

Article

Spatial Prediction of Groundwater Potentiality in Large Semi-Arid and Karstic Mountainous Region Using Machine Learning Models

Mustapha Namous ¹, Mohammed Hssaisoune ^{2,3}, Biswajeet Pradhan ^{4,5,*}, Chang-Wook Lee ^{6,*},
Abdullah Alamri ⁷, Abdenbi Elaloui ⁸, Mohamed Edahbi ⁹, Samira Krimissa ¹, Hasna Eloudi ²,
Mustapha Ouayah ¹, Hicham Elhimer ¹⁰ and Tarik Tagma ¹¹

- ¹ Laboratory of Biotechnology and Sustainable Development of Natural Resources, Polydisciplinary Faculty, Sultan Moulay Slimane University, Mghila B.P. 592, Beni Mellal 23000, Morocco; mustapha.namous@usms.ma (M.N.); samira_krimissa@yahoo.fr (S.K.); mustaphaouayah@gmail.com (M.O.)
 - ² Applied Geology and Geoenvironment Laboratory, Faculty of Sciences, Ibn Zohr University, Agadir 80000, Morocco; m.hssaisoune@uiz.ac.ma (M.H.); hasnaeloudi@gmail.com (H.E.)
 - ³ Faculty of Applied Sciences, Ibn Zohr University, B. O. 6146, Ait Melloul 86153, Morocco
 - ⁴ Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia
 - ⁵ Earth Observation Center, Institute of Climate Change, Universiti Kebangsaan Malaysia (UKM), Bangi 43600, Selangor, Malaysia
 - ⁶ Division of Science Education, Kangwon National University, Chuncheon-si 24341, Gangwon-do, Korea
 - ⁷ Department of Geology & Geophysics, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia; amsamri@ksu.edu.sa
 - ⁸ Water and Remote Sensing Team (GEVARET), Faculty of Sciences and Techniques, Sultan Moulay Slimane University, Beni Mellal 23000, Morocco; a.elaloui@usms.ma
 - ⁹ Higher School of Technology of Fkih Ben Salah, Sultan Moulay Slimane University, Beni Mellal 23000, Morocco; edahbimohamed@gmail.com
 - ¹⁰ Laboratory of Geostuctures, Geomaterials and Water Resources, Faculty of Sciences Semlalia, Cadi Ayyad University, Marrakesh 44000, Morocco; elhimer_h@yahoo.fr
 - ¹¹ Laboratoire Multidisciplinaire de Recherche et d'Innovation (LAMRI), Equipe Ingénierie des Ressources Naturelles et Impacts Environnementaux (IRNIE), Polydisciplinary Faculty of Khouribga, Sultan Moulay Slimane University, Khouribga 25000, Morocco; tariktagma@usms.ma
- * Correspondence: Biswajeet.Pradhan@uts.edu.au (B.P.); cwlee@kangwon.ac.kr (C.-W.L.)



Citation: Namous, M.; Hssaisoune, M.; Pradhan, B.; Lee, C.-W.; Alamri, A.; Elaloui, A.; Edahbi, M.; Krimissa, S.; Eloudi, H.; Ouayah, M.; et al. Spatial Prediction of Groundwater Potentiality in Large Semi-Arid and Karstic Mountainous Region Using Machine Learning Models. *Water* **2021**, *13*, 2273. <https://doi.org/10.3390/w13162273>

Academic Editors: Evangelos Tziritis and Andreas Panagopoulos

Received: 13 June 2021

Accepted: 16 August 2021

Published: 19 August 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: The drinking and irrigation water scarcity is a major global issue, particularly in arid and semi-arid zones. In rural areas, groundwater could be used as an alternative and additional water supply source in order to reduce human suffering in terms of water scarcity. In this context, the purpose of the present study is to facilitate groundwater potentiality mapping via spatial-modelling techniques, individual and ensemble machine-learning models. Random forest (RF), logistic regression (LR), decision tree (DT) and artificial neural networks (ANNs) are the main algorithms used in this study. The preparation of groundwater potentiality maps was assembled into 11 ensembles of models. Overall, about 374 groundwater springs was identified and inventoried in the mountain area. The spring inventory data was randomly divided into training (75%) and testing (25%) datasets. Twenty-four groundwater influencing factors (GIFs) were selected based on a multicollinearity test and the information gain calculation. The results of the groundwater potentiality mapping were validated using statistical measures and the receiver operating characteristic curve (ROC) method. Finally, a ranking of the 15 models was achieved with the prioritization rank method using the compound factor (CF) method. The ensembles of models are the most stable and suitable for groundwater potentiality mapping in mountainous aquifers compared to individual models based on success and prediction rate. The most efficient model using the area under the curve validation method is the RF-LR-DT-ANN ensemble of models. Moreover, the results of the prioritization rank indicate that the best models are the RF-DT and RF-LR-DT ensembles of models.

Keywords: drinking and irrigation water scarcity; groundwater potential mapping; machine learning; remote sensing; GIS; karstic mountainous aquifers; Morocco

1. Introduction

Mountainous areas cover more than 20% of the Earth's land surface where 25% of the global population lives [1]. The mountainous areas are well known to provide 50% of freshwater [2] compared to the other critical resources (i.e., food and wood). These areas constitute the main recharge areas of several porous and continuous aquifers in downstream lowland regions [3–5]. Nevertheless, understanding the details of groundwater functioning in a mountain massif requires a comprehensive knowledge of the most semi-arid areas [6].

Furthermore, mountainous areas assume the deep dynamics of groundwater, denying us direct access to groundwater outcrops that aid in hydrogeological exploration. An additional complication arises because mountainous regions are frequently fractured (e.g., the Atlas Mountains) and contain a discontinuous aquifer (e.g., a karstic aquifer). The groundwater potential in mountainous aquifers is governed by several parameters (i.e., lithology, geomorphology, topography, secondary porosity, geological structures, fracture density, permeability, drainage pattern and density, groundwater recharge, piezometric level, slope, land use/cover and climatic conditions, and their interrelationships) [7].

Overall, there is insufficient geodatabase related to groundwater, notably in fractured and karstic bedrock aquifers [8]. The knowledge gaps in terms of geodatabase make the development of numerical models difficult and consequently understand the aquifer functioning in mountainous areas [9]. In order to resolve this issue, various statistical models and machine learning algorithms have been employed for groundwater potential modelling and mapping using inventories of springs for dependent variables (i.e., binary logistic regression (LR) [10], certainty factor (CF) [11], weights-of-evidence (WE) [12], artificial neural networks (ANNs), random forest (RF), support vector machines (SVMs), naïve Bayes (NB) and decision tree (DT) [13]). Generally, machine learning methods have shown more robustness and stability during modelling, and thus, they have been popular and cost-effective in predicting groundwater potentiality. Furthermore, the application of machine learning algorithms remains more known in the prediction of natural disasters such as landslides, floods [14] and gully erosion. Recently, several researchers have tested machine learning ensemble models to improve the performance of prediction. Furthermore, other machine learning ensemble methods have been tested to delineate groundwater potential zones based on spring or well location inventories [15]. The purpose of this research is to apply four models individually (random forest (RF), logistic regression (LR), decision tree (DT) and artificial neural network (ANN)) and test different possible combinations using two, three and four models in each ensemble to produce groundwater potential maps in a large mountainous area.

In the semi-arid Oum Er-Rbia catchment located in the central part of Morocco, water resources are threatened by climate change. As a result, the economy and population requirements will be increased in the future [16–18]. Groundwater resources are derived from two system types: (1) the multi-layered system of the Tadla plains; and (2) the karst aquifer of the Atlas Mountains. The first system is overexploited and polluted by anthropogenic activities [5,19,20]. The second, located in the High Atlas Mountains of Beni Mellal, is characterized by the emergence of several springs (more than 370). The high flow rate of these springs indicates the presence of important groundwater reserves in this mountainous area. Several studies [21–25] using chemical and isotopic tools have been conducted to determine the water quality, hydrodynamic functioning and recharge area. In addition, the groundwater potential mapping of this aquifer has not yet worked out. In fact, mapping potential groundwater areas allow to determine zoning areas, permitting the identification of new sites that can provide a water supply for both drinking and irrigation. The results of this paper can be used by water managers and stakeholders to find potential

water supplies and manage the water resources which are more vulnerable to climate change and anthropogenic impacts.

The main objectives of this study are the following: (1) Evaluate the performance of individual and ensemble models in the prediction of groundwater potentiality in karstic mountainous areas; (2) compare the performance, robustness and stability of these models using several statistical and validation techniques; (3) test the importance of using a maximum of groundwater influencing factors and (4) produce reliable maps of the spatial distribution of groundwater potential in the study area.

2. Study Area

2.1. Geographic and Climatic Context

The study area is located in the Oum Er Rbia catchment, specifically in the High Atlas Mountains of Beni Mellal. It is bounded by the Tadla plain in the North and in the West, by the El Aabid River in the South and by the Oum Er-Rbia River in the East (Figure 1).

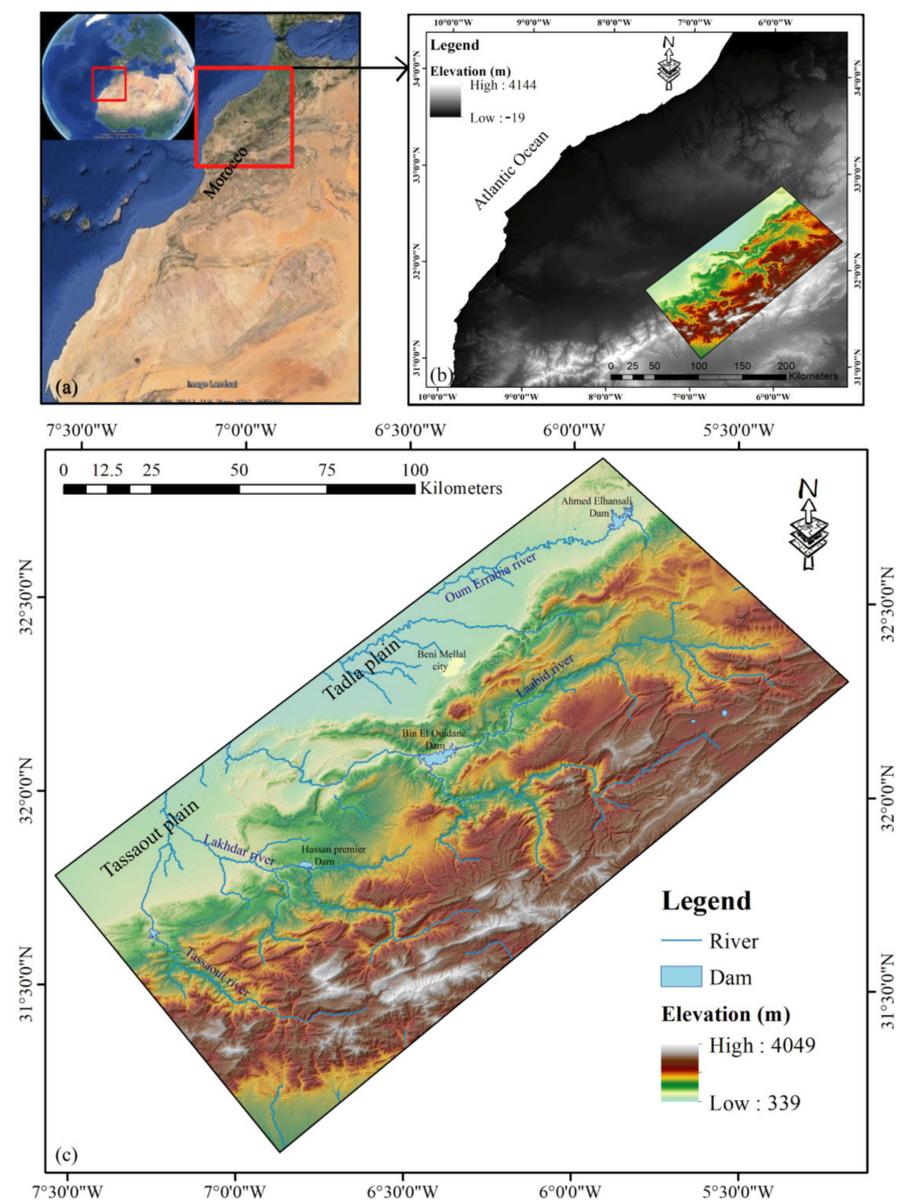


Figure 1. Geographical situation of study area at (a) national scale, (b) regional scale and (c) digital elevation model showing altitudes variability of the study area.

The climate ranges from semi-arid at the borders to sub-humid at high elevations, with a dominance in both cases of two distinct wet and dry seasons [18]. This Mediterranean climate is characterized by an annual rainfall varied between 300 to 750 mm and poorly distributed throughout the year. The average annual temperature is 18 °C (with peak periods of over 40 °C in August and 0 to 4 °C in January). The snow appears from 900 m of altitude and the prevailing wind is the Chergui in the summer period.

2.2. Geological and Hydrogeological Setting

The Atlas of Beni Mellal is composed of a Liassic and Middle Jurassic massive dolomitic and/or calcareous facies [5]. The other formations are composed of (1) Cenomanian conglomerates and sandstones interbedded with clays, gypsum marls and limestones, (2) limestones and karstic dolomites interlayered with marl horizons of the Turonian, (3) marls and limestones with evaporitic characters of the Senonian and (4) a complex of limestones, marls and phosphate sandstones forming the phosphate series (Figure 2) [26]. The geological structure of the basin implies a continuation of the north Atlas thrust fault toward the Tadla plain [5,27] (Figure 2).

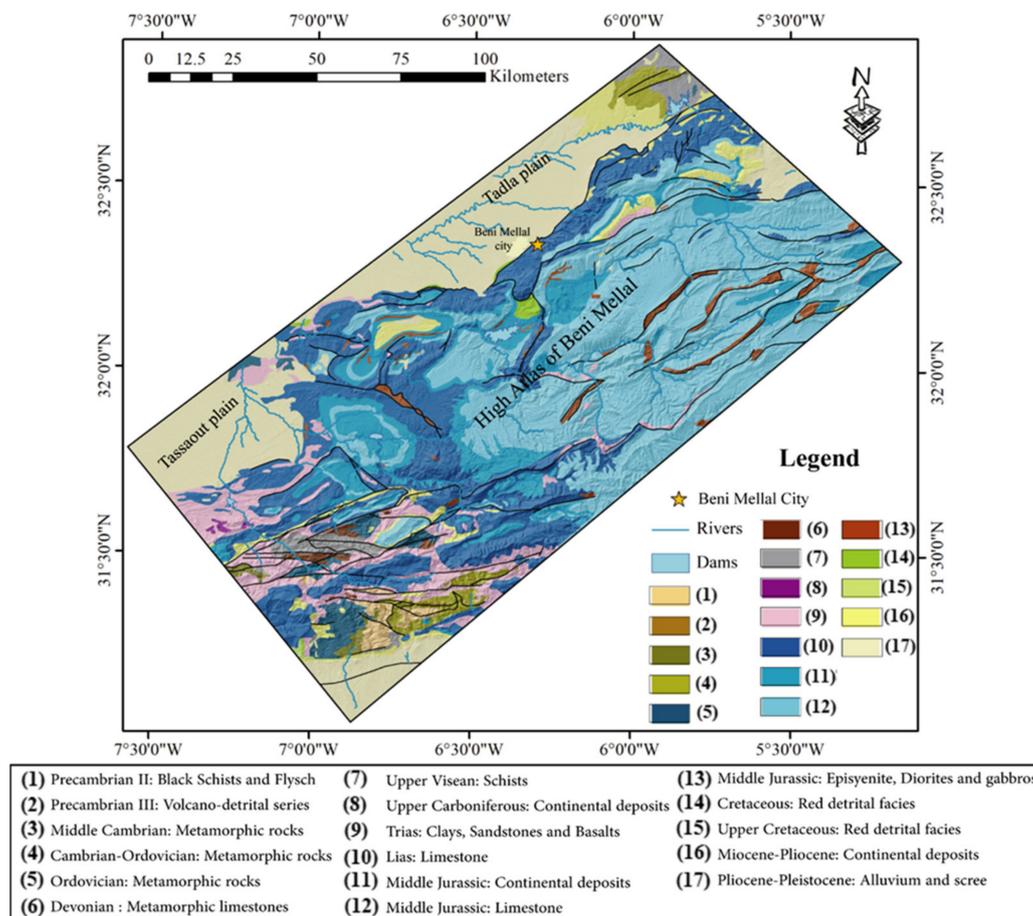


Figure 2. Geological setting of the study area.

