

Article

Groundwater-Potential Mapping Using a Self-Learning Bayesian Network Model: A Comparison among Metaheuristic Algorithms

Sadegh Karimi-Rizvandi ¹, Hamid Valipoori Goodarzi ², Javad Hatami Afkoeieh ³, Il-Moon Chung ⁴,
Ozgun Kisi ⁵, Sungwon Kim ⁶ and Nguyen Thi Thuy Linh ^{7,8,*}

¹ Department of Geomatics Engineering, K. N. Toosi University of Technology, Tehran 19697, Iran; Sadeghk2gis@yahoo.com

² Department of Mining Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran; Hamid.valipoorigoodarzi@gmail.com

³ Control and Information Processing from Department of Mechanics and Mechatronics, Academy of Engineering, Peoples' Friendship University of Russia (RUDN University), 117198 Moscow, Russia; khatamiafkuiekh-d@rudn.ru

⁴ Department of Land, Water and Environmental Research, Korea Institute of Civil Engineering and Building Technology, Goyang 10223, Korea; imchung@kict.re.kr

⁵ Civil Engineering Department, Ilia State University, Tbilisi 0162, Georgia; ozgur.kisi@iliauni.edu.ge

⁶ Department of Railroad Construction and Safety Engineering, Dongyang University, Yeongju 36040, Korea; swkim1968@dyu.ac.kr

⁷ Institute of Research and Development, Duy Tan University, Danang 550000, Vietnam

⁸ Faculty of Environmental and Chemical Engineering, Duy Tan University, Danang 550000, Vietnam

* Correspondence: nguyentthuylinh58@duytan.edu.vn



Citation: Karimi-Rizvandi, S.; Goodarzi, H.V.; Afkoeieh, J.H.; Chung, I.-M.; Kisi, O.; Kim, S.; Linh, N.T.T. Groundwater-Potential Mapping Using a Self-Learning Bayesian Network Model: A Comparison among Metaheuristic Algorithms. *Water* **2021**, *13*, 658. <https://doi.org/10.3390/w13050658>

Academic Editor: Fi-John Chang

Received: 27 December 2020

Accepted: 12 February 2021

Published: 28 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Owing to the reduction of surface-water resources and frequent droughts, the exploitation of groundwater resources has faced critical challenges. For optimal management of these valuable resources, careful studies of groundwater potential status are essential. The main goal of this study was to determine the optimal network structure of a Bayesian network (BayesNet) machine-learning model using three metaheuristic optimization algorithms—a genetic algorithm (GA), a simulated annealing (SA) algorithm, and a Tabu search (TS) algorithm—to prepare groundwater-potential maps. The methodology was applied to the town of Baghmalek in the Khuzestan province of Iran. For modeling, the location of 187 springs in the study area and 13 parameters (altitude, slope angle, slope aspect, plan curvature, profile curvature, topography wetness index (TWI), distance to river, distance to fault, drainage density, rainfall, land use/cover, lithology, and soil) affecting the potential of groundwater were provided. In addition, the statistical method of certainty factor (CF) was utilized to determine the input weight of the hybrid models. The results of the OneR technique showed that the parameters of altitude, lithology, and drainage density were more important for the potential of groundwater compared to the other parameters. The results of groundwater-potential mapping (GPM) employing the receiver operating characteristic (ROC) area under the curve (AUC) showed an estimation accuracy of 0.830, 0.818, 0.810, and 0.792, for the BayesNet-GA, BayesNet-SA, BayesNet-TS, and BayesNet models, respectively. The BayesNet-GA model improved the GPM estimation accuracy of the BayesNet-SA (4.6% and 7.5%) and BayesNet-TS (21.8% and 17.5%) models with respect to the root mean square error (RMSE) and mean absolute error (MAE), respectively. Based on metric indices, the GA provides a higher capability than the SA and TS algorithms for optimizing the BayesNet model in determining the GPM.

Keywords: groundwater-potential mapping; Bayesian network model; metaheuristic algorithms; geographic information system (GIS); receiver operating characteristic; area under the curve

1. Introduction

Rapid population growth and industrial development in most countries will pose a shortage of freshwater by 2025 [1]. Groundwater is one of the natural resources that affects biosphere and geosphere processes [2], and is one of the most considerable freshwater resources in the world; the daily needs of 2.5 billion people depend on these resources [3]. Groundwater resources are also considered the primary water resources in arid and semi-arid regions [4]. Globally, the use of groundwater is about 20%; it is a very efficient source for drinking water, forestry, industry, livestock, and agriculture. It also is less exposed to environmental contamination than surface water [5,6]. In many parts of the world, especially Iran, the uncontrolled abstraction of groundwater has led to the depletion of these resources [7]. Increased demand for groundwater resources around the world is due to components such as insufficient rainfall, surface-water scarcity, and population growth [8]. Improper management of groundwater resources reduces their quality, which will lead to the deterioration of water resources [9]. Iran is facing a shortage of fresh water due to desertification of about two-thirds of its lands, which are free of any forests and green pastures. This has led to the high use of groundwater [10].

The existence of groundwater is dynamic from place to place and highly variable. The occurrence of groundwater in a given area is not accidental, and is due to the interaction of many causes, such as natural geography, climate, hydrology, geology, ecology, topography, type of soil and underlying soil layers, fracture density, slope of the ground, and land use [11,12]. Given that these factors are interdependent, the reliability of the estimates for a given area is diminished by considering a single factor to describe the recharge process. Four assumptions were made to guide the selection of variables that mainly affect aquifer recharge: groundwater potential increases with (1) increasing groundwater recharge (precipitation level is capable of infiltrating), (2) higher soil and rock permeability (geological and lithological units), (3) higher lineament density (geological structures), and (4) flat slopes [12]. To manage groundwater resources effectively, it is essential to identify the areas with high groundwater storage potential. One of the most effective approaches that can help administrators to adopt management plans is to map the potential of groundwater [13,14]. Groundwater-potential mapping (GPM) includes the traditional approaches: geographic information systems (GISs) and remote sensing (RS). Despite their accurate estimation of the potential of groundwater, the traditional methods, due to time and cost, as well as the unavailability of data and technology for extraction in many parts of the world, are a barrier to its use [13]. The GIS-based methods provide the ability to store, manipulate, and analyze data in various formats and scales, which ultimately lead to the production of thematic maps [15–19]. Statistical, machine-learning, and hybrid models integrated with GISs have been used over the last decade to achieve more accurate models for assessing the potential of groundwater [20]. Statistical models also were used for groundwater-potential mapping, including frequency ratio (FR) [21], logistic regression (LR) [22], Shannon entropy (SE) [23], weight of evidence (WOE) [24], and the evidential belief function (EBF) [25].

GIS-based groundwater-potential mapping based on machine-learning models has become very attractive among researchers recently. The selection of suitable parameters that affect the potential of groundwater and appropriate models for identifying its potential areas is one of the most valuable criteria for evaluating the potential of groundwater. Since these models can predict the nonlinear relationship effectively, the diverse variables affect the potential of groundwater [26]. In groundwater-potential mapping, the machine-learning models used so far are classified as random forest (RF) [27], support vector machine (SVM) [28], boosted regression trees (BRT) [29], classification and regression trees (CART) [30], linear discriminant analysis (LDA) [31], multivariate adaptive regression spline (MARS) [32], and naïve Bayes (NB) [33], respectively.

Despite the high accuracy of machine-learning models in predicting nonlinear data, the overfitting problem is one of the drawbacks that affects the accuracy of the model [34]. In this regard, to achieve higher accuracy in preparing the groundwater-potential mapping, different hybrid models obtained by optimizing the base model with different algo-

gorithms [34,35] or group techniques [36–38] have recently been used. The Bayesian network (BayesNet) model has been successfully used in hazard modeling so far [39,40], but one of the disadvantages of this model is the inability to determine the optimal network structure.

The main objective of this research is to develop GPM using a BayesNet model combined with genetic algorithm (GA), simulated annealing (SA), and Tabu search (TS) algorithms to determine the optimal network structure of the BayesNet in the Baghmalek Plain, Iran. Second, the study aims to compare the prediction capabilities of the different metaheuristic algorithms. For this purpose, the GA, SA, and TS algorithms have been used to optimize the network structure of the BayesNet model to prepare a groundwater-potential mapping. To our knowledge, the BayesNet-GA, BayesNet-SA, and BayesNet-TS models have not been investigated to address this issue yet.

2. Materials and Methods

This research was accomplished in five levels. In the first level, a spatial database was created, consisting of the locations of the springs and the criteria affecting the potential of groundwater. In the second level, the important criteria affecting the potential of groundwater was determined using the OneR technique. In the third level, the spatial relationship between the effective criteria and locations of the springs was achieved using the certainty factor (CF) method. In the fourth level, spatial modeling was performed using ensembles of the BayesNet model and metaheuristic (i.e., GA, SA, and TS) algorithms. In the last level, the GPM was evaluated using metric indicators. A flowchart of our research is shown in Figure 1.

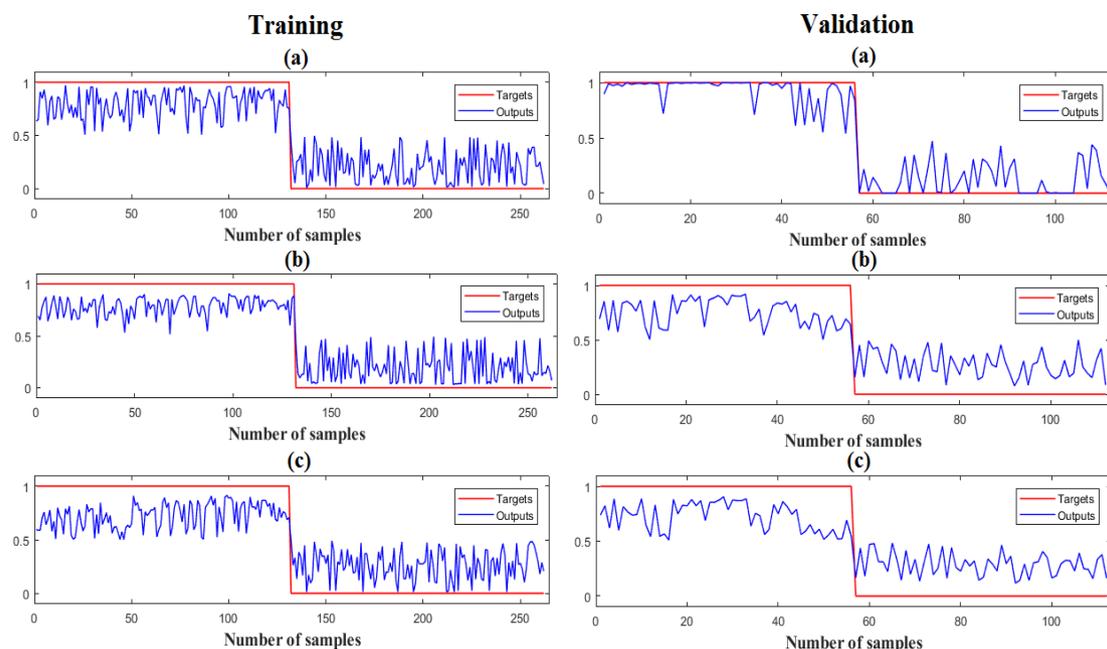


Figure 1. Research methodology. TWI: topographic wetness index, GA: genetic algorithm, SA: simulated annealing, TS: Tabu search, RMSE: root mean square error, MAE: mean absolute error, ROC: receiver operating characteristic, AUC: area under the curve.

