training (success rate) and testing (prediction rate) for both models by using ROC module in IDRISI software, Figs. 14 and 15. The success rate is important to explain how well the resulting GWPI map classified the area of existing borehole locations. The success rate results were obtained by comparing the training borehole locations (47) with the two GWPI maps. The AUC for FR and Shannon's model was 0.874 and 0.844, respectively implying that FR performs better than Shannon's model. On the other hand, the prediction rate used a measure of performance of a predictive rule

(Yesilnacar and Topal 2005; Pradhan et al. 2010). It only used the testing data set to explore the predictive capability of the model. The AUC for prediction rate is shown in Figs. 14 and 15, for both models. The FR model had slightly better predictive capability than Shannon's entropy model where AUC for FR and Shannon's was 0.869 and 0.824, respectively. The prediction accuracy for FR was $\approx 87\%$ while for Shannon's entropy was $\approx 82\%$. It can be seen that both models were capable to prospect GWPI with very good results, but FR was better in terms of success and prediction rates. This conclusion supports the use of this very simple method to demarcate groundwater productivity zones instead of using more complicated models such as Shannon's entropy model.

Fig. 13 Groundwater potential index map (Shannon’s entropy model)

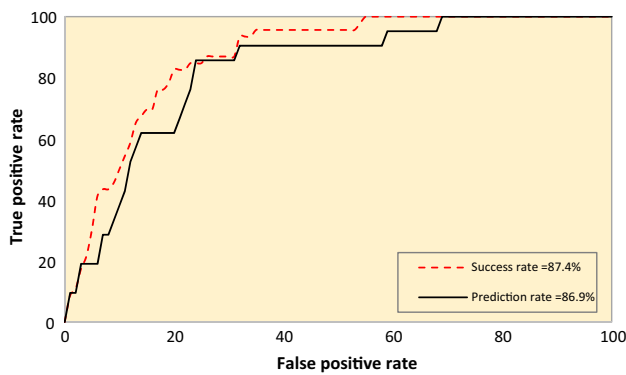
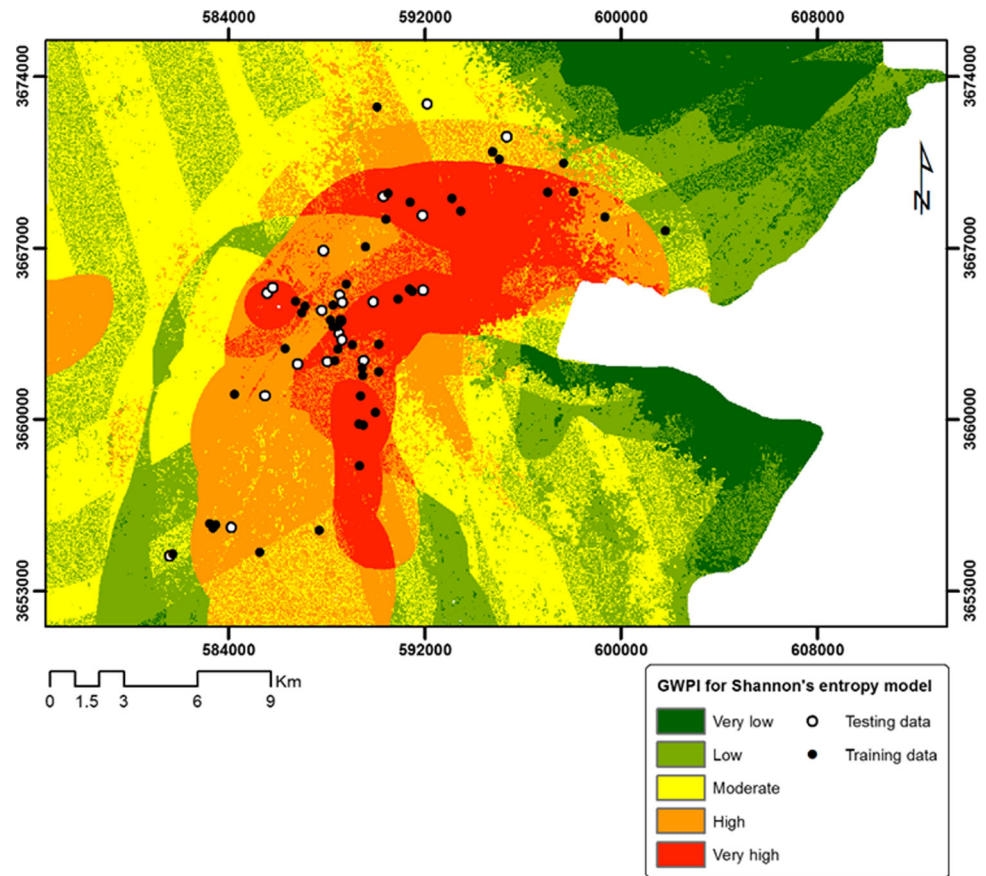