In the study area, the groundwater resources are derived from two system types [5,22]: (1) the shallow and deep aquifers of the Tadla plain and (2) the karst aquifer of the Atlas Mountains. The first system is composed of four aquifers: (1) Mio-Plio-Quaternary, (2) Eocene, (3) Senonian and (4) Turonian which is the main productive aquifer in the region and separated by impermeable or semi-permeable horizons. The groundwater of the second system (the subject of this study) is contained in the aquifer calcareous rocks—of Liassic age, which are mainly karstic system which possess high storage and conductive capacities.

These formations are exposed on the mountains and favor water percolation (rainwater and snowmelt), which constitutes the natural replenishment of the aquifer [22,28].

From the areas of recharge in the mountains, the groundwater flows in the aquifer to the points of discharge outlets: natural springs, underground seepage and pumping wells in the plain area [5]. The major karst outlets are located along the northern High Atlas accident of the High Atlas Mountains from Timoulilt in the southwest to Zaouit Cheikh in the northeast; the most important is Ain Asserdoune, with an average flow rate of 700 L/s. It is important to note that, in this environment, most karst eminences are probably underground and are therefore not visible, particularly at the mountain's transition of the Tadla plain.

3. Materials and Methods

A spring can be defined as a window through which groundwater flows from an aquifer to the Earth's surface [29]. Based on this characteristic, the emergence of springs reflects the groundwater potentiality. To assess the relationship between source occurrence and factors controlling groundwater flow, the groundwater potential mapping (GPM) a tool has been used to provide spatial information [29].

The methodology of this study is summarized in the flowchart (Figure 3). The main steps are as follows: (1) preparation of the data for modelling (preparation of the spring inventory map and preparation of the conditioning factors datasets). Two methods were applied for the factors selection which contribute to springs emergence (*IG*) and variance inflation factor (*VIF*). (2) A frequency ratio method was applied to determine the spatial relationships between spring occurrence and its predisposing factors. (3) RF, LR, DT and ANN models were applied for mapping groundwater potential; then, different ensemble of models were tested in order to find the best rate of prediction, in addition to the production of various groundwater potentiality maps. (4) Several statistical parameters were applied to test the results of the models application, and a general comparison was carried out based on a compound factor (*CF*) method and priority rank (*PR*). Geographic Information System environments and statistics software were used during the current study for database preparation and groundwater potential mapping, and R packages for machine learning algorithm modelling were also used (randomForest, C50, neuralnet and calibrateBinary). Table 1 highlights the spatial datasets used in this study.

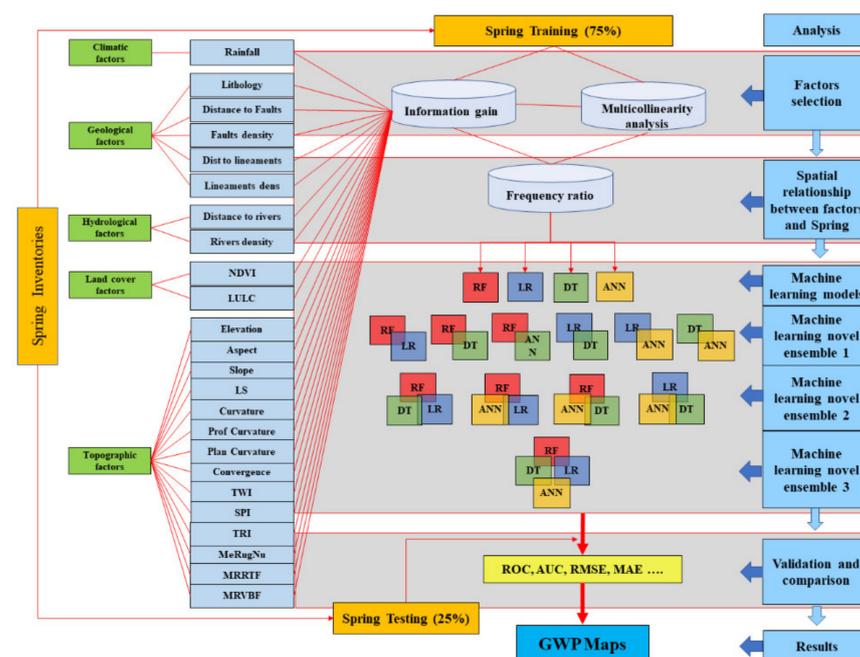


Figure 3. Methodology used in this research.

Table 1. Spatial database of the study area.

Factors	Data Layers	Data Provider
Spring inventory		Previous studies Field investigation Topographic maps
Topographic factors	Elevation Aspect Slope Curvature Profile curvature Plan curvature Convergence TWI SPI TRI MeRugNu MRRTF MRVBF LS	SRTM-DEM from (http://gdex.cr.usgs.gov/gdex/ (accessed on 10 January 2020)) pixel size of 30 m × 30 m.
Geologic factors	Lithology Distance to Faults Faults Density Distance to lineaments Lineaments density	Geological map of Morocco at the scale 1:500,000 Geological map of Morocco at the scale 1:500,000 LANDSAT satellite image at 30 m from (https://earthexplorer.usgs.gov/ (accessed on 10 January 2020))
Hydrologic factors	Distance to rivers Rivers density	DEM at 30 m
Land cover factors	NDVI LULC	LANDSAT satellite image at 30 m from (https://earthexplorer.usgs.gov/ (accessed on 10 January 2020))
Climatic factors	Rainfall	Climatic stations data from hydraulic basin agency of Oum Erabia Tropical Rainfall Measuring Mission

3.1. Groundwater Springs Inventory (SI)

The spring inventory map was developed using an extensive field data. A total of 374 springs were identified, where 280 springs (75%) have been randomly selected for the training dataset. The remaining springs (25%) were used for the validation dataset (Figure 4). The springs' discharge values vary between 0.1 and 1450 L/s.

For spatial modelling, several researchers have recommended the use of equal proportions of spring and non-spring pixels, while, others have suggested to use a high number of non-spring compared to the number of spring pixels when the study area is generally large and it cannot ensure a good spatial representativeness of the springs. Due to the large size of the study area, a high number of non-spring pixels were selected for this study. As a result, randomly mapped of 840 non-spring points for training data and 282 non-spring points for testing data were tested (three times the number of spring pixels) (Figure 4).

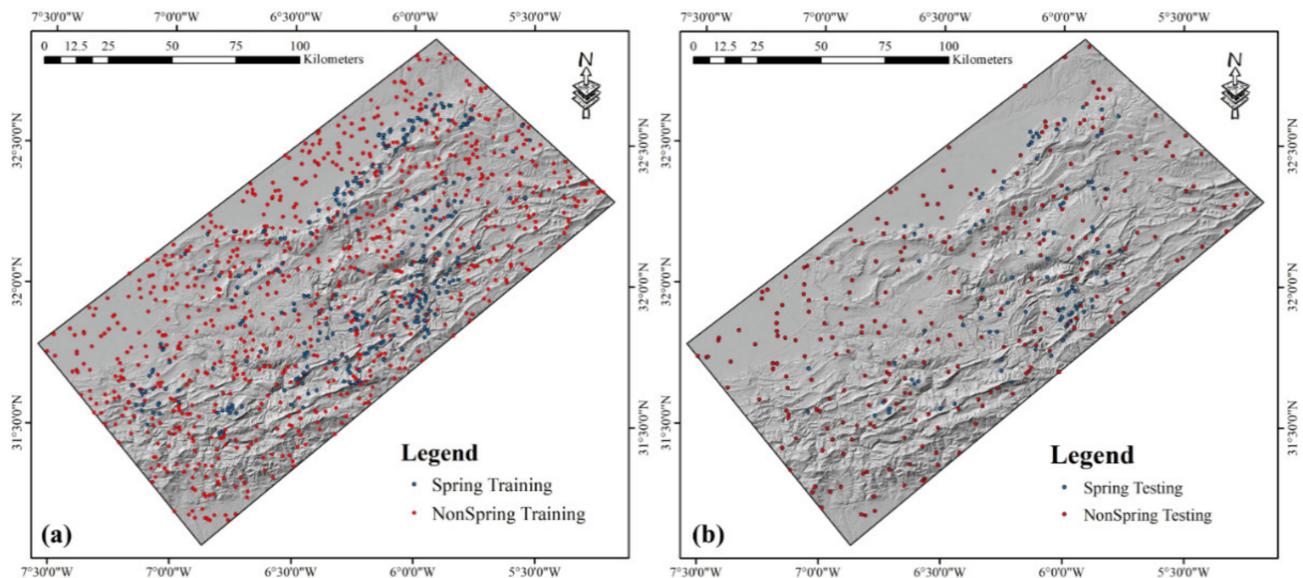


Figure 4. Location of springs and non-springs in the study area, (a) Training datasets and (b) Testing datasets.

3.2. Groundwater Influencing Factors (GIFs)

The selection of the groundwater potentiality influencing factors is very challenging due to the complexity of the groundwater functioning phenomenon. Moreover, this choice is very difficult since there are no exact standard norms. For the present study, our challenging aim was to combine many factors as possible that may have an influence on the groundwater potentiality. Consequently, we have prepared a total of 24 geological, hydrological, climatic, topographic and land cover/use factors (Figure 5).

3.2.1. Climatic Factors

Climate is a key factor directly involved in groundwater availability. In that sense, rainfall permits and directly encourages the recharge of aquifers. Annual precipitation data were obtained from the Tropical Rainfall Measuring Mission (TRMM) between 1998 and 2016 (validation of estimated TRMM rainfall data by [30]). According to the rainfall map produced, the annual average rainfall varies between 119 and 889 mm/year in the study area. The most significant values are located in the northern part, while in the south, the precipitation decreases intensely (Figure 5u).

3.2.2. Hydrological Factors

The hydrological factors chosen are the distance from rivers and the density of rivers. The distance to rivers was calculated by the Euclidean distance method in ArcGIS environment for the purpose of determining the distance of the spring from the drainage system (Figure 5n), while river density helps us recognize the spatial distribution of streams in the study area (Figure 5o). The maps show that the distances to the rivers vary between 0 and 5077 m, and that the rivers are more concentrated in the northwest part (plain area) than in the southeast part (mountainous area).

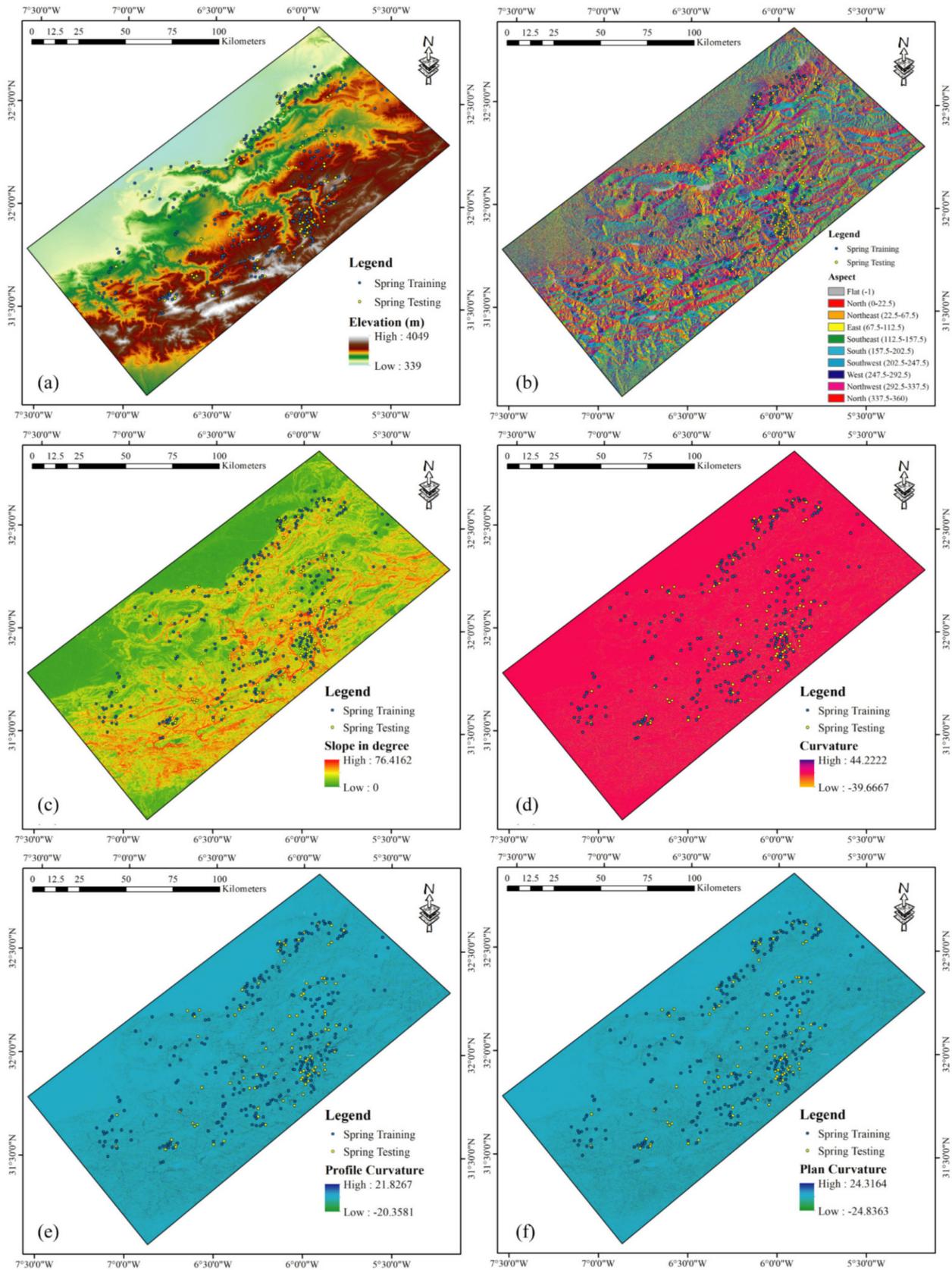


Figure 5. Cont.

