water

MDPI

Article

Mapping of Groundwater Spring Potential in Karst Aquifer System Using Novel Ensemble Bivariate and Multivariate Models

Viet-Ha Nhu [1,2], Omid Rahmati [3], Fatemeh Falah [4], Saeed Shojaei [5], Nadhir Al-Ansari [6,*], Himan Shahabi [7,8], Ataollah Shirzadi [9], Krzysztof Górski [10], Hoang Nguyen [11,*] and Baharin Bin Ahmad [12]

1 Geographic Information Science Research Group, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam; nhuvietha@tdtu.edu.vn
2 Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam
3 Soil Conservation and Watershed Management Research Department, Kurdistan Agricultural and Natural Resources Research and Education center, AREEO, Sanandaj 36311-66169, Iran; o.rahmati@areeo.ac.ir
4 Department of Watershed Management, Faculty of Agriculture and Natural Resources, Lorestan University, Lorestan 68151-44316, Iran; falahfateme69@gmail.com
5 Young Researchers and Elite Club, Zahedan Branch, Islamic Azad University, Zahedan 98167-43545, Iran; sshojaei@ut.ac.ir
6 Department of Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, 971 87 Lulea, Sweden
7 Department of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran; h.shahabi@uok.ac.ir
8 Board Member of Department of Zrebar Lake Environmental Research, Kurdistan Studies Institute, University of Kurdistan, Sanandaj 66177-15175, Iran
9 Department of Rangeland and Watershed Management, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran; a.shirzadi@uok.ac.ir
10 Faculty of Mechanical Engineering, Kazimierz Pulaski University of Technology and Humanities in Radom, Chrobrego 45 Street, 26-200 Radom, Poland; krzysztof.gorski@uthrad.pl
11 Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
12 Department of Geoinformation, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia (UTM), 81310 Johor Bahru, Malaysia; baharinahmad@utm.my
* Correspondence: nadhir.alansari@ltu.se (N.A.-A.); nguyenhoang23@duytan.edu.vn (H.N.)

Received: 3 March 2020; Accepted: 27 March 2020; Published: 31 March 2020;
Corrected: 3 April 2026



Abstract: Groundwater is an important natural resource in arid and semi-arid environments, where discharge from karst springs is utilized as the principal water supply for human use. The occurrence of karst springs over large areas is often poorly documented, and interpolation strategies are often utilized to map the distribution and discharge potential of springs. This study develops a novel method to delineate karst spring zones on the basis of various hydrogeological factors. A case study of the Bojnourd Region, Iran, where spring discharge measurements are available for 359 sites, is used to demonstrate application of the new approach. Spatial mapping is achieved using ensemble modelling, which is based on certainty factors (*CF*) and logistic regression (LR). Maps of the *CF* and LR components of groundwater potential were generated individually, and then, combined to prepare an ensemble map of the study area. The accuracy (A) of the ensemble map was then assessed using area under the receiver operating characteristic curve. Results of this analysis show that LR (A = 78%) outperformed *CF* (A = 67%) in terms of the comparison between model predictions and known occurrences of karst springs (i.e., calibration data). However, combining the *CF* and LR results through ensemble modelling produced superior accuracy (A = 85%) in terms of spring potential mapping. By combining *CF* and LR statistical models through ensemble modelling, weaknesses in *CF* and LR methods are offset, and therefore, we recommend this ensemble approach for similar karst

mapping projects. The methodology developed here offers an efficient method for assessing spring discharge and karst spring potentials over regional scales.

Keywords: ensemble model; karst springs; certainty factor; logistic regression; GIS

1. Introduction

Groundwater has been considered the most important resource in arid and semi-arid areas. Population increases and economic growth have resulted in water shortages around the globe, especially in developing countries, leading to greater exploitation and reliance on groundwater resources [1,2]. Among the various forms of subsurface freshwater, karst groundwater is a major resource for water supply in many regions, including China, Turkey, Iran, Mexico, and the United States [3–7]. Karst aquifers are relatively high-yielding, heterogeneous systems that are characterized by complex groundwater flow paths [8–10]. They store and supply drinking water for approximately 20–25% of the world's population, while constituting only small areas of the global land surface [11]. The characterization of karst aquifers is hampered by their complex character, and the often limited number of wells available to support hydrogeological observations and geological descriptions [12].

Groundwater models of karst aquifer systems invariably simplify karst features—and the spatial heterogeneity more generally—because it is impossible to acquire sufficient data to properly delineate the main flow features [13]. Additionally, specialized models are required to simulate the conduit flow that occurs within karst aquifers, at least for situations where equivalent porous medium approaches are not valid [14]. However, physically based models of karst aquifers are limited by both data availability and the heavy computational burden of conduit-matrix simulation. Thus, strategies to characterize the hydrogeology of karst aquifers, and in particular, to quantify spring hydrology, are essential precursors to the effective management of groundwater in these systems [15].

Geographical information systems (GIS) and remote sensing (RS) are commonly used to evaluate the hydrogeology of karst aquifers, including the spatial distribution of spring potential [16,17]. Recently, Chen et al. [18] undertook a global-scale project—the World Karst Aquifer Mapping (WOKAM)—to map karst aquifers and carbonate rock settings around the world. Karst springs, wells, caves and other water abstraction structures were also identified, although the large scale of the project meant that many regions have only limited information regarding point features within the landscape.

'Spring groundwater potential' was defined as the possibility of spring occurrence (with sufficient water) in an area and helps planners to estimate groundwater productivity [19,20]. There are myriad approaches to groundwater potential mapping using GIS-based statistical models. These include: frequency ratio (FR) [16,21], weight-of-evidence [22], evidential belief function (EBF) [23], logistic regression (LR) [24,25], certainty factor (*CF*) [26], analytical hierarchy process (AHP) [27], Shannon's entropy [28], maximum entropy [29], support vector machine (SVM) [30] and boosted regression tree [31]. Groundwater potential mapping highlights the locations of potential groundwater zones including low, moderate and high for the optimal use of resources [32]. In general, multivariate statistical models (e.g., LR) cannot analyze the classes of independent factors (e.g., lithology, land use/land cover, soil, etc.) because they just determine the weight of predictive factors. In addition, bivariate statistical models (e.g., *CF*) are limited in their capacity to recognize the importance of independent variables [26] because they use a class-based procedure in modelling. To overcome these problems, scientists have developed ensemble statistical approaches [33–39].

Ensemble modelling is the combination of multiple predictive models that are used conjunctively to provide an evaluation of the same dataset [39]. Ensemble modelling approaches have been widely used in different fields, including groundwater [40–42], flood [43,44], landslide hazard [45], land/ground subsidence, gully erosion [46], dust storm [47], wildfire [48], sinkhole [49], droughtiness [50], earthquake [51] and species distribution. These studies provide clear evidence that the application of

ensemble models can potentially result in improved capability (over individual methods) of GIS-based statistical models. However, previous studies aimed at the prediction of karst spring occurrence have not investigated the potential of ensemble models that integrate bivariate and multivariate statistical models to support the construction of spring potential maps [24,25].

The main aim of this study is to explore different statistical models, both single and ensemble, to assess spring groundwater potential in a karst area, namely Bojnourd Region, northeast Iran. Although various researchers have tried to delineate groundwater spring potential zones [30,52,53], the evaluation of other modeling approaches is highly necessary. To address this, a new ensembling technique is proposed to improve the accuracy of the statistical models and to identify groundwater potential zones. The novel approach of ensemble certainty factor and logistic regression models has not previously been touched upon in the literature. Therefore, the principal motivation of this study is to use this ensemble model to delineate prone areas to karst springs in Bojnourd region, Khorasan Shomali, Iran. It is anticipated that the outputs of this study will identify the strengths and weaknesses of the applied models, thereby, providing improved insights into model choice in the spatio-temporal analysis of karst springs discharge using statistical models.

2. Study Area

The study area, Bojnourd Region, is located in the north east of Khorasan Shomali province, Iran, between latitudes 37°15′ to 37°35′ N and longitudes 57°03′ to 57°40′ E (Figure 1), with minimum and maximum elevation of 875 and 2968 m, respectively. The mean precipitation is 272 mm/year and temperature ranges between 6.8 to 19.7 °C [54]. There are six main lithological formations in the study area including Aitamir (Cretaceous), Sarcheshmeh (Early Cretaceous), Shurijeh (Jurassic-Cretaceous), Upper-Red (Miocene), Niur (Silurian) and Shemshak (Triassic-Jurassic). Lithological characteristics of these formations are summarized in Table 1. From a geomorphological point of view, the study area is dominated by a mountainous landscape. In addition, there are numerous folds and faults in the study area which cause the occurrence of springs [55]. The major faults are oriented NE-SW and the others NW-SE and N-S. Inceptisols and entisols are dominant soil types of the study area. The depth of soil varies from one to three meters, although there are numerous surficial rocks (lithological outcrops) in steep areas of upstream zones. According to the reports of Iranian Department of Water Resources Management (IDWRM), hydraulic conductivities in this mountainous region range between 10^{-4} and 10^{-5} m/s.

Nearly 88% of water demand in Khorasan Shomali province is supplied by groundwater resources [56]. Therefore, it has become necessary to predict groundwater potential zone in this region. The discharge of springs ranges between 0.5 and 15.2 SI. The size of karstic areas in this province is over 17,000 km^2, which are used as karst water resources to provide drinking water for the Khorasan Shomali and Khorasan Razavi provinces. Furthermore, Khorasan Shomali province covers approximately an area of 248,000 km^2, which is inhabited by more than 6 million people and where agriculture is the major activity.