Fig. 14 ROC analysis of FR results

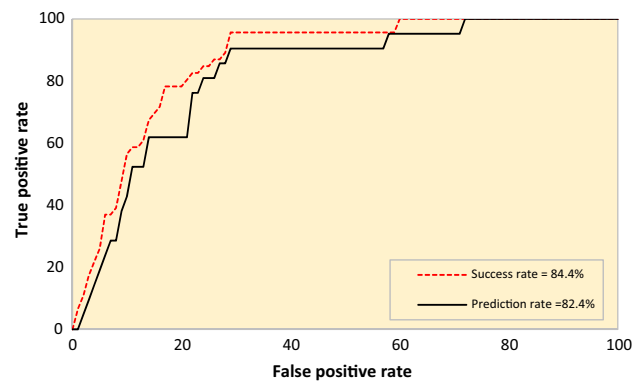


Fig. 15 ROC analysis of Shannon’s entropy model

Conclusions

Demarcation of groundwater prospective zones of an area plays an increasingly significant role for sustainable management of groundwater resource across the world. In this study, an effort made to delineate groundwater productivity at northeastern Wasit governorate using probabilistic ratio and Shannon’s entropy models. The first one is popular in the analysis of relationship between groundwater reservoir productivity and groundwater occurrence influential factors. Only few number of studies deal with application of

the second method in the groundwater studies. In order to prepare the groundwater productivity map by using these two methods, eight factors that are believed to have influence on the groundwater occurrence within the study area were selected and used as the input data. These factors were elevation (m), slope angle (degree), geology, soil, aquifer transmissivity (m²/d), specific storage (dimensionless), distance to river (m), and distance to faults (m). The total boreholes used in analysis were 68. 47 boreholes (70 %) were used as training data and the rest 21 (30 %) were used for validation. The two GWPI maps were

validated using reservoir operating characteristics curves. The AUC curve for training and testing (success rate and prediction rate) showed that the two models show similar performance. The FR model was slightly better than Shannon's entropy (success rate, 87.4 %; prediction rate, 86.9 % for FR; success rate, 84.4 %; prediction rate, 82.4 % for Shannon's entropy). The final conclusion was that both models were capable to produce groundwater prospective zones with very good accuracy. Results of this study could be helpful for better management of groundwater reserve in the study area and give planners and decision makers an opportunity to prepare appropriate groundwater investment plans.

