International Journal of Engineering & Technology

Research paper

Groundwater Site Prediction Using Remote Sensing, GIS and Statistical Approaches: A Case Study in the Western Desert, Iraq

Fadhil M. Shnewer [1]*, Alauldeen A. Hasan [2], Mudhaffar S. AL-Zuhairy[3]

[1]University of Misan , Engineering College, Civil Department, Misan , Iraq
[2]Middle Technical University, College of Technical Engineering, Surveying Department, Baghdad, Iraq
[3] Southern Technical University, Civil Department, Basra , Iraq
*Corresponding author E-mail: fadhil.mz87@uomisan.edu.iq

Abstract

Combination of remote sensing data and geographical information system (GIS) for the investigation of groundwater has become an advance approach in the researches of groundwater. The purpose of this research is to apply statistical models such as Evidential Belief Function (EBF) and Logistic Regression (LR) for mapping groundwater potential sites at Iraqi western desert (located at Al-Ramadi and Shithatha). The potential of the groundwater areas were determined depending on the spatial relationship between groundwater wells and different conditioning factors. These factors include altitude, curvature, aspect, slope, soil, normalized difference vegetation index (NDVI), topographic wetness index, fault, rainfall, stream density, stream power index, and lithology. The algorithms were used to model all layers of groundwater conditioning factors to generate groundwater probability areas. Then, the final maps included five potential classes i.e., very high, high, moderate, low and very low susceptible zones were generated. The final outcomes were validated using Area Under the Curve (AUC) algorithm. The values of success rates were 76.5% and 71.5% for EBF an LR respectively. The prediction rates for the same methods were 73.7% and 70%, respectively. The thematic maps attained from the present study indicated the capability of EBF and LR methods in groundwater potential mapping.

Keywords: *Evidential Belief Function (EBF); GIS; Groundwater; Iraqi western desert; Logistic Regression (LR).*

1. Introduction

Groundwater is one of the most vital source of water that offers to the requirements in all climatic areas in the world and is the most valuable and dependable source of water [1]. It is formed by snowmelt- water or rain-water that leaks down over the soil and into the underlying rocks [2]. Groundwater meets the need in the absence of surface- water and is more preserved from the contagions and pollutants in comparison with the surface –water. In addition, it is much regularly extent over large regions, is less exposed to permanent and seasonal fluctuations, and has a superior quality [3]. Iraqi lands have suffered from water shortages for a long time due to lack of rain-water. The absence of best employment of water resource by making more advanced irrigation system is additional reason for this problem. Another reason regard to external affairs is that Turkey and Iran have built many dams in the upstream of rivers (Euphrates and Tigris) which flow through Iraq, One of the serious problems in Iraq at the present time is how to explore and develop the water resource. The western desert of Iraq will be highly impacted by the conditions revealed above because it will stop any future plans for irrigation system growth through the use of Euphrates River. The rising need for groundwater requires a new and rapid method for searching the groundwater where the traditional approaches such as geophysical, hydrological and geological required high economic investments with high cost. The current study focuses on predicting and mapping of groundwater in an important site of Iraq using GIS modeling based statistical methods. This approach has appeared as a necessary baseline in the perdition of groundwater sites.

With respect to the previous researches, few studies have been conducted concerning groundwater prediction in Iraqi western desert using the approach of remote sensing and GIS [4, 5]. More recently, several global studies have been conducted to assess the groundwater sites using statistical models. Some of studies have been used index- based model [6, 7, 8]. Other studies have been applied probabilistic models such as weights-of evidence [9, 10], Analytical Hierarchy Process (AHP) [11, 12], multi-criteria–decision-analysis [13, 14], and frequency ratio [15]. In recent years, some soft computing methods have been applied for mapping the potential zones of groundwater, for example numerical modeling and decision tree (DT) [16], and fuzzy logic [17, 18].

Two statistical models were applied in this paper: (1) evidential belief function (EBF); and (2) logistic regression (LR). Each model built the ground-water potential map based on the spatial relation between the dependent variable (DV) which indicates the ground-water wells scattered in the region of interest, and independent variables which represent the groundwater conditioning aspects. The conditioning elements chosen for this research are; altitude, slope, aspect, curvature, lithology, fault, soil, normalized difference vegetation index (NDVI), rainfall, topographic wetness index (TWI), density of stream and stream power index (SPI).

2. Study Area

The area under study locates in western desert of Iraq (Shithatha and Al-Ramadi) with an area about 5391.65 km². It lies between longitude (42º 02' 12" E to 42º 59' 06" E) and altitude (32º 43' 55" N to 33º 25' 45" N), as shown in fig. (1). Numerous large valleys drain in the study area from the west into the east such as Mohammadi valley which drains into Euphrates river, as well as to Ghadaf and Thumail valley which drains into Al-Razazza lake [19]. Investigation of groundwater in this study area is necessary especially there is no surface water. Another reason is that the region was considered as a promising area for extraction of ground water [20] as well as there are a sufficient number of wells in this region. These will help to conduct ground water studies. The research area receives an average of (100 -150 mm) annual rain-fall with a maximum and a minimum temperature of 50°C and 16°C, respectively [21]. From topographic point of view, the study area is commonly classified as a flat zone, increasing in elevation from east to west.

3. Data used

Twelve available variables were selected to estimate the sites of ground-water in the study area. These variables are: rainfall, soil, elevations, lithology, curvature, topographic wetness index (TWI), aspect, faults, slope, stream power index (SPI), normalized difference vegetation index (NDVI) and drainage density. Most of these parameters have not been used in local groundwater studies; in addition, many of the previous global studies recommend the use of these elements; these are the reasons behind the selection of these parameters.