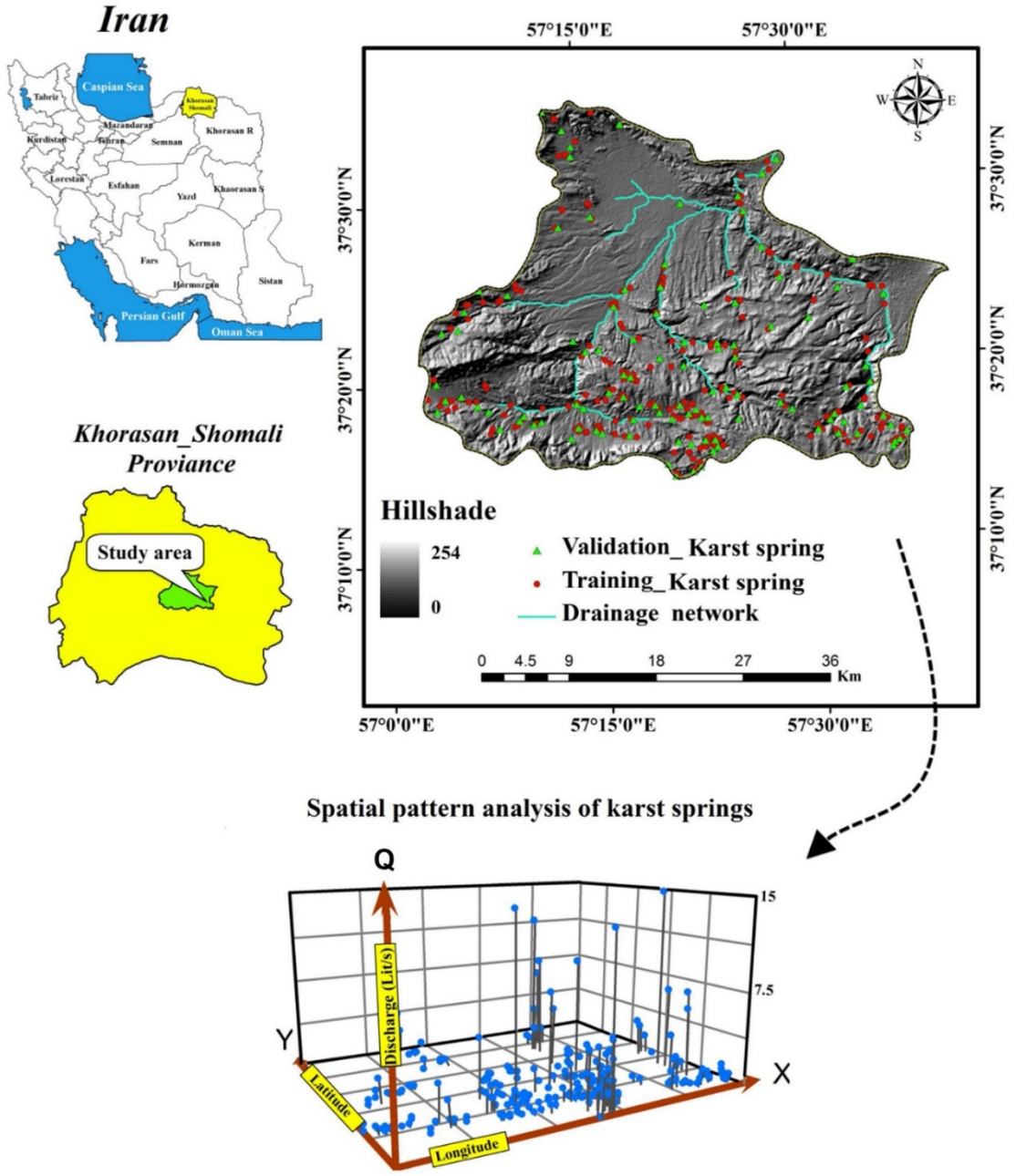


Figure 1. Karst spring location with hill-shaded map of the study area, Khorasan Shomali, Iran.

Table 1. Geological formations in the study area.

Geology (Group)	Formation	Lithology
Cretaceous (C)	Aitamir	Olive green glauconitic sandstone and shale
Early Cretaceous (EC)	Sarcheshmeh	Ammonite bearing shale with interaction of orbitolin limestone
Jurassic-Cretaceous (JC)	Shurijeh	Pale red argillaceous limestone, sandstone and conglomerate
Miocene (M)	Upper-Red	Red marl, gypsiferous marl, sandstone and conglomerate
Quaternary (Q)	-	Low level piedmont fan and valley terrace deposits
Silurian (S)	Niur	Coral limestone and dolomite, shale, sandstone
Triassic-Jurassic (TJ)	Shemshak	Subordinate sandy limestone, dark grey shale and sandstone

3. Methodology

The adopted methodology (Figure 2) included: preparation of data, construction of models and production of karst spring potential map (KSPM), validation of results and evaluation of their efficiency.

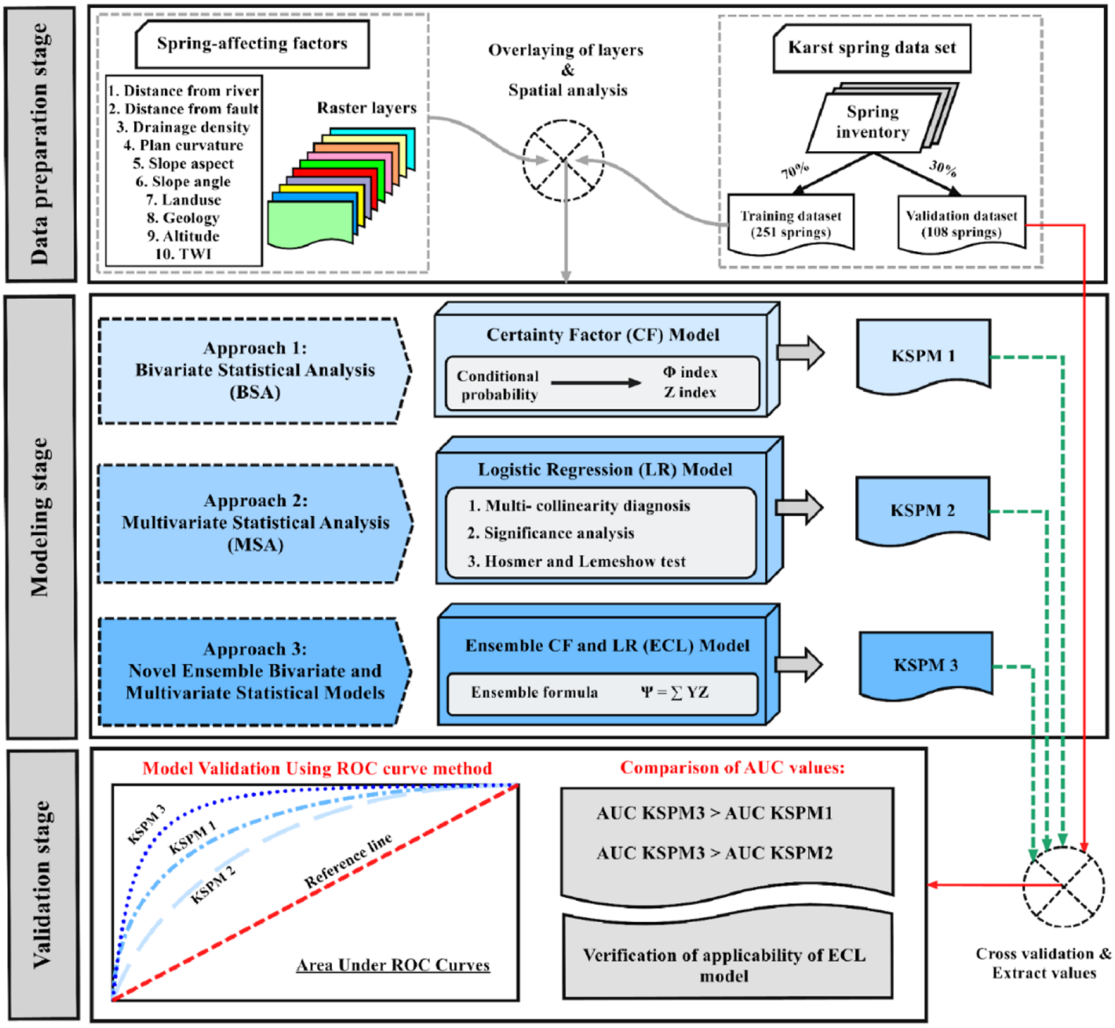


Figure 2. Methodology flowchart used in the study.

3.1. Data Used

In order to analyze karst spring potential, springs locations and 10 spring-affecting factors were considered as dependent and independent variables, respectively. All data were obtained from the Iranian Department of Water Resources Management (IDWRM).

3.1.1. Karst Spring Inventory

A total of 359 karst springs was identified and recorded by the Iranian Department of Water Resources Management (IDWRM) in the study area, of which 251 (70%) were randomly selected for calibration/training of models and the remaining 108 (30%) for validation purpose. The distribution of training and validation karst springs is shown in Figure 1.

3.1.2. Karst Spring-Affecting Factors

As mentioned by Crozier [57], preparation of corresponding thematic data layers and the selection of factors are important components of any model. In this study, 10 karst spring-affecting factors

including distance from fault, land use/land cover, geology, topographic elevation, slope angle, slope aspect, plan curvature, topographic wetness index (*TWI*), drainage density and distance from river were utilized for producing KSPMs.

Distance from fault

Distance from fault is among those factors that has been widely used as an important groundwater influential factor [58]. The faults were extracted from a geological map of the study area as line features in ArcGIS 10.2 and a map of 'distance from faults' was developed using the Euclidean distance tool and classified (Figure 3a).

Land use/ land cover

Land use/land cover is a significant factor in controlling groundwater recharge processes [59]. The land use/land cover map of the study area (Figure 3b) shows that rangeland covers most of study area, followed by agriculture, forest and urban classes (Figure 3b).

Geology

Since the hydraulic conductivity and porosity of soils and rocks are affected by lithological variations, groundwater occurrence is highly influenced by geology [53,60]. In this study, the geologic map of study area with seven lithological groups was extracted from 1:100,000 scale geologic map of Iran (Figure 3c, Table 1).

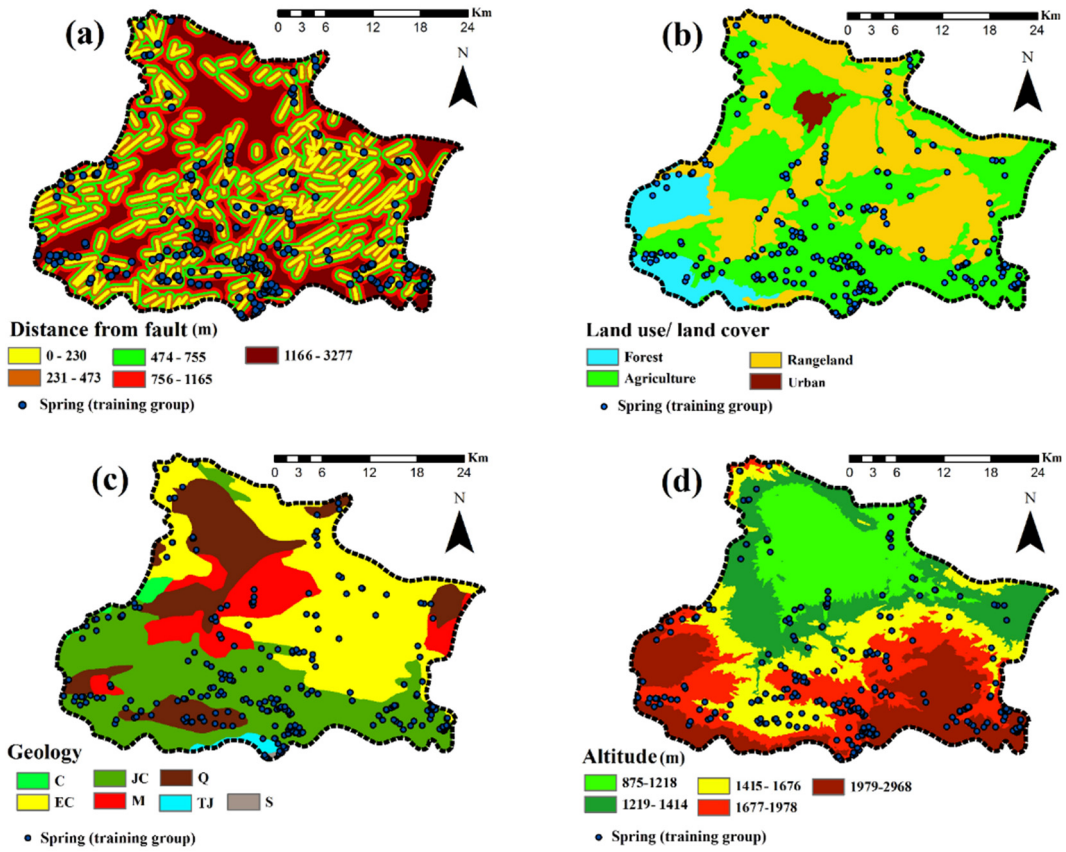


Figure 3. *Cont*.