Acknowledgments The author sincerely acknowledges the efforts of Mr. Fadil Al-Aqabi from General Commission of Groundwater/Missan Governorate for the great collaboration and very warm hospitality. Great and deep appreciation for Mr. Hazim Abbas Naser from Geology Department, College of Science, Basra University for his support and review of the English language of the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abdalla F (2012) Mapping of groundwater prospective zones using remote sensing and GIS techniques: a case study from the Central Eastern Desert, Egypt. *J Afr Earth Sci* 70:8–17
- Al-Abadi AM (2015) Groundwater potential mapping at northeastern Wasit and Missan governorates, Iraq using a data-driven weights of evidence technique in framework of GIS. *Environ Earth Sci*. doi:10.1007/s12665-015-4097-0
- Al-Abadi AM, Al-Shamma'a A (2014) Groundwater potential mapping of the major aquifer in Northeastern Missan Governorate, South of Iraq by using analytical hierarchy process and GIS. *J Environ Earth Sci* 10:125–149
- Al-Sayab A, Al-Ansari N, Al-Rawi D, Al-Jassim J, Al-Omari F, Al-Shaikh Z (1982) *Geology of Iraq*. Mosul University (In Arabic), Mosul
- Al-Shammery SH (2006) *Hydrogeology of Galal Basin, Wasit, east of Iraq*. PhD Thesis, Baghdad University, Iraq (unpublished)
- Bednarik M, Magulova B, Matys M, Marschalko M (2010) Landslide susceptibility assessment of the Kral ovany–Liptovský Mikuláš railway case study. *Phys Chem 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–Sereď landslide area in south-west Slovakia. *Nat Hazards* 64:547–575
- Bonham-Carter GF (1994) *Geographic information systems for geoscientists, modeling with GIS*. Pergamon Press, Oxford
- Chen J, Zhang Y, Chen Z, Nie Z (2014) Improving assessment of groundwater sustainability with analytic hierarchy process and information entropy method: a case study of the Hohhot Plain. *Environ Earth Sci, China*. doi:10.1007/s12665-014-3583-0
- Chi K, Lee BJ (1994) Extracting potential groundwater area using remotely sensed data and GIS techniques. In: *Proceedings of the Arab J Geosci Author's personal copy Regional Seminar on Integrated Applications of Remote Sensing and GIS for Land and Water Resources Management*. Bangkok (Bangkok: Economic and Social Commission for Asia and the Pacific), pp 64–69
- Chowdary V, Chakraborty D, Jeyaram A, Murthy YK, Sharma J, Dadhwal V (2013) Multi-Criteria decision making approach for watershed prioritization using analytic hierarchy process technique and GIS. *Water Resour Manage* 27:1–17
- Chowdhury A, Jha MK, Chowdary VM (2010) Delineation of groundwater recharge zones and identification of artificial recharge sites in West Medinipur District, West Bengal using RS, GIS and MCDM techniques. *Environ Earth Sci* 59:1209–1222. doi:10.1007/s12665-009-0110-9
- Chung CF, Fabbri AG (2003) Validation of spatial prediction models for landslide hazard mapping. *Nat Hazards* 30:451–472
- Constantin M, Bednarik M, Jurchescu MC, Vlaicu M (2011) Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). *Environ Earth Sci* 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
- Dar AI, Sankar K, Dar MA (2010) Remote sensing technology and geographic information system modeling: an integrated approach towards the mapping of groundwater potential zones in Hardrock terrain, Mamundiyyar basin. *J Hydrol* 394:285–295. doi:10.1016/j.jhydrol.08.022
- Devkota KC, Regmi AD, Pourghasemi HR, Youshid K, Pradhan B, In Ryu, Dhital MR, Althuwanee OF (2013) Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling-Narayanghat road section in Nepal Himalaya. *Nat Hazards* 65:135–165
- Edet A, Okereke CS, Teme SC, Esu EO (1998) Application of remote-sensing data to groundwater exploration: a case study of the Cross River State, Southeastern Nigeria. *Hydrogeo J* 6:394–404. doi:10.1007/s100400050162
- Elmahdy SI, Mohamed MM (2014) Probabilistic frequency ratio model for groundwater potential mapping in Al Jaww plain, UAE. *Arab J Geosci*. doi:10.1007/s12517-014-1327-9
- Ganapuram S, Vijaya Kumar GT, Murali Krishna IV, Kahya E, Cüneyd Demirel M (2009) Mapping of groundwater potential zones in the Musi basin using remote sensing data and GIS. *Adv Eng Softw* 40:506–518
- Gumma MK, Pavelic P (2013) Mapping of groundwater potential zones across Ghana using remote sensing, geographic information systems, and spatial modeling. *Environ Monit Assess* 185:3561–3579. doi:10.1007/s10661-012-2810-y
- Ihara S (1993) *Information theory for continuous systems*. World Scientific Pub Co Inc, Hackensack
- Jaafari A, Najafi A, Pourghasemi HR, Rezaeian J, Sattarian A (2013) GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran. *Int J Environ Sci Technol* 11:909–926. doi:10.1007/s13762-013-0464-0
- Jaiswal RK, Mukherjee S, Krishnamurthy J, Saxena R (2003) Role of remote sensing and GIS techniques for generation of groundwater prospect zones towards rural development-an approach. *Int J Remote Sens* 24:993–1008
- Jassim SZ, Goff JC (2006) *Geology of Iraq*. Dolin, Prague and Moravian Museum, Brno, p 431