The digital elevation model (DEM) of the study area was obtained from Aster-date with spatial resolution of 30 m. DEM is basically defined as a spatially geo-referenced data set that is a common technique for representing the topography for environmental modeling objectives [22]. DEM is used to derive the following topographic factors: elevations, aspect, slope, curvature. It is also used to compute several water- related factors like, TWI, drainage density and SPI. Soil map (of 1:1,000,000 scale) was gained from ministry of agriculture- Iraq as a raster map then digitized in GIS to gain the soil layers. Ministry of science and technology-agricultural research department provided geological map (of 1:250,000 scale) as a raster form, then digitized to obtain the faults and geological layers. Rainfall data (1982–2013 years) was prepared from Iraqi Meteorological Organization and Seismology in excel file, and then the data was interpolated to produce rainfall map through GIS using Inverse Distance Weighted (IDW) algorithm. NDVI is a numerical index that usages the near-infrared and visible (red) bands of electromagnetic range, and is implemented to analyze the measurements of remote sensing and evaluate whether the feature being detected comprises living green vegetation or not [23]. The information of NDVI can be derived through concentrating on the bands of satellite that are more sensitive to vegetation information (red and near-infrared). The values of NDVI are ranged from +1.00 to -1.00. In a practical sense, the low NDVI value (0.1 or less) represents barren ground, snow, water bodies or sand. The moderate values (0.2 to 0.5) denote to sparse vegetation like grasslands, shrubs or senescing

crops. Finally, the high values (0.6 to 0.9) relate to the vegetation dense. The NDVI values were derived using remote sensing imagery (Landsat 8 ETM+) by applying equation (1) as follow [23]:

$$NDVI=(NIR-VIS)/(NIR+VIS) \qquad (1)$$

Where: *VIS and NIR: Represent the spectral reflectance measurements attained in the visible (red) and near-infrared regions, respectively.*

The image was geometrically corrected by the source and no other preprocessing was required. The image was masked by the study area to be similar to the other layers.

Using GIS package, the thematic map for each factor was prepared. All the data used in this paper are presented in fig. (2). In this research the groundwater sites potential were estimated depending on the strong relation between factors mentioned above and ground-water wells locations. Ministry of Water Resources-General Commission for Groundwater provided the date of ground water wells (43 wells). It is randomly separated in to 30 and 13 wells as training and testing dataset, respectively. A training data-set will be used in models analysis, and testing data-set will be used for validation of results.
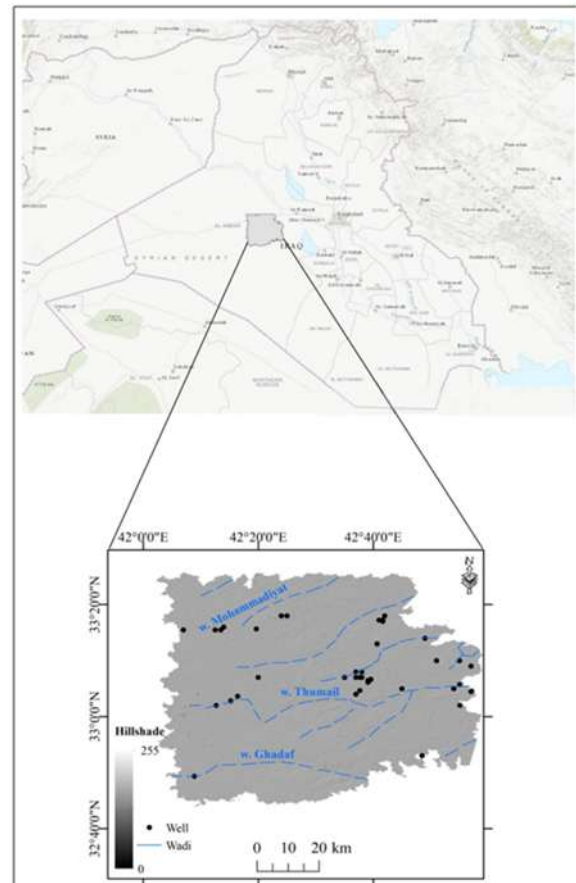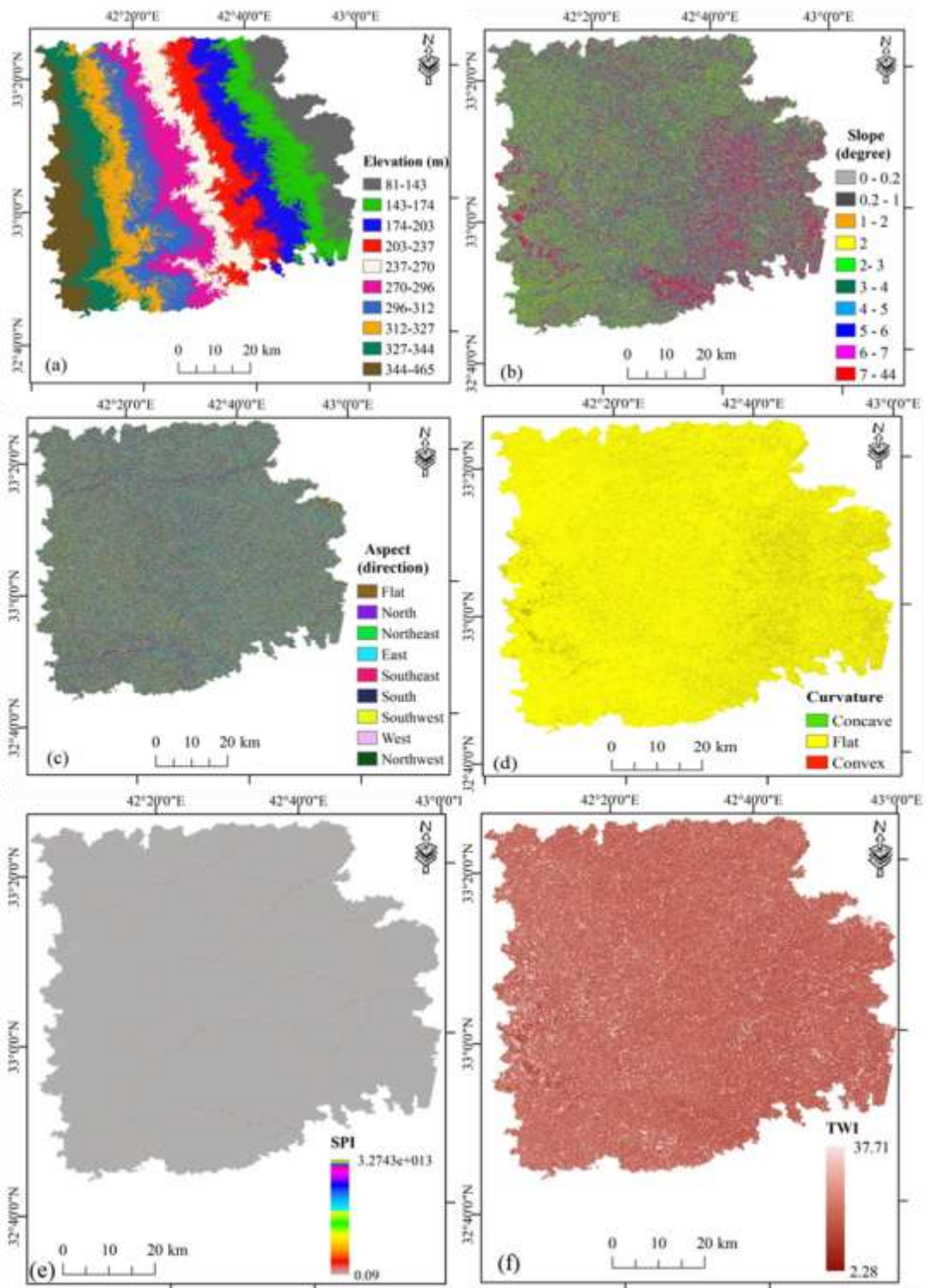