2.1. Study Area

The town of Baghmalek is in the east of Khuzestan Province, Iran. It is geographically located at $49^{\circ} 53' E$ and $31^{\circ} 31' N$, at an altitude of 917 m above sea level. The temperature in the center of the study area in summer reaches a high of $42^{\circ} C$, and in winter reaches a low of $0^{\circ} C$. The average annual rainfall is 500 mm. Due to the high use of groundwater in this area for agriculture, in recent years there has been a clear decrease in groundwater levels. The location of the study area, along with the distribution of the springs, is shown in Figure 2.

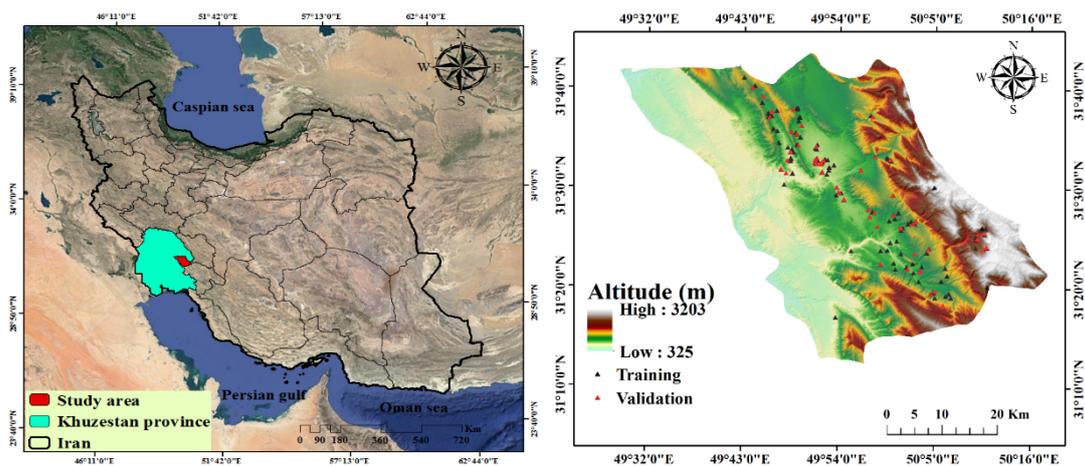


Figure 2. Study area with spring locations.

2.2. Spring Inventory Map

Our general assumption was that the presence of groundwater in regions with high groundwater yields was highly likely [26]. A total of 187 springs were chosen, with a yield greater than $11 \text{ m}^3/\text{h}$ and a mean pH and electrical conductivity (EC) of 7.1 and $490 \text{ }\mu\text{mhos}/\text{cm}$, respectively. In this study, the locations of the springs were provided by the Water Resources Management Organization (WRMO) of Iran. Google Earth was used to verify the locations of the springs. Of the locations, 70% (131 locations) were randomly used for modeling (i.e., training), and 30% (56 locations) were randomly used for the modeling evaluation (i.e., validation).

2.3. Criteria Affecting the Potential of Groundwater

The topography, meteorology, hydrology, human activities, and environmental features from the earth are major criteria affecting the potential of groundwater among the different elements [13,22]. To determine the GPM, 13 factors were selected in our research. Considering the previous articles, these criteria included: altitude, slope angle, slope aspect, plan curvature, profile curvature, topography wetness index (TWI), distance to river, distance to fault, drainage density, rainfall, land use/cover, lithology, and soil [26,35]. ArcGIS 10.3 and SAGA GIS software were used to prepare and classify the effective criteria. The following is a description of each of the effective criteria. The method for preparing and classifying the effective criteria is summarized in Appendix A Table A1.

Altitude: Because of differences in soil and vegetation caused by altitude differences, groundwater is usually found at low altitudes. The shuttle radar topography mission (STRM) digital elevation model (DEM) was downloaded from <https://earthexplorer.usgs.gov> (accessed on 28 February 2021) at a spatial resolution of $30 \text{ m} \times 30 \text{ m}$. Topographic parameters were obtained from the DEM in ArcGIS10.3 and SAGA GIS 2.1.2 software, including altitude, profile curvature, slope aspect, slope angle, and plan curvature (Figure 3a) [41].

Slope angle: The slope angle affects vertical water infiltration and surface runoff velocity. In areas with lower slopes, runoff recharge is also higher (Figure 3b) [42].

Slope aspect: The amount of water in the soil is affected by the slope aspect, and the physiographic trends and the direction of precipitation are related to the slope aspect (Figure 3c) [22].

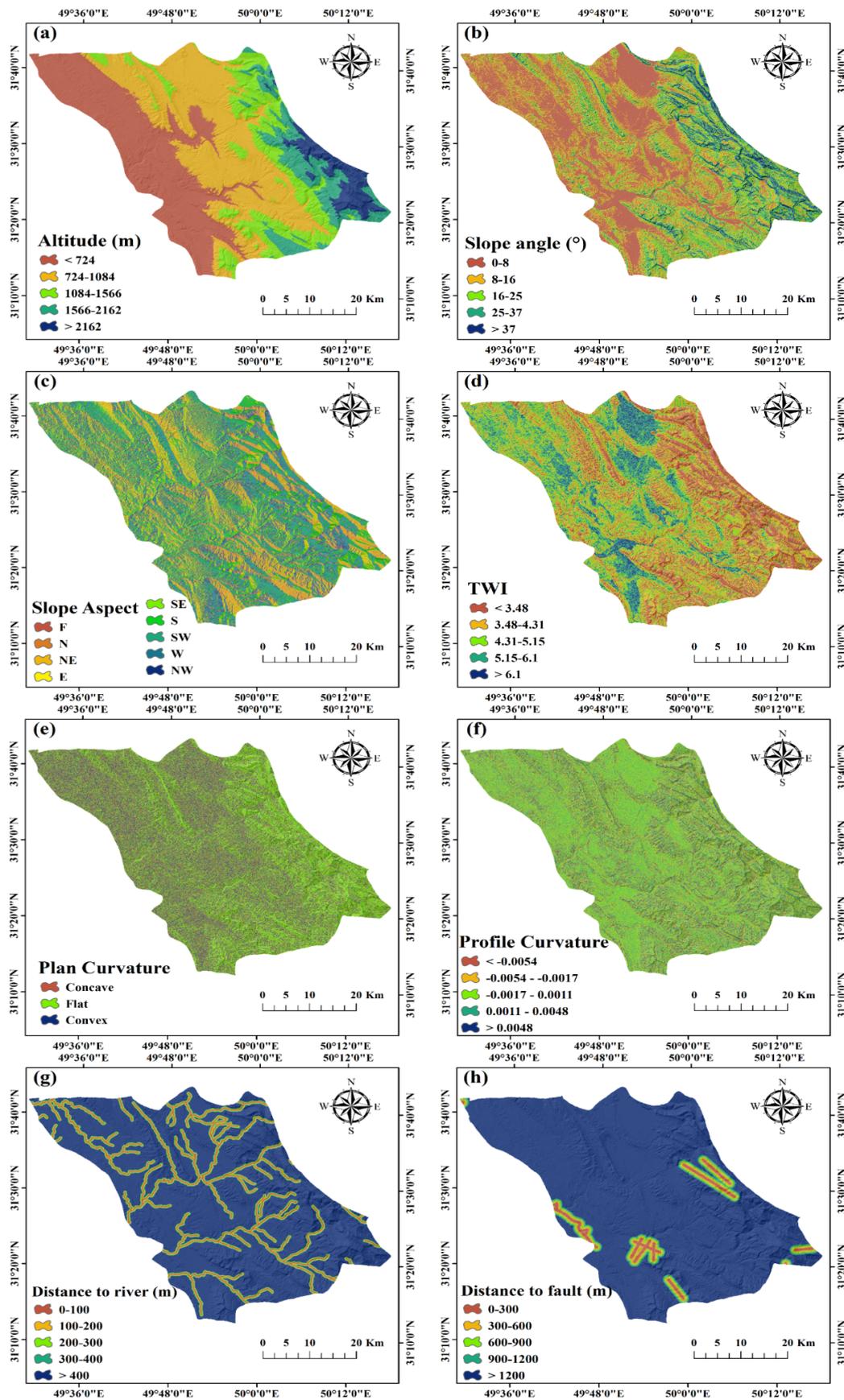


Figure 3. Cont.