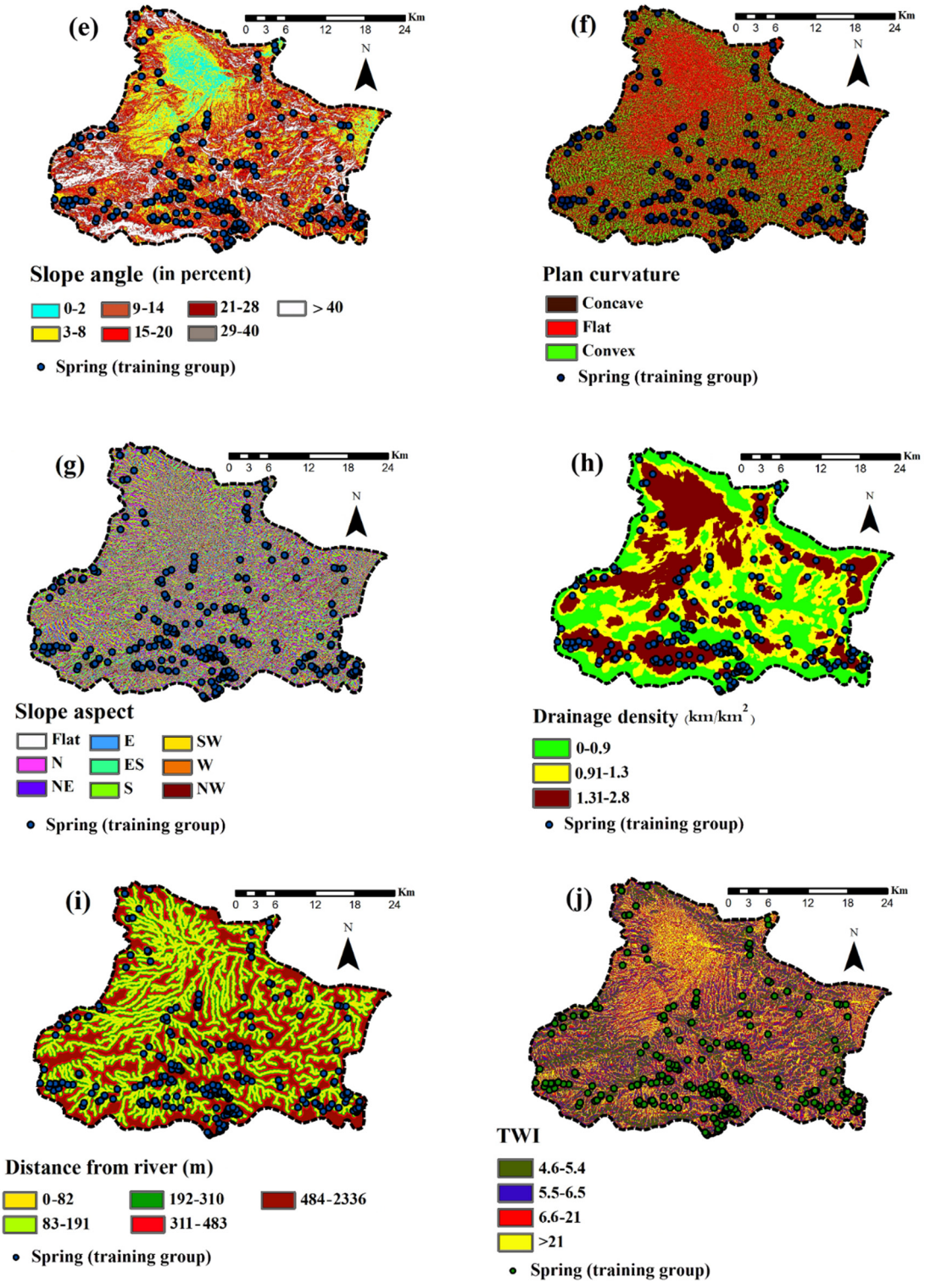


Figure 3. Karst spring inventory: (**a**) distance from fault, (**b**) land use/land cover, (**c**) geology, (**d**) topographic elevation, (**e**) slope angle, (**f**) plan curvature, (**g**) slope aspect, (**h**) drainage density, (**i**) distance from river and (**j**) *TWI*.

Topographic factors

Surface topography controls the runoff rate and direction [61,62]. Therefore, topographic elevation has been frequently used by researchers as a controlling parameter for groundwater occurrence [31,63]. The digital elevation model (DEM) map of the study area was obtained from the IDWRM and we

classified it using an equal-interval classification scheme (Figure 3d). IDWRM originally produced the DEM from aerial photographs and topographical lines was obtained from the Iranian Department of Water Resources Management (IDWRM) at 20 m cell resolution. Slope angle affects the infiltration, subsurface flow and duration of overland flow [64,65]. Therefore, the groundwater recharge process is strongly controlled by slope angle [66]. According to Magesh et al. [64], due to low infiltration and high run-off in steeper slope, these areas have poor potential to groundwater accumulation. The slope angle map of study area was extracted from the DEM and classified into seven classes based on IDWRM classification (Figure 3e). Plan curvature represents the rate of slope aspect or slope gradient change. Negative and positive values of plan curvatures characterize concavity and convexity in the downslope direction, respectively. Zero curvature also represents flat areas [67]. As mentioned by Lee and Pradhan [67], concave slope probably contains more water. The plan curvature map of study area was produced from DEM in concave, flat and convex classes (Figure 3f). The slope aspect layer was derived from DEM and divided into nine classes including eight main directions (Figure 3g).

Water-related factors

There is a significant relationship between distance from river and groundwater spring occurrence [31]. The maps of drainage density and distance from rivers were created from the topographic data using Kernel density and Euclidean distance tools, respectively. The maps of drainage density and distance from river are shown in Figure 3h,i, respectively. The soil moisture conditions are reflected in values of *TWI*, which is defined as:

$$TWI = ln\left(\frac{A_s}{tan\beta}\right) \quad (1)$$

where A_s is the specific catchment area (m^2/m) and β is slope gradient [68]. *TWI* map was shown in Figure 3j. Drainage density is defined as ratio of total length of a stream to the area of the drainage basin. In general, low drainage density areas are more suitable for groundwater development in comparison with high drainage density [69].

3.2. Statistical Methods

A bivariate statistical analysis (certainty factor) and a multivariate statistical analysis (logistic regression) were applied individually and as an ensemble to delineate karst spring potential. Finally, the ensemble map was compared with each individual map using the receiver operating characteristic (ROC) analysis to find if there is any improvement in combining the models.

3.2.1. Certainty Factor Model

Shortliffe and Buchanan [70] proposed and defined certainty factor (*CF*) as a probability function that was later modified by Heckerman [71]. Different researchers have applied this model in their studies for landslide susceptibility mapping [72] and groundwater potential mapping [26]. The formula to calculate *CF* is [71]:

$$CF = \begin{cases} \frac{P_a - P_b}{P_a(1-P_b)} \; if \; P_a \geq P_b \\ \frac{P_a - P_b}{P_b(1-P_a)} \; if \; P_a < P_b \end{cases} \quad (2)$$

where P_a is conditional probability of having a number of karst spring occurrence in class a, P_b is prior probability of having the total number of karst spring occurrence in the study area. The *CF* value for different attribute classes in each conditioning factor were calculated from their relationship to the spring occurrence. The *CF* values ranges from −1 to +1. A positive value of *CF* means that the groundwater potential is higher, while a negative value illustrates a decrease in the certainty of groundwater occurrence [73]. In order to delineate the combination of two different layers of

information, which is denoted as Z, the CF values of the spring-affecting factors are pairwise combined using the integration rules as follow [74].

$$Z = \begin{cases} CF_X + CF_Y - CF_X CF_Y & CF_X \text{ and } CF_Y \geq 0 \\ \frac{CF_X + CF_Y}{1 - min(|CF_X| \text{ and } |CF_Y|)} & CF_X \text{and } CF_Y \text{ Opposite sign} \\ CF_X + CF_Y + CF_X CF_Y & CF_X \text{ and } CF_Y < 0 \end{cases} \quad (3)$$

where CF_X and CF_Y are CF values of two different spring-affecting factors.

The pairwise combination is carried out in different iterations until all the spring-affecting layers are added to prepare the KSPM. The final map was categorized into four classes: low, medium, high and very high [29,75].

3.2.2. Logistic Regression Model

LR aims at finding the best-fit model to express the relationship among a binary dependent variable (groundwater occurrence) and some independent variables [76]. This model is helpful to predict the absence or presence of outcome based on predictive variables (Lee 2005). The LR model is widely applied in different fields of geosciences [24]. However, it has some limitations to determine the weight of classes of independent variables [77]. Hosmer and Lemeshow [78] and Kleinbaum et al. [79] also explained the detailed information of this technique. The general form of logistic regression is as follows:

$$Z = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_m x_m \quad (4)$$

where Z is a linear combination function of the explanatory variables, $x_1, x_2, \ldots, x_m$ are independent variables and $b_1, b_2, \ldots, b_m$ are the coefficients of the logistic regression to be estimated. The relationship between the occurrence and its dependency on several variables is defined as:

$$P = \frac{e^z}{1 + e^z} \quad (5)$$

where P and e are the probability of spring occurrence (varying from 0 to 1) and the Neperian number, respectively [80]. In addition, Z factor was defined above Equation (4). Value of 1 means the presence of a spring, and value 0 indicates the absence of a spring [24]. After creating the KSPM by LR model, we classified it using natural break classification method. As mentioned by Hosmer and Lemeshow [78], independent variables in LR model is sensitive to co-linearity for model fitting, while more than two variables in LR model is often called 'multi-collinearity'. Multi-collinearity refers to a linear relation among two or more variables [81]. LR model was used to do statistical analysis and determine whether there was a significant effect on spring occurrence. Tolerance (TOL) and the variance inflation factor (VIF) are significant indicators for determining multi-collinearity among the independent variables [82] that were applied (Table 2). Based on this diagnosis method, variables with TOL < 0.1 and/or VIF > 10 shall be excluded from further analysis [83]. Moreover, the overall statistic of the LR model for all 10 independent variables was assessed using χ^2 value of the Hosmer–Lemeshow test, Cox and Snell R^2, Nagelkerke R^2 and pseudo-R^2 mesaures.3.2.3. Ensemble model.