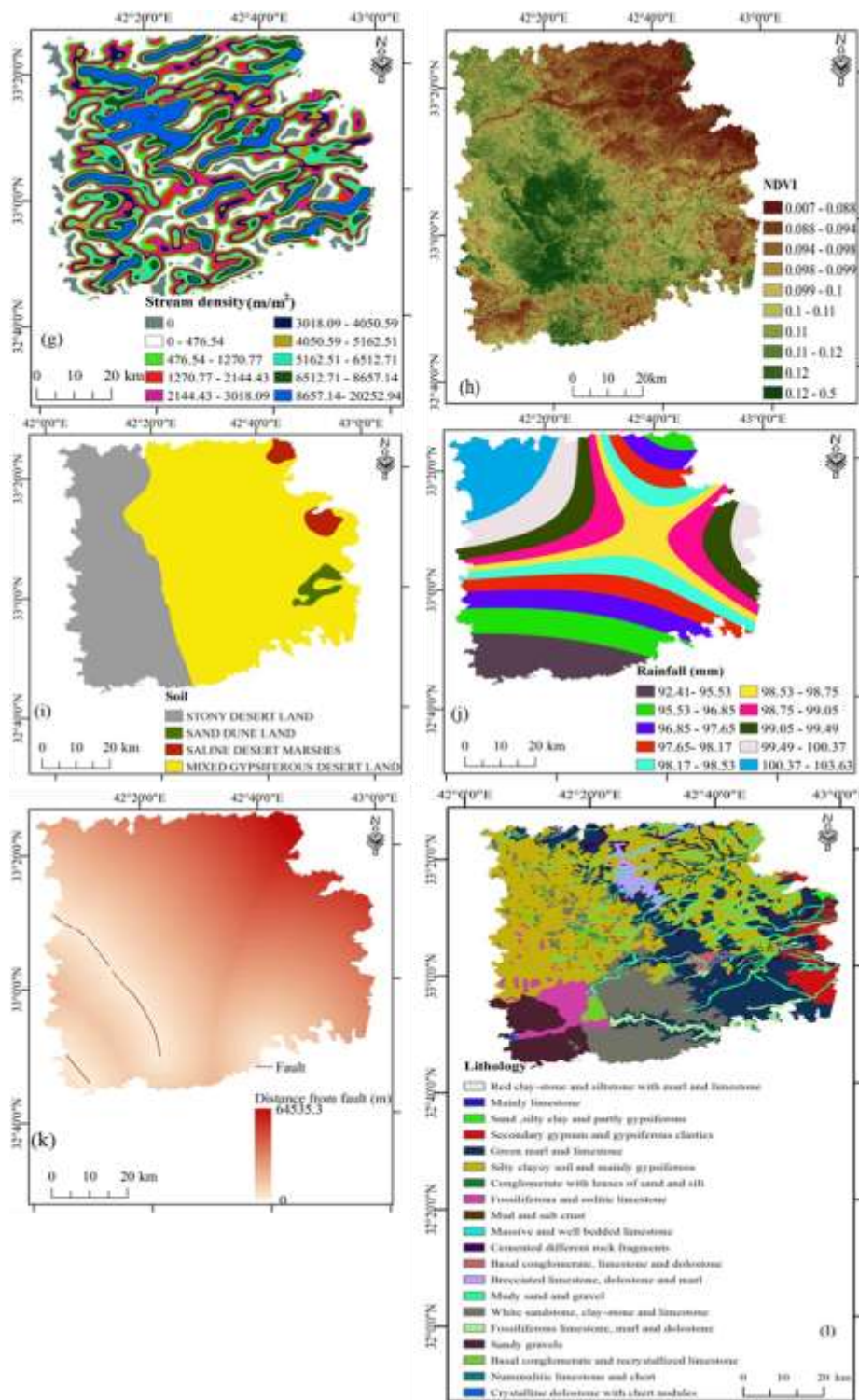**Fig. 1:** Location map of the study area

Fig. 2: Groundwater conditioning factors; (a) altitude,(b) slope, (c) aspect, (d) curvature, (e) SPI, (f) TWI, (g) stream density, (h) NDVI, (i) soil, (j) rainfall,(k) fault, and (l) lithology

4. Methodology overview

The following steps must be done to estimate the groundwater sites potential using the statistical models [9]: (a) collection of real data and building of geo-database concerning to groundwater occurrence, (b) finding the strong relation between the groundwater occurrence factors and the groundwater wells, (c) the final outcomes should be validated, (d) interpretation and description of the results obtained. The general methodology that considered in this research is shown in fig. (3)

4.1. Evidential Belief Function (EBF) model

The conversion of the groundwater conditioning parameters into evidential layers is a priority for the application of EBF model. The next step, compute the quantitative statistics of the spatial association among the evidential layers and groundwater wells to combine of the thematic layers and create the probability map of groundwater potential index (GWPI). There is a scale for the EBF approach that recognized by Hale and Carranza, (2003) [24], this scale comprises belief degree (Bel), disbelief degree (Dis), uncertainty degree (Unc) and plausibility degree (Pls) in range of [0,1]. According to (Dempster, 1967) [25], the principal elements of the theory are displayed by Pls and Bel as upper and lower likelihood, respectively; in addition, the basic probability

assignment function (bpa or m) designates a mapping of power set to (0 to 1). Equation (2) displays the theory of Dempster-Shafer of evidence which is created from Hale and Carranza (2003).

$$m: P(H) = \{0, 1\} \tag{2}$$

Where: *H is referred to all the hypotheses.*

Belief degree represents the probability of occurrence the event (in this case is groundwater), and plausibility indicates to highest degree of belief [26]. Based on mass or bpa function, the Bel and Pls function can be derived by the following equations.

$$Bel(H) = \sum_{A \sqsubseteq H} m(A) \tag{3}$$

$$Pls(H) = \sum_{A \cap H \neq \emptyset} m(A) \tag{4}$$

$$Pls(H) - Bel(H) = Unc \text{ (unawareness or doubt)} \tag{5}$$

When the value of Unc=0, then Pls = Bel.