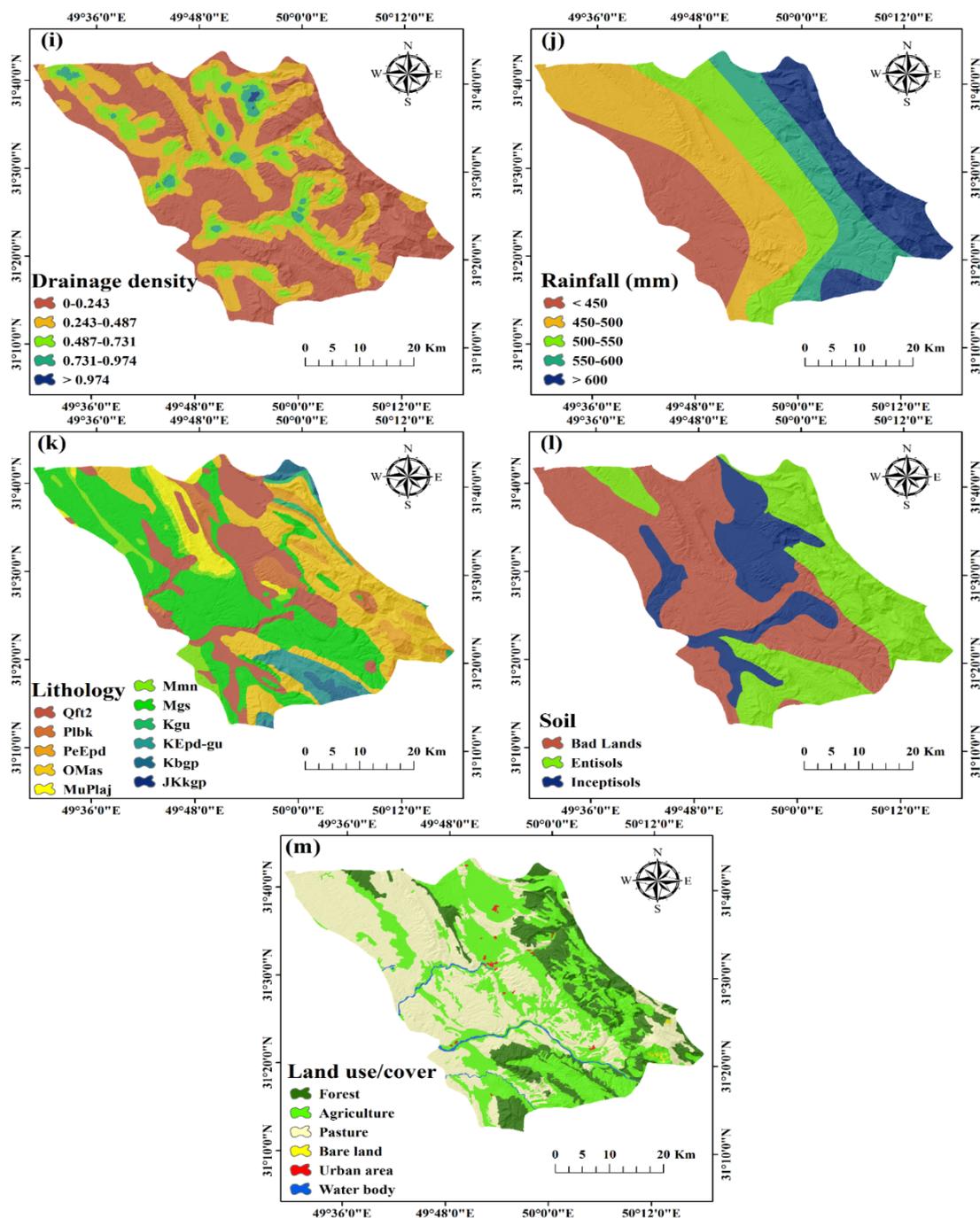


Figure 3. Maps of affecting factors: (a) altitude, (b) slope angle, (c) slope aspect, (d) topographic wetness index (TWI), (e) plan curvature, (f) profile curvature, (g) distance to river, (h) distance to fault, (i) drainage density, (j) rainfall, (k) lithology, (l) soil and (m) land use/cover.

TWI: The TWI shows the location and size of saturated source areas, and the production of surface runoff is affected by topographic/morphological conditions. The TWI is calculated using Equation (1) (Figure 3d) [43].

$$TWI = \ln\left(\frac{A_s}{\tan \alpha}\right) \tag{1}$$

where A_s . is the specific surface of the catchment and α is the slope gradient.

Plan and profile curvature: This controls the speed of water flow. The values of the profile curvature represent the morphology of the topography, with a positive curvature attesting to an upward concave curvature, a negative curvature attesting to an upward convex curvature, and a value of zero attesting to a flat surface. For a longer period of rainfall, a concave slope retains more water (Figure 3e,f) [26].

Distance to river: This is the primary source of groundwater recharge in semiarid regions that affects the potential of groundwater (Figure 3g) [42].

Distance to fault: The movements of groundwater springs over an area can be controlled by various faults (Figure 3h) [44].

Drainage density: The drainage characteristics of each basin control the subsurface hydrological status of each area. The drainage density varies in different regions due to land use effects, different climatic regimes, and natural landscapes (Figure 3i) [45].

Rainfall: One of the most essential and natural sources of groundwater recharge that impacts the infiltration for the potential of groundwater is the amount of rainfall. A 30-year average rainfall dataset for 22 meteorological stations obtained by the WRMO was used to prepare the rainfall map. The rainfall map was constructed in ArcGIS 10.3 by using the kriging interpolation method (Figure 3j) [17].

Lithology: Permeability and porosity are two criteria that affect the movement and occurrence of groundwater as its efficiencies increase with growing permeability and porosity. Lithology plays a crucial role in permeability and porosity (Figure 3k) [46].

Soil: Surface-water permeability is controlled by soil properties and is directly related to permeability (Figure 3l) [47].

Land use/cover: Groundwater movements are affected by diversity in land use/cover. Evapotranspiration, surface runoff, and groundwater recharge are affected by land use/cover. For the preparation of the land use/cover map, Landsat-7 images from 2013 and 2019 were used. To this purpose, 420 training points obtained via a Global Positioning System (GPS) were used to create a land use/cover map, of which 70% were used for training, and the remaining 30% were used to test the accuracy of the map. The land use/cover maps were produced at an accuracy of 90% using the maximum likelihood algorithm in ENVI 4.8 software (Figure 3m) [48].

2.4. CF Method

To solve for uncertainty and heterogeneity of the input data for modeling, the CF method was implemented. This method was developed by Shortliffe and Buchanan [49]. The weight of each category of criteria that affected the potential of groundwater was calculated using Equation (2) [49].

$$\left\{ \begin{array}{l} cf = \frac{pp_a - pp_s}{pp_a(1 - pp_s)} \text{ if } pp_a > pp_s \\ cf = \frac{pp_a - pp_s}{pp_s(1 - pp_a)} \text{ if } pp_a < pp_s \end{array} \right. \quad (2)$$

where cf is the weight of each category, pp_a is the conditional probability of the existing number of springs in a category, and pp_s is the initial probability of the existence of all springs in study area. The favorability values (pp_a , pp_s) were obtained from overlaying the existing spring distribution layer in GIS with each data layer and measuring the frequency of occurrence of spring. The value of pp_a is given by the ratio of the spring area falling into a certain category and the total area of that category. The value of pp_s can be determined by dividing the total area of the spring with the total area of the study. The value in the CF method is between -1 and $+1$. Positive values indicate an increase in certainty, and negative values indicate a decrease in certainty. In this research, the CF method was used in order to match and take into account the effect of different classes of independent parameters in the modeling [49].

2.5. Bayesian Network (BayesNet) Machine-Learning Model

The BayesNet machine-learning model is a knowledge representation tool that enables effective management of dependency/independence relationships between the random variables that make up the modeling-problem domain [50]. This model consists of two components, structure and parameter. The structure of the BayesNet machine-learning model, $\vec{V} = (\vec{V}, E)$, is a directional graph with no rotation where $\vec{V} = \{x_1, x_2, \dots, x_n\}$ is a set of nodes that represent system variables and $E \subset (\vec{V} \times \vec{V})$ is a collection of edges. It describes the relationship of direct dependence between these variables [39]. For each variable $\vec{V} x_i \in$ is a set of conditional distributions of $p(x_i | pa(x_i))$ that $pa(x_i)$ represents the set of parents of the variable x_i . The cumulative distribution on \vec{V} can be calculated using Equation (3) [18,50].

$$p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p(x_i | pa(x_i)) \quad (3)$$

2.6. Genetic Algorithm (GA)

A GA that mimics biological evolution can be used to solve optimization problems with and without restrictions. This algorithm was introduced by Holland in 1975 as one of the first metaheuristic algorithms [51,52]. In this algorithm, each individual solution is called a chromosome, and consists of single arrays called genes. This algorithm repeatedly modifies a population of solutions. At each step, the GA randomly selects individual solutions from the present population and uses them as parent chromosomes to create offspring chromosomes for the new generation. Genes are transformed from current chromosomes into new generation chromosomes using operators. After creating successive generations, the population moves toward an optimal solution [53,54].

2.7. Simulated Annealing (SA) Algorithm

The SA algorithm is a simple and effective metaheuristic search used to solve hybrid optimization problems [55]. The SA algorithm simulates a gradual refrigeration process to solve an optimization problem. The objective function of such a problem is similar to the energy of a substance, which must be minimized by defining a virtual temperature. The SA algorithm is a probabilistic algorithm in which a solution is proposed to exit the local optimization. Also, it is a memoryless algorithm in the sense that there is no mechanism for storing information during the search [56,57].

2.8. Tabu Search (TS) Algorithm

The TS algorithm is a direct search algorithm for optimizing complex nonlinear problems. In this algorithm, the transition from a possible initial solution to a possible secondary solution is called a move [58]. The list of the last points to be examined is called the taboo list, which is proposed based on human memory. The purpose of addressed list is to prevent relocation to areas that have already been surveyed. This can expand the algorithm search area. First, the algorithm randomly selects an initial answer, then all neighboring points are identified, and the value of objective function is calculated [59].

2.9. Hybrid Model

One of the main components in a BayesNet machine-learning model is the structure of the network. The traditional methods for determining the best network structure in their search space can acquire local optimization and ignore global optimization [60]. For this purpose, metaheuristic algorithms are an effective solution to determine the best network structure. In our research, to optimize the structure of the BayesNet machine-learning model, metaheuristic (i.e., GA, SA, and TS) algorithms were used. The objective

function of optimizing the BayesNet structure (B_s) for the D dataset using metaheuristic algorithms is defined in Equation (4):

$$Q_{\text{Bayes}}(B_s, D) = P(B_s) \prod_{i=0}^n \prod_{j=1}^{q_i} \frac{\Gamma(N'_{ij})}{\Gamma(N'_{ij} + N_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(N'_{ijk} + N_{ijk})}{\Gamma(N'_{ijk})} \quad (4)$$

where $P(B_s)$ is the previous network structure, $\Gamma(\cdot)$ is the gamma function, and N'_{ij} and N'_{ijk} are Dirichlet distribution indices for a set of parameters. The hybrid model was applied using the Waikato Environment for Knowledge Analysis (WEKA 3) software. To implement the models, the data was prepared in ArcGIS 10.3 software and then transferred to the WEKA 3 software for modeling. Finally, the output of WEKA 3 software was transferred to ArcGIS 10.3, and the final groundwater potential map was prepared (Figure 4).