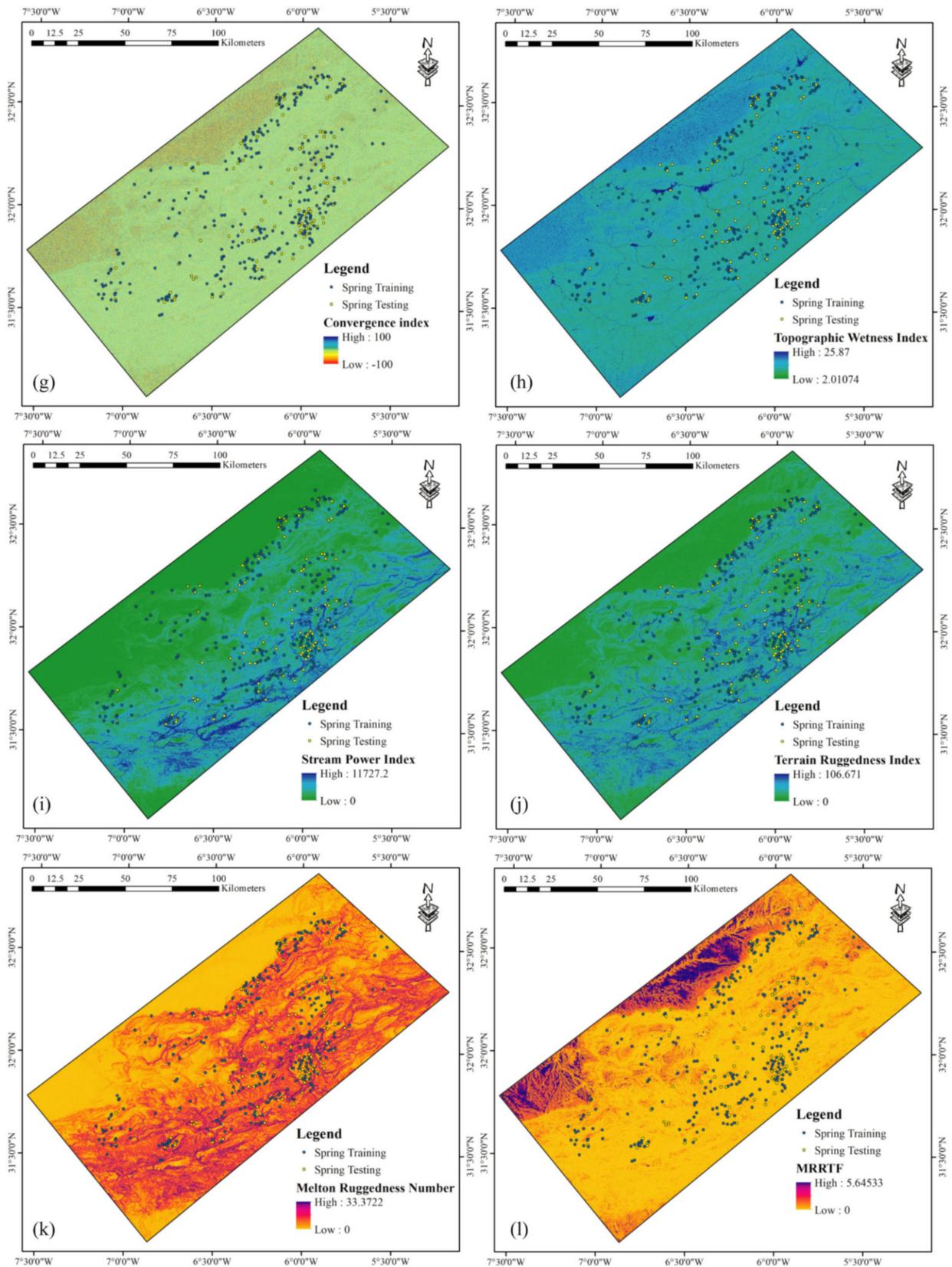


Figure 5. Cont.

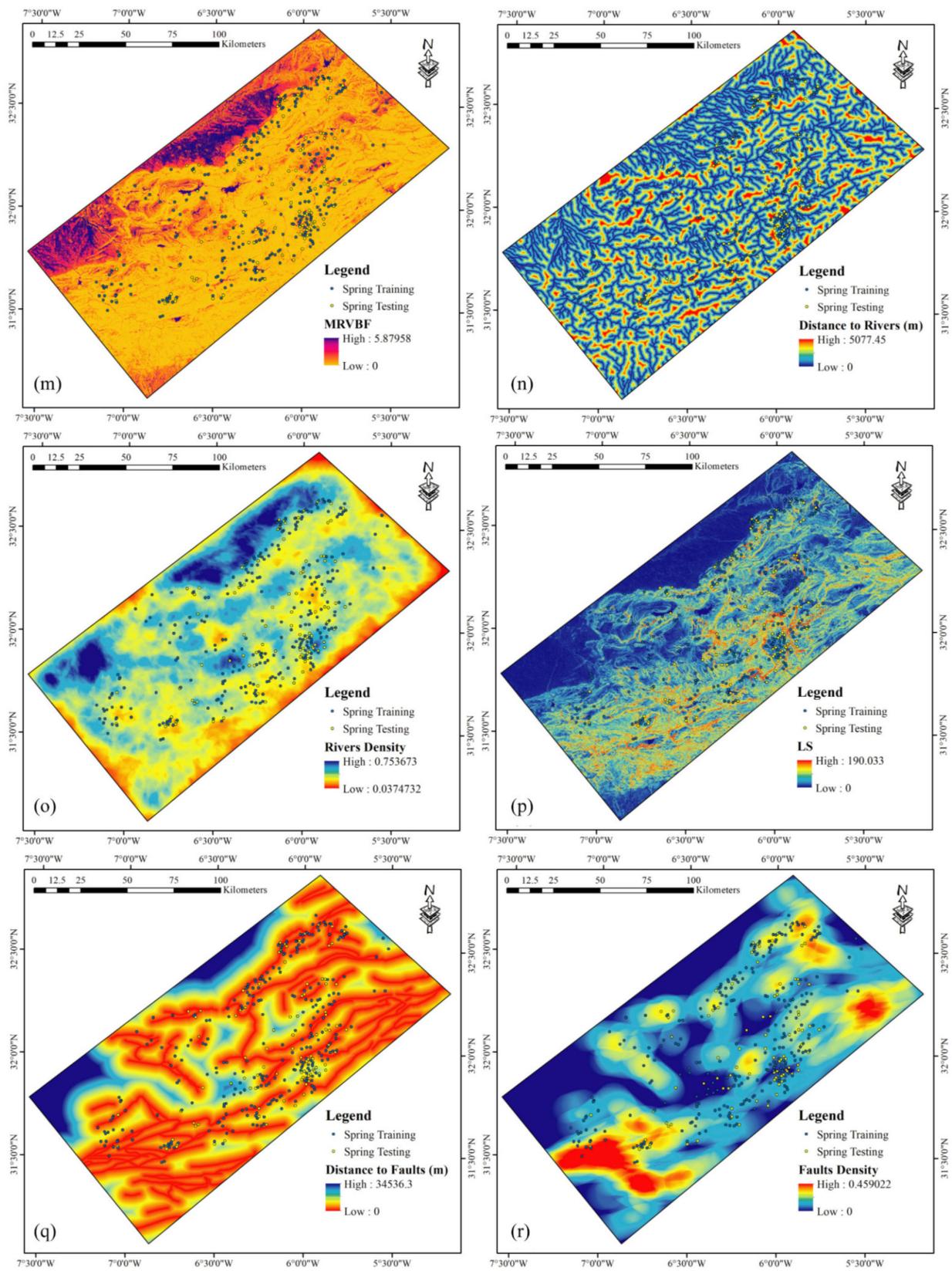


Figure 5. Cont.

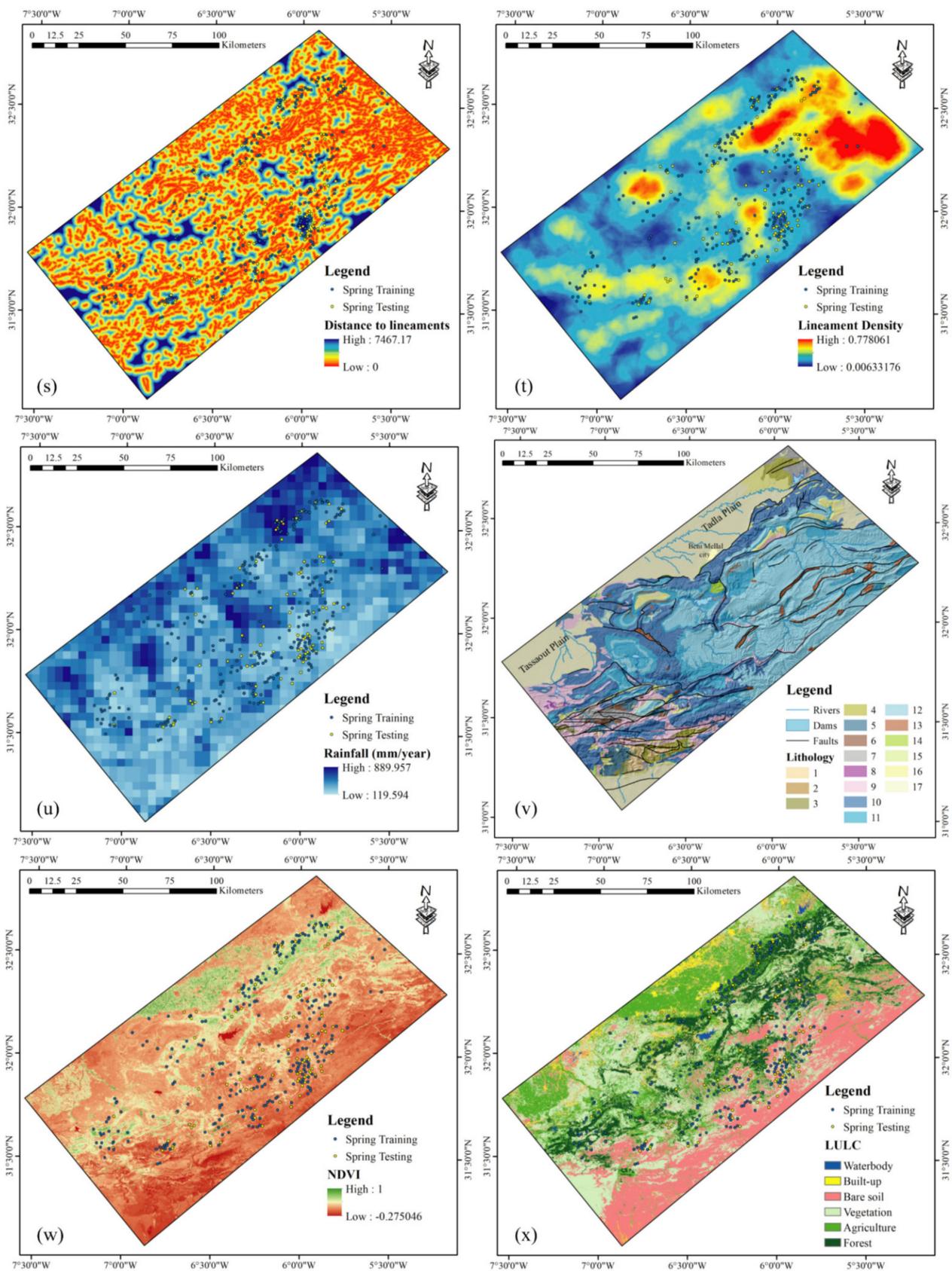


Figure 5. Groundwater influencing factors considered in the present study. (a) Elevation, (b) Aspect, (c) Slope, (d) Curvature, (e) Profile Curvature, (f) Plane Curvature, (g) Convergence, (h) TWI, (i) SPI, (j) TRI, (k) MeRugNu, (l) MRRTF, (m) MRVBF, (n) Distance to Rivers, (o) Density of Rivers, (p) LS, (q) Distance to Faults, (r) Density of Faults, (s) Distance to lineament, (t) Density of lineament, (u) Rainfall, (v) Lithology, (w) NDVI, (x) LULC.

3.2.3. Geological Factors

There are many geological factors, and they play an essential role in the formation, availability and recharge of groundwater. The main geological factor is lithology, since rock types determine aquifer formation and its continuous recharge by controlling the permeability and water circulation [31]. The 1/500,000 geological map of Morocco is used to digitalize the main lithological units in this study; the results are presented in Figures 2 and 5v.

In addition, the faults influence the presence/recharge of groundwater and the emergence of springs (secondary permeability of rocks). In our case, the Beni Mellal High Atlas is largely fractured by a dominant northeast–southwest oriented fault network, which gives great importance to the investigation of the fault-groundwater potentiality spatial relationship. Thus, two fault factor maps are produced, one representing the distance to faults and the other represent the fault density (Figure 5q,r, respectively).

The lineament is very important in the third type of environment, as it will give an idea of the spatial distribution and density of fractures in the karst landscape. The fractures frequently participate in groundwater recharge and spring emergence [32]. In this study, lineaments were detected through the interpretation of Landsat OLI imagery; the maps of distance to lineaments and lineament density are shown in Figure 5s,t.

3.2.4. Topographic Factors

Topographic factors play an essential role in controlling hydrological conditions, such as the flow of groundwater and soil moisture. In this study, 14 topographic factors were used: elevation, aspect, slope, curvature, profile curvature, plan curvature, convergence, topographic wetness index (TWI), sediment power index (SPI), terrain ruggedness index (TRI), Melton ruggedness number (MeRugNu), multi-resolution ridge top flatness (MRRTF), multi-resolution valley bottom flatness (MRVBF) and slope length (LS). These topographic factors are shown in Figure 5.

The elevation and slope factors generally negatively control the groundwater potential where in flat areas and low elevation the rainwater has much more time to infiltrate and recharge groundwater [33]. For the aspect factor, the exposure of the slopes favors more water infiltration on slopes exposed to humid winds and protected from solar radiation [15]. In this study, the potential of groundwater will be higher on the north- and northwest-facing slopes. The curvature influences groundwater recharge. The same influence may also be related to the LS factor. The MRRTF and MRVBF indicate the flatness and size of valley bottoms; low values show a smaller possibility of the existence of a groundwater aquifer, and high values indicate the potential zones [15]. In this study, the highest values for these two parameters are calculated in the plains bordering the Atlas Mountains towards the North. The TWI factor shows the influence of topography on runoff generation and flow accumulation; a high groundwater potentiality is favored when TWI values increase. The SPI expresses the erosive influence and power of water flow [34], and the TRI indicates the difference in elevation between adjacent cells of a digital elevation grid [35].

3.2.5. Land Use/Cover Factors

Land use/cover factors (LULC) strongly influence hydrological processes, such as infiltration, evapotranspiration and surface runoff, and they consequently play a significant role in groundwater potentiality. Two factors in relation to land cover have been prepared: land use/land cover (LU/LC) and normalized difference vegetation index (NDVI).