Disbelief value represents that the event is not occur. For more information about the theory of Dempster–Shafer and its functionality in EBF, follow the paper [27]. It shows the steps on how the model was derived and applied.
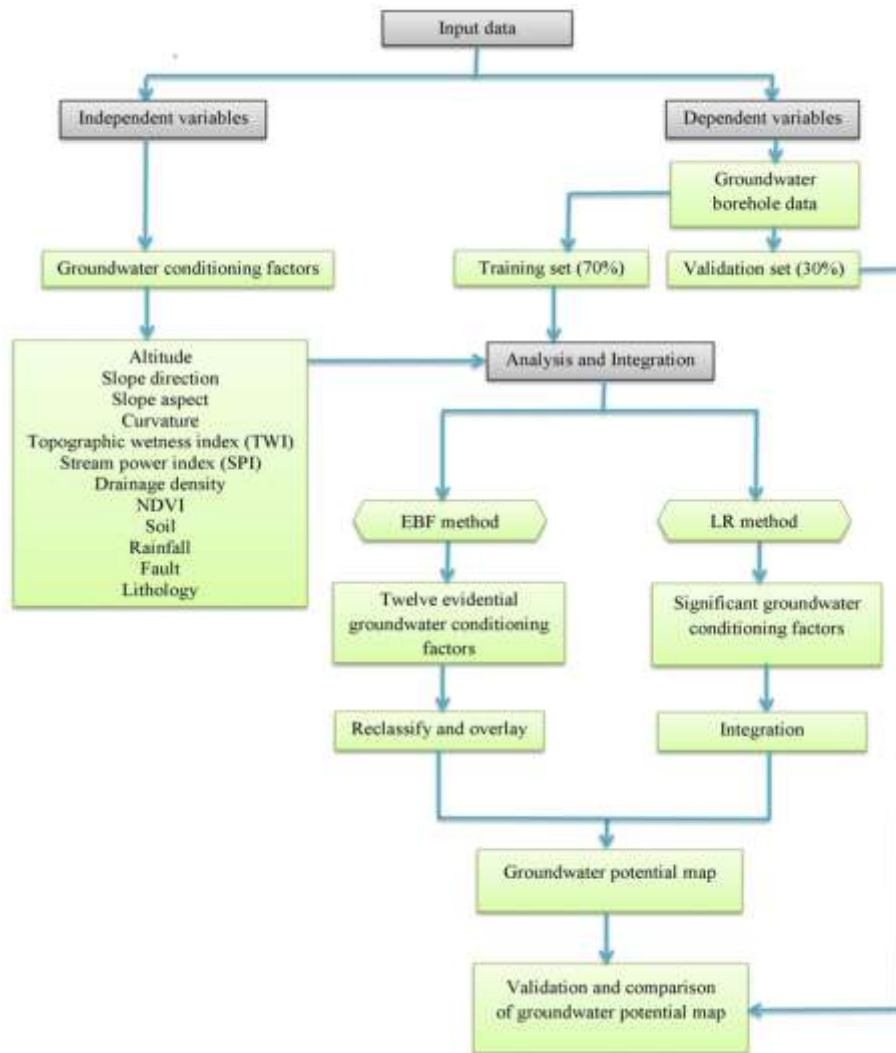

Fig.3: Flowchart of methodology used in this research

4.2. Logistic Regression (LR) model

LR model (multivariate model) is used to find an appropriate method to express the correlation among a dependent variable (the event) and conditioning influences which caused the happening of that event [3]. Through this correlation, each variable will obtain on a coefficient. These coefficients represent weights in algorithm. Then the algorithm is used for creating a map showing the probability of occurrence the event (groundwater) using GIS. The LR includes a multivariate regression between dependent variable (DV) and numerous independent variables. The existence or nonexistence the phenomena of groundwater occurrences within the specified land forms the foundation of such analyses [3]. The independent variables have a binary value (1 and 0) with LR model. 1 value denotes the incidence of locations, as a result, viewing the probability of the occurrences. 0 value shows the non-appearance of locations, hence, there is no expectation of

occurrence. The LR model agrees to the general linear model that can be calculated by following equation [9]:

$$P = e^z / (1+e^z) \tag{6}$$

Where: *P: is the probability of an occurrence, Z: is a value from −∞ to +∞ and can be obtained from the following equation:*

$$Z = a + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n \tag{7}$$

Where: *Z: is a function of linear combination of conditioning factors indicating a linear relation, a: is the cut off of model, n: shows the number of conditioning factors, $\beta_1, \beta_2, \ldots, \beta_n$: are coefficients which reflect the contribution of each conditioning factors $X_1, X_2, \ldots, X_n$.*

5. Results and discussion

5.1. Estimation and classification of EBF for groundwater potential mapping

The probability of the occurrence of groundwater for each conditioning element within EBF approach rely on the values of Belief and Dis-belief which shown in table (1). These values can be derived depending on equations bel, dis, pls, unc. Each controlling factor is classified into different classes using quantile method [27] through GIS, and each class has EBF values. There is a strong relation between groundwater probability with high Bel value, and conversely with the Dis, where Low Bel value denotes a lower probability of groundwater potential and vice versa. Table (1) displays the assessed EBF values for each element. By summation of Bel+ Dis + Unc, the result was 100; this equation is used to test the precision of EBF method. Before starting the analysis, each factor was changed to a raster grid (30×30 m) cells. The total number of columns and rows is 3024 and 2635, respectively (total number of cells is 5990719). In this paper, 30 groundwater wells were used in the analysis.

As shown in table (1), the areas with lower altitude, moderate slope, north-northwest direction and concave –flat curvatures, indicates a high value of Bel and low of Dis, this demonstrates that these areas with higher value of groundwater potential because the lower run-off and higher infiltration. With respect to TWI, the values of Dis and Bel show that the increases in value of TWI causes to increase in groundwater incidence probability, but as noted in the range between (21.73 - 37.71) was recorded the lowest Bel value although of the highest TWI value, this may be because the negative association between this area and groundwater probability or may be due to the effect of another conditioning factor. With respect to SPI, the relation between it and groundwater probability is very low where the high SPI value refers to low probability of groundwater occurrence. The range between (0.09 - 128403920984.43474) was recorded the highest Bel value with zero Dis, this indicates the high groundwater incidence probability in this area. For stream density, the range between (8657.14 - 20252.94) m/m2 has low Bel value and high Dis value; this indicates the increases in drainage density lead to decrease the probability of groundwater occurrence. For NDVI values below 0.1, a high Bel and low Dis values were recorded, and for NDVI values ≥ 0.1, a low Bel and high Dis values were recorded. For the fault, according to the values of Bel, the area which located far from faults, the chance of groundwater occurrence is increased. With respect to the lithology element, the class of (sand, silty clay and partly gypsiferous) has the higher occurrence probability of groundwater. Concerning the soil type, the highest Bel value was observed in (sand dune land) class and this indicates a higher probability of groundwater occurrence, while the lower probability was expected with (stony desert land) type where it was recorded the lowest Bel value. Moreover, the values of Dis and Bel showed that the regions which have high rainfall have the higher probability of groundwater occurrence.