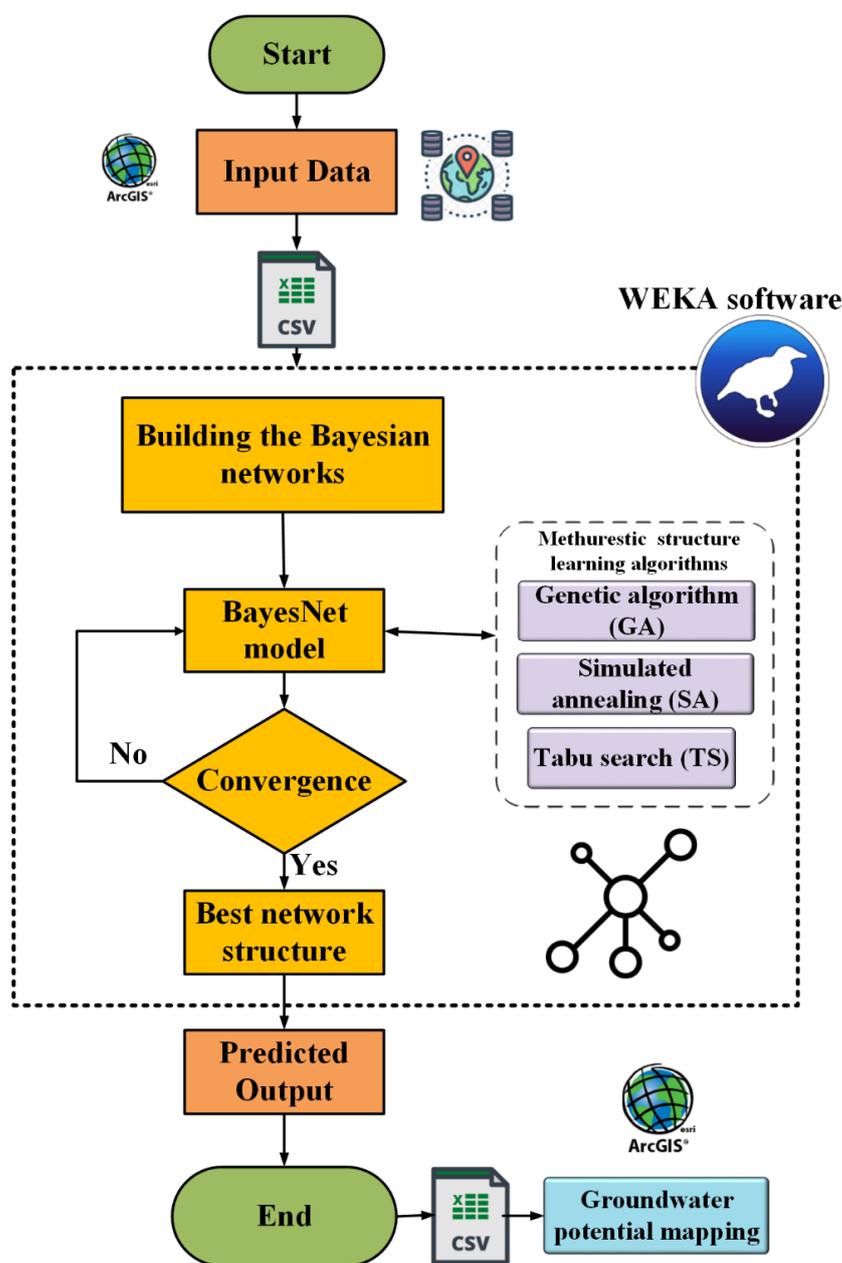


Figure 4. Flowchart of the hybrid model.

2.10. OneR Technique

In this study, the OneR technique was used to determine the feature selection and select the most effective criterion in modeling the groundwater-potential map [61]. The OneR technique is a one-tier decision tree that contains a set of rules. It is simple and often provides good rules for characterizing data structures. It also uses the computational error ratio and several rules to gain the weight of each effective criterion [62].

2.11. Validation Indices

To evaluate the hybrid models, Precision, Recall, Specificity, Sensitivity, Kappa, receiver operating characteristic (ROC)–area under the curve (AUC), root mean square error (RMSE), and mean absolute error (MAE) indices were used. The ROC curve is a graph in which the occurrence or nonoccurrence of springs correctly predicted by the model is plotted on the horizontal axis (1-Specificity) versus incorrectly predicted pixels (Sensitivity) on the vertical axis. The area below the ROC curve is called AUC. Its value varies between 0.5 and 1; the closer it is to one, the higher the modeling efficiency is. These indicators were calculated using Equations (5)–(11) [61,63].

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (7)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

$$\text{Kappa} = \frac{(\text{TP} + \text{TN}) - ((\text{TP} + \text{FN})(\text{TP} + \text{FP}) + (\text{FP} + \text{TN})(\text{FN} + \text{TN}))}{1 - ((\text{TP} + \text{FN})(\text{TP} + \text{FP}) + (\text{FP} + \text{TN})(\text{FN} + \text{TN}))} \quad (9)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (Y - Y')^2}{n}} \quad (10)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y - Y'| \quad (11)$$

where TP is the number of spring values that were correctly predicted, FP is the number of spring values that were incorrectly detected, TN is the number of spring values that were rejected correctly, FN is the number of spring values that were incorrectly rejected, Y is the real value, Y' is the predicted value, and N is the number of samples. The RMSE and MAE indices were used to measure the error between the actual and predicted values [64,65].

3. Results and Discussion

3.1. Determining the Important Criteria

The importance we determined for the criteria that affect the potential of groundwater is shown in Figure 5. Based on the results of the OneR technique, the addressed criteria of altitude (64.12), lithology (62.59), drainage density (62.21), TWI (62.1), rainfall (61.06), distance to river (59.54), plan curvature (59.16), profile curvature (57.25), land use/cover (55.72), distance to fault (54.88), slope angle (51.145), slope aspect (50.76), and soil (48.09) were determined for the potential of groundwater in the study area. Based on the results of OneR technique, the criteria of altitude, lithology, and drainage density had the most significant impact on the occurrence of groundwater in study area, which is consistent with previous articles [66,67].

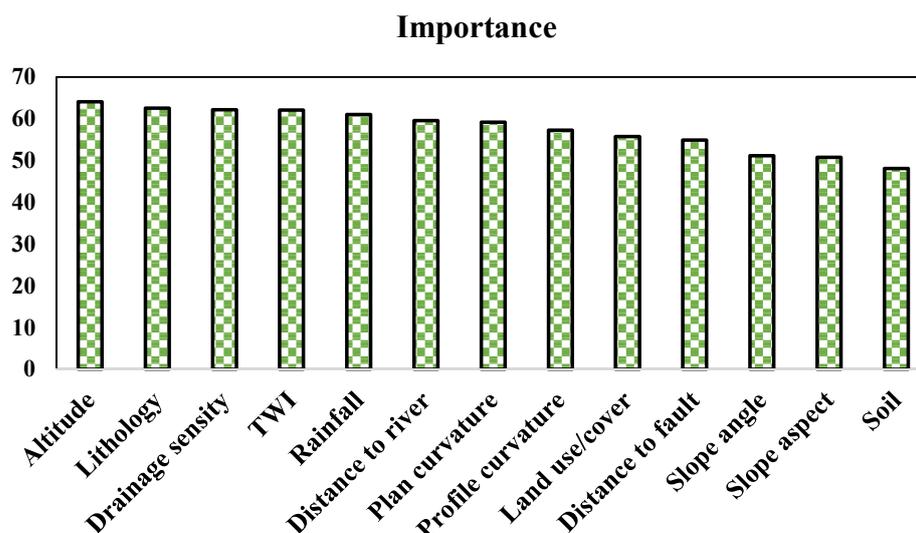


Figure 5. The importance of the effective criteria determined using the OneR technique.

3.2. Results of the CF Method

The spatial relationship using the CF method between the criteria affecting the potential of groundwater and the location of springs is presented in Table 1. Based on the results of altitude criterion, the class of 724–1084 m provided the highest correlation with the presence of groundwater (0.355). As the altitude increased, the potential of groundwater decreased. Also, the reason for this phenomena can be expressed as the reduction of aquifer thickness at high altitudes [26]. Considering the results of slope angle criterion, the slope angle class of less than 8° supplied the highest weight (0.220). The potential of groundwater decreased with an increasing slope. The main reason for this behavior is that the nature of water tends to descend from points at high altitudes to lower altitudes, and also to move in the direction of the slope and accumulate in the lowlands [14]. In the results for the slope aspect criterion, the south direction with a weight of 0.392 supported the highest importance. Due to receiving more rainfall and humidity than other areas, it played a more significant role in controlling the formation of groundwater [41]. Regarding the results of plan curvature criterion, the concave class yielded the highest weight and the most significant impact on the occurrence of groundwater (CF = 0.427) [23]. Recognizing the results of profile curvature criterion, the class of less than -0.0054 produced the highest weight, equal to 0.608. With an increasing TWI, the speed of water passage from high altitudes will decrease, and water will have more opportunity to penetrate the ground. Therefore, the potential of groundwater will increase [68]. But in this study, the middle classes of TWI had a greater impact on the occurrence of groundwater. One of the reasons for the negative weight of the category larger than 6.1 is that a small number of springs (5) are located in this category. The results show that values between 4 and 6 for this index are more important in the occurrence of groundwater, which is consistent with [4]. For the results of distance to river, the class of 0–100 m yielded the highest weight (CF = 0.744) [69]. For distance to fault, the class of 0–300 m reported the highest value (CF = 0.474). Faults can transfer groundwater in hard formations. The presence of a fault facilitates infiltration conditions. Therefore, at distances closer to the fault, the potential of groundwater is higher [14]. In the results for the drainage density criterion, the maximum weight was related to the class of 0.487–0.731, with a value of 0.618. As the drainage density increased, the area's permeability increased, and thus provided suitable conditions for groundwater recharge [16]. For the results of rainfall criterion, the class of 550–500 mm had the highest weight (CF = 0.490). With increasing rainfall, the probability of groundwater occurring increased. Since the study area has a semiarid climate, the average rainfall of 500 mm had a more significant impact on the occurrence of groundwater. In the study area, more rainfall occurred in areas with a high altitude and slope, which are not prone to groundwater occurrence;

therefore, the rainfall criterion was affected by these two criteria, and high amounts of rainfall (>550 mm) had negative weight. Regarding the results for the lithology criterion, the MuPlaj unit showed the highest weight (CF = 0.639). The MuPlaj unit, owing to its high permeability, had the most significant impact on the potential of groundwater in the study area [14]. Contemplating the results of the land use/cover criterion, the agricultural class provided the highest weight (CF = 0.347) and the most significant impact on the occurrence of groundwater. Agricultural irrigation water, mainly supplied from river water, penetrates the saturated aquifer and feeds the aquifer during irrigation [70]. Finally, for the soil criterion, the Inceptisols class suggested the highest weight (CF = 0.370). The reason for this is that the increased permeability of Inceptisols provides the necessary conditions for the formation of aquifers in areas covered with this type of soil [26].