Table 2. Indices of multi-collinearity diagnosis for independent variables used in logistic regression (LR) modelling.

Spring-Affecting Factors	TOL	VIF
Plan curvature	0.89	1.18
Slope aspect	0.96	1.20
Slope angle	0.86	1.05
Distance from river	0.94	1.20
Drainage density	0.84	1.08
TWI	0.83	1.17
Topographic elevation	0.93	1.06
Geology	0.97	1.13
Land use/ land cover	0.87	1.02
Distance from fault	0.82	1.05

Ensemble modeling is a technical process of synthesizing the results of single models into a single incorporated model for enhancing the prediction [84]. This type of modeling approach tends to overcome their individual disadvantages while keeping their individual advantages. Weighted linear combination technique is proposed here as an ensemble method to combine the result of *CF* and LR:

$$KSPM = (DF_r DF_w) + (L_r L_w) + (G_r G_w) + (A_r A_w) + (TWI_r TWI_w) + (DD_r DD_w) + (DR_r DR_w) + (S_r S_w) + (AS_r AS_w) + (PC_r PC_w) \quad (6)$$

where *DF* is distance from fault, *L* is land use/ land cover, *G* is geology, *A* is topographic elevation, *TWI* is topographic wetness index, *DD* is drainage density, *DR* is distance from river, *S* is slope angle, *AS* is slope aspect, *PC* is plan curvature and the subscripts 'r' and 'w' indicate ratings and weights for spring-affecting factors, respectively. Ratings are obtained from the *CF* model (i.e., the weights of classes of a given spring-affecting factor), whereas w was calculated based on the spatial sensitivity analysis (SSA) of LR model. The SSA allows to determine the influence of predictive variables on the model output [16,85]. In this study, this was conducted by excluding each spring-affecting factor and determining w through the relative decrease of *AUC* values Equation (7) [29]:

$$w = \frac{[AUC_{all} - AUC_i]}{AUC_{all}} \times 100 \quad (7)$$

where AUC_{all} and AUC_i indicate the *AUC* values obtained from the LR model using all spring-affecting factors and the prediction when the ith spring-affecting factor was excluded, respectively.

3.2.3. Performance Assessment

Performance assessment is required in any predictive analysis [86]. Nampak et al. [75] also stated that there is no merit in modeling with no scientific significance; hence, validity and uncertainty of models need to be assessed. Furthermore, many researchers have validated their results using different models. Naghibi et al. [31], Pourghasemi and Beheshtirad [58] and Rahmati et al. [3] have utilized area under ROC curve (*AUC*) to assess the accuracy of groundwater potential maps. This method illustrates to what extent the occurrence or non-occurrence of a considered occurrence is correctly predicted [29]. In the ROC curve, false and true positive rates are plotted in X- and Y-axis, respectively. The resultant *AUC* value ranges from 0.5 (i.e., a random prediction) to 1 (i.e., an excellent prediction) [87].

4. Results

4.1. CF Modelling

CF values of spring-affecting factors are presented in Figure 4. As depicted in the figure, negative and positive relationships between factors and spring occurrence are well illustrated. Obviously,

negative relationship exists between spring potential and distance from faults parameter. Areas with more distance from fault have lower potential to spring occurrence. In the case of geology factor, some lithologic formations including Silurian (S), Jurassic-Cretaceous (JC) and Triassic-Jurassic (TJ) had positive influence in spring occurrence with *CF* values of 0.94, 0.42 and 0.21, respectively.

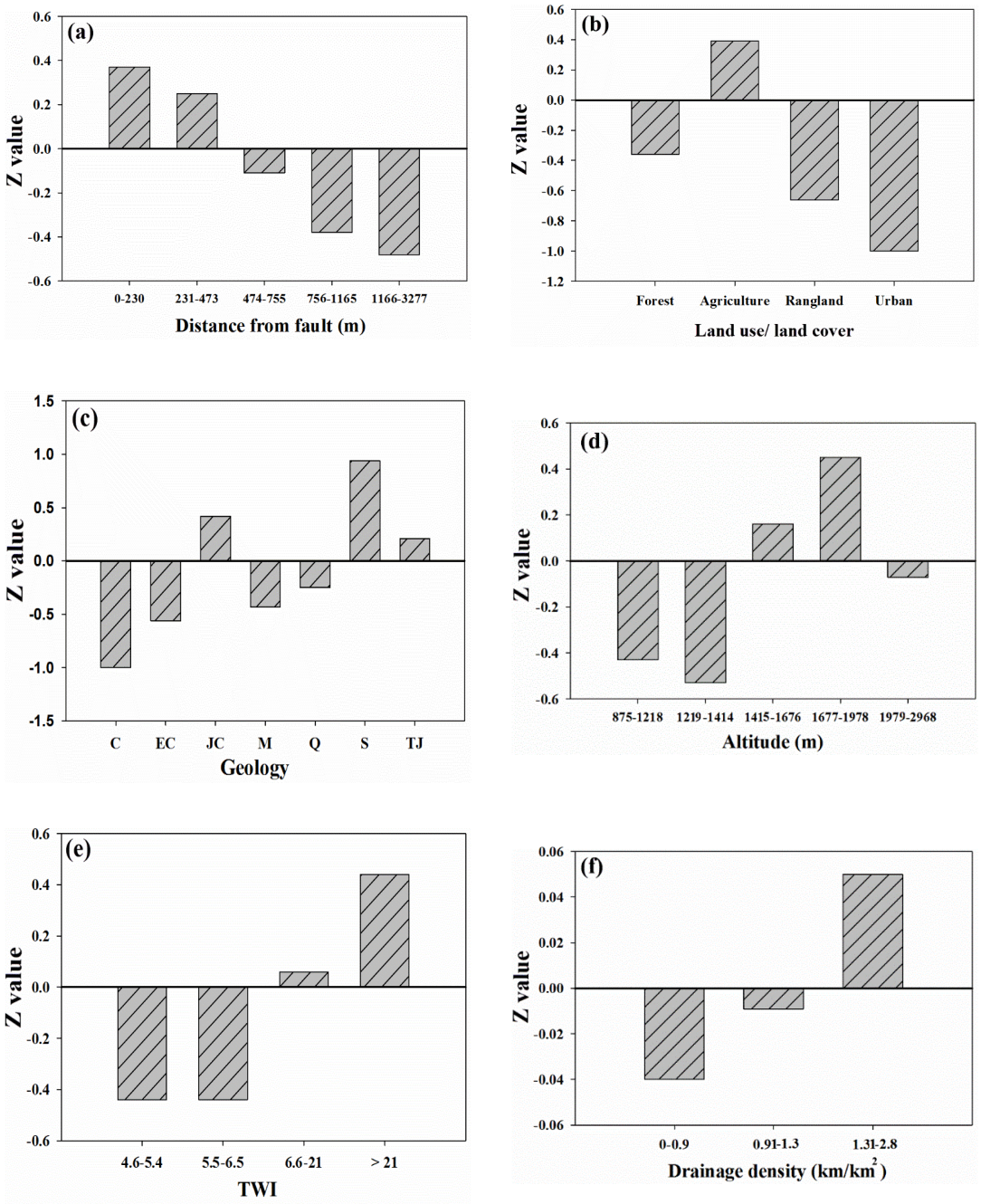


Figure 4. *Cont.*

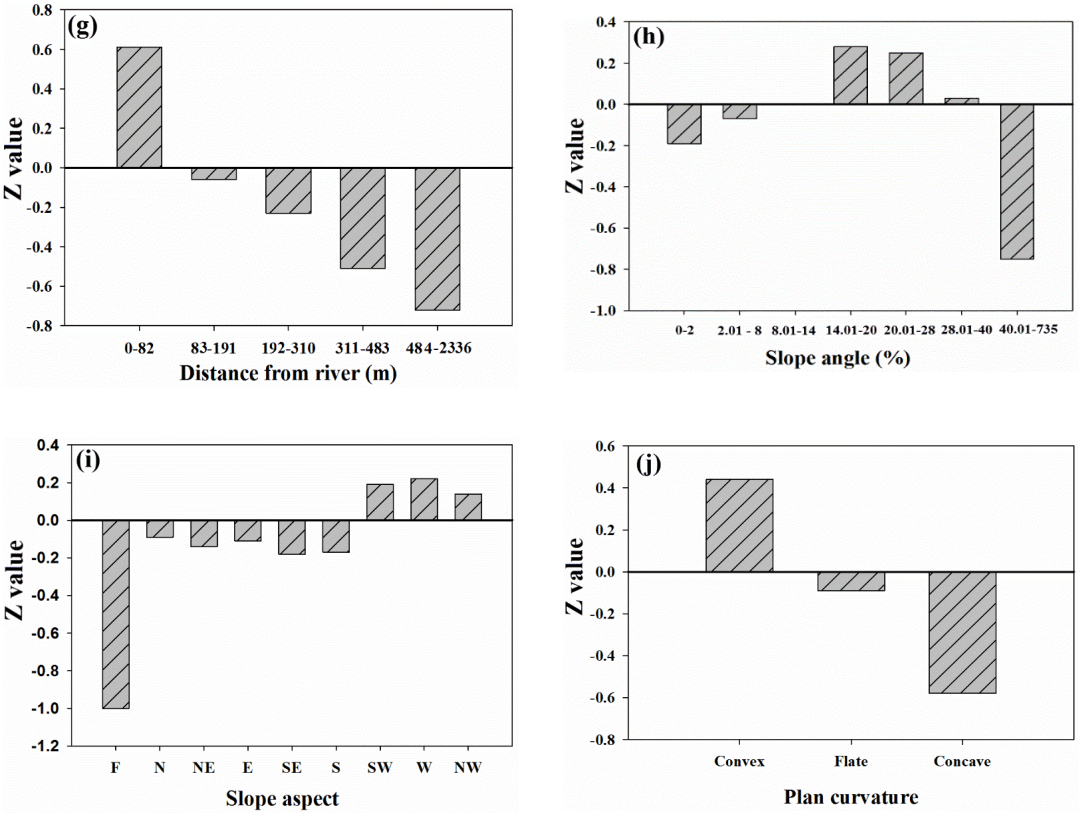


Figure 4. Results of the certainty factors (*CF)* modelling per factor class.