- Jenks GF (1967) The data model concept in statistical mapping. *Int Yearb Cartogr* 7:186–190
- Jha MK, Chowdhary VM, Chowdhury A (2010) Groundwater assessment in Salboni Block, West Bengal (India) using remote sensing, geographical information system and multi-criteria decision analysis techniques. *Hydrogeol J* 18:1713–1728. doi:10.1007/s10040-010-0631-z
- Kamaraju MV, Bhattacharya A, Reddy GS, Rao GC, Murthy GS, Rao TC (1995) Groundwater potential evaluation of West Godavari District, Andhra Pradesh State, India- a GIS approach. *Ground Water* 34:318–325
- Khodaei K, Nassery HR (2011) Groundwater exploration using remote sensing and geographic information systems in a semi-arid area (Southwest of Urmieh Northwest of Iran). *Arab J Geosci* 6:1229–1240. doi:10.1007/s12517-011-0414-4
- Krishnamurthy J, Srinivas G (1995) Role of geological and geomorphological factors in groundwater exploration: a study using IRS LISS data. *Int J Remote Sens* 16:2595–2618
- Krishnamurthy J, Venkatesa K, Jayaraman V, Manuvel M (1996) An approach to demarcate groundwater potential zones through remote sensing and a geographical information system. *Int J Remote Sens* 17:1867–1884
- Lee S, Oh H, Park N (2006) Mineral potential assessment of sedimentary deposit using frequency ration and logistic regression of Gangreung area, Korea. In: *IEEE International Conference on Geoscience and Remote Sensing Symposium*, pp 1576–1579. doi:10.1109/IGARSS.2006.406
- Lee S, Kim YS, Oh HJ (2012) Application of a weight-of-evidence method and GIS to regional groundwater productivity potential mapping. *J Environ Manage* 96:91–105. doi:10.1016/j.jenvman.2011.09.016
- Machiwal D, Madan KJ, Bimal CM (2010) Assessment of groundwater potential in a semi-arid region of India using remote sensing, GIS and MCDM techniques. *Water Resour Manage* 25:1359–1386
- Madrucci V, Taioli F, de Araújo CC (2008) Groundwater favorability map using GIS multicriteria data analysis on crystalline terrain, São Paulo State, Brazil. *Hydrogeol J* 357:153–173
- Malczewski J (1999) *GIS and multicriteria decision analysis*. Wiley, New York
- Manap MA, Sulaiman WN, Ramli MF, Pradhan B, Surip N (2011) A knowledge-driven GIS modeling technique for groundwater potential mapping at the Upper Langat Basin, Malaysia. *Arab J Geosci* 6:1621–1637. doi:10.1007/s12517-011-0469-2
- Moghaddam DD, Rezaei M, Pourghasemi HR, Pourtaghi ZS, Pradhan B (2013) Groundwater spring potential mapping using bivariate statistical model and GIS in the Taleghan Watershed, Iraq. *Arab J Geosci*. doi:10.1007/s12517-013-1161-5
- Naghbi SA, Pourghasemi HR, Pourtaghi ZS, Rezaei A (2014) Groundwater qanat potential mapping using frequency ratio and Shannon's entropy models in the Moghan watershed, Iraq. *Earth Sci Inform*. doi:10.1007/s12145-014-0145-7
- Nampak H, Pradhan B, Manap MA (2014) Application of GIS based data driven evidential belief function model to predict groundwater potential zonation. *J Hydrol* 513:283–300
- Oh HJ, Kim YS, Choi JK, Park E, Lee S (2011) GIS mapping of regional probabilistic groundwater potential in the area of Pohang City, Korea. *J Hydrol* 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). *J Hydrol* 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. *J Hydrol* 411:290–308
- Pandey VP, Shrestha S, Kazama F (2013) A GIS-based methodology to delineate potential areas for groundwater development: a case study from Kathmandu Valley, Nepal. *Appl Water Sci* 3:453–465. doi:10.1007/s13201-013-0094-1
- Parsons RM (1956) *Groundwater resources of Iraq*. Khanaqin-Jassan area, vol 1. Development Board, Ministry of Development, Government of Iraq. ILL/Joseph R. Skeen Library, New Mexico Institute of Mining and Technology, Socorro, NM 87801