Table 1. Results of the CF model.

Class	No. of Pixels in Domain	No. of Springs	CF	Class	No. of Pixels in Domain	No. of Springs	CF
Altitude (m)				Distance to river (m)			
<724	864,512	39	−0.076	0–100	151,724	29	0.744
724–1084	962,838	73	0.355	100–200	137,908	22	0.693
1084–1566	432,096	4	−0.810	200–300	142,610	7	0.0031
1566–2162	271,106	14	0.053	300–400	126,125	3	−0.513
>2162	150,608	1	−0.864	>400	211,880	70	−0.324
Slope angle				Distance to fault (m)			
0–8	973,112	61	0.220	0–300	53,716	5	0.474
8–16	691,532	38	0.110	300–600	52,757	1	−0.612
16–25	528,894	18	−0.303	600–900	51,893	0	−1.000
25–37	349,366	11	−0.355	900–1200	53,661	0	−1.000
>37	138,256	3	−0.555	>1200	246,514	125	0.035
TWI				Drainage density			
<3.48	655,123	12	−0.625	0–0.243	130,811	27	−0.570
3.48–4.31	775,835	28	−0.261	0.243–0.487	936,236	51	0.101
4.31–5.15	645,477	51	0.381	0.487–0.731	366,305	47	0.618
5.15–6.1	457,665	35	0.361	0.731–0.974	62,963	6	0.486
>6.1	147,060	5	−0.304	>0.974	3556	0	−1.000
Profile curvature				Rainfall (mm)			
<−0.0054	160,359	20	0.608	< 450	466,347	0	−1.000
−0.0054–0.0017	601,468	35	0.160	450–500	818,925	49	0.180
−0.0017–0.0011	101,300	46	−0.070	500–550	497,838	48	0.490
0.0011–0.0048	716,132	27	−0.220	550–600	426,536	18	−0.137
>0.0048	190,193	3	−0.670	>600	468,081	16	−0.301
Plan curvature				Soil			
Concave	644,059	55	0.427	Bad lands	123,049	61	0.014
Flat	136,304	60	−0.099	Inceptisols	579,991	45	0.370
Convex	674,060	16	−0.514	Entisols	871,261	25	−0.410
Lithology				Slope aspect			
Qft2	484,268	43	0.449	F	1623	0	−1.000
MuPlaj	118,043	16	0.639	N	312,384	12	−0.213
Mgs	953,888	32	−0.313	NE	324,921	10	−0.370
Kgu	35,391	0	−1.000	E	242,263	11	−0.070
OMas	598,244	14	−0.520	SE	243,474	9	−0.243
Kbgp	53,728	0	−1.000	S	373,092	30	0.392
Mmn	1,07201	3	−0.420	SW	482,906	24	0.016
Plbk	114,957	12	0.530	W	392,328	21	0.087
PeEpd	114,465	11	0.490	NW	308,169	14	−0.070
JKkgp	399	0	−1.000				
KEpd-gu	101,159	0	−1.000				

Table 1. Cont.

Class	No. of Pixels in Domain	No. of Springs	CF	Class	No. of Pixels in Domain	No. of Springs	CF
Land use/cover							
Forest	468,449	18	−0.213				
Agriculture	117,252	75	0.347				
Pasture	100,168	36	−0.371				
Bare land	7285	0	−1.000				
Urban area	4071	0	−1.000				
Water body	27,672	2	0.324				

3.3. Results of the Hybrid Models

To optimize the BayesNet machine-learning model, a spatial database provided the points of presence of spring (value 1) and the absence of spring (value 0). The weights were obtained from all criteria by the CF method. Nonoccurrence points were randomly created as equal to the training and validation points. The control parameters in the metaheuristic algorithms are presented in Table 2. The diagrams for the comparison of training and validation data with target data (occurrence and nonoccurrence of springs) are shown in Figure A1. The evaluation results of the hybrid models using metric indicators for the training and validation data are summarized in Table 3. As observed in Table 3, the BayesNet-GA model provided the best approximation in training with the lowest RMSE (0.2633) and MAE (0.2178), and the highest Kappa (0.8850), Precision (0.9470), Recall (0.9430), and ROC–AUC (0.9870), followed by the BayesNet-SA model. Furthermore, the BayesNet-GA model supplying the lowest RMSE (0.334) and MAE (0.288), and the highest Kappa (0.7500), Precision (0.8770), Recall (0.8750), and ROC–AUC (0.9480) in validation performed superior to the BayesNet-SA, BayesNet-TS, and BayesNet models in mapping groundwater. The BayesNet-GA model improved the RMSE and MAE accuracies by 4.6% and 7.5% compared to the second-best model (BayesNet-SA), and these improvements with respect to the BayesNet-TS model were 21.8% and 17.5%, respectively.

Table 2. Parameters of the metaheuristic algorithms.

GA	SA	TS
Iterations = 200	Iterations = 200	Iterations = 200
Population Size = 50	Subiterations = 10	Tabu list = 5
Mutation Rate = 0.2	Initial temperature = 10	Maximum number of parents = 2
	Temperature reduction rate = 0.99	

Table 3. Results for metric indicators.

Index	BayesNet-GA		BayesNet-SA		BayesNet-TS		BayesNet	
	Training	Validation	Training	Validation	Training	Validation	Training	Validation
RMSE	0.2633	0.334	0.2844	0.3502	0.3415	0.427	0.3621	0.44
MAE	0.2178	0.288	0.2238	0.3112	0.297	0.349	0.3123	0.3721
Kappa	0.885	0.75	0.8168	0.7321	0.725	0.5	0.711	0.5
Precision	0.947	0.877	0.909	0.867	0.863	0.787	0.8321	0.7468
Recall	0.943	0.875	0.908	0.866	0.863	0.75	0.8321	0.7212
ROC-AUC	0.987	0.948	0.974	0.939	0.944	0.892	0.9212	0.8732

3.4. GPM Using Hybrid Models

After training (calibrating) the hybrid models, to generalize the hybrid models for the entire study area, the entire area first was converted into points, and the values produced by the hybrid models were prepared for the entire study area. We used ArcGIS 10.3 software to prepare the GPM using hybrid models. The GPM was divided into five categories: very

low potential, low potential, medium potential, high potential, and very high potential, using the natural breaks classification method. The GPM using four models is shown in Figure 6. According to the BayesNet-GA model of GPM, the percentages for the very low, low, medium, high, and very high potential categories were 2%, 24%, 37%, 27%, and 10%, respectively. The GPM using the BayesNet-SA model showed that the percentages for the very low, low, medium, high, and very high potential categories respectively were 18%, 21%, 24%, 24%, and 13%, respectively. In the BayesNet-TS model, the percentages for the very low, low, medium, high, and very high potential categories were 15%, 24%, 25%, 24%, and 12%, respectively. The GPM using the BayesNet model showed percentages for the very low, low, medium, high, and very high potential categories of 20%, 21%, 25%, 20%, and 14%, respectively.