The LULC map of the study area was prepared using supervised classification and a maximum likelihood algorithm in ArcGIS environment from three merged images of the Landsat Operational Land Imager (OLI); their dates of acquisition are July 9 and 18, 2019. From the same merged images, we calculate the NDVI to determine the density of the vegetation. The LULC map contained six different classes: Waterbody, Built-up, Bare soil, Vegetation, Agriculture and Forest (Figure 5x). The NDVI map is shown in Figure 5w.

3.3. Groundwater Influencing Factors (GIFs) Analysis

3.3.1. Multicollinearity Analysis and Confusion Matrix

Multicollinearity analysis was used in statistics to detect the linearity between the conditioning factors of a given phenomenon, and detect and quantify information redundancies between the parameters that may have a negative impact on the model performance. Multicollinearity refers to the non-independence of conditioning factors that may occur in datasets. It is widely used in the prediction of several phenomena, such as landslides, gully erosion and groundwater potentiality. In this study, the multicollinearity for the groundwater influencing factors was identified using confusion matrix, tolerances and variable inflation factor (VIF) methods, according to Equations (1) and (2):

$$Tolerance = 1 - R_j^2 \quad (1)$$

$$VIF = \left[\frac{1}{Tolerance} \right] \quad (2)$$

where:

R_j^2 is the coefficient of determination.

When $VIF \geq 10$, there are linear relationships between conditioning factors.

3.3.2. Selection of Groundwater Influencing Factors

The ability to estimate groundwater potentiality depends on the factors introduced into the model. Indeed, some factors can decrease this capacity. As a result, a preliminary selection of the factors is necessary. To meet that condition, we used the information gain method to select GIFs. The information gain (IG) value for a groundwater influencing factor X_i and a class Y is calculated using Equations (3)–(5):

$$IG(Y, X_i) = H(Y) - H(Y|L_i) \quad (3)$$

where:

$$H(Y) = - \sum_i P(Y_i) \text{Log}_2(P(Y_i)) \quad (4)$$

$$H(Y_i|L_i) = - \sum_i P(Y_i) \sum_j P(Y_i|L_i) \text{Log}_2((P(Y_i|L_i))) \quad (5)$$

where:

$H(Y)$ is the entropy value of Y_i ;

$H(Y_i|L_i)$ is the entropy of Y after associating the values of the landslide conditioning factor L_i ;

$P(Y_i)$ is the prior probability of the out-class Y ;

$P(Y_i|L_i)$ are the posterior probabilities of Y given the values of the conditioning factor L_i .

Factors that have a negative IG are considered to have no effect on groundwater potential, and therefore, they will be eliminated from the analysis.

3.3.3. Weight of the Groundwater Influencing Factors

In order to assign a weight of each class of factor before the modelling phase, several researchers recommend the use of the frequency ration (FR) method. The FR method helps to determine the spatial relationship between the predisposition factors and the dependent factor [36]. Each factor is segmented into several classes; Fr index value is calculated for each class of factors using the following equation (Equation (6)):

$$Fr = \frac{PS_i}{PD_i} = \frac{\left(\frac{NS_i}{NST} \right) \times 100}{\left(\frac{NA_i}{NAt} \right) \times 100} \quad (6)$$

where:

PSi denotes the percentage of spring pixels for each class i of influencing factors, relative to the total number of spring pixels in the study area;

PDi is the percentage of each class i of influencing factors, relative to the total area;

NSi is the number of spring pixels in a thematic class i ;

NSt is the number of pixels of all springs;

NAi is the total number of pixels in a thematic class i ;

NAt is the total number of all pixels.

The results obtained represent the correlation between each class of influencing factors and the groundwater spring areas. The final step is the standardization of the FR to give equal importance to the different factors. The method used is to arrange the values of FR between 0.01 and 0.99 by the max-min normalization method according to Equation (7):

$$FRN = \frac{FR - Max(FR)}{Max(FR) - Min(FR)} \times (0.99 - 0.01) + 0.01 \quad (7)$$

where:

FRN is the normalized FR matrix;

FR is the original data matrix.

3.4. Methods

3.4.1. Random Forest (RF) Model

Random forest model has been developed based on the classification and regression trees (CARTs) [37]. The objective of the method is to evaluate the relationships between all factors: the springs are dependent factors and the GIFs are independent factors. The purpose is to identify the most appropriate model to build the GPM map and determine the weight of each factor. The approach is based on the creation of several decision trees, and for each tree, there will be a random selection of a set of predictive factors that will be used at each node to improve the prediction [38]. Consequently, all trees' prediction results are averaged to build the final set of model predictions [39], and the data that are not involved in the analysis are defined as the out-of-bag (OOB) error.

The basic parameters to build the RF model are $mtry$ and $ntree$. $mtry$ designates the number of factors to be considered in each tree-building process, and $ntree$ designates the number of trees. The advantage of the RF model is that $mtry$ and $ntree$ can be changed and varied to test different possibilities to choose the best performance pathways and the minimum OOB error. In addition, the RF method also allows to classify the factors according to their importance. The calculation of the weights is done by measuring the mean decrease in prediction accuracy.

3.4.2. Logistic Regression (LR) Model

Logistic regression is a machine learning method developed to solve classification problems. The LR predictive analysis algorithm accepts continuous or discrete variables for model input, and it does not require that they have a normal distribution [40]. The approach is based on the concept of using probability to determine the relationship between independent factors (GIFs) and the dependent factor (SI). To achieve this objective, the dependent factor is coded as 1's and 0's (binary variable), where important groundwater potential is coded as a 1 and weak groundwater potential is coded as a 0. Nevertheless, independent factors can be continuous or categorical. In this study, we chose to classify all GIFs into numerical values representing their weights based on the frequency ratio method.

The groundwater probability potentiality is calculated according to Equations (8) and (9):

$$P = \frac{1}{1 + e^{-z}} \quad (8)$$

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (9)$$

where:

P is the probability;

Z is the linear combination of the independent variables;

β_0 is the intercept of the model;

$\beta_1, \beta_2 \dots \beta_n$ are the coefficients of the logistic regression model;

$x_1, x_2 \dots x_n$ are the independent variables;

n is the number of independent variables.

3.4.3. Decision Tree (DT) Model (C 5.0)

A decision tree is used to classify future observations based on an already classified data set. The base of the tree corresponds to a root. Then, a series of branches whose intersections are called nodes end in leaves that each correspond to one of the classes to be predicted. Each node of the decision tree makes a binary decision that separates one class, or several classes, from the other classes [41]. In this study, we have chosen to use the C 5.0 classification algorithm, which is more efficient than its predecessor C 4.5 and offers similar results with smaller decision trees [42]. The algorithm uses the adaptive boosting method to improve the model accuracy, and it is based on the concept of entropy. A calculation of the information gain of the variables is carried out beforehand to classify them according to the maximum values; this helps to eliminate the leaves of null or weak values, which improves the classification accuracy [39,43].

3.4.4. Artificial Neural Network (ANN) Model

In this study, the multi-layer perceptron (MLP) architecture was chosen, which contains three layers connected by several neurons: the input layer, the hidden layer and the output. For the input layer, which has one input and several output pathways for each neuron, each node is connected with the different determining factors (GIFs). Hidden nodes, where there are several inputs and output connections for each neuron, use weighted connections to learn and process the problem; weights can take positive or negative values. Usually, for the modelling phase, the ANN method starts with the adjustment of the weights of the different connections between neurons during the training phase; then, the output prediction stage is based on the constructed models [44]. In the ANN method, the input (x_i) and output (y_i) layers can be expressed by the following equations (Equations (10) and (11)):

$$net = \sum_{i=0}^n wixi \quad (10)$$

$$yi = f(net) \quad (11)$$

where:

x_i are the inputs;

w_i are the corresponding weights;

y_i is the output.

3.5. Ensembles of Models

To improve the performance and accuracy of model prediction, ensembles of models have been used and tested by many researchers. They have confirmed their effectiveness and efficiency in landslide prediction and soil erosion assessment. However, those who have tested the method for assessing groundwater potential remain limited. The choice of ensemble in this study was based on a weighted aggregation of the individual RF, LR, DT and ANN models to determine the best possible combination. Three different combinations were tested: two models, three models and four models. The equation used is:

$$EM = \frac{\sum_{i=1}^n (AUCSi \times Mi)}{\sum_{i=1}^n AUCS} \quad (12)$$

where:

EM is the ensemble of models;
 $AUCSi$ is the area under the success rate curve for the model Mi ;
 Mi is the individual model.

3.6. Performance Metrics and Comparison

Validation of the results in modelling is an essential step to confirm the validity of the results and the performance of the models. However, when the database partition changes, it is important to assess the stability of the computations. To do this work, we have examined the success and prediction rate gains of the following sample divisions: 25/75%, 50/50% and 75/25%. Then, statistical metrics and the area under the receiver operating characteristic curve (AUC) were used for the testing dataset to evaluate the performance of the RF, LR, DT and ANN models and the different ensembles.

3.6.1. Statistical Metrics

The validation approach is based on the calculation of four parameters, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). Their determination is based on the calculation of spring pixels which are correctly or incorrectly classified as springs in the training and testing datasets. The sensitivity is the proportion of spring pixels that are correctly classified as spring occurrences, while the specificity is the proportion of the non-spring pixels that are correctly classified as non-spring [45]. In addition, other parameters have been calculated to improve the comparison between the models: accuracy, precision, FP-Rate, MCC, RMSE, MAE and the Kappa index. Higher values of sensitivity, specificity, accuracy, precision, FP-Rate and MCC indicate better performance of a model, especially if the RMSE and MAE values are close to 0. A Kappa index value of 1 indicates a perfect model, whereas -1 represents a non-reliable model. All the equations used in the calculation of these parameters are written below:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (13)$$

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (14)$$

$$\text{Accuracy} = \frac{TN + TP}{TP + FP + TN + TP} \quad (15)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (16)$$

$$\text{FPRate} = \frac{FP}{FP + TN} \quad (17)$$

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (18)$$

$$\text{Kappa} = \frac{\text{Accuracy} - B}{1 - B} \quad (19)$$

where:

$$B = \frac{(TP + FN)(TP + FP) + (FP + TN)(FN + TN)}{\sqrt{TP + TN + FN + FP}} \quad (20)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_P - X_A)^2} \quad (21)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |(X_P - X_A)| \quad (22)$$

3.6.2. ROC Curve

In terms of the excellence and the performance of machine learning models, the ROC curve represents the most useful way to validate the results [45]. For this method, a comparison was done between the groundwater potentiality map, the training, and validation of spring inventory maps. The receiver operating characteristics curve is a graphical representation that plots the true positive percentage in the y -axis and the cumulative false positive percentage in the x -axis [46]. Finally, the area under the curve was calculated (AUC) from the ROC curve, and the precision of the model was evaluated. The area under the ROC curve varies between 0 and 1; it can be categorized as low (0.5–0.6), medium (0.6–0.7), good (0.7–0.8), very good (0.8–0.9) and excellent (0.9–1.0) [47,48].

3.6.3. Model Prioritization Using Compound Factor

The compound factor (CF) method was used to rank the different models and to compare their performance and accuracy. The method is based first of all on the ranking of all models and ensembles of models with respect to AUC values and statistical metrics. Then, to find the best fit model for producing GPMs, the CF—based prioritization was calculated in terms of the accuracy, precision, specificity, sensitivity, FP-Rate, MCC, Kappa index, AUC, MAE and RMSE values among the 15 models. The CF calculation is performed according to the equation below (Equation (23)):

$$CF = \frac{1}{n} \sum_{i=1}^n R \quad (23)$$

where:

R is the variable rank;

n is the number of variables.

4. Results

4.1. GIF Selection and Analysis

After the first step of the analysis, which included the realization of an inventory map of springs and non-springs that constituted the basic document to start the modelling phase, an analysis of the influencing factors was undertaken to select the most useful GIFs and eliminate those that have no effect or those that have a multicollinearity.

Primarily, the multicollinearity analyses of the 24 groundwater influencing factors show that tolerance values vary between 0.231 for the TRI factor to 0.983 for the aspect factor. In the same way, the VIF values fluctuate between 1.017 for the aspect factor and 4.323 as the maximum value for the TRI factor (Table 2). These results are acceptable: the tolerance values are greater than 0.1 and the VIF values are less than 10, which confirms that all of the selected GIFs have no multicollinearity. However, the confusion matrix diagram results indicate a linear relationship between several variables, including TRI (0.75), slope (0.74), LS (0.75), faults density (0.69), distance to faults (0.69), LULC (0.49) and NDVI (0.49). To avoid data redundancy, we conducted many experiments, including the elimination of redundant factors, by testing the success and prediction rates for all models. Results show that there is no significant impact on the rates of learning and prediction of the various models except DT model, where the prediction rate has been decreased significantly from 0.738 to 0.609. These findings indicate that the redundancy of data from some factors has only a little impact on performance, indicating that all factors must be taken into account in this analysis.

Afterwards, the results of the analysis using the information gain method show that the lithology, fault density and distance to faults factors have the highest values (0.073, 0.030 and 0.029, respectively), followed by the rainfall (0.025), MRRTF (0.020), MRVBF (0.019), elevation (0.017), and lineament density (0.016) (Table 2 and Figure 6). The minimum IG values were calculated for the distance to rivers and MeRugNu factors (0.003 and 0.002,

respectively). However, the 24 GIFs had a positive information gain, and for that reason, all of them were included in this analysis.

Table 2. Multicollinearity diagnosis, average information gain and parametric statistics (LR model) for the groundwater influencing factors.

<i>Influencing Factors</i>	<i>Information Gain</i>	<i>Collinearity Statistics</i>		<i>LR Model</i>
	<i>Average Merit</i>	<i>Tolerance</i>	<i>VIF</i>	β
<i>Elevation</i>	0.017	0.530	1.886	0.967
<i>Aspect</i>	0.008	0.983	1.017	2.758
<i>Slope</i>	0.012	0.411	2.434	0.021
<i>Curvature</i>	0.004	0.667	1.499	−0.751
<i>Profile Curvature</i>	0.003	0.696	1.436	−0.116
<i>Plan Curvature</i>	0.006	0.773	1.294	0.389
<i>Convergence</i>	0.008	0.754	1.327	−2.774
<i>TWI</i>	0.004	0.907	1.103	−0.081
<i>SPI</i>	0.011	0.471	2.122	0.522
<i>TRI</i>	0.009	0.231	4.323	−0.965
<i>MeRugNu</i>	0.002	0.707	1.414	−0.947
<i>MRRTF</i>	0.020	0.615	1.625	1.687
<i>MRVBF</i>	0.020	0.546	1.832	1.232
<i>Distance to Rivers</i>	0.002	0.977	1.024	1.971
<i>Density of Rivers</i>	0.019	0.899	1.112	2.136
<i>LS</i>	0.010	0.330	3.034	0.749
<i>Distance to Faults</i>	0.029	0.407	2.456	0.510
<i>Density of Faults</i>	0.030	0.503	1.988	1.380
<i>Distance to lineament</i>	0.006	0.975	1.026	1.962
<i>Density of lineament</i>	0.016	0.949	1.054	2.979
<i>Rainfall</i>	0.025	0.829	1.207	1.836
<i>Lithology</i>	0.073	0.741	1.350	1.519
<i>NDVI</i>	0.013	0.681	1.469	2.313
<i>LULC</i>	0.008	0.681	1.469	−0.540
Constant				−12.0977

The rank of the weights of the different GIF classes by the FR analysis shows a positive correlation between the spring potentiality and high values of TWI, since the class 15.39–25.87 holds the highest FR weight (1.811). This class is closely followed by class 10 of lithology relating to Lias limestones, with a value of 1.808. Even if the slope factor has a very low GI value, its class 34.16–76.42 shows a high FR value of 1.792, followed by the class 0.99–24.32 of the curvature factor and the class 19.66–106.67 of the TRI factor. The values of FR which follow are on the order of 1.696, 1.592, 1.583, 1.565 and 1.546, characterizing the classes 3484.68–7467.17 (distance to lineaments factor), Forest (LULC factor), 13.41–26.83 (LS factor), 0.21–0.30 (NDVI factor) and 1379.67–2299.45 (SPI factor), respectively. Finally, the lowest (null) FR values are calculated for the classes 2.30–3.45 (MRRTF), 3.45–5.65 (MRRTF), 4.04–5.88 (MRVBF), 26.83–190.03 (LS), 18,419.36–34,536.29 (distance to faults), (1, 2, 3, 4, 5, 6, 8, 14 and 15) lithology and waterbody (LULC), indicating their minimal effect on groundwater potentiality.