In the case of land use/land cover, urban areas and agriculture have the least and most impact on groundwater, respectively. The impervious surface areas in urban areas tend to reduce recharge and hence groundwater availability. As the agriculture areas as a consequence factor need to high water requirement, these areas have a higher potential to groundwater recharge [42]. It was also found that there was no significant correlation between spring occurrence and topographic elevation classes. Areas with elevation ranges of 1676–1978 m and 1414–1676 m have positive influence on spring occurrence probability with *CF* values of 0.45 and 0.16 which indicate a good karst spring occurrence probability. In another study in this region, Rahmati et al. [40] indicated that spring occurrence probability will reduce as elevation decreases which is in agreement with the results of our study. It is noted that, groundwater is more probable to occur in areas with high *TWI* values because it is well apparent in Figure 4e that positive correlation exists between increasing spring occurrence and high *TWI* value.

Conversely, Razandi et al. [26] stated that there is not positive correlation between groundwater occurrence and *TWI*, which shows a higher groundwater potential over a decreasing *TWI* value. The reason of this difference between their results and our achievements refer to geomorphological characteristics of study areas (i.e., their study area were plain whereas our study area is a mountainous region). In areas where drainage density ranges from 1.3 to 2.8 km/km^2 (*CF* = 1.82), groundwater potential is higher indicating that springs are more probable in areas with denser drainage density. The assessment of distance from river demonstrated that areas with lower distance from river have more potential for spring occurrence and as this distance increase, the probability to groundwater occurrence decreases. Lee et al. [88] proved these results by indicating that drainage density and distance from the river have positive and negative influence on groundwater yield. Results indicated that the areas with 8–40° slope angles have positive influences with highest value in 14–20° and 20–28° with *CF* values of 0.28 and 0.25. Manap et al. [21] also mentioned that flat to moderate slopes have higher groundwater probabilities.

In the case of slope aspect, only west (0.22), southwest (0.19) and northwest (0.14) have a positive influence on groundwater occurrence. It is difficult to explain this result, but it might be related to the differences of aspects in terms of vegetation, soil depth and recharge potential. Plan curvature behavior also illustrates that springs are more probable to occur in convex areas. Moghadam et al. [63] also stated that concave areas are more favorable for groundwater compared to the convex slope. Finally, according to the *CF* value for each influencing factor, karst spring probability map was produced and then categorized into four classes; namely low, medium, high and very high using the natural break classification method (Figure 5), as suggested by Naghibi et al. [31]. Areas with high and very high karst spring potential are seen to be near rivers and faults.

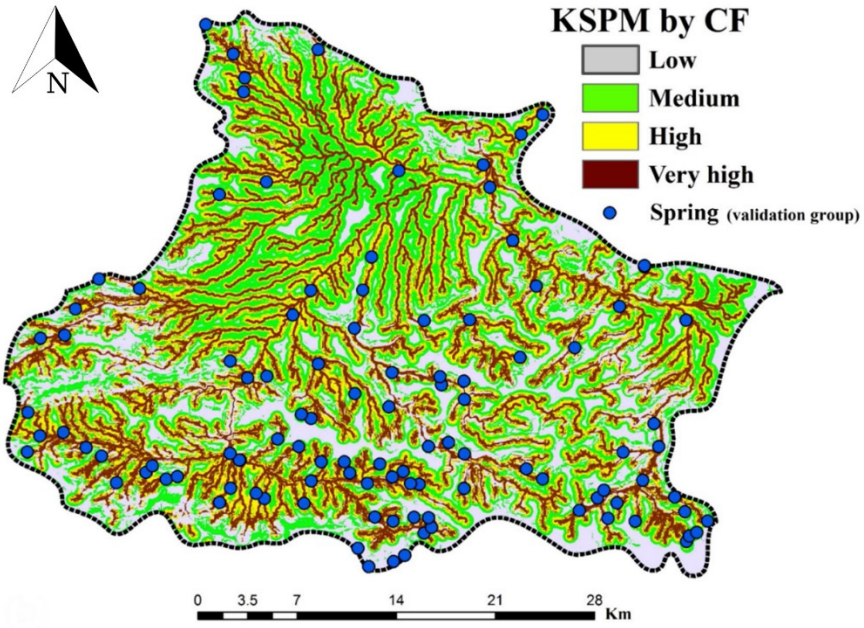


Figure 5. Karst spring potential map produced by the *CF* modelling.

4.2. LR Modelling

According to results of TOL and VIF analysis, no severe multi-collinearity was found between independent variables (Table 2).

As shown in Table 3, results indicated that all applied factors have significant effects on spring occurrence ($p < 0.05$). Similar results can be found in many previous studies [58].

Table 3. Coefficients of each spring-affecting factors used in logistic regression (LR) model.

Spring-Affecting Factors	β1	S.E2	Wald3	Df4	Sig.5
Geology (Cretaceous)	−12.022	30.011	0.160	1	0.000
Geology (Early Cretaceous)	−3.483	31.228	0.012	1	0.000
Geology (Jurassic-Cretaceous)	8.550	29.001	0.087	1	0.000
Geology (Miocene)	−2.110	33.328	0.004	1	0.000
Geology (Quaternary)	−1.930	30.943	0.004	1	0.000
Geology (Silurian)	10.211	34.550	0.087	1	0.001
Geology (Triassic-Jurassic)	4.685	31.783	0.022	1	0.001
Distance from river (m)	−0.003	0.019	0.025	1	0.000
Slope angle (%)	−0.199	0.021	89.798	1	0.000
Slope aspect (North)	3.340	55.313	0.004	1	0.001

Table 3. *Cont.*

Spring-Affecting Factors	β1	S.E2	Wald3	Df4	Sig.5
Slope aspect (Northeast)	2.011	51.066	0.002	1	0.015
Slope aspect (East)	4.516	63.014	0.005	1	0.000
Slope aspect (Southeast)	3.782	59.099	0.004	1	0.000
Slope aspect (South)	2.080	58.318	0.001	1	0.001
Slope aspect (Southwest)	13.652	61.021	0.050	1	0.000
Slope aspect (West)	5.318	54.510	0.010	1	0.000
Slope aspect (Northwest)	3.401	50.709	0.004	1	0.001
TWI	−0.990	0.938	1.114	1	0.000
Drainage density (km/km^2)	−2.663	0.737	13.056	1	0.000
Altitude (m)	0.374	0.019	387.468	1	0.000
Distance from fault (m)	−0.319	0.044	52.563	1	0.000
Plan curvature	0.044	0.026	2.864	1	0.000
Landuse (Forest)	11.408	73.990	0.024	1	0.012
Landuse (Agriculture)	−0.913	80.428	0.000	1	0.018
Landuse (Rangland)	3.325	72.565	0.002	1	0.011
Constant	−16.33	16.710	0.955	1	0.001

β1= logistic coefficient; S.E.2 = standard error of estimate; Wald3 = Wald chi-square values; Df4 = degree of freedom; Sig.5 = significance.

Table 4 showed the overall statistic of the LR model for all 10 independent variables. The results indicated that the goodness-of-fit of the equation could be accepted because the significance of χ^2 is larger than 0.05 [83]. The values of Cox and Snell R^2 and Nagelkerke R^2 showed that the independent variables could explain the dependent variables [83].

Table 4. Overall statistics of the LR model with 10 independent variables.

Hosmer and Lemeshow Test		Cox and Snell R^2	Nagelkerke R^2	Pseudo R^2
χ^2	Significance			
563.31	0.294	0.699	0.821	0.710

Using the coefficient values obtained from the final output of the LR analysis, the LR equation was obtained, as shown in Equation (8):

$$\begin{aligned} Z = C_{Geology} + \quad & (-0.003 \times Distance\ from\ river) + (-0.199 \times Slope\ angle) + C_{Slope\ aspect} \\ & + (-0.990 \times TWI) + (-2.663 \times Drainage\ density) \\ & + (0.374 \times Topographic\ elevation) + (-0.319 \times Distance\ from\ fault) \\ & + (0.044 \times Plan\ curvature) + C_{Land\ use} - 16.33 \end{aligned} \quad (8)$$

where $C_{Geology}$ is the LR coefficient for geology factor, whereas $C_{Land\ use}$ is the LR coefficient for land use factor. All coefficients of Equation (8) were obtained from Table 3. The final generated map using LR model was then classified into low, medium, high and very high classes using the natural break classification method (Figure 6), as suggested by Naghibi et al. [31]. Areas with high and very high potential to karst springs are located in southwest to southeast of the study area.

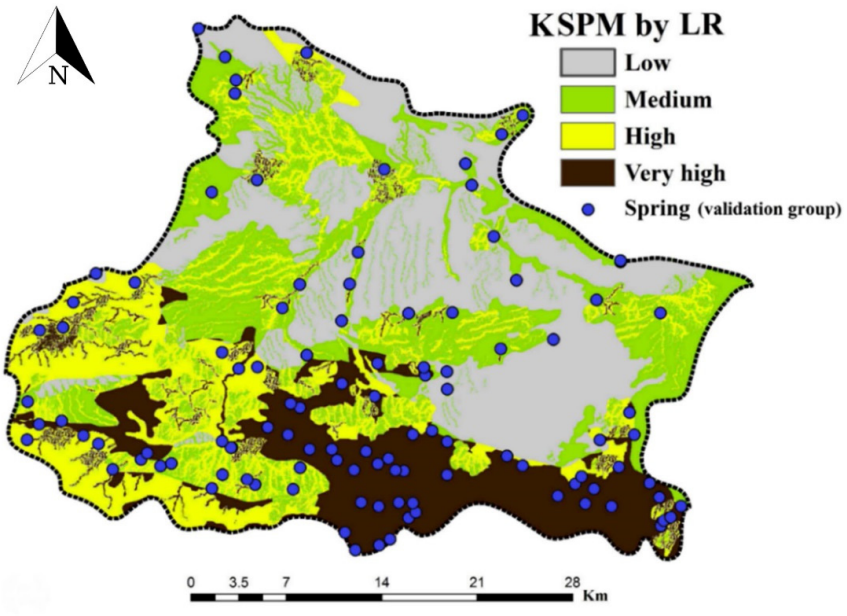


Figure 6. Karst spring potential map produced by the LR modelling.

4.3. Ensemble Modelling

Although LR was capable to implement multivariate statistical analysis, it had some weakness in analyzing the classes of each occurrences conditioning factor. Conversely, the *CF* model could evaluate the influence of classes of each spring-affecting factor on an occurrence. However, the relationship between karst spring locations and spring-affecting factors was mostly neglected by LR model. As a result, it can be clearly seen that *CF* and LR models had some weakness when applied individually. The generated map of ensemble *CF* and LR (ECL) model was prepared and then classified into low, moderate, high and very high classes using the natural break classification method (Figure 7), as suggested by Naghibi et al. [31].