- Pourtaghi ZS, Pourghasemi HR (2014) GIS-based groundwater spring potential assessment and mapping in the Birjand Township, southern Khorasan Province, Iran. *Hydrogeol J* 22:643–662. doi:10.1007/s10040-013-1089-6
- Pradeep Kumar GN, Srinivas P, Jaya Chandra K, Sujatha P (2010) Delineation of groundwater potential zones using remote sensing and GIS techniques: a case study of Kurmapalli Vagu basin in Andhra Pradesh, India. *Int J Water Resour Environ Eng* 2:70–78
- Pradhan B, Lee S, Buchroithner MF (2010) Remote sensing and GIS-based landslide susceptibility analysis and its cross-validation in three test areas using a frequency ratio model. *Photogramm Fernerkun* 1:17–32. doi:10.1127/1432-8364/2010/0037
- Rahmati O, Nazari Samani A, Mahdavi M, Pourghasemi HR, Zeinivand H (2014) Groundwater potential mapping at Kurdistan region of Iran using analytic hierarchy process and GIS. *Arab J Geosci*. doi:10.1007/s12517-014-1668-4
- Rao YS, Jugran DK (2003) Delineation of groundwater potential zones and zones of groundwater quality suitable for domestic purposes using remote sensing and GIS. *Hydrol Sci J* 48:821–833
- Ravi Shankar MN, Mohan G (2006) Assessment of the groundwater potential and quality in Bhatsa and Kalu river basins of Thane district, western Deccan Volcanic Province of India. *Environ Geol J* 49:990–998
- Sahu P, Sikdar PK (2011) Groundwater potential zoning of a pre-urban wetland of south Bengal Basin, India. *Environ Monit Assess* 174:119–134. doi:10.1007/s10661-010-1443-2
- Sander P, Chesley MM, Minor TB (1996) Groundwater assessment using remote sensing and GIS in a rural groundwater project in Ghana: lessons learned. *Hydrogeol J* 4:40–49
- Saraf AK, Choudhury PR (1998) Integrated remote sensing and GIS for groundwater exploration and identification of artificial recharge sites. *Int J Remote Sens* 19:1825–1841
- Satty TL (1980) *The analytic hierarchy process*. McGraw-Hill, New York
- Sener E, Davraz A, Ozcelik M (2005) An integration of GIS and remote sensing in groundwater investigations: a case study in Burdur, Turkey. *Hydrogeol J* 13:826–834
- Shahid S, Nath SK, Roy J (2000) Groundwater potential modeling in a GIS. *Int J Remote Sens* 21:1919–1924
- Shahid S, Nath SK, Kamal AS (2014) GIS integration of remote sensing and topographic data using fuzzy logic for ground water assessment in Midnapur District, India. *Geocarto Int* 17:69–74. doi:10.1080/10106040208
- Sikdar PK, Chakraborty S, Adhya E, Paul PK (2004) Land use/land cover changes and groundwater potential zoning in and around Raniganj coal mining area, Bardhaman District, West Bengal: a GIS and remote sensing approach. *Spat Hydrol J* 4:1–24
- Solomon S, Quiel F (2006) Groundwater study using remote sensing and geographic information systems (GIS) in the central highland of Eritrea. *Hydrogeol J* 14:729–741
- Suja Rose RS, Krishnan N (2009) Spatial analysis of groundwater potential using remote sensing and GIS in the Kanyakumari and Nambiyar basins, India. *J Indian Soc Remote Sens* 37:681–692
- Swets JA (1988) *Measuring the accuracy of diagnostic systems*. Science 240:1285–1293
- Todd DK, Mays LW (2005) *Groundwater hydrology*. Wiley, New York, p 652

- United States Department of Agriculture, Soil Conservation Service (USDA) (1986) Urban hydrology for small watersheds, Technical release no. 55, 2nd edn, Washington, DC
- Yesilnacar EK (2005) The application of computational intelligence to landslide susceptibility mapping in Turkey. PhD Thesis. University of Melbourne, Australia, p 423
- 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). Eng Geol 79:251–266
- Yufeng S, Fengxiang J (2009) Landslide stability analysis based on generalized information entropy. Int Conf Environ Sci Inf Appl Technol 2:83–85