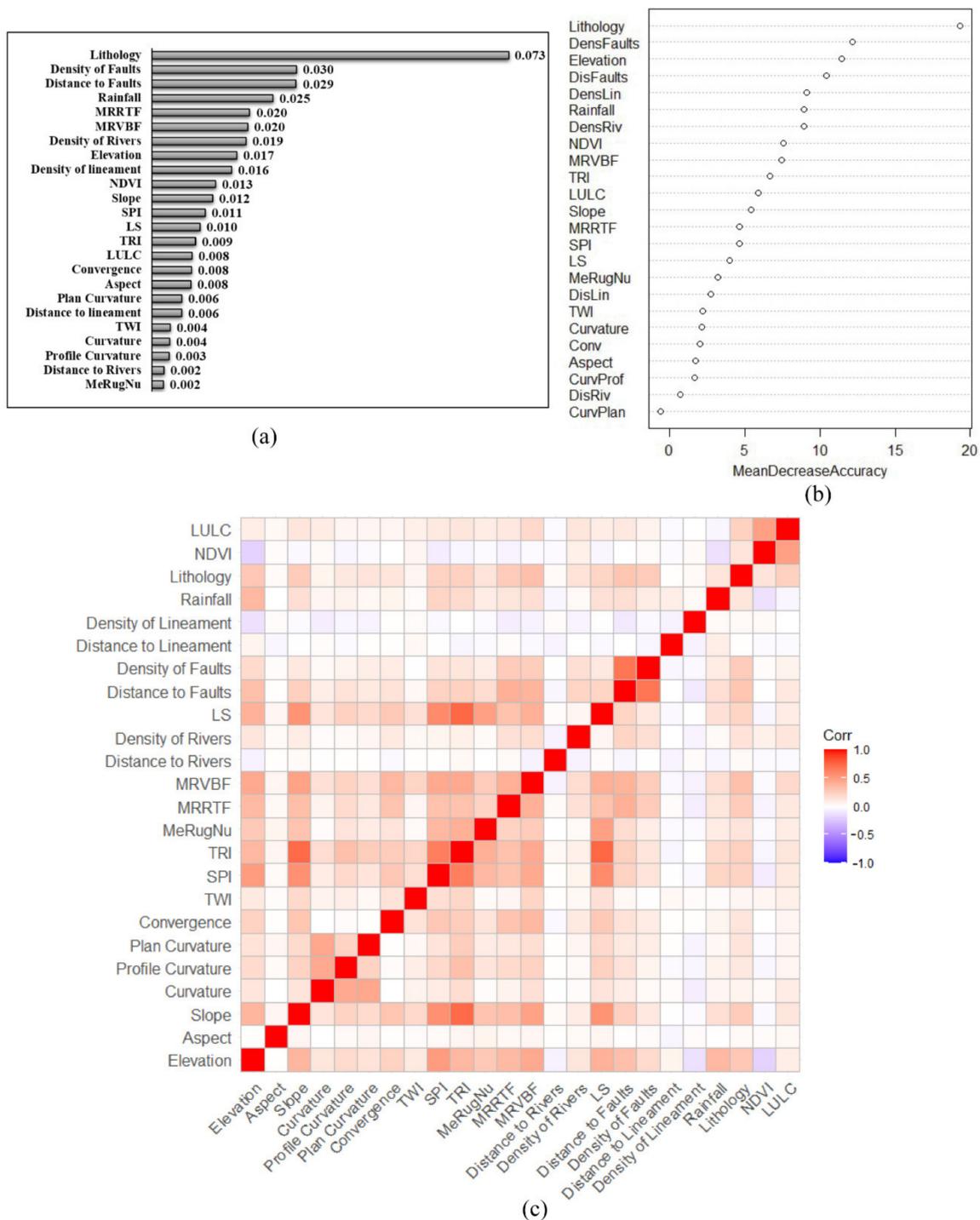


Figure 6. Predictive capabilities using information gain method: (a) importance of GIFs derived from Random Forest model; and (b) correlation matrix of the twenty-four GIFs (c).

Following the developed methodology, the random forest method was employed for estimating the importance of the GPM related factors that were selected by the IG method (Figure 6). The values of RF ranged between -0.598 and 19.313 , with the lowest value corresponding to the variable plane curvature and the highest to variable lithology. In decreasing order of RF importance value, after the lithology factor, we have fault density, elevation, distance to faults, lineament density, rainfall and river density, with values of 12.187 , 11.419 , 10.446 , 9.095 , 8.921 and 8.913 , respectively. The factors of least importance are the distance to rivers, plane curvature, profile curvature, aspect, convergence and curvature.

4.2. Models Building and Hyperparameters Tuning

This is a critical step in modeling using machine learning algorithms. To do this, we randomly subdivided the data Training (75%) and Testing (25%) dataset, then the cross-validation method is applied on the Training data to minimize the splitting error. The findings show that the RF model has an average accuracy of 0.71 and a kappa value of 0.16. The optimization of the RF parameters was applied using a random search and based on the OOB error rate (Figure 7a). Then, accuracy was used to select the optimal model using the largest value (Table 3). The final value used for the RF model was $mtry = 13$. These results confirm that the subdivision of the data does not have a great impact on accuracy, which allowed us to apply the other models without fearing a degradation of the accuracy for the different partitions. For the ANN model, after several tests, the optimum number of hidden layers is equal 2 (Figure 7b).

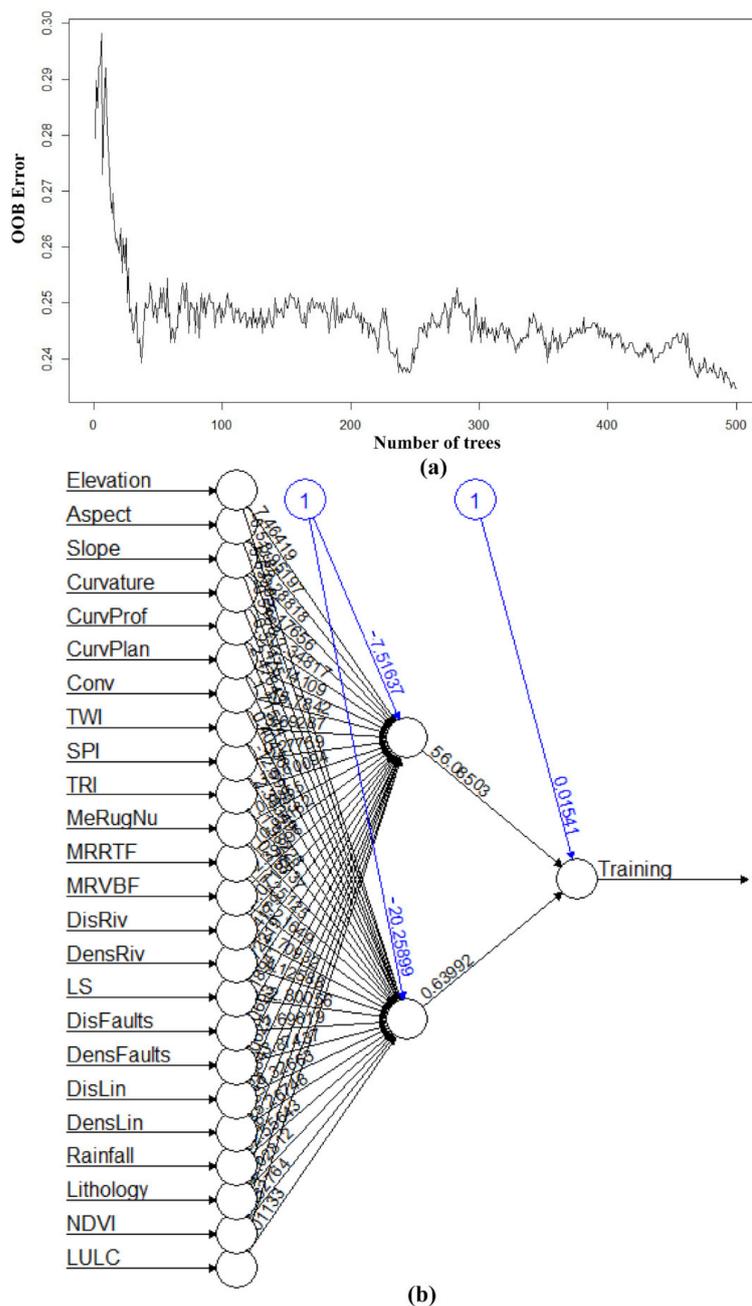


Figure 7. Number of trees optimized based on OOB estimates of the error rate in RF model (a) and diagram of ANN model (b).

Table 3. Resampling results across tuning parameters using RF model. Accuracy was used to select the optimal model using the largest value. The final value used for the model was $mtry = 13$.

	Accuracy	Kappa	Resample
1	0.7321429	0.21052632	Fold02
2	0.6785714	0.10000000	Fold01
3	0.7017544	0.17637059	Fold04
4	0.7079646	0.12196845	Fold03
5	0.6902655	0.09309791	Fold06
6	0.6696429	−0.01369863	Fold05
7	0.7232143	0.19480519	Fold08
8	0.7678571	0.31578947	Fold07
9	0.7142857	0.20000000	Fold10
10	0.7232143	0.25301205	Fold09

4.3. Groundwater Potential Mapping

The main objective of this study is to produce groundwater potentiality maps (GPMs) using individual and ensemble machine learning models. Fifteen GPMs were produced based on the application of four models, both individually and in ensembles: random forest, logistic regression, decision tree and artificial neural network. The GPMs produced were divided into five classes based on Jenk's natural breaks classification method (calibration results). The five classes are very low, low, moderate, high and very high. The results are shown in Figure 8. Accordingly, the spatial distribution are: (1) very low and low classes which are dominant; (2) the areas with the highest potential are located in the middle of the study area, and in the northern parts, especially in the mountains–plains transition zone; and (3) the regions with the least potential are those located in the southern, western and eastern parts of the study area.

For the individual models four groundwater potential maps are shown in Figure 8. The rate of springs in each class showed that all of the training and validation springs were identified in the very high class for the RF, DT and ANN models, with a maximum training value for the RF GPM (100% of springs in the high class) and a maximum validation value identified for the ANN GPM (46.81%). In the case of the GPM based on the LR model (Figure 8b), the high class makes up the major part of the training (35.00%) and validation (32.98%) springs, followed by the moderate class (25.36% for training and 26.60% for validation).

Following the completion of the GPMs for the different models individually, the different possible combinations were tested. The first sets of models tested involved two models, so six ensembles were considered: RF-LR, RF-DT, RF-ANN, LR-DT, LR-ANN and DT-ANN. The GPMs that were derived and the statistical results of the spatial distribution of the different classes are shown in Figure 8. On first examination, it is clear that the very low and low classes cover the major part of the territory; the percentages of the very low class vary from 47.92% for the DT-ANN ensemble (the maximum value) to 25.01% for RF-LR (the minimum value). For the low classes, the values vary from 32.23% (RF-DT) to 15.70% (DT-ANN). On the other hand, the area covered by the high and very high classes represent the lowest percentages for all model ensembles. From the point of view of the spring percentage, it has been observed that the training springs largely cover the high and very high classes. Indeed, except for the LR-ANN ensemble, where the high class holds the major part of the training springs (33.93%), all of the other ensembles show very high class which contains the most important percentages of training springs: 92.14%, 86.07%, 93.21%, 50.36% and 67.14% for RF-LR, RF-DT, RF-ANN, LR-DT and DT-ANN, respectively. To finish, the highest percentage of validation springs are located at the very high-class level, with 30.85% for RF-DT, 41.49% for RF-ANN, 26.60% for LR-DT, 29.79% for LR-ANN and 35.11% for DT-ANN. Only for the RF-LR ensemble does the moderate class contain the majority of the validation springs (32.98%), followed by the very high class (31.91%).

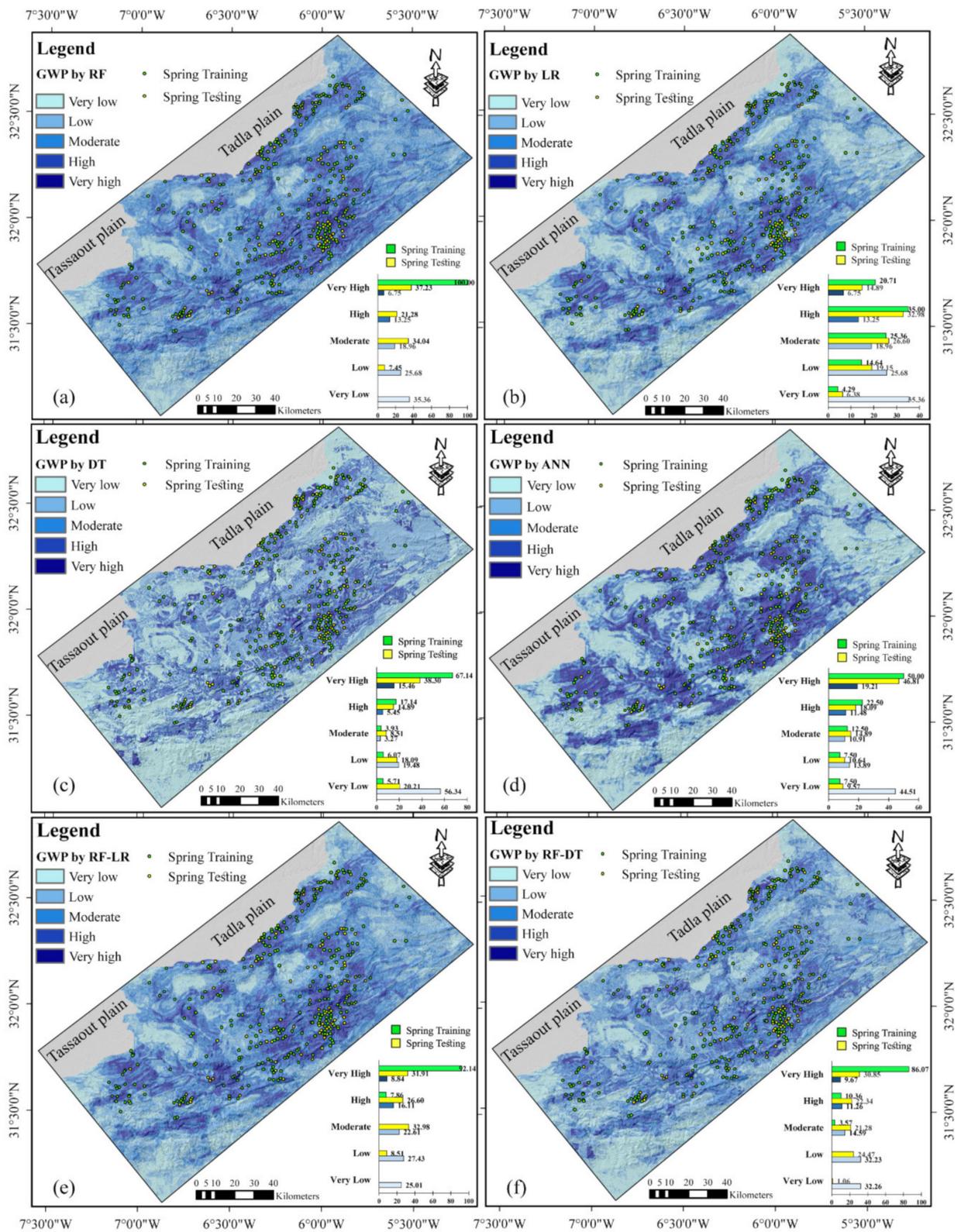


Figure 8. Cont.