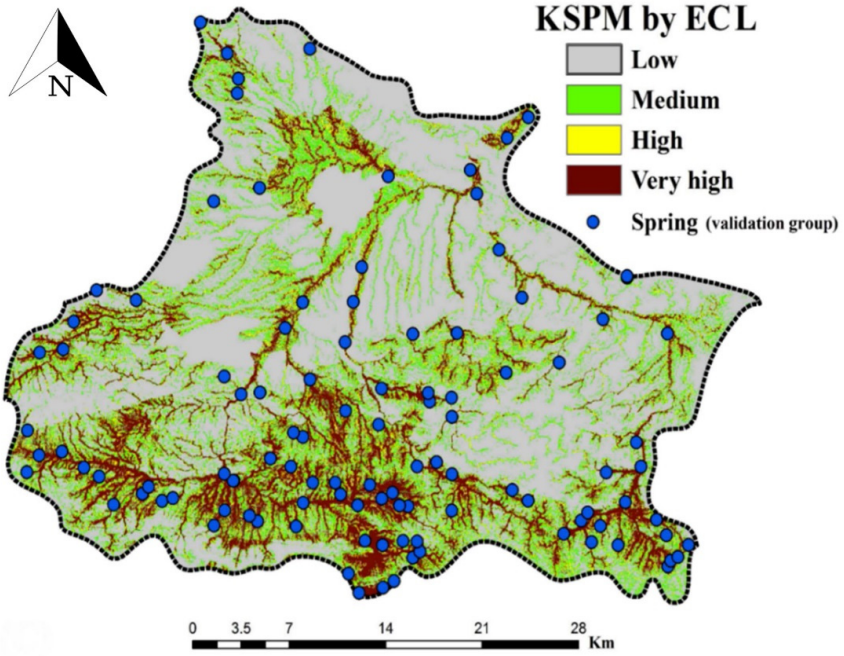


Figure 7. Karst spring potential map produced by the generated map of ensemble *CF* and LR (ECL) modelling.

4.4. Validation of Generated KSPMs

The ensemble map was compared to each individual (i.e., *CF* and LR) map using ROC curve method to find if there was any improvements in combining the individual models. The generated SPMs were overlaid with 30% of spring locations for this phase, which has not been used for model application. According to Figure 8, validation outcomes indicated that LR model could satisfactorily generate groundwater potential map (*AUC* = 78.4%) and performed better than *CF* (AUC=66.7%). Moreover, the ECL model had the best performance with an *AUC* value of 84.7%. It was shown that the ensemble method allowed for a better prediction of groundwater potentiality.

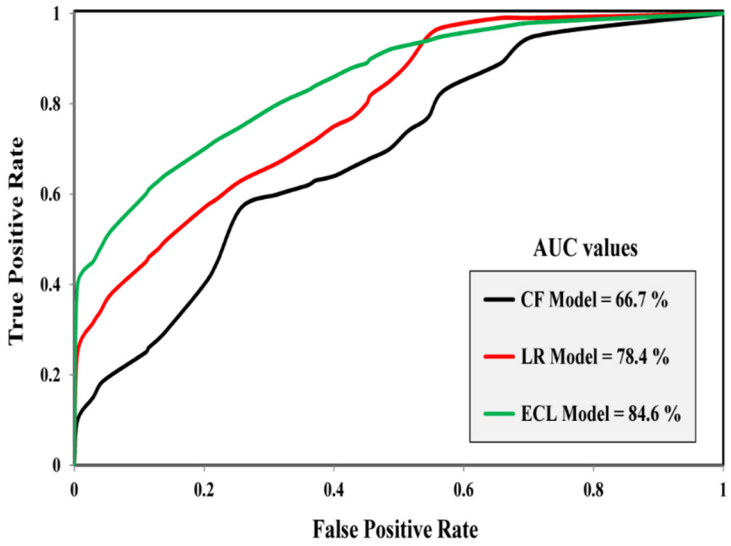


Figure 8. Receiver operating characteristic (ROC) curves for the karst spring potential maps.

5. Discussion

5.1. Spatial Modeling of Karst Spring Potential

Although several studies have been conducted on construction of karst spring potential maps using statistical models, there are different drawbacks about them, especially defining strict assumptions prior to study. As stated by Dreiseitl and Ohno-Machadoand Ozdemir [89], the LR model is usually flexible (especially when it includes only the original covariates) and may be generally desirable. On the other hand, the weak point of bivariate statistical methods such as *CF* is the neglect of relationship between the groundwater springs and independent variables as observed in this study (i.e., it cannot determine the weight of independent variables). In other word, the *CF* model just determines the weight of classes of independent variables. Therefore, to overcome these shortcomings, we developed a GIS-based ensemble model that can link the strengths of LR and *CF* together.

The statistical analysis of LR (Table 3) demonstrated a high contribution of slope aspect (southwest), geology (Cretaceous, Silurian and Jurassic-Cretaceous), land use/land cover (forest) to the modelling process. Prior studies have noted the importance of slope aspect effects on sunshine, air temperature, soil transformation and soil moisture content [90]. According to the literature, structural and lithological variations often lead to differences in the permeability and hydraulic properties [8,91,92]. In the case of land use/land cover, it usually influences infiltration and recharge rates as well as the relationship between surface water and groundwater [93]. This is also in agreement with earlier reports, which showed that there is a strong relationship between land use/land cover factor and groundwater recharge [94,95].

5.2. Performance of Ensemble and Individual Models

In Section 4.4, models' results were analyzed and validated using the validation dataset and the ROC curve method. The results of validation revealed differences between all three approaches, namely including *CF* (i.e., bivariate statistical model), LR (i.e., multivariate statistical model) and their ensemble.

It should be emphasized that the ensemble model (ECL) provides the best simulation with a good accuracy. This result showed that the accuracy of the prepared KSPM was improved (6.3–17.9%) with the use of the ensemble of bivariate certainty factor and multivariate logistic regression. Therefore, the ensembling process leads to successfully overcome the limitations of LR and *CF* models. In fact, results shown in this study highlight a remarkable decrease of error through the ensemble approach. This result is in agreement with Arabameri et al. [96] in which the ensemble of LR with evidential belief function (EBF) model can extend an efficient model to reinforce the accuracy of spatial predictions. According to results of this study, the aggregation of multiple prediction of the same variable may lead to better performance, more exact learning and generalization than using a single model prediction. The findings of the current study are consistent with that of Naghibi et al. [30] who found that frequency ratio data mining ensemble model had the best performance among all the single models for groundwater potential mapping. Study in Razandi et al. [26] also compared the performance of different data-mining models and stated that frequency ration (FR) and weight-of-evidence (WOE) could perform better than *CF* model for groundwater potential mapping in Varamin plain, Iran. This is in line with our finding that the *CF* model is incapable of analyzing the relation between karst groundwater potential and geo-environmental factors.

6. Conclusions

Springs, as groundwater resources, are important for many sectors, such as domestic consumption and agriculture in arid and semi-arid areas of the world. Sometimes, several families or villages depend heavily on occurrence of a spring; therefore, their spatial modeling is necessary. We designed a new ensemble approach called ECL to deal with this issue. In conclusion, the novel ensemble model gives better predictive performance than could be obtained from any of the constituent models. Therefore, the strength of the proposed methodology is to improve model accuracy by using this novel ensembling technique and to efficiently produce a KSPM for sustainable water recourses management programs. Decision makers and water resources planners can use the proposed approach as an efficient tool to assess the karst spring potential, to establish relationship between karst springs and geo-environmental setting, and to plan and implement effective water resources management in drought condition. Future studies on the current topic are therefore recommended.

Author Contributions: V.-H.N., O.R., F.F., S.S., N.A.-A., H.S., A.S., K.G., H.N. and B.B.A. contributed equally to the work. O.R., F.F. and S.S. collected field data and conducted the groundwater spring potential in karst aquifer and analysis. O.R., F.F., S.S., N.A.-A., H.S., and K.G. wrote the manuscript. V.-H.N., N.A.-A., A.S., K.G., H.N. and B.B.A. provided critical comments in planning this paper and edited the manuscript. All the authors discussed the results and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wada, Y.; Van Beek, L.; Bierkens, M.F. Nonsustainable groundwater sustaining irrigation: A global assessment. *Water Resour. Res.* **2012**, *48*, 1–18. [CrossRef]
2. Gorelick, S.M.; Zheng, C. Global change and the groundwater management challenge. *Water Resour. Res.* **2015**, *51*, 3031–3051. [CrossRef]
3. Rahmati, O.; Samani, A.N.; Mahdavi, M.; Pourghasemi, H.R.; Zeinivand, H. Groundwater potential mapping at kurdistan region of iran using analytic hierarchy process and gis. *Arab. J. Geosci.* **2015**, *8*, 7059–7071. [CrossRef]