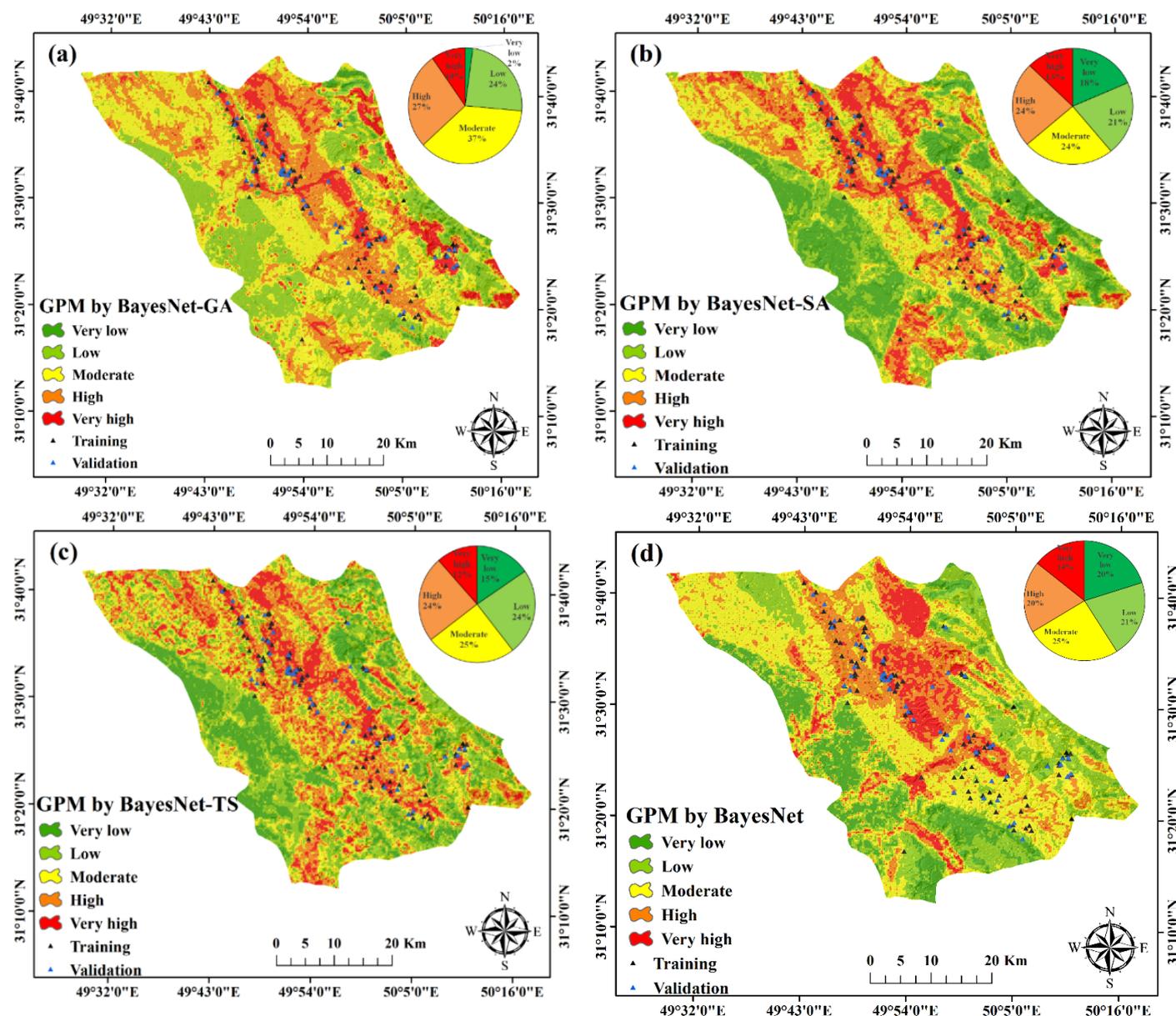


Figure 6. GPM by: (a) BayesNet-GA, (b) BayesNet-SA, (c) BayesNet-TS, (d) BayesNet models.

3.5. Validation of GPM

To evaluate the results using the ROC–AUC, 30% of the occurrence data and the data of the absence were used for the individual springs. The results of the GPM evaluation with

three hybrid models using the ROC–AUC are shown in Figure 7 and Table 4. According to the results, the ROC–AUC's values for the BayesNet-GA, BayesNet-SA, BayesNet-TS, and BayesNet models were 0.830, 0.818, 0.810, and 0.792, respectively. The results showed that the BayesNet-GA model was more accurate than the other three models.

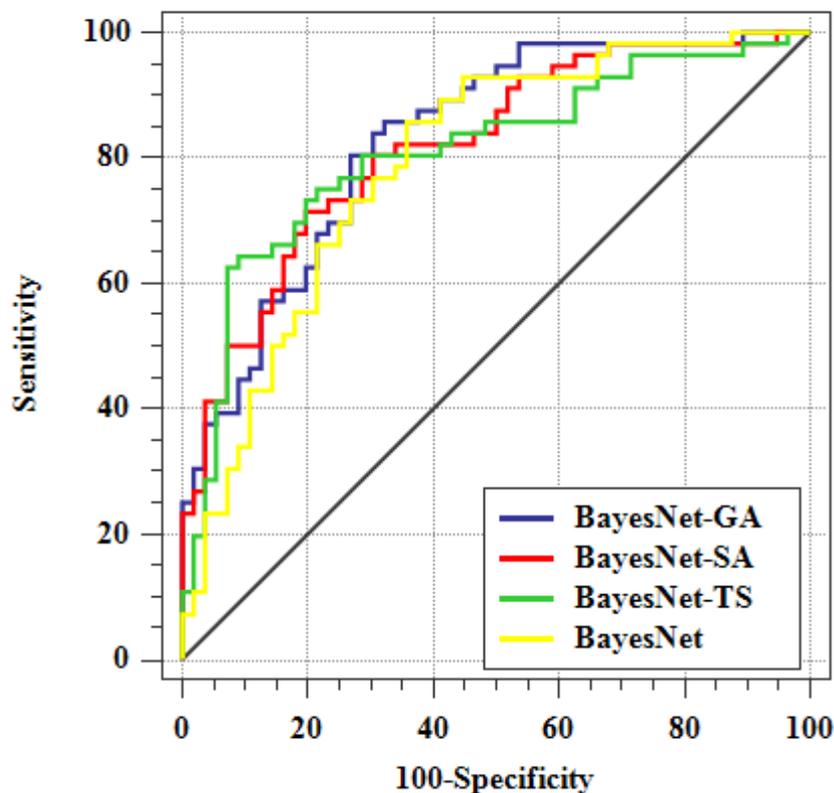


Figure 7. Graph of the ROC–AUC.

Table 4. Results of the ROC–AUC for the hybrid models.

Model	ROC–AUC	SE	95% CI
BayesNet-GA	0.830	0.0378	0.748 to 0.895
BayesNet-SA	0.818	0.0393	0.734 to 0.884
BayesNet-TS	0.810	0.0419	0.725 to 0.878
BayesNet	0.792	0.0428	0.705 to 0.863

The results of the three hybrid models indicated a superior accuracy in preparing the GPM. Therefore, it can be concluded that metaheuristic algorithms (i.e., GA, SA, and TS) were very useful in optimizing BayesNet model to prepare GPM. Among the metaheuristic algorithms, the GA was more accurate than the other two algorithms. The use of objective-function values to perform the optimization process without the need for additional information such as function derivatives, the simplicity of the search process, and the flexibility can be demonstrated as the best advantages of the GA application [71]. After the GA, the SA algorithm had the next-highest accuracy in optimizing the BayesNet model. The advantage of the SA algorithm was very low memory consumption and the ability to pass local optimization due to a randomized process [72]. However, this algorithm was less accurate in optimizing the BayesNet model than the GA due to its high dependence on the initial values of parameters and the lack of proper determination for the initial temperature parameter [73]. The TS algorithm may stop at a local optimization. Still, it is not clear whether the result obtained is a local optimization or a global optimization, and the local optimization obtained may depend on the initial response [74].

Therefore, the TS algorithm was less accurate in optimizing the BayesNet model than the other two algorithms.

4. Conclusions

Groundwater plays a vital role in human activities in various sectors, such as agriculture and industry. Therefore, determining the potential of groundwater using a machine-learning model is an effective step in groundwater management. To achieve higher accuracy in determining the potential of groundwater and solving the BayesNet model's problem, metaheuristic algorithms were used. Based on the results, metaheuristic algorithms improved groundwater-potential mapping with the BayesNet model. Based on the results of metric indices and the ROC–AUC values for the GPM, the GA had a higher accuracy than the SA and TS algorithms. According to the RMSE and MAE criteria, the BayesNet-GA model improved the GPM estimation accuracy by 4.6% and 7.5% (21.8% and 17.5%) compared to the BayesNet-SA and BayesNet-TS models. The results of the CF method showed that the criteria of distance to river and lithology had a higher weight in modeling compared to the other criteria. Also, the criteria of altitude, lithology, and drainage density were of higher importance in groundwater in study area. Prepared maps can help manage and control groundwater resources in arid and semiarid regions.

Author Contributions: Conceptualization, S.K.-R., N.T.T.L.; data curation, S.K.-R., J.H.A.; formal analysis, H.V.G., S.K.-R.; methodology, S.K.; resources, J.H.A.; supervision, O.K., S.K., N.T.T.L.; validation S.K., H.V.G.; visualization, S.K., H.V.G.; writing—original draft, S.K.-R., S.K.; writing—review and editing, I.-M.C., O.K., H.V.G., N.T.T.L.; funding acquisition, I.-M.C., N.T.T.L.; project administration, S.K.; software, S.K.-R. All authors have read and agreed to the published version of the manuscript.

Funding: In this research, Il-Moon Chung was funded by the Korea Ministry of Environment (MOE) as a Demand Responsive Water Supply Service Program (146515).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All of the relevant data is detailed in the article. The data presented in this study are available on request from the corresponding author.

Acknowledgments: The fourth author, Il-Moon Chung, would like to acknowledge the funding from the Korea Ministry of Environment (MOE) as a Demand Responsive Water Supply Service Program (146515).

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

Appendix A

Table A1. Database of the conditioning factors.

Factors	Source	Scale	Classification Method		
Altitude	SRTM DEM	1:60,000	Natural breaks		
Slope angle			Natural breaks		
Slope aspect			Manual		
Plan curvature			Manual		
Profile curvature			Natural breaks		
TWI			Natural breaks		
Distance to river			Manual		
Drainage density			Natural breaks		
Lithology			Geological Survey of Iran	1:100,000	Lithological units
Distance to fault					Manual
Soil	Khuzestan Natural Resources Organization	1:100,000	Soil units		
Land use/cover	Landsat-8 image	1:60,000	Land use/cover units		
Rainfall	The annual average of 22 meteorological stations in Khuzestan	1:60,000	Natural breaks		

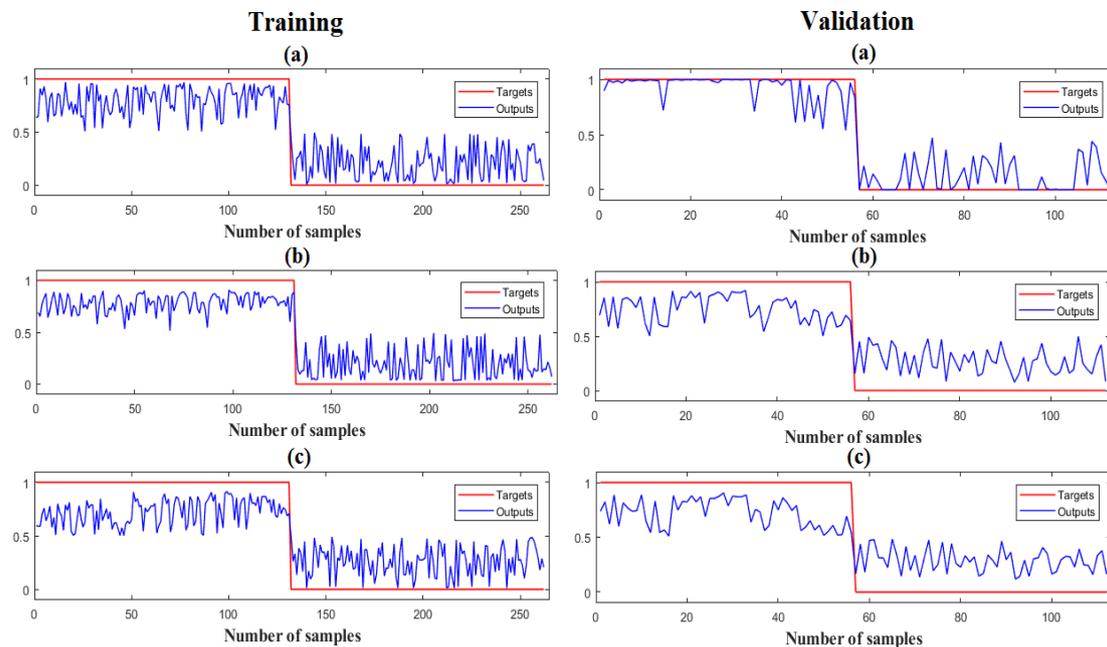


Figure A1. Results of the hybrid models: (a) BayesNet-GA, (b) BayesNet-SA, (c) BayesNet-TS.