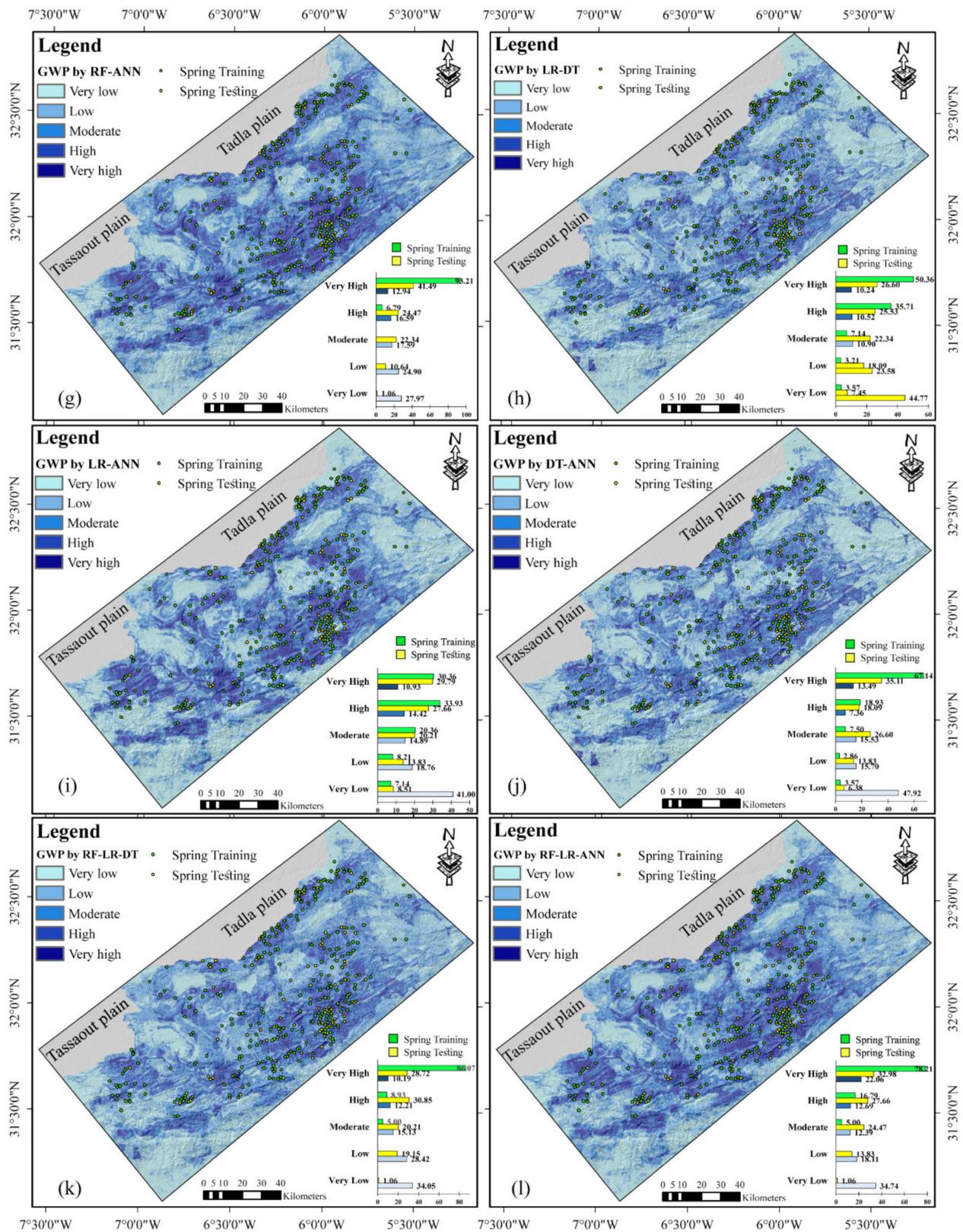


Figure 8. Cont.

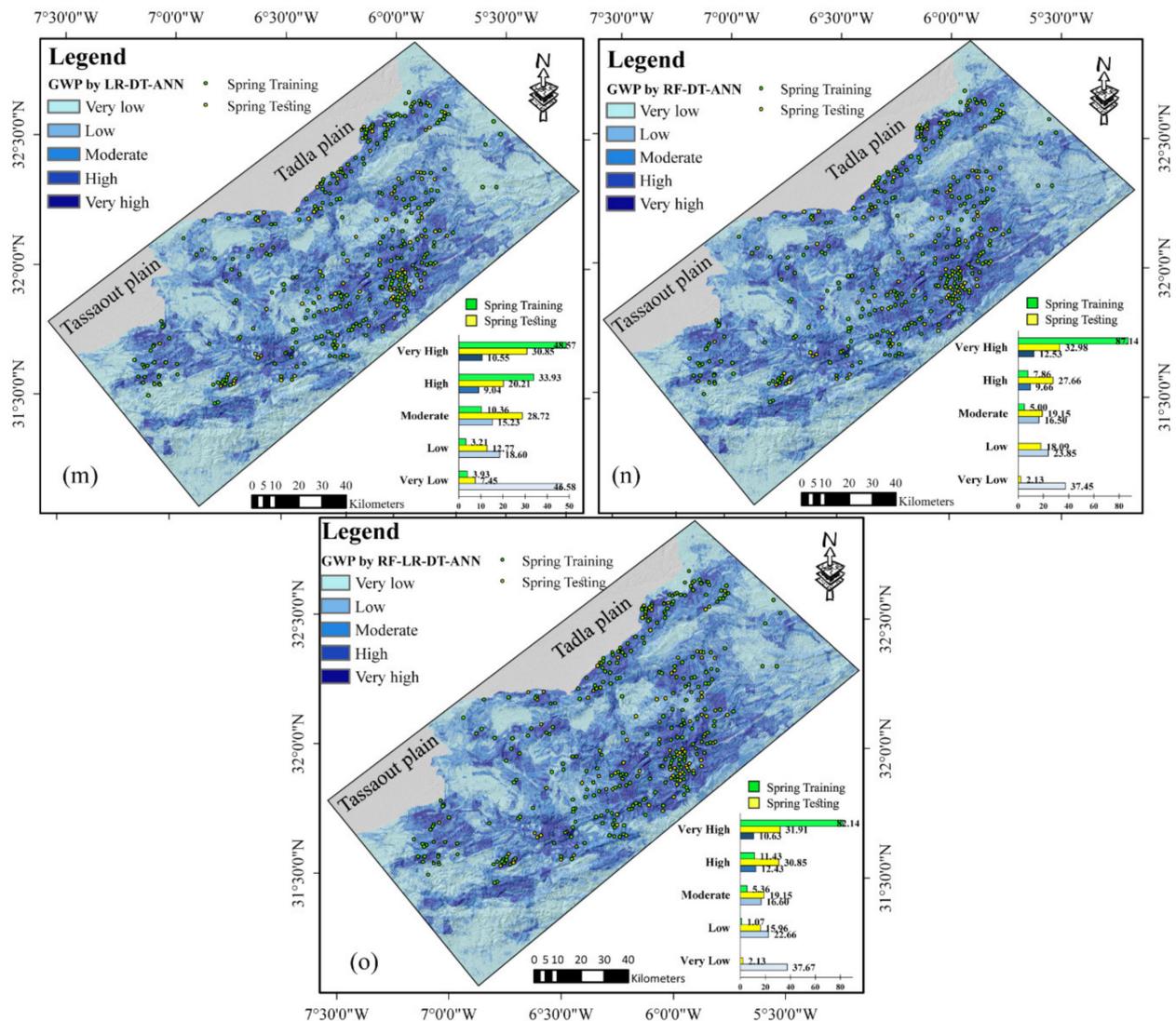


Figure 8. GSPM obtained from RF (a), LR (b), DT (c), ANN (d), RF-LR (e), RF-DT (f), RF-ANN (g), LR-DT (h), LR-ANN (i), DT-ANN (j), RF-LR-DT (k), RF-LR-ANN (l), RF-DT-ANN (m), LR-DT-ANN (n) and RF-LR-DT-ANN (o) models.

Combining three models at the same time has allowed to produce four new GPMs: RF-LR-DT, RF-LR-ANN, RF-DT-ANN and LR-DT-ANN. The results and the spatial distribution of the different classes of potentiality are presented in Figure 8. Generally, the percentages covered by the different potentiality classes' decrease from the very low class to the very high class. From the point of view of the spatial distribution of the springs, all ensembles of three models show that the training springs are mostly located in the very high class, with values of 87.14%, 86.07%, 78.21% and 48.57% for the RF-DT-ANN, RF-LR-DT, RF-LR-ANN and LR-DT-ANN ensembles, respectively. For the validation springs, the very high classes of the RF-LR-ANN and RF-DT-ANN ensembles contain the highest percentage (32.98%), followed by the LR-DT-ANN ensemble (30.85%) and the high class (30.85%) of the RF-LR-DT ensemble.

Finally, a set of all four models was used to produce one more GPM. The map produced and the statistics relating to the spaces occupied by the different potentiality classes are shown in Figure 8. Given the spatial distribution, very low is the dominant class, with a percentage of 37.67%, followed by the low class (22.66%), the moderate class (16.60%), the high class (12.43%) and finally, the very high class (10.63%). Inversely, the percentage of training and validation springs increases from the very low to the very high classes.

4.4. Performance Metrics and Comparison

In this part, a comparison was realized to test values of success and prediction rate 25/75%, 50/50% and 75/25% (Table 4). The table reveals a considerable stability in the overall performance in all tests (25%, 50% and 75%) except the prediction rate of LR and DT, where a slight instability was observed. For other individual or ensemble models, the success and prediction rates increase generally from 25% to 75% for the sample.

Table 4. Gain in success and prediction rate using the partition Training/Testing progresses from 25/75% to 75/25% with redundant factors elimination.

Models	25/75% of Overall Sample		50/50% of Overall Sample		75/25% Of Overall Sample		Elimination of Redundant Factors	
	Success Rate	Prediction Rate	Success Rate	Prediction Rate	Success Rate	Prediction Rate	Success Rate	Prediction Rate
RF	1.000	0.719	1.000	0.767	1.000	0.786	1.000	0.780
LR	0.729	0.780	0.780	0.755	0.784	0.744	0.781	0.746
DT	0.939	0.612	0.911	0.596	0.964	0.738	0.946	0.609
ANN	0.788	0.693	0.776	0.743	0.784	0.744	0.782	0.743
RF-LR	0.999	0.773	0.999	0.775	0.999	0.779	0.999	0.773
RF-DT	0.996	0.722	0.997	0.742	0.998	0.787	0.997	0.725
RF-ANN	0.999	0.722	1.000	0.760	1.000	0.775	1.000	0.770
LR-DT	0.920	0.759	0.904	0.739	0.949	0.779	0.931	0.702
LR-ANN	0.783	0.753	0.796	0.750	0.794	0.749	0.784	0.746
DT-ANN	0.922	0.697	0.919	0.714	0.954	0.771	0.933	0.700
RF-LR-DT	0.995	0.757	0.995	0.758	0.996	0.790	0.995	0.732
RF-LR-ANN	0.985	0.757	0.981	0.766	0.983	0.772	0.980	0.765
RF-DT-ANN	0.995	0.723	0.995	0.748	0.997	0.789	0.996	0.730
LR-DT-ANN	0.915	0.745	0.905	0.739	0.943	0.780	0.920	0.713
RF-LR-DT-ANN	0.987	0.750	0.984	0.756	0.989	0.791	0.985	0.738

The success rate of all samples is stable. In order to verify the obtained results, the GPMs and the spring inventory locations were compared (Figure 8). For the four individual models, we see that the majority of the validation and training springs fall into the high and very high susceptibility classes. In addition, the very low susceptibility class either has very weak or no spring occurrence in all GPMs. Even better, for the model sets, the results show an increase in the percentages of validation and training springs in high and very high classes, especially for RF-LR, DT-ANN and RF-ANN. It is clear from these results that the field-recorded spring locations have a better fit with the RF, DT, ANN, RF-LR, DT-ANN and RF-ANN maps than with the other GPMs.

The results of validation techniques and accuracy prioritization based on the training datasets are shown in Table 5, and the results based on the testing datasets are shown in Table 6. Ten parameters have been calculated to improve the comparison between the models: accuracy, sensitivity, specificity, precision, FP-Rate, MCC, KAPPA, AUC, RMSE and MAE. For the training datasets, the accuracy ranges from 0.754 to 1.000; the highest value is identified for the RF model, and the lowest is from the LR-ANN and ANN models. Like accuracy, the maximum sensitivity value (1.000) was found for the RF individual model and the minimum for the LR model (0.257). Specificity varies between 0.797 (ANN model) to 1.000 (RF and RF-ANN models). For precision, the values range from 0.311 (ANN) to 1.000 (RF). The FP-Rate calculation indicates that the best result is always for the RF model (0.334), and the worst result is for LR (0.086). In addition, for the other parameters, RF is better, with values of 1.000 for Kappa and AUC. The minimum values are Kappa = 0.233 (LR-ANN) and AUC = 0.784 (LR and ANN models). For the reliability, which has been assessed by applying the MAE and RMSE methods, the training results indicate minimum values for the RF model (0.000) and maximum values for LR (0.137) and ANN (0.496).

Table 5. Results of validation techniques and accuracy prioritization based on training datasets.