Table 1: EBF values for classes of groundwater conditioning factors

Conditioning factor	Range	No. of class pixel	Pixels in domain	Bel (*100)	Dis (*100)	Unc (*100)	Pls (*100)
Elevation (m)	81 - 143	612576	58	23	8	69	92
	143 - 174	618940	35	13	9	78	91
	174 - 203	600017	45	18	9	73	91
	203 - 237	603743	35	14	9	77	91
	237 - 270	596476	9	3	10	87	90
	270 - 296	605201	0	0	11	89	89
	296 - 312	622480	9	3	10	87	90
	312 - 327	615701	35	14	9	77	91
	327 - 344	563942	11	4	10	86	90
	344 - 465	551643	8	3	10	87	90
Slope (degree)	0 – 0.2	520123	10	4	10	86	90
	0.2 – 1	769523	46	15	9	76	91
	1 – 2	770570	34	11	9	80	91
	2	503433	23	11	9	80	91
	2 – 3	627551	35	14	9	77	91
	3 – 4	606524	25	10	9	81	91
	4 – 5	620278	15	6	10	84	90
	5 – 6	533879	17	7	10	83	90
	6 – 7	537160	25	11	9	80	91
	7 - 44	501678	15	7	10	83	90
Aspect (direction)	Flat	633574	16	6	11	83	89
	Table 1: Continued						
	North	638755	33	14	10	76	90
	Northeast	731546	33	12	10	78	90
	East	682682	23	9	11	80	89
	Southeast	721179	31	11	11	78	89
	South	722636	34	12	10	78	90
	Southwest	638544	24	10	11	79	89
	West	575574	17	8	11	81	89
	Northwest	646229	34	14	10	76	90
Curvature	Concave	318921	15	40	33	27	67
	Flat	5289777	219	35	31	34	69
	Convex	382021	11	24	34	42	66
TWI	2.28 - 4.5	451798	14	7	10	83	90
	4.5- 5.06	757168	25	8	10	82	90
	5.06 - 5.62	711250	27	9	10	81	90
	5.62 - 7.56	705032	41	14	9	77	91
	7.56 – 12	507764	19	9	10	81	90

	12 - 13.12	560017	26	11	9	80	91
	13.12 - 14.51	589966	26	10	9	81	91
	14.51 - 16.17	579851	24	10	9	81	91
	16.17 - 21.73	582979	35	14	9	77	91
	21.73 - 37.71	544894	8	3	10	87	90
SPI	0 - 0.09	25805	0	0	11	89	89
	0.09 - 128403920984.4	5946559	245	100	0	0	100
	128403920984.4 - 256807841968.8	6574	0	0	11	89	89
	256807841968.8 - 385211762953.1	2836	0	0	11	89	89
	385211762953.1 - 642019604921.8	2676	0	0	11	89	89
	642019604921.8 - 1027231367874.8	1864	0	0	11	89	89
	1027231367874.8 - 1540847051812.2	1276	0	0	11	89	89
	1540847051812.2 - 2439674498702.6	1116	0	0	11	89	89
	2439674498702.6 - 4237329392483.5	1016	0	0	11	89	89
	4237329392483.5 - 32614595930023.6	997	0	0	11	89	89
Stream density (m/m^2)	0	336200	9	6	10	84	90
	0 - 476.54	961978	9	2	11	87	89
	476.54 - 1270.77	628723	42	15	9	76	91
	1270.77 - 2144.43	589546	2	0	10	90	90
	2144.43 - 3018.09	577318	40	16	9	75	91
	3018.09 - 4050.59	613777	27	10	9	81	91
	4050.59 - 5162.51	580057	20	8	10	82	90
	5162.51 - 6512.71	571878	52	21	8	71	92
	6512.71 - 8657.14	577614	26	10	9	81	91
	8657.14 - 20252.94	553628	18	7	10	83	90
NDVI	0.006955 - 0.088047	410375	31	18	9	73	91
	0.088047 - 0.093839	808968	51	15	9	76	91
	0.093839 - 0.097701	872672	47	13	9	78	91
	0.097701 - 0.099631	522783	22	10	9	81	91
	0.099631 - 0.101562	575961	28	12	9	79	91
	0.101562 - 0.103493	612723	24	9	10	81	90
	0.103493 - 0.105424	597311	17	7	10	83	90
	0.105424 - 0.107354	507933	4	1	10	89	90
	0.107354 - 0.111216	641873	4	1	11	88	89
	0.111216 - 0.497369	440120	17	9	10	81	90
Distance from fault (m)	0 - 3796.19	585896	18	7	10	83	90
	3796.19 - 7339.31	629286	0	0	11	89	89
	7339.31 - 10376.26	626130	8	3	10	87	90
	10376.26- 13919.38	603145	18	7	10	83	90
	13919.38 - 18980.97	616263	18	7	10	83	90
	18980.97 - 24042.56	604771	9	3	10	87	90
	24042.56 - 29357.23	581782	27	11	9	80	91
	29357.23 - 35937.30	598723	71	28	7	65	93
	35937.30 - 44035.85	570816	34	14	9	77	91
	44035.85 - 64535.30	576479	42	17	9	74	91
Lithology	Red clay-stone and siltstone with marl and limestone	5426	0	0	5	95	95
	Mainly limestone	7250	0	0	5	95	95
	Sand ,silty clay and partly gypsiferous	28634	9	43	4	53	96
	Secondary gypsum and gypsiferous clastics	208177	23	15	4	81	96
	Green marl and limestone	1254501	40	4	5	91	95
	Silty clayey soil and mainly gypsifereos	2175056	129	8	3	89	97
	Conglomerate with lenses of sand and silt	35777	0	0	5	95	95
Table 1: Continued							
	Fossiliferous and oolitic limestone	216599	0	0	5	95	95
	Mud and salt crust	3164	0	0	5	95	95
	Massive and well bedded limestone	1018	0	0	5	95	95
	Cemented different rock fragments	106733	5	6	5	89	95
	Basal conglomerate, limestone and dolostone	34474	0	0	5	95.	95
	Brecciated limestone, dolostone and marl	123830	0	0	5	95	95
	Mudy sand and gravel	435341	31	9	4	87	96
	White sandstone, clay-stone and limestone	880591	0	0	5	95	95
	Fossiliferous limestone, marl and dolostone	91128	8	12	4	84	96
	Sandy gravels	335520	0	0	5	95	95
	Basal conglomerate and recrystallized limestone	39265	0	0	5	95	95
	Nummulitic limestone and chert	4692	0	0	5	95	95
	Crystalline dolostone with chert nodules	3543	0	0	5	95	95
Soil	Stony desert land	2007303	45	9	31	60	69
	Sand dune land	92176	9	41	25	34	75
	Saline desert marshes	124069	8	27	25	48	75