4. Chen, W.; Li, H.; Hou, E.; Wang, S.; Wang, G.; Panahi, M.; Li, T.; Peng, T.; Guo, C.; Niu, C. Gis-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. *Sci. Total Environ.* **2018**, *634*, 853–867. [CrossRef] [PubMed]
5. White, W.B.; Herman, E.; Rutigliano, M.; Herman, J.; Vesper, D.; Engel, S. Karst groundwater contamination and public health. *Leesburg VA Karst Waters Inst. Spec. Publ.* **2016**, *19*, 1–347.
6. Zhang, L.; Qin, X.; Tang, J.; Liu, W.; Yang, H. Review of arsenic geochemical characteristics and its significance on arsenic pollution studies in karst groundwater, southwest china. *Appl. Geochem.* **2017**, *77*, 80–88. [CrossRef]
7. Yuan, J.; Xu, F.; Deng, G.; Tang, Y. Using stable isotopes and major ions to identify hydrogeochemical characteristics of karst groundwater in xide country, sichuan province. *Carbonates Evaporites* **2018**, *33*, 223–234. [CrossRef]
8. Bakalowicz, M. Karst groundwater: A challenge for new resources. *Hydrogeol. J.* **2005**, *13*, 148–160. [CrossRef]
9. Papadopoulou-Vrynioti, K.; Bathrellos, G.D.; Skilodimou, H.D.; Kaviris, G.; Makropoulos, K. Karst collapse susceptibility mapping considering peak ground acceleration in a rapidly growing urban area. *Eng. Geol.* **2013**, *158*, 77–88. [CrossRef]
10. Ebrahimi, O.; Ahmadi, M.; Shahabi, H.; Asgari, S. Evaluation of karst features using principal component analysis (pca): A case from zarneh and kergan, western iran. *Carbonates Evaporites* **2018**, *33*, 625–635. [CrossRef]
11. Andreo, B. *Introductory Editorial: Advances in Karst Hydrogeology*; Springer: Berlin/Heidelberg, Germany, 2012.
12. Cuchí, J.A.; Chinarro, D.; Villarroel, J.L. Linear system techniques applied to the fuenmayor karst spring, huesca (spain). *Environ. Earth Sci.* **2014**, *71*, 1049–1060. [CrossRef]
13. Fleury, P.; Plagnes, V.; Bakalowicz, M. Modelling of the functioning of karst aquifers with a reservoir model: Application to fontaine de vaucluse (south of france). *J. Hydrol.* **2007**, *345*, 38–49. [CrossRef]
14. Chebud, Y.A.; Melesse, A.M. Numerical modeling of the groundwater flow system of the gumera sub-basin in lake tana basin, ethiopia. *Hydrol. Process. Int. J.* **2009**, *23*, 3694–3704. [CrossRef]
15. Kuniansky, E.L.; Weary, D.J.; Kaufmann, J.E. The current status of mapping karst areas and availability of public sinkhole-risk resources in karst terrains of the united states. *Hydrogeol. J.* **2016**, *24*, 613–624. [CrossRef]
16. Oh, H.-J.; Kim, Y.-S.; Choi, J.-K.; Park, E.; Lee, S. Gis mapping of regional probabilistic groundwater potential in the area of pohang city, korea. *J. Hydrol.* **2011**, *399*, 158–172. [CrossRef]
17. Naghibi, S.A.; Pourghasemi, H.R. A comparative assessment between three machine learning models and their performance comparison by bivariate and multivariate statistical methods in groundwater potential mapping. *Water Resour. Manag.* **2015**, *29*, 5217–5236. [CrossRef]
18. Chen, Z.; Auler, A.S.; Bakalowicz, M.; Drew, D.; Griger, F.; Hartmann, J.; Jiang, G.; Moosdorf, N.; Richts, A.; Stevanovic, Z. The world karst aquifer mapping project: Concept, mapping procedure and map of europe. *Hydrogeol. J.* **2017**, *25*, 771–785. [CrossRef]
19. Corsini, A.; Cervi, F.; Ronchetti, F. Weight of evidence and artificial neural networks for potential groundwater spring mapping: An application to the mt. Modino area (northern apennines, italy). *Geomorphology* **2009**, *111*, 79–87. [CrossRef]
20. Jha, M.K.; Chowdary, V.; Chowdhury, A. Groundwater assessment in salboni block, west bengal (india) using remote sensing, geographical information system and multi-criteria decision analysis techniques. *Hydrogeol. J.* **2010**, *18*, 1713–1728. [CrossRef]
21. Manap, M.A.; Nampak, H.; Pradhan, B.; Lee, S.; Sulaiman, W.N.A.; Ramli, M.F. Application of probabilistic-based frequency ratio model in groundwater potential mapping using remote sensing data and gis. *Arab. J. Geosci.* **2014**, *7*, 711–724. [CrossRef]
22. Sahoo, S.; Munusamy, S.B.; Dhar, A.; Kar, A.; Ram, P. Appraising the accuracy of multi-class frequency ratio and weights of evidence method for delineation of regional groundwater potential zones in canal command system. *Water Resour. Manag.* **2017**, *31*, 4399–4413. [CrossRef]
23. Ghorbani Nejad, S.; Falah, F.; Daneshfar, M.; Haghizadeh, A.; Rahmati, O. Delineation of groundwater potential zones using remote sensing and gis-based data-driven models. *Geocarto Int.* **2017**, *32*, 167–187. [CrossRef]
24. Ozdemir, A. Using a binary logistic regression method and gis for evaluating and mapping the groundwater spring potential in the sultan mountains (aksehir, turkey). *J. Hydrol.* **2011**, *405*, 123–136. [CrossRef]

25. Pourtaghi, Z.S.; Pourghasemi, H.R. Gis-based groundwater spring potential assessment and mapping in the birjand township, southern khorasan province, iran. *Hydrogeol. J.* **2014**, *22*, 643–662. [CrossRef]
26. Razandi, Y.; Pourghasemi, H.R.; Neisani, N.S.; Rahmati, O. Application of analytical hierarchy process, frequency ratio, and certainty factor models for groundwater potential mapping using gis. *Earth Sci. Inform.* **2015**, *8*, 867–883. [CrossRef]
27. Fenta, A.A.; Kifle, A.; Gebreyohannes, T.; Hailu, G. Spatial analysis of groundwater potential using remote sensing and gis-based multi-criteria evaluation in raya valley, northern ethiopia. *Hydrogeol. J.* **2015**, *23*, 195–206. [CrossRef]
28. Naghibi, S.A.; Pourghasemi, H.R.; Pourtaghi, Z.S.; Rezaei, A. Groundwater qanat potential mapping using frequency ratio and shannon's entropy models in the moghan watershed, iran. *Earth Sci. Inform.* **2015**, *8*, 171–186. [CrossRef]
29. Rahmati, O.; Pourghasemi, H.R.; Melesse, A.M. Application of gis-based data driven random forest and maximum entropy models for groundwater potential mapping: A case study at mehran region, iran. *Catena* **2016**, *137*, 360–372. [CrossRef]
30. Naghibi, S.A.; Ahmadi, K.; Daneshi, A. Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping. *Water Resour. Manag.* **2017**, *31*, 2761–2775. [CrossRef]
31. Naghibi, S.A.; Pourghasemi, H.R.; Dixon, B. Gis-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in iran. *Environ. Monit. Assess.* **2016**, *188*, 44. [CrossRef]
32. Anbazhagan, S.; Jothibasu, A. Geoinformatics in groundwater potential mapping and sustainable development: A case study from southern india. *Hydrol. Sci. J.* **2016**, *61*, 1109–1123. [CrossRef]
33. Althuwaynee, O.F.; Pradhan, B.; Lee, S. Application of an evidential belief function model in landslide susceptibility mapping. *Comput. Geosci.* **2012**, *44*, 120–135. [CrossRef]
34. Umar, Z.; Pradhan, B.; Ahmad, A.; Jebur, M.N.; Tehrany, M.S. Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in west sumatera province, indonesia. *Catena* **2014**, *118*, 124–135. [CrossRef]
35. Youssef, A.M.; Pradhan, B.; Jebur, M.N.; El-Harbi, H.M. Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in fayfa area, saudi arabia. *Environ. Earth Sci.* **2015**, *73*, 3745–3761. [CrossRef]
36. Aghdam, I.N.; Varzandeh, M.H.M.; Pradhan, B. Landslide susceptibility mapping using an ensemble statistical index (wi) and adaptive neuro-fuzzy inference system (anfis) model at alborz mountains (iran). *Environ. Earth Sci.* **2016**, *75*, 553. [CrossRef]
37. Bui, D.T.; Ho, T.-C.; Pradhan, B.; Pham, B.-T.; Nhu, V.-H.; Revhaug, I. Gis-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with adaboost, bagging, and multiboost ensemble frameworks. *Environ. Earth Sci.* **2016**, *75*, 1101.
38. Mojaddadi, H.; Pradhan, B.; Nampak, H.; Ahmad, N.; Ghazali, A.H. Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and gis. *Geomat. Nat. Hazards Risk* **2017**, *8*, 1080–1102. [CrossRef]
39. Termeh, S.V.R.; Kornejady, A.; Pourghasemi, H.R.; Keesstra, S. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. *Sci. Total Environ.* **2018**, *615*, 438–451. [CrossRef]
40. Rahmati, O.; Naghibi, S.A.; Shahabi, H.; Bui, D.T.; Pradhan, B.; Azareh, A.; Rafiei-Sardooi, E.; Samani, A.N.; Melesse, A.M. Groundwater spring potential modelling: Comprising the capability and robustness of three different modeling approaches. *J. Hydrol.* **2018**, *565*, 248–261. [CrossRef]
41. Barzegar, R.; Moghaddam, A.A.; Deo, R.; Fijani, E.; Tziritis, E. Mapping groundwater contamination risk of multiple aquifers using multi-model ensemble of machine learning algorithms. *Sci. Total Environ.* **2018**, *621*, 697–712. [CrossRef]
42. Kordestani, M.D.; Naghibi, S.A.; Hashemi, H.; Ahmadi, K.; Kalantar, B.; Pradhan, B. Groundwater potential mapping using a novel data-mining ensemble model. *Hydrogeol. J.* **2019**, *27*, 211–224. [CrossRef]
43. Bui, D.T.; Panahi, M.; Shahabi, H.; Singh, V.P.; Shirzadi, A.; Chapi, K.; Khosravi, K.; Chen, W.; Panahi, S.; Li, S. Novel hybrid evolutionary algorithms for spatial prediction of floods. *Sci. Rep.* **2018**, *8*, 15364. [CrossRef]

44. Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Bui, D.T.; Pham, B.T.; Khosravi, K. A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environ. Model. Softw.* **2017**, *95*, 229–245. [CrossRef]
45. Pham, B.T.; Prakash, I.; Dou, J.; Singh, S.K.; Trinh, P.T.; Tran, H.T.; Le, T.M.; Van Phong, T.; Khoi, D.K.; Shirzadi, A. A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. *Geocarto Int.* **2019**, *14*, 1–25. [CrossRef]
46. Tien Bui, D.; Shirzadi, A.; Shahabi, H.; Chapi, K.; Omidavr, E.; Pham, B.T.; Talebpour Asl, D.; Khaledian, H.; Pradhan, B.; Panahi, M. A novel ensemble artificial intelligence approach for gully erosion mapping in a semi-arid watershed (iran). *Sensors* **2019**, *19*, 2444. [CrossRef]
47. Rahmati, O.; Panahi, M.; Ghiasi, S.S.; Deo, R.C.; Tiefenbacher, J.P.; Pradhan, B.; Jahani, A.; Goshtasb, H.; Kornejady, A.; Shahabi, H. Hybridized neural fuzzy ensembles for dust source modeling and prediction. *Atmos. Environ.* **2020**, *224*, 117320. [CrossRef]
48. Jaafari, A.; Zenner, E.K.; Panahi, M.; Shahabi, H. Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability. *Agric. For. Meteorol.* **2019**, *266*, 198–207. [CrossRef]
49. Taheri, K.; Shahabi, H.; Chapi, K.; Shirzadi, A.; Gutiérrez, F.; Khosravi, K. Sinkhole susceptibility mapping: A comparison between bayes-based machine learning algorithms. *Land Degrad. Dev.* **2019**, *30*, 730–745. [CrossRef]
50. Roodposhti, M.S.; Safarrad, T.; Shahabi, H. Drought sensitivity mapping using two one-class support vector machine algorithms. *Atmos. Res.* **2017**, *193*, 73–82. [CrossRef]
51. Lee, S.; Panahi, M.; Pourghasemi, H.R.; Shahabi, H.; Alizadeh, M.; Shirzadi, A.; Khosravi, K.; Melesse, A.M.; Yekrangnia, M.; Rezaie, F. Sevucas: A novel gis-based machine learning software for seismic vulnerability assessment. *Appl. Sci.* **2019**, *9*, 3495. [CrossRef]
52. Chen, W.; Zhao, X.; Tsangaratos, P.; Shahabi, H.; Ilia, I.; Xue, W.; Wang, X.; Ahmad, B.B. Evaluating the usage of tree-based ensemble methods in groundwater spring potential mapping. *J. Hydrol.* **2020**, *583*, 124602. [CrossRef]
53. Al-Abadi, A.M. 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.* **2015**, *74*, 1109–1124. [CrossRef]
54. Shamsi, A.; Kazemi, G. A review of research dealing with isotope hydrology in iran and the first iranian meteoric water line. *Geopersia* **2014**, *4*, 73–86.
55. Rahmati, O.; Kornejady, A.; Samadi, M.; Nobre, A.D.; Melesse, A.M. Development of an automated gis tool for reproducing the hand terrain model. *Environ. Model. Softw.* **2018**, *102*, 1–12. [CrossRef]
56. Vali-Khodjeini, A. Human impacts on groundwater resources in iran. *IAHS Publ. Ser. Proc. Rep. Intern Assoc Hydrol. Sci.* **1995**, *230*, 141–146.
57. Crozier, M.J. *Landslides: Causes, Consequences & Environment*; Taylor & Francis: Abingdon-on-Thames, UK, 1986.
58. Pourghasemi, H.R.; Beheshtirad, M. Assessment of a data-driven evidential belief function model and gis for groundwater potential mapping in the koohrang watershed, iran. *Geocarto Int.* **2015**, *30*, 662–685. [CrossRef]
59. Gebere, S.B.; Alamirew, T.; Merkel, B.J.; Melesse, A.M. Land use and land cover change impact on groundwater recharge: The case of lake haramaya watershed, ethiopia. In *Landscape Dynamics, Soils and Hydrological Processes in Varied Climates*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 93–110.
60. Yalcin, A. Gis-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in ardesen (turkey): Comparisons of results and confirmations. *Catena* **2008**, *72*, 1–12. [CrossRef]
61. Regmi, N.R.; Giardino, J.R.; Vitek, J.D. Modeling susceptibility to landslides using the weight of evidence approach: Western colorado, USA. *Geomorphology* **2010**, *115*, 172–187. [CrossRef]
62. Xu, C.; Xu, X.; Dai, F.; Xiao, J.; Tan, X.; Yuan, R. Landslide hazard mapping using gis and weight of evidence model in qingshui river watershed of 2008 wenchuan earthquake struck region. *J. Earth Sci.* **2012**, *23*, 97–120. [CrossRef]
63. Moghaddam, D.D.; Rezaei, M.; Pourghasemi, H.; Pourtaghie, Z.; Pradhan, B. Groundwater spring potential mapping using bivariate statistical model and gis in the taleghan watershed, iran. *Arab. J. Geosci.* **2015**, *8*, 913–929. [CrossRef]
64. Magesh, N.; Chandrasekar, N.; Soundranayagam, J.P. Delineation of groundwater potential zones in theni district, tamil nadu, using remote sensing, gis and mif techniques. *Geosci. Front.* **2012**, *3*, 189–196. [CrossRef]

65. Selvam, S.; Magesh, N.; Sivasubramanian, P.; Soundranayagam, J.P.; Manimaran, G.; Seshunarayana, T. Deciphering of groundwater potential zones in tuticorin, tamil nadu, using remote sensing and gis techniques. *J. Geol. Soc. India* **2014**, *84*, 597–608. [CrossRef]
66. Prasad, J. A ground water brochure of jhansi district. *Uttar Pradesh* **2008**, *8*, 1–19.
67. Conforti, M.; Pascale, S.; Robustelli, G.; Sdao, F. Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the turbolo river catchment (northern calabria, italy). *Catena* **2014**, *113*, 236–250. [CrossRef]
68. Beven, K.J.; Kirkby, M.J. A physically based, variable contributing area model of basin hydrology/un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrol. Sci. J.* **1979**, *24*, 43–69. [CrossRef]
69. Shekhar, S.; Pandey, A.C. Delineation of groundwater potential zone in hard rock terrain of india using remote sensing, geographical information system (gis) and analytic hierarchy process (ahp) techniques. *Geocarto Int.* **2015**, *30*, 402–421. [CrossRef]
70. Shortliffe, E.H.; Buchanan, B.G. A model of inexact reasoning in medicine. *Math. Biosci.* **1975**, *23*, 351–379. [CrossRef]
71. Heckerman, D. *Probabilistic Interpretation of Mycin's Certainty Factors w Pracy Kanal ln, Lemmers jf (Eds): Uncertainty in Artificial Intelligence*; Elsevier: Amsterdam, The Netherlands; London, UK; New York, NY, USA, 1986.
72. Devkota, K.C.; Regmi, A.D.; Pourghasemi, H.R.; Yoshida, K.; Pradhan, B.; Ryu, I.C.; Dhital, M.R.; Althuwaynee, O.F. 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* **2013**, *65*, 135–165. [CrossRef]
73. Pourghasemi, H.R.; Pradhan, B.; Gokceoglu, C.; Mohammadi, M.; Moradi, H.R. Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at haraz watershed, iran. *Arab. J. Geosci.* **2013**, *6*, 2351–2365. [CrossRef]
74. Binaghi, E.; Luzi, L.; Madella, P.; Pergalani, F.; Rampini, A. Slope instability zonation: A comparison between certainty factor and fuzzy dempster–shafer approaches. *Nat. Hazards* **1998**, *17*, 77–97. [CrossRef]
75. Nampak, H.; Pradhan, B.; Manap, M.A. Application of gis based data driven evidential belief function model to predict groundwater potential zonation. *J. Hydrol.* **2014**, *513*, 283–300. [CrossRef]
76. Ohlmacher, G.C.; Davis, J.C. Using multiple logistic regression and gis technology to predict landslide hazard in northeast kansas, USA. *Eng. Geol.* **2003**, *69*, 331–343. [CrossRef]
77. Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Spatial prediction of flood susceptible areas using rule based decision tree (dt) and a novel ensemble bivariate and multivariate statistical models in gis. *J. Hydrol.* **2013**, *504*, 69–79. [CrossRef]
78. Hosmer, D.; Lemeshow, S. *Applied Logistic Regression Second Edition Wiley*; John Wiley & Sons: New York, NY, USA, 2000.
79. Kleinbaum, D.G.; Dietz, K.; Gail, M.; Klein, M.; Klein, M. *Logistic Regression*; Springer: Berlin/Heidelberg, Germany, 2002.
80. Dai, F.; Lee, C. Landslide characteristics and slope instability modeling using gis, lantau island, hong kong. *Geomorphology* **2002**, *42*, 213–228. [CrossRef]
81. Alin, A. Multicollinearity. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 370–374. [CrossRef]
82. Menard, S. *Applied Logistic Regression Analysis*; Sage: Newcastle Upon Tyne, UK, 2002; Volume 106.
83. Bai, S.-B.; Wang, J.; Lü, G.-N.; Zhou, P.-G.; Hou, S.-S.; Xu, S.-N. Gis-based logistic regression for landslide susceptibility mapping of the zhongxian segment in the three gorges area, china. *Geomorphology* **2010**, *115*, 23–31. [CrossRef]
84. Rokach, L. Ensemble-based classifiers. *Artif. Intell. Rev.* **2010**, *33*, 1–39. [CrossRef]
85. Park, N.-W. Using maximum entropy modeling for landslide susceptibility mapping with multiple geoenvironmental data sets. *Environ. Earth Sci.* **2015**, *73*, 937–949. [CrossRef]
86. Mancini, F.; Ceppi, C.; Ritrovato, G. GIS and statistical analysis for landslide susceptibility mapping in the daunia area (Italy). *Nat. Hazards Earth Syst. Sci.* **2010**, *10*, 1851–1864. [CrossRef]
87. Cervi, F.; Berti, M.; Borgatti, L.; Ronchetti, F.; Manenti, F.; Corsini, A. Comparing predictive capability of statistical and deterministic methods for landslide susceptibility mapping: A case study in the northern apennines (reggio emilia province, Italy). *Landslides* **2010**, *7*, 433–444. [CrossRef]

88. Lee, S.; Oh, H.-J. Ensemble-based landslide susceptibility maps in jinbu area, korea. In *Terrigenous Mass Movements*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 193–220.
89. Ozdemir, A. 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.* **2011**, *411*, 290–308. [CrossRef]
90. Bi, H.; Zhang, J.; Zhu, J.; Lin, L.; Guo, C.; Ren, Y.; Yun, L.; Ma, N. Spatial dynamics of soil moisture in a complex terrain in the semi-arid loess plateau region, china 1. *JAWRA J. Am. Water Resour. Assoc.* **2008**, *44*, 1121–1131. [CrossRef]
91. Worthington, S.R. Diagnostic hydrogeologic characteristics of a karst aquifer (kentucky, USA). *Hydrogeol. J.* **2009**, *17*, 1665. [CrossRef]
92. Ford, D.; Williams, P.D. *Karst Hydrogeology and Geomorphology*; John Wiley & Sons: New York, NY, USA, 2013.
93. Kazakis, N.; Kougias, I.; Patsialis, T. Assessment of flood hazard areas at a regional scale using an index-based approach and analytical hierarchy process: Application in rhodope–evros region, greece. *Sci. Total Environ.* **2015**, *538*, 555–563. [CrossRef]
94. Ghayoumian, J.; Saravi, M.M.; Feiznia, S.; Nouri, B.; Malekian, A. Application of gis techniques to determine areas most suitable for artificial groundwater recharge in a coastal aquifer in southern iran. *J. Asian Earth Sci.* **2007**, *30*, 364–374. [CrossRef]
95. Ehlers, L.; Herrmann, F.; Blaschek, M.; Duttmann, R.; Wendland, F. Sensitivity of mgrowa-simulated groundwater recharge to changes in soil and land use parameters in a mediterranean environment and conclusions in view of ensemble-based climate impact simulations. *Sci. Total Environ.* **2016**, *543*, 937–951. [CrossRef] [PubMed]
96. Arabameri, A.; Pradhan, B.; Rezaei, K.; Yamani, M.; Pourghasemi, H.R.; Lombardo, L. Spatial modelling of gully erosion using evidential belief function, logistic regression, and a new ensemble of evidential belief function–logistic regression algorithm. *Land Degrad. Dev.* **2018**, *29*, 4035–4049. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).