	<i>Training</i>																		<i>Rank Total</i>				
	Ac	R	Sen	R	Sp	R	Pr	R	FPR	R	MCC	R	Ka'1	R	AUC	R	MAE	R	RMSE	R	RT	CF	PR
<i>RF</i>	1.000	1	1.000	1	1.000	1	1.000	1	0.334	1	1.000	1	1.000	1	1.000	1	0.000	1	0.000	1	10	1	1
<i>LR</i>	0.765	13	0.257	15	0.935	13	0.567	13	0.086	15	0.262	13	0.234	14	0.784	14	0.137	15	0.485	13	138	13.8	13
<i>DT</i>	0.942	8	0.861	5	0.969	11	0.903	12	0.287	5	0.843	8	0.843	8	0.964	9	0.012	2	0.241	9	77	7.7	7
<i>ANN</i>	0.754	15	0.512	13	0.797	15	0.311	15	0.092	13	0.256	15	0.244	13	0.784	14	0.098	13	0.496	15	141	14.1	15
<i>RF-LR</i>	0.956	6	0.825	7	1.000	1	1.000	1	0.275	7	0.883	6	0.876	6	0.999	3	0.044	10	0.209	7	54	5.4	6
<i>RF-DT</i>	0.968	2	0.886	2	0.995	6	0.984	6	0.295	2	0.914	2	0.911	2	0.998	4	0.025	3	0.179	2	31	3.1	2
<i>RF-ANN</i>	0.963	5	0.850	6	1.000	1	1.000	1	0.283	6	0.900	5	0.895	5	1.000	1	0.038	7	0.194	5	42	4.2	4
<i>LR-DT</i>	0.929	9	0.793	9	0.974	9	0.910	9	0.264	10	0.804	9	0.801	9	0.949	11	0.038	8	0.194	6	89	8.9	9
<i>LR-ANN</i>	0.763	14	0.264	14	0.929	14	0.552	14	0.088	14	0.257	14	0.233	15	0.794	13	0.130	14	0.487	14	140	14	14
<i>DT-ANN</i>	0.929	10	0.793	10	0.974	9	0.910	9	0.264	10	0.804	10	0.801	10	0.954	10	0.032	6	0.267	10	94	9.4	10
<i>RF-LR-DT</i>	0.967	3	0.882	3	0.995	6	0.984	6	0.294	3	0.911	3	0.909	3	0.996	6	0.026	4	0.182	3	40	4	3
<i>RF-LR-ANN</i>	0.910	11	0.646	11	0.998	4	0.989	5	0.215	11	0.754	11	0.728	11	0.983	8	0.087	12	0.300	11	95	9.5	11
<i>RF-DT-ANN</i>	0.966	4	0.871	4	0.998	4	0.992	4	0.290	4	0.909	4	0.906	4	0.997	5	0.030	5	0.184	4	42	4.2	5
<i>LR-DT-ANN</i>	0.893	12	0.639	12	0.977	8	0.904	11	0.213	12	0.700	12	0.683	12	0.943	12	0.073	11	0.327	12	114	11.4	12
<i>RF-LR-DT-ANN</i>	0.950	7	0.814	8	0.955	12	0.983	8	0.271	8	0.865	7	0.859	7	0.989	7	0.043	9	0.224	8	81	8.1	8

Table 6. Results of validation techniques and accuracy prioritization based on testing dataset.

	<i>Testing</i>																		<i>Rank Total</i>				
	Ac	R	Sen	R	Sp	R	Pr	R	FPR	R	MCC	R	Ka	R	AUC	R	MAE	R	RMSE	R	RT	CF	PR
RF	0.753	7	0.213	14	0.933	5	0.513	6	0.071	14	0.207	15	0.181	15	0.786	5	0.141	9	0.497	8	98	9.8	15
LR	0.756	5	0.223	11	0.933	6	0.525	5	0.074	11	0.220	13	0.193	14	0.744	13	0.143	10	0.494	6	94	9.4	14
DT	0.737	11	0.511	2	0.813	13	0.475	8	0.170	1	0.316	1	0.316	2	0.738	15	0.019	3	0.512	12	68	6.8	3
ANN	0.756	6	0.520	1	0.792	15	0.277	15	0.080	9	0.245	8	0.227	7	0.744	14	0.117	8	0.494	6	89	8.9	11
RF-LR	0.761	3	0.234	10	0.936	4	0.550	4	0.078	10	0.240	10	0.211	10	0.779	7	0.143	11	0.489	3	72	7.2	6
RF-DT	0.737	12	0.468	4	0.827	11	0.473	9	0.155	3	0.296	3	0.296	4	0.787	4	0.003	1	0.512	12	63	6.3	1
RF-ANN	0.761	4	0.213	14	0.943	2	0.556	3	0.071	15	0.230	11	0.197	12	0.775	9	0.154	13	0.489	3	86	8.6	9
LR-DT	0.735	13	0.500	3	0.813	14	0.470	11	0.166	2	0.306	2	0.306	3	0.779	7	0.154	14	0.489	3	72	7.2	7
LR-ANN	0.764	2	0.223	11	0.943	3	0.568	2	0.074	12	0.243	9	0.209	11	0.749	12	0.151	12	0.486	2	76	7.6	8
DT-ANN	0.729	15	0.436	5	0.827	12	0.456	14	0.145	4	0.267	4	0.267	5	0.771	11	0.010	2	0.520	15	87	8.7	10
RF-LR-DT	0.748	8	0.351	7	0.880	8	0.493	7	0.117	6	0.261	5	0.255	6	0.790	2	0.072	6	0.502	9	64	6.4	2
RF-LR-ANN	0.767	1	0.223	11	0.947	1	0.583	1	0.074	13	0.251	6	0.215	9	0.772	10	0.154	15	0.483	1	68	6.8	4
RF-DT-ANN	0.735	14	0.340	8	0.866	9	0.457	13	0.113	7	0.229	12	0.225	8	0.789	3	0.064	5	0.515	14	93	9.3	13
LR-DT-ANN	0.740	9	0.362	6	0.866	10	0.472	10	0.120	5	0.250	7	0.339	1	0.780	6	0.058	4	0.510	10	68	6.8	5
RF-LR-DT-ANN	0.740	10	0.277	9	0.894	7	0.464	12	0.092	8	0.208	14	0.197	13	0.791	1	0.101	7	0.510	10	91	9.1	12

In the case of testing datasets (Table 6), validation results indicate that the RF-LR-ANN ensemble is the best performing model in terms of accuracy (0.767), specificity (0.947), precision (0.583) and RMSE (0.483). For sensitivity, the maximum value is identified for the ANN model (0.520), and for the FP-Rate and MCC parameters, the best results were calculated for DT (FPR = 0.170, MCC = 0.316). The calculated Kappa index values show a maximum value of 0.339 for the ensemble of three LR-DT-ANN models, and the minimum value of the MAE parameter was calculated for the RF-DT ensemble model (MAE = 0.003). Finally, the most efficient model in terms of AUC was the ensemble of four models, RF-LR-DT-ANN (AUC = 0.791).

Additionally, the estimation of prediction capability for the fifteen models is obtained by comparing the spring training and validation inventories with the GPMs. Then, the rate curves were created (ROC), and the areas under each curve (AUCs) were calculated (Figure 9). For training datasets, the prediction-rate curve showed that the maximum AUC values were 1.000, 0.999 and 0.998 for the RF and RF-ANN, RF-LR and RF-DT models, respectively. Moreover, in the prediction-rate curve obtained by comparing the spring validation data with GPMs, it was observed that all models present tolerable performance for groundwater potentiality mapping (AUC > 0.7). The RF-LR-DT-ANN ensemble model achieved the best performance (AUC = 0.791), followed by the RF-LR-DT ensemble model (AUC = 0.790), the RF-DT-ANN model (AUC = 0.789) and the RF-DT model (AUC = 0.787).

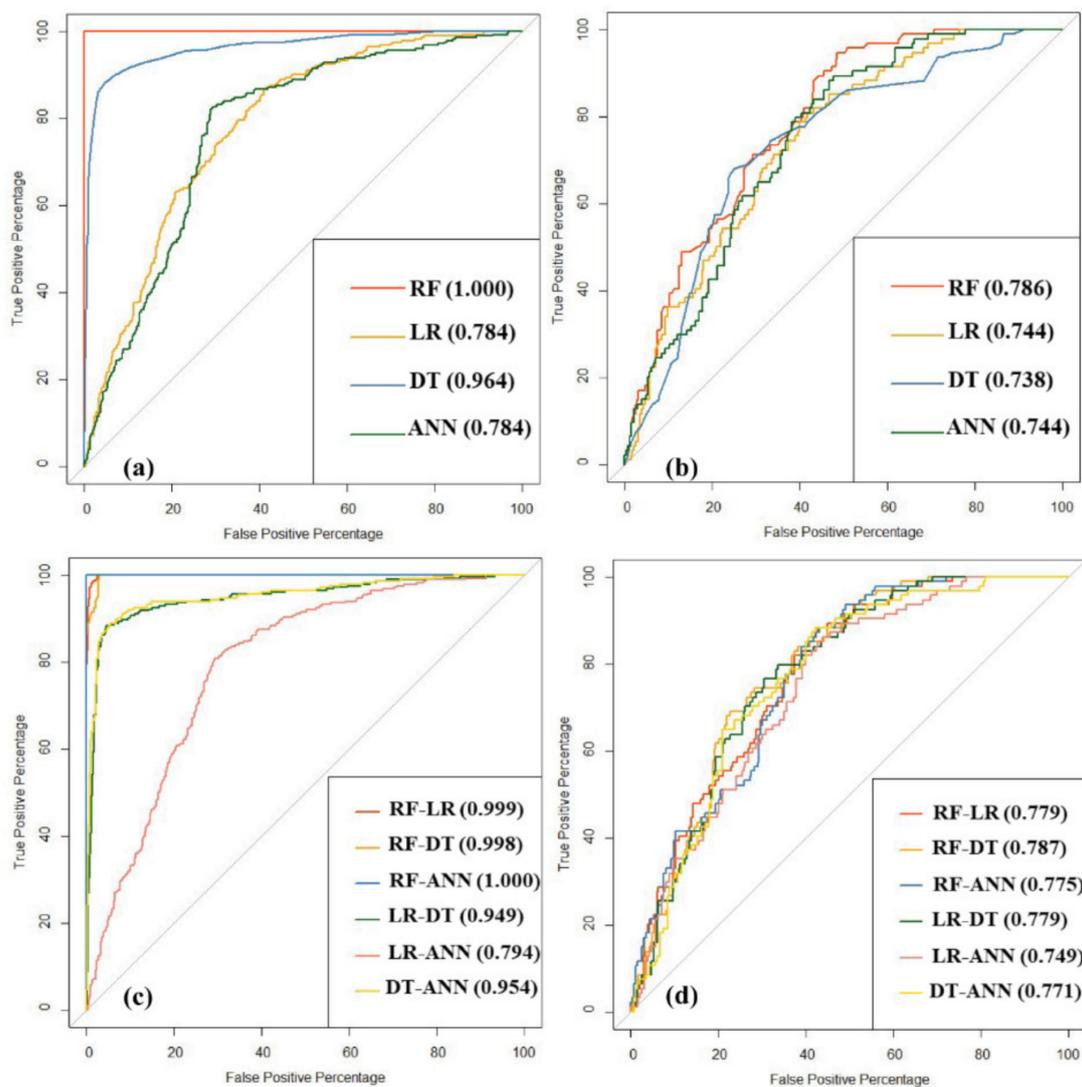


Figure 9. Cont.

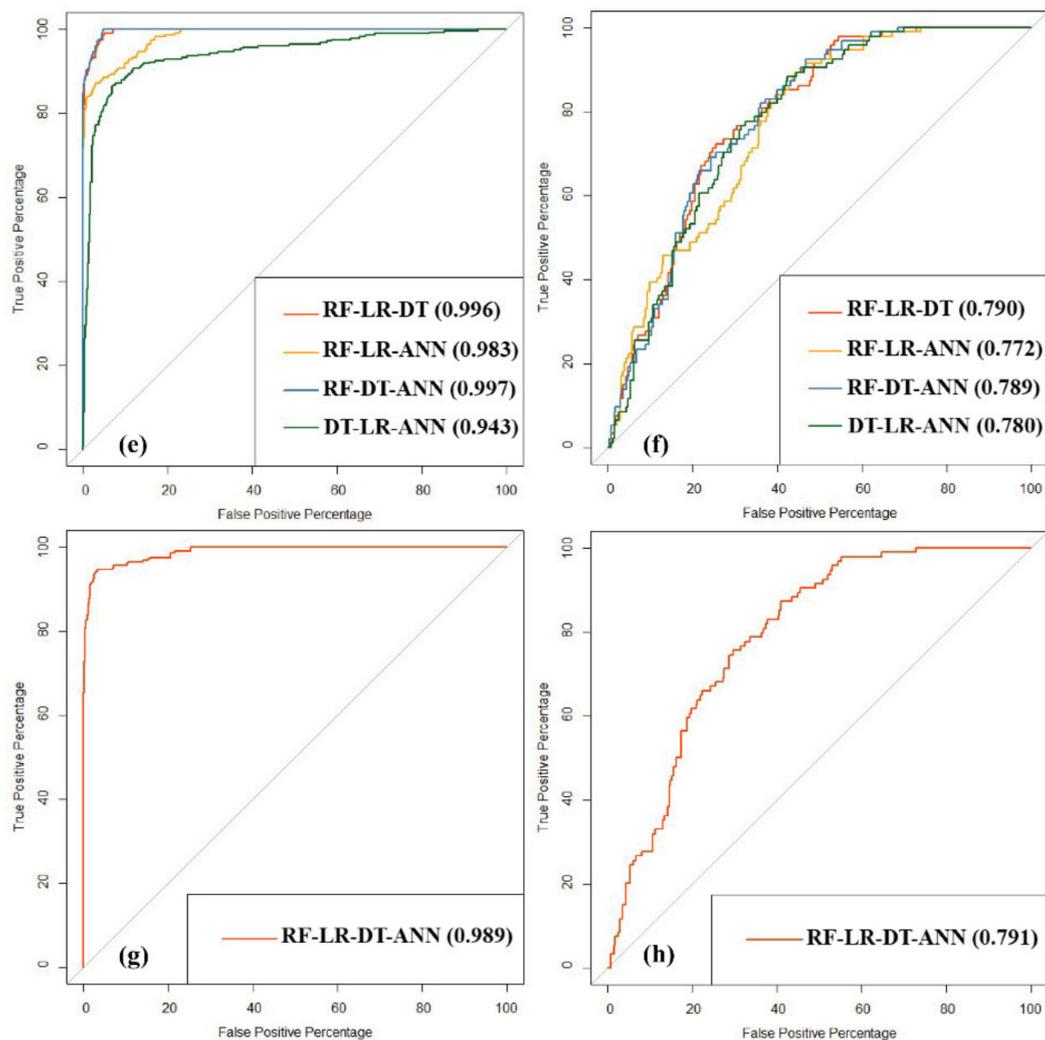


Figure 9. ROC curve: (a) success rate of individual models, (b) prediction rate of individual models, (c) success rate of ensemble of two models, (d) prediction rate of ensemble of two models, (e) success rate of ensemble of three models, (f) prediction rate of ensemble of three models, (g) success rate of ensemble of four models and (h) prediction rate of ensemble of four models.