	Mixed gypsiferous desert land	3769743	183	20	17	63	83
Rainfall	92.406418 - 95.532164	587792	0	0	11	89	89
(mm)	95.532164 - 96.852902	605028	0	0	11	89	89
	96.852902 - 97.645345	602767	0	0	11	89	89
	97.645345 - 98.17364	595939	34	14	9	77	91
	98.17364 - 98.525837	602180	35	14	9	77	91
	98.525837 - 98.74596	648324	70	26	8	66	92
	98.74596 - 99.054132	634288	18	6	10	84	90
	99.054132 - 99.494378	589991	17	7	10	83	90
	99.494378 - 100.37487	576611	44	18	9	73	91
	100.37487 - 103.63269	547799	27	12	9	79.	91

After calculating all values of Bel, Dis, Unc and Pls for each element depending on the association among conditioning elements and groundwater wells locations, then these values will be integrated to obtain the map of each one (Bel, Dis, Unc and Pls). The integrated outcomes are displayed in fig. (4).In this paper, high correlation was observed between groundwater wells and the belief map; in contrast, an inverse relation was noticed with the other EBF functions. As shown in fig. (4), the contrasting distribution for the belief map was noted in the dis- belief map. The uncertainty map designated either there is a gap in the

evidential data layers for valuation of groundwater potential or lack of information. For the plausibility map, although it is similar to the belief map, but there is an ambiguity between higher and lower values. Consequently, the belief map was viewed to be correspondent to the ground-water potential map. For more interpretation, the map of probability index was classified in to five susceptibility classes using quantile method [27] as follow: very low (41%), (20.1%) for low potential, moderate areas with (19.8%), (9.7%) for high and (9.4%) for very high potential, fig. (5).
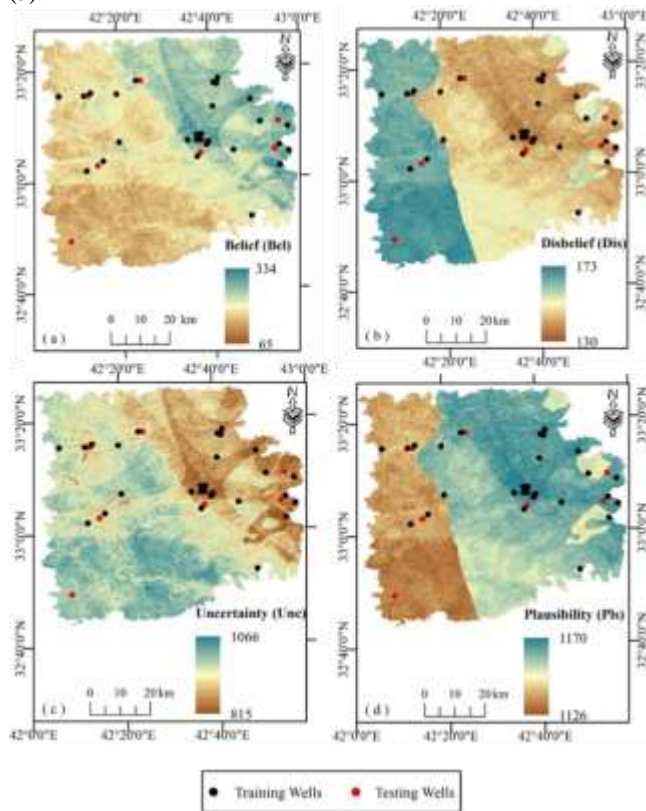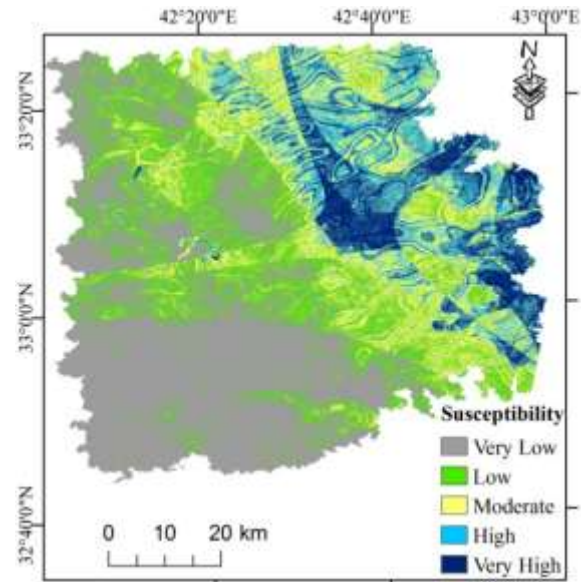

Fig. 5: Groundwater potential zones using EBF model

Fig. 4: Integration maps of EBF result

5.2. Application and classification of LR method for mapping the groundwater potential

In order to apply LR model, the number of pixels containing the wells and the number of pixels that do not have wells shall be selected. In the analysis of groundwater occurrence probability, this model should deal with the dependent variable (DV) either 0 or 1. (0) value indicates the absence of wells in a given area, and (1) value indicates the presence of wells. The ratios of each conditioning element are estimated by calculating the coefficients of this model. The linear model designed is then a logistic regression of existence or non- existence of groundwater happenings on the independent variables. Before starting the analysis, multi-collinearity was applied and its results proved the low correlation between variables. The next step was calculating the variable significance (by calculating P-value for each conditioning element). If it is ≤ 0.05, this means that the model is statistically acceptable, but if it goes further, the model is unacceptable [28].Table (2) displays the coefficient and p-value for all independent variables depending on LR model.