References

1. Nguyen, P.T.; Ha, D.H.; Jaafari, A.; Nguyen, H.D.; Van Phong, T.; Al-Ansari, N.; Prakash, I.; Van Le, H.; Pham, B.T. Groundwater Potential Mapping Combining Artificial Neural Network and Real AdaBoost Ensemble Technique: The DakNong Province Case-study, Vietnam. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2473. [[CrossRef](#)] [[PubMed](#)]
2. Doke, A.; Pardeshi, S.D.; Das, S. Drainage morphometry and groundwater potential mapping: Application of geoinformatics with frequency ratio and influencing factor approaches. *Environ. Earth Sci.* **2020**, *79*, 1–17. [[CrossRef](#)]
3. Díaz-Alcaide, S.; Martínez-Santos, P. Review: Advances in groundwater potential mapping. *Hydrogeol. J.* **2019**, *27*, 2307–2324. [[CrossRef](#)]
4. Termeh, S.V.R.; Khosravi, K.; Sartaj, M.; Keesstra, S.D.; Tsai, F.T.-C.; Dijkma, R.; Pham, B.T. Optimization of an adaptive neuro-fuzzy inference system for groundwater potential mapping. *Hydrogeol. J.* **2019**, *27*, 2511–2534. [[CrossRef](#)]
5. Kim, J.-C.; Jung, H.-S.; Lee, S. Spatial Mapping of the Groundwater Potential of the Geum River Basin Using Ensemble Models Based on Remote Sensing Images. *Remote Sens.* **2019**, *11*, 2285. [[CrossRef](#)]
6. Lee, S.; Hyun, Y.; Lee, S.; Lee, M.-J. Groundwater potential mapping using remote sensing and gis-based machine learning techniques. *Remote Sens.* **2020**, *12*, 1200. [[CrossRef](#)]
7. Moghaddam, D.D.; Rahmati, O.; Haghizadeh, A.; Kalantari, Z. A Modeling Comparison of Groundwater Potential Mapping in a Mountain Bedrock Aquifer: QUEST, GARP, and RF Models. *Water* **2020**, *12*, 679. [[CrossRef](#)]
8. Panahi, M.; Sadhasivam, N.; Pourghasemi, H.R.; Rezaie, F.; Lee, S. Spatial prediction of groundwater potential mapping based on convolutional neural network (CNN) and support vector regression (SVR). *J. Hydrol.* **2020**, *588*, 125033. [[CrossRef](#)]
9. Chen, W.; Li, Y.; Tsangaratos, P.; Shahabi, H.; Ilia, I.; Xue, W.; Bian, H. Groundwater Spring Potential Mapping Using Artificial Intelligence Approach Based on Kernel Logistic Regression, Random Forest, and Alternating Decision Tree Models. *Appl. Sci.* **2020**, *10*, 425. [[CrossRef](#)]
10. 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.* **2014**, *8*, 7059–7071. [[CrossRef](#)]
11. Abd Manap, M.; Sulaiman, W.N.A.; Ramli, M.F.; Pradhan, B.; Surip, N. A knowledge-driven gis modeling technique for groundwater potential mapping at the upper langat basin, malaysia. *Arab. J. Geosci.* **2013**, *6*, 1621–1637. [[CrossRef](#)]
12. Taghizadeh-Mehrjardi, R.; Schmidt, K.; Toomanian, N.; Heung, B.; Behrens, T.; Mosavi, A.; Band, S.S.; Amirian-Chakan, A.; Fathabadi, A.; Scholten, T. Improving the spatial prediction of soil salinity in arid regions using wavelet transformation and support vector regression models. *Geoderma* **2021**, *383*, 114793. [[CrossRef](#)]
13. Nguyen, P.T.; Ha, D.H.; Avand, M.; Jaafari, A.; Nguyen, H.D.; Al-Ansari, N.; Van Phong, T.; Sharma, R.; Kumar, R.; Van Le, H.; et al. Soft Computing Ensemble Models Based on Logistic Regression for Groundwater Potential Mapping. *Appl. Sci.* **2020**, *10*, 2469. [[CrossRef](#)]
14. Razavi-Termeh, S.V.; Sadeghi-Niaraki, A.; Choi, S.-M. Groundwater Potential Mapping Using an Integrated Ensemble of Three Bivariate Statistical Models with Random Forest and Logistic Model Tree Models. *Water* **2019**, *11*, 1596. [[CrossRef](#)]

15. 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](#)]
16. Oikonomidis, D.; Dimogianni, S.; Kazakis, N.; Voudouris, K. A GIS/Remote Sensing-based methodology for groundwater potentiality assessment in Tirnavos area, Greece. *J. Hydrol.* **2015**, *525*, 197–208. [[CrossRef](#)]
17. 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](#)]
18. Band, S.S.; Janizadeh, S.; Chandra Pal, S.; Saha, A.; Chakraborty, R.; Melesse, A.M.; Mosavi, A. Flash flood susceptibility modeling using new approaches of hybrid and ensemble tree-based machine learning algorithms. *Remote Sens.* **2020**, *12*, 3568. [[CrossRef](#)]
19. Band, S.S.; Janizadeh, S.; Chandra Pal, S.; Saha, A.; Chakraborty, R.; Shokri, M.; Mosavi, A. Novel ensemble approach of deep learning neural network (dlmn) model and particle swarm optimization (pso) algorithm for prediction of gully erosion susceptibility. *Sensors* **2020**, *20*, 5609. [[CrossRef](#)]
20. Al-Fugara, A.K.; Pourghasemi, H.R.; Al-Shabeeb, A.R.; Habib, M.; Al-Adamat, R.; Al-Amoush, H.; Collins, A.L. A comparison of machine learning models for the mapping of groundwater spring potential. *Environ. Earth Sci.* **2020**, *79*, 1–19. [[CrossRef](#)]
21. 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](#)]
22. Park, S.; Hamm, S.-Y.; Jeon, H.-T.; Kim, J. Evaluation of Logistic Regression and Multivariate Adaptive Regression Spline Models for Groundwater Potential Mapping Using R and GIS. *Sustainability* **2017**, *9*, 1157. [[CrossRef](#)]
23. Zabihi, M.; Pourghasemi, H.R.; Pourtaghi, Z.S.; Behzadfar, M. GIS-based multivariate adaptive regression spline and random forest models for groundwater potential mapping in Iran. *Environ. Earth Sci.* **2016**, *75*, 1–19. [[CrossRef](#)]
24. Madani, A.; Niyazi, B. Groundwater potential mapping using remote sensing techniques and weights of evidence GIS model: A case study from Wadi Yalamlam basin, Makkah Province, Western Saudi Arabia. *Environ. Earth Sci.* **2015**, *74*, 5129–5142. [[CrossRef](#)]
25. Khoshtinat, S.; Aminnejad, B.; Hassanzadeh, Y.; Ahmadi, H. Groundwater potential assessment of the Sero plain using bivariate models of the frequency ratio, Shannon entropy and evidential belief function. *J. Earth Syst. Sci.* **2019**, *128*, 152. [[CrossRef](#)]
26. Razavi-Termeh, S.V.; Khosravi, K.; Sadeghi-Niaraki, A.; Choi, S.-M.; Singh, V.P. Improving groundwater potential mapping using metaheuristic approaches. *Hydrol. Sci. J.* **2020**, *65*, 2729–2749. [[CrossRef](#)]
27. Al-Abadi, A.M.; Shahid, S. Spatial mapping of artesian zone at Iraqi southern desert using a GIS-based random forest machine learning model. *Model. Earth Syst. Environ.* **2016**, *2*, 1–17. [[CrossRef](#)]
28. 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 Res. Manag.* **2017**, *31*, 2761–2775. [[CrossRef](#)]
29. 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 Res. Manag.* **2015**, *29*, 5217–5236. [[CrossRef](#)]
30. Choubin, B.; Rahmati, O.; Soleimani, F.; Alilou, H.; Moradi, E.; Alamdari, N. Regional Groundwater Potential Analysis Using Classification and Regression Trees. In *Spatial Modeling in Gis and R for Earth and Environmental Sciences*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 485–498. [[CrossRef](#)]
31. Naghibi, S.A.; Pourghasemi, H.R.; Abbaspour, K. A comparison between ten advanced and soft computing models for groundwater qanat potential assessment in Iran using R and GIS. *Theor. Appl. Clim.* **2017**, *131*, 967–984. [[CrossRef](#)]
32. Golkarian, A.; Naghibi, S.A.; Kalantar, B.; Pradhan, B. Groundwater potential mapping using C5.0, random forest, and multivariate adaptive regression spline models in GIS. *Environ. Monit. Assess.* **2018**, *190*, 149. [[CrossRef](#)]
33. Miraki, S.; Zanganeh, S.H.; Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Pham, B.T. Mapping Groundwater Potential Using a Novel Hybrid Intelligence Approach. *Water Resour. Manag.* **2018**, *33*, 281–302. [[CrossRef](#)]
34. 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.* **2018**, *27*, 211–224. [[CrossRef](#)]
35. Khosravi, K.; Panahi, M.; Bui, D.T. Spatial prediction of groundwater spring potential mapping based on an adaptive neuro-fuzzy inference system and metaheuristic optimization. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 4771–4792. [[CrossRef](#)]
36. Chen, W.; Tsangaratos, P.; Ilia, I.; Duan, Z.; Chen, X. Groundwater spring potential mapping using population-based evolutionary algorithms and data mining methods. *Sci. Total. Environ.* **2019**, *684*, 31–49. [[CrossRef](#)] [[PubMed](#)]
37. Banadkooki, F.B.; Ehteram, M.; Ahmed, A.N.; Teo, F.Y.; Fai, C.M.; Afan, H.A.; Sapitang, M.; El-Shafie, A. Enhancement of Groundwater-Level Prediction Using an Integrated Machine Learning Model Optimized by Whale Algorithm. *Nat. Resour. Res.* **2020**, *29*, 3233–3252. [[CrossRef](#)]
38. Prasad, P.; Loveson, V.J.; Kotha, M.; Yadav, R. Application of machine learning techniques in groundwater potential mapping along the west coast of India. *GISci. Remote Sens.* **2020**, 1–18. [[CrossRef](#)]
39. Lee, S.; Lee, M.-J.; Jung, H.-S.; Lee, S. Landslide susceptibility mapping using naïve bayes and bayesian network models in Umyeonsan, Korea. *Geocarto Int.* **2020**, *35*, 1665–1679. [[CrossRef](#)]
40. Zhang, S.; Li, C.; Zhang, L.; Peng, M.; Zhan, L.; Xu, Q. Quantification of human vulnerability to earthquake-induced landslides using Bayesian network. *Eng. Geol.* **2020**, *265*, 105436. [[CrossRef](#)]

41. Chen, W.; Li, H.; Hou, E.; Wang, S.; Wang, G.; Panahi, M.; Li, T.; Peng, T.; Guo, C.; Niu, C.; et al. 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](#)]
42. Maskooni, E.K.; Naghibi, S.A.; Hashemi, H.; Berndtsson, R. Application of Advanced Machine Learning Algorithms to Assess Groundwater Potential Using Remote Sensing-Derived Data. *Remote Sens.* **2020**, *12*, 2742. [[CrossRef](#)]
43. Moore, I.D.; Grayson, R.B.; Ladson, A.R. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrol. Process.* **1991**, *5*, 3–30. [[CrossRef](#)]
44. Bui, D.T.; Shirzadi, A.; Chapi, K.; Shahabi, H.; Pradhan, B.; Pham, B.T.; Singh, V.P.; Chen, W.; Khosravi, K.; Bin Ahmad, B.; et al. A Hybrid Computational Intelligence Approach to Groundwater Spring Potential Mapping. *Water* **2019**, *11*, 2013. [[CrossRef](#)]
45. Thilagavathi, N.; Subramani, T.; Suresh, M.; Karunanidhi, D. Mapping of groundwater potential zones in Salem Chalk Hills, Tamil Nadu, India, using remote sensing and GIS techniques. *Environ. Monit. Assess.* **2015**, *187*. [[CrossRef](#)]
46. Jothibas, A.; Anbazhagan, S. Spatial mapping of groundwater potential in Ponnaiyar River basin using probabilistic-based frequency ratio model. *Model. Earth Syst. Environ.* **2017**, *3*, 33. [[CrossRef](#)]
47. Dar, I.A.; Sankar, K.; Dar, M.A. Remote sensing technology and geographic information system modeling: An integrated approach towards the mapping of groundwater potential zones in Hardrock terrain, Mamundiyyar basin. *J. Hydrol.* **2010**, *394*, 285–295. [[CrossRef](#)]
48. Kumar, V.A.; Mondal, N.C.; Ahmed, S. Identification of Groundwater Potential Zones Using RS, GIS and AHP Techniques: A Case Study in a Part of Deccan Volcanic Province (DVP), Maharashtra, India. *J. Indian Soc. Remote Sens.* **2020**, *48*, 497–511. [[CrossRef](#)]
49. Shortliffe, E.H.; Buchanan, B.G. A model of inexact reasoning in medicine. *Math. Biosci.* **1975**, *23*, 351–379. [[CrossRef](#)]
50. Perlich, C.; Provost, F.; Simonoff, J.S. Tree induction vs. Logistic regression: A learning-curve analysis. *J. Mach. Learn. Res.* **2003**, *4*, 211–255.
51. Holland, J. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Application to Biology*; MIT Press: Cambridge, MA, USA, 1975.
52. Ecer, F.; Ardabili, S.; Band, S.S.; Mosavi, A. Training multilayer perceptron with genetic algorithms and particle swarm optimization for modeling stock price index prediction. *Entropy* **2020**, *22*, 1239. [[CrossRef](#)] [[PubMed](#)]
53. Aras, N.; Yumusak, S.; Altmel, I.K. *Solving the Capacitated Multi-Facility Weber Problem by Simulated Annealing, Threshold Accepting and Genetic Algorithms In Metaheuristics*; Operations Research/Computer Science Interfaces Series; Springer: Boston, MA, USA, 2007; Volume 39. [[CrossRef](#)]
54. Jabeen, T.; Ashraf, H.; Khatoon, A.; Band, S.S.; Mosavi, A. A lightweight genetic based algorithm for data security in wireless body area networks. *IEEE Access* **2020**, *8*, 183460–183469. [[CrossRef](#)]
55. Van Laarhoven, P.J.M.; Aarts, E.H.L.; Lenstra, J.J.K. Job Shop Scheduling by Simulated Annealing. *Oper. Res.* **1992**, *40*, 113–125. [[CrossRef](#)]
56. Aydin, M.E.; Fogarty, T.C. A simulated annealing algorithm for multi-agent systems: A job-shop scheduling application. *J. Intell. Manuf.* **2004**, *15*, 805–814. [[CrossRef](#)]
57. Yu, C.; Chen, M.; Cheng, K.; Zhao, X.; Ma, C.; Kuang, F.; Chen, H. Sgoa: Annealing-behaved grasshopper optimizer for global tasks. *Eng. Comput.* **2021**, 1–28. [[CrossRef](#)]
58. Zhang, C.; Li, P.; Guan, Z.; Rao, Y. A tabu search algorithm with a new neighborhood structure for the job shop scheduling problem. *Comput. Oper. Res.* **2007**, *34*, 3229–3242. [[CrossRef](#)]
59. Gallego, R.; Romero, R.; Monticelli, A. Tabu search algorithm for network synthesis. *IEEE Trans. Power Syst.* **2000**, *15*, 490–495. [[CrossRef](#)]
60. Kotthoff, L.; Thornton, C.; Hoos, H.H.; Hutter, F.; Leyton-Brown, K. Auto-weka 2.0: Automatic model selection and hyperparameter optimization in weka. *J. Mach. Learn. Res.* **2017**, *18*, 826–830.
61. Muda, Z.; Yassin, W.; Sulaiman, M.N.; Udzir, N.I. Intrusion detection based on k-means clustering and OneR classification. In Proceedings of the 7th International Conference on Information Assurance and Security (IAS), Melaka, Malaysia, 5–8 December 2011; pp. 192–197. [[CrossRef](#)]
62. Razavi-Termeh, S.V.; Sadeghi-Niaraki, A.; Choi, S.-M. Gully erosion susceptibility mapping using artificial intelligence and statistical models. *Geomat. Nat. Hazards Risk* **2020**, *11*, 821–844. [[CrossRef](#)]
63. Nhu, V.-H.; Mohammadi, A.; Shahabi, H.; Bin Ahmad, B.; Al-Ansari, N.; Shirzadi, A.; Clague, J.J.; Jaafari, A.; Chen, W.; Nguyen, H. Landslide Susceptibility Mapping Using Machine Learning Algorithms and Remote Sensing Data in a Tropical Environment. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4933. [[CrossRef](#)] [[PubMed](#)]
64. Razavi-Termeh, S.V.; Sadeghi-Niaraki, A.; Choi, S.-M. Ubiquitous GIS-Based Forest Fire Susceptibility Mapping Using Artificial Intelligence Methods. *Remote Sens.* **2020**, *12*, 1689. [[CrossRef](#)]
65. Razavi-Termeh, S.V.; Sadeghi-Niaraki, A.; Choi, S.-M. Asthma-prone areas modeling using a machine learning model. *Sci. Rep.* **2021**, *11*, 1–16. [[CrossRef](#)] [[PubMed](#)]
66. 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](#)]

67. 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.* **2015**, *188*, 1–27. [[CrossRef](#)] [[PubMed](#)]
68. Moghaddam, D.D.; Rezaei, M.; Pourghasemi, H.R.; Pourtaghie, Z.S.; Pradhan, B. Groundwater spring potential mapping using bivariate statistical model and GIS in the Taleghan Watershed, Iran. *Arab. J. Geosci.* **2013**, *8*, 913–929. [[CrossRef](#)]
69. Villeneuve, S.; Cook, P.G.; Shanafield, M.; Wood, C.; White, N. Groundwater recharge via infiltration through an ephemeral riverbed, central Australia. *J. Arid Environ.* **2015**, *117*, 47–58. [[CrossRef](#)]
70. Zeng, R.; Cai, X. Analyzing streamflow changes: Irrigation-enhanced interaction between aquifer and streamflow in the Republican River basin. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 493–502. [[CrossRef](#)]
71. Conkey, D.B.; Brown, A.N.; Caravaca-Aguirre, A.M.; Piestun, R. Genetic algorithm optimization for focusing through turbid media in noisy environments. *Opt. Express* **2012**, *20*, 4840–4849. [[CrossRef](#)] [[PubMed](#)]
72. Albash, T.; Lidar, D.A. Demonstration of a Scaling Advantage for a Quantum Annealer over Simulated Annealing. *Phys. Rev. X* **2018**, *8*, 031016. [[CrossRef](#)]
73. Selim, S.Z.; Alsultan, K. A simulated annealing algorithm for the clustering problem. *Pattern Recognit.* **1991**, *24*, 1003–1008. [[CrossRef](#)]
74. Lin, Y.; Bian, Z.; Liu, X. Developing a dynamic neighborhood structure for an adaptive hybrid simulated annealing–tabu search algorithm to solve the symmetrical traveling salesman problem. *Appl. Soft Comput.* **2016**, *49*, 937–952. [[CrossRef](#)]