Table 2: LR coefficient and p-value for all independent parameters

groundwater conditioning factor	Logistic regression coefficient (β)	P-value
Altitude	-0.003020305	0.000483334
Slope	-0.005946783	0.315021513
Aspect	-0.000257638	0.063947
Curvature	-0.017148764	0.54129
SPI	-1.65123E-11	0.644766687
TWI	-0.005307313	0.126925457
Stream density	1.5915E-05	0.003719501
NDVI	1.182176656	0.689828313
Distance from fault	-1.18306E-05	8.94207E-05
Geology	-0.003735753	0.394205923
Soil	0.146407081	1.18307E-10
Rainfall	0.102275215	4.73953E-13

| Intercept | -8.89842052 | 8.32143E-10 |

It is clear from table (2), that p-value for both slope, aspect, curvature, SPI, TWI, NDVI and geology is > 0.05. Consequently, these factors are considered statistically unacceptable according to LR model. Therefore, they were neglected during the analysis (considered to be ineffective on the occurrence of the event which is groundwater in this case). The coefficients of significant factors (only five factors) were shown in table (3), and the probability of groundwater occurrence based on these coefficients is generated through applying the equations of LR (6) and (7) which mentioned in methodology section.

Table 3: LR coefficient and p-value for significant parameters.

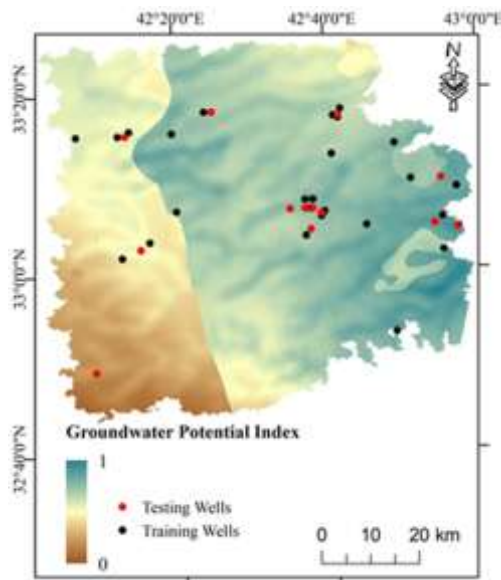| Significant groundwater conditioning factor	Logistic regression coefficient (β)	P-value
Altitude	-0.002883849	2.13985E-10
Stream density	1.5329E-05	0.004155257
Distance from fault	-1.22122E-05	4.54251E-06
Soil	0.152566764	5.09241E-13
Rainfall	0.112493412	2.57788E-24
Intercept	-9.984892083	1.98052E-20

Fig. 6: Groundwater potential index using LR model

After computing the coefficients of LR of the five groundwater conditioning elements, the map of groundwater potential is obtained using a linear summation of the values (by raster calculator in GIS) as follow:

Z= (-0.002883849 * "altitude") + (1.5329E-05 * "stream_den") + (-1.22122E-05* "dis_fault") + (0.152566764 * "soil") + (0.112493412*"rainfall")-9.984892083 (8)

Fig.(6) shows the map of probability index for groundwater occurrence created from LR approach, it ranges from zero to one. The map of probability index was divided in to five susceptibility classes using quantile approach as a method for classification [27], as follow: (40.4%) very low, (21.2%) low, (20.4%) moderate, (9.4%) high and (8.6%) from the whole area was very high, fig. (7).
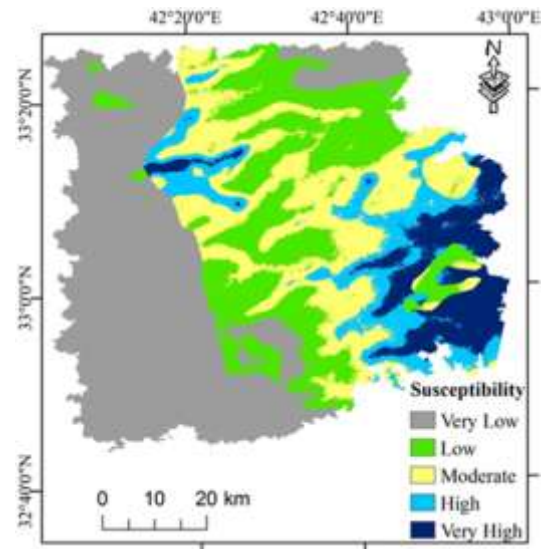

Fig. 7: Groundwater potential zones using LR model

5.3. Validation of outcomes

After obtaining the ground-water potential maps using EBF and LR method, the results must be validated. The final outcomes without validation have no scientific sense [29]. In this paper the "Area Under the Curve" algorithm was applied to assess the success and prediction-rates for the suggested models. The success-rate depends on the location of training groundwater-wells which were used during the analysis (30 wells), while the prediction-rate relies on the testing-wells which allocated for the validation (13 wells). By overlapping the maps of ground-water potential with the training and the testing wells locations, accumulative percentage of the ground-water occurrence beginning from the maximum to the minimum values of groundwater potential index map were calculated. Then, the curves of the success and prediction-rates were attained. As shown in fig. (8), the validation results showed that the success accuracy for EBF and LR are 76.5% and 71.5%, respectively. The prediction accuracy for the same methods are73.7% and 70% , respectively. The above values indicate that EBF model is relatively well estimator for groundwater potential occurrence in the region of interest
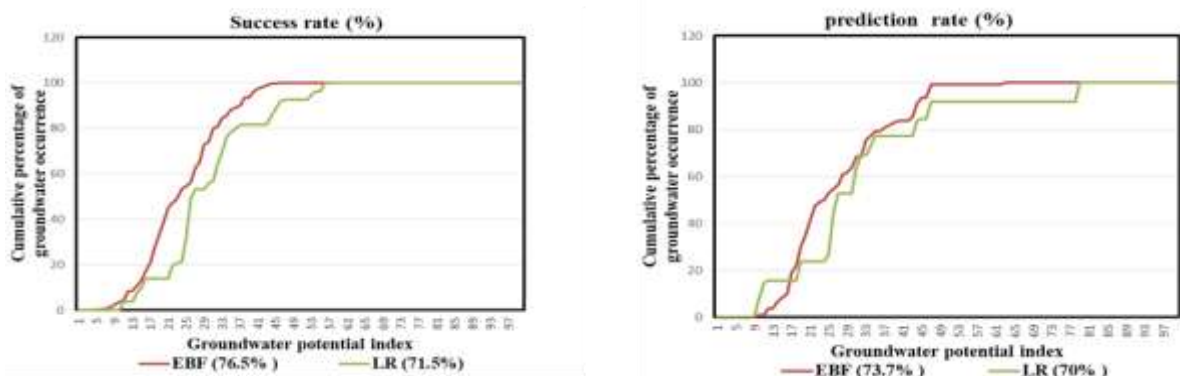

Fig. 8: The success and prediction rates of each model.

6. Conclusions

The conclusion of this paper could be summarized to the followings:

1- The investigation of ground-water in area under study is good solution for future challenge.

2-For estimation of groundwater sites potential, the major step was the creation of a geo-database including the factors that influence on ground-water occurrence.

3-Applying the AUC method is very significant procedure to assess the model efficiency.

4-In this paper, among the two statistical approaches used, the EBF approach has been relatively successful in predicting groundwater in the study area. The verification outcomes showed that the success ratio for this model was 76.5% with prediction ratio of 73.7%.

5-This study indicates that the combination of remote sensing data and GIS-based EBF and LR methods can be offered a powerful tool for delineation of ground-water resources.

References

[1] Todd, D. K. and Mays, L. W. (2005). Groundwater Hydrology. 3rd edition, John Wiley and Sons, NJ, pp. 636.

[2] Banks, D., Robins, N., & Robins, N. (2002). An introduction to groundwater in crystalline bedrock (p. 64). Trondheim: Norges geologiske undersøkelse.

[3] Nampak, H. (2014). Development of Optimized Model Based on Evidential Belief Function for Groundwater Mapping. Universiti Putra Malaysia, MSc. Thesis.

[4] Al-Mankoshy, H.J.(2008). Integration of Remote Sensing Data and GIS to Determine Potential Ground Water in Ubaiydh and Ghadaf Valleys, PhD. Thesis, Baghdad University, college of science .

[5] Al-Muqdadi, S. W. (2012). Groundwater investigation and modeling-western desert of Iraq. Technische Universität Bergakademie Freiberg,PhD. Thesis.

[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. Environ. Geol. 55 (3), 467–475.

[7] Dar, I.A., Sankar, K., Dar, M.A., 2010. Remote sensing technology and geographic information system modeling: an integrated approach towards the mapping of groundwater potential zones in Hardrock terrain, Mamundiyar basin. J. Hydrol. 394, 285–295.

[8] Nag, A., Ghosh, S., Biswas, S., Sarkar, D., Sarkar, P.P., 2012. An image steganography technique using X-box mapping. In: Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on pp. 709–713. IEEE.

[9] Ozdemir, A. (2011). 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(3-4):290-308.

[10] Lee, S., Kim, Y.S., Oh, H.J., 2012. Application of a weights-of-evidence method and GIS to regional groundwater productivity potential mapping. J. Environ. Manage. 96 (1), 91–105.

[11] Pradhan, B., 2009. Groundwater potential zonation for basaltic watersheds using satellite remote sensing data and GIS techniques. Cent. Eur. J. Geosci. 1 (1), 120– 129.

[12] Mahalingam, B., &Vinay, M. (2015). Identification of ground water potential zones using GIS and Remote Sensing Techniques: A case study of Mysore taluk-Karnataka. International Journal of Geomatics and Geosciences, 5(3), 393.

[13] Murthy, K.S.R., Mamo, A.G., 2009. Multi-criteria decision evaluation in groundwater zones identification in Moyale-Teltele sub-basin, South Ethiopia. Int. J. Remote Sens. 30 (11), 2729–2740.

[14] Gupta, M., Srivastava, P.K., 2010. Integrating GIS and remote sensing for identification of groundwater potential zones in the hilly terrain of Pavagarh,Gujarat, India. Water Int. 35 (2), 233–245.

[15] Manap, M. A., Nampak, H., Pradhan, B., Lee, S., Sulaiman, W. N. A., and Ramli, M. F. (2012). Application of probabilistic-based frequency ratio model in groundwater potential mapping using remote sensing data and GIS. Arabian Journal of Geosciences, 7(2), 711-724.

[16] Chenini, I., Mammou, A.B., Moufida, E.M., 2010. Groundwater recharge zone mapping using GIS-based multi-criteria analysis: a case study in Central Tunisia (Maknassy Basin). Water Resour. Manage 24 (5), 921–939.

[17] Ghayoumian, J., MohseniSaravi, M., Feiznia, S., Nouri, B., Malekian, A., 2007. Application of GIS techniques to determine areas most suitable for artificial groundwater recharge in a coastal aquifer in southern Iran. J. Asian Earth Sci. 30 (2), 364–374.

[18] Shahid, S., Nath, S. K., & Maksud Kamal, A. S. M. (2002). GIS integration of remote sensing and topographic data using fuzzy logic for ground water assessment in Midnapur District, India. Geocarto International, 17(3), 69-74.

[19] Al-Bassam, K.S. and Karim, S.A.(1997). The Akashat Formation: A new name for the Paleocene Lithostratigraphic Unit in the Western Desert of Iraq. Iraqi Geol.Jour. Vol. 30, No.1, p. 22 – 35.

[20] Consortium-Yugoslavia, (1977). Hydrogeological explorations and hydro-technical work-Western Desert of IRAQ. Blook7, Directorate of western desert development projects. Republic of Iraq .

[21] Iraqi Meteorological Organization and Seismology, www.meteoseism.gov.iq/.

[22] Toz, G., & Erdogan, M. (2008). DEM (digital elevation model) production and accuracy modeling of DEMs from 1: 35.000 scale aerial photographs. In The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, ISPRS Congress (Vol. 21, pp. 775-780).

[23] Roderick, M., Smith, R., & Lodwick, G. (1996). Calibrating long-term AVHRR-derived NDVI imagery. Remote Sensing of Environment, 58(1), 1-12.

[24] Carranza, E.J.M., and Hale, M. (2003). Evidential belief functions for data-driven geologically constrained mapping of gold potential, Baguio district, Philippines. Ore Geology Reviews, 22 (1), 117–132.

[25] Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. The Annals of Mathematical Statistics, 325-339.

[26] Park, N.W.(2011). Application of Dempster–Shafer theory of evidence to GIS-based landslide susceptibility analysis. Environmental Earth Sciences, 62 (2), 367–376.

[27] Nampak, H., Pradhan, B., and 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.

[28] Proudlove, N., Moxham, C., & Boaden, R. (2008). Lessons for lean in healthcare from using six sigma in the NHS. Public Money and Management, 28(1), 27-34.

[29] Chung, C.J., and Fabbri, A.G., 2003. Validation of spatial prediction models for landslide hazard mapping. Natural Hazards, 30: 451-472.