Furthermore, we have calculated the prioritization rank based on all evaluation criteria; the results are shown in Tables 5 and 6 and Figure 10. The prioritization results by compound factor (CF) analysis using the training GPMs of all models found that the RF and RF-DT models had the best success rates; it ranked RF at 1 and RF-DT at 2. Moreover, for prediction aptitude using testing GPMs, the prioritization analysis indicates that the best models are RF-DT and RF-LR-DT. Taking into account the results in terms of success and prediction rate, the two best models for mapping the groundwater potential in our mountainous study area are RF-DT, followed by RF-LR-DT.

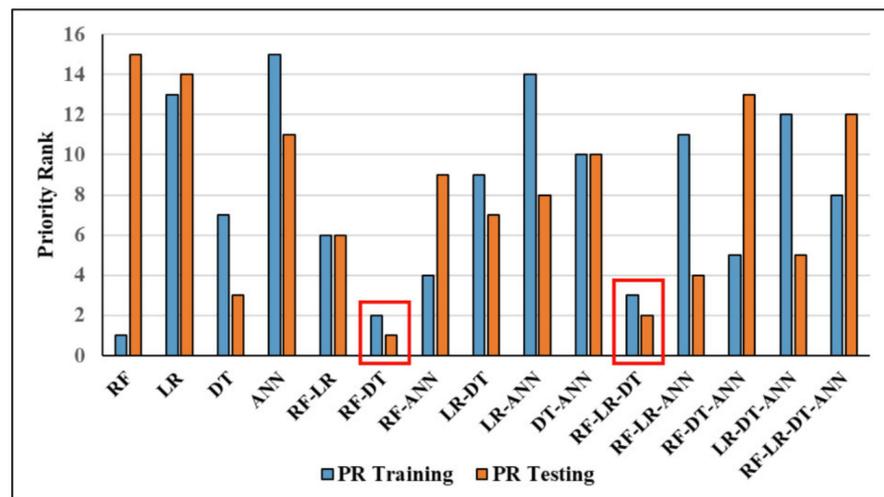


Figure 10. Prioritization results of all models using compound factor (CF) analysis used in the training and testing processes.

5. Discussion

The discussion will focus on three main points: (1) the analysis and selection of GIFs; (2) role of factors in the occurrence of karst springs, and (3) the performance of individual and ensemble machine learning algorithms in GWP mapping in large-scale areas.

5.1. GIF Selection and Analysis

According to IG analysis, lithology, fault density, distance to faults and rainfall were identified as factors that were highly predictive of the presence of groundwater; moreover, MeRugNu, distance to rivers, profile curvature and curvature were identified as the least predictive. This seems logical given the very important role of the rock type in the genesis and recharge of aquifers, in addition to the structuring factors (faults and fractures) that facilitate recharge and the emergence of springs at the same time. According to FR analysis results, groundwater potentiality is more likely to be found within Liasic limestone areas, in areas with a high wetness index and within steep slopes. These last two factors underline the importance of topographic control on hydrological processes and on GPMs in mountain areas in particular. The results, showing the importance of high TWI values in the prediction of groundwater potential.

As in the IG factor rankings, geological factors were classified as the best predictor variables of GPM in our study according to the RF importance value. The importance given to the elevation factor in this study may be related to the fact that the lower elevation zones correspond to the Tadla and Tassaout plains, where the emergence of springs is very rare due to their low slopes (<10%). Elevation is a key factor in the production of groundwater potential maps in our region and in other mountainous regions around the world. The rainfall factor also remains a very good predictor, since it is the source of aquifer recharge in addition to the snow cover that characterizes these mountainous areas, especially during the winter and spring seasons [28], which is in line with several previous studies [49,50]. Another factor that presents a robustness of prediction under the RF importance calculation is the river density, since rivers represent an important source of exchange between aquifer systems and surface hydrology. The importance of the predictors such as rainfall and river density in GPM has been seen in other semi-arid regions around the world [11,51].

According to the results obtained from this study on GIF selection and analysis, it is recommended to take into account many factors as possible in the analysis and in the modelling processes, especially in regional studies where GIFs vary spatially and influence groundwater potentiality from one location to another.

5.2. Role of Factors in the Occurrence of Karst Springs

Geological factors were identified as the top predictor variables of GPM based on the RF importance value and IG results. Certainly, lithology is the most relevant factor, followed by fault density and distance to faults. Indeed, the karstic environment is largely composed of the Lias limestones [22] which explains the importance of faults and lineaments in the karstification process. These findings are in accordance with results reported by previous studies in similar geological contexts, which detect higher groundwater potentiality in limestone-dominated karstic areas with severe fracturing [52].

According to the results of this study in the Beni Mellal, Atlas Mountains, the factors controlling the karst springs development can be divided into four categories: lithology, geological structure, topography and climate condition: The liasic and Middle Jurassic limestone contains 85% of the inventoried springs in the research region. The remaining springs are found in Middle Jurassic and Pliocene–Pleistocene continental and alluvium deposits. This confirms the important role of carbonated lithology in the emergence of these springs by the effect of underground dissolution by rainwater. In addition, near to the faults the probability of emerging sources increases, the same result is observed in areas where the density of the faults is important. Regarding the topographic factors, the altitude controls the most of this phenomenon, which mainly explains the emergence of most springs in the transition area between the mountains and the Tadla plain (Dir). Finally, the non-homogeneous appearance of the springs in this large mountain region appears strongly controlled by the spatial variation of the precipitation, this will be mainly related to the importance of the rains on the northern slopes compared to those of the south.

5.3. Machine Learning Algorithm Performance

Generally, the use of the RF, LR, DT and ANN individual learning machine models has given good results, even though there is little stability between the success and the prediction rate of some models (e.g., RF and DT models). Indeed, the RF and DT models have a very high success rate (AUC = 1.000 for RF and AUC = 0.964 for DT), but the prediction rate has decreased significantly for both models (AUC = 0.786 for RF and AUC = 0.738 for DT).

In order to improve the performance and prediction rates of the models, combinations of these four models were used. Different combinations were tested, and the prioritization rank method was used to select the best models for groundwater potentiality mapping in mountainous areas by integrating the maximum number of influencing factors in the analysis. Thus, a significant difference between the individual models and ensembles of models based on the predictive performance can be perceived. Indeed, the average prediction rate using AUC values showed an interesting progression. The average prediction rate for the four individual models was 0.753; then, it increased to 0.773 for the ensembles of two models. Then, it increased to 0.783 on average for the ensembles of three models, and finally, the AUC recorded its maximum value, which was reached by the ensemble based on all four models, RF-LR-DT-ANN (0.791). This is confirmed by the compound factor method, which has allowed us to draw up a general ranking of individual models and ensembles of models using several statistics metrics applied to the training and validation datasets. The results of the prioritization clearly indicate that the ensembles of models improve performance and reduce some errors related to data preparation or modelling processes. The best set of models for groundwater potentiality mapping in our study are the RF-DT and RF-LR-DT ensembles. This is also the situation in several studies that compare the predictive performance of either optimized, hybrid or ensemble models. This is also confirmed by previous studies [53] that indicated that hybrid models show better accuracies than individual models. However, in our study, while the majority of the model ensembles performed better, others were marked by a decrease in performance: for example, LR-ANN's success rate and RF-DT-ANN's prediction rate. This requires us to make the maximum number of groupings possible in order to select the best ensemble of models.

6. Conclusions

In the present study, a wide variety of methodologies were applied based on GIS, remote sensing and the use of individual and ensemble machine learning algorithms to evaluate GP in large-scale mountainous areas. In addition, the novel aspect of this study is that we have tried to integrate many of groundwater potentiality influencing variables as possible, which include geological, topographical, hydrological, climatic and land cover factors; 24 factors have been considered. Then, after a test of multicollinearity and an information gain calculation, all of the factors were retained to produce groundwater potentiality maps. In the same way, the importance of GIFs has been estimated from the random forest method. Lithology represents the factor that most influences groundwater potentiality in our karstic zone, followed by tectonic factors (faults and lineaments) and a climatic factor (rainfall). Alternatively, several machine learning algorithms were used for GP mapping. The RF, LR, DT and ANN models were chosen due to their satisfactory results in other regions of the world. The application of individual models indicates that RF represented the best model in terms of success and prediction rate. To improve the performance and robustness of the prediction, ensemble models based on combinations of the RF, LR, DT and ANN models were developed to investigate their capability to predict groundwater potentiality in our large-scale mountainous area. According to the results, the ensembles of models showed better performance, especially from a prediction rate point of view. The AUC recorded an interesting progression: the maximum value was calculated for the RF-LR-DT-ANN ensemble model, with AUC = 0.791. Furthermore, to test the performance and reliability of different models and ensembles of models, several statistics metrics were applied, and a prioritization rank was carried out based on the compound factor method. Consequently, the best results were obtained for the RF-DT and RF-LR-DT ensembles. Finally, the methodology developed in this study may be useful for detecting groundwater potential zones, especially in mountainous areas with difficult access and where the application of geophysical methods of exploration remains costly and difficult to initiate for very large areas.

Author Contributions: Conceptualization, M.N., M.H., B.P., A.A., C.-W.L. and T.T.; methodology, M.N., M.H., B.P., A.E. and S.K.; software, M.N., B.P. and M.E.; validation, M.N. and H.E. (Hicham Elhimer); formal analysis, M.N., A.E., M.E., S.K. and H.E. (Hicham Elhimer); investigation, M.N., M.H., M.E. and T.T.; resources, M.H., H.E. (Hasna Eloudi) and M.O.; data curation, M.N., M.H., A.E., M.E., S.K., M.O. and H.E. (Hicham Elhimer); writing—original draft preparation, M.N., M.H., B.P., C.-W.L., S.K., H.E. (Hasna Eloudi), M.O., and T.T.; writing—review and editing, M.N., M.H., B.P., A.A., C.-W.L., and T.T.; visualization, A.A., A.E., M.E., H.E. (Hicham Elhimer) and H.E. (Hasna Eloudi); supervision, B.P., A.A. and T.T.; project administration, B.P., A.A. and T.T.; funding acquisition, B.P., C.-W.L. and A.A.; All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), University of Technology Sydney (UTS). This research was also supported by grants from the National Research Foundation of Korea, provided by the Korea government (No. 2019R1A2C1085686) and King Saud University, Riyadh, Saudi Arabia (grant number RSP-2021/14).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors express their gratitude to the Hydraulic Basin Agency of Haouz and Oum Er-Rabia for their help and collaboration.

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

References

1. Price, M.F.; Byers, A.C.; Friend, D.A.; Kohler, T.; Price, L.W. (Eds.) *Mountain Geography: Physical and Human Dimensions*; University of California Press: Berkeley, CA, USA, 2013.
2. Kohler, T.; Giger, M.; Hurni, H.; Ott, C.; Wiesmann, U.; Von Dach, S.W.; Maselli, D. Mountains and Climate Change: A Global Concern. *Mt. Res. Dev.* **2010**, *30*, 53–55. [[CrossRef](#)]
3. FAO. *State of the World's Forests 2003*; FAO: Rome, Italy, 2003.
4. Bouchaou, L.; Michelot, J.; Qurtobi, M.; Zine, N.; Gaye, C.; Aggarwal, P.; Marah, H.; Zerouali, A.; Taleb, H.; Vengosh, A. Origin and residence time of groundwater in the Tadla basin (Morocco) using multiple isotopic and geochemical tools. *J. Hydrol.* **2009**, *379*, 323–338. [[CrossRef](#)]
5. Bouchaou, L.; Michelot, J.; Vengosh, A.; Hsissou, Y.; Qurtobi, M.; Gaye, C.; Bullen, T.; Zuppi, G. Application of multiple isotopic and geochemical tracers for investigation of recharge, salinization, and residence time of water in the Souss–Massa aquifer, southwest of Morocco. *J. Hydrol.* **2008**, *352*, 267–287. [[CrossRef](#)]
6. Bouimouass, H.; Fakir, Y.; Tweed, S.; Leblanc, M. Groundwater recharge sources in semiarid irrigated mountain fronts. *Hydrol. Process.* **2020**, *34*, 1598–1615. [[CrossRef](#)]
7. Rathay, S.; Allen, D.; Kirste, D. Response of a fractured bedrock aquifer to recharge from heavy rainfall events. *J. Hydrol.* **2018**, *561*, 1048–1062. [[CrossRef](#)]
8. Voekler, H.; Allen, D.M. Estimating regional-scale fractured bedrock hydraulic conductivity using discrete fracture network (DFN) modeling. *Hydrogeol. J.* **2012**, *20*, 1081–1100. [[CrossRef](#)]
9. Shenga, Z.D.; Baroková, D.; Šoltész, A. Numerical modeling of groundwater to assess the impact of proposed railway construction on groundwater regime. *Pollack Period.* **2018**, *13*, 187–196. [[CrossRef](#)]
10. 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](#)]
11. 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](#)]
12. 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](#)]
13. 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](#)] [[PubMed](#)]
14. Ganguly, K.K.; Nahar, N.; Hossain, B.M.M. A machine learning-based prediction and analysis of flood affected households: A case study of floods in Bangladesh. *Int. J. Disaster Risk Reduct.* **2019**, *34*, 283–294. [[CrossRef](#)]
15. Rizeei, H.M.; Pradhan, B.; Saharkhiz, M.A.; Lee, S. Groundwater aquifer potential modeling using an ensemble multi-adoptive boosting logistic regression technique. *J. Hydrol.* **2019**, *579*, 124172. [[CrossRef](#)]
16. Hssaisoune, M.; Bouchaou, L.; Sifeddine, A.; Bouimetarhan, I.; Chehbouni, A. Moroccan Groundwater Resources and Evolution with Global Climate Changes. *Geosciences* **2020**, *10*, 81. [[CrossRef](#)]
17. Milewski, A.; Seyoum, W.M.; Elkadiri, R.; Durham, M. Multi-Scale Hydrologic Sensitivity to Climatic and Anthropogenic Changes in Northern Morocco. *Geosciences* **2019**, *10*, 13. [[CrossRef](#)]
18. Ouatiki, H.; Boudhar, A.; Ouhinou, A.; Arioua, A.; Hssaisoune, M.; Bouamri, H.; Benabdelouahab, T. Trend analysis of rainfall and drought over the Oum Er-Rbia River Basin in Morocco during 1970–2010. *Arab. J. Geosci.* **2019**, *12*, 128. [[CrossRef](#)]
19. Barakat, A.; Hilali, A.; El Baghdadi, M.; Touhami, F. Assessment of shallow groundwater quality and its suitability for drinking purpose near the Béni-Mellal wastewater treatment lagoon (Morocco). *Hum. Ecol. Risk Assess. Int. J.* **2020**, *26*, 1476–1495. [[CrossRef](#)]
20. Barakat, A.; Mouhtarim, G.; Saji, R.; Touhami, F. Health risk assessment of nitrates in the groundwater of Beni Amir irrigated perimeter, Tadla plain, Morocco. *Hum. Ecol. Risk Assess. Int. J.* **2019**, *26*, 1864–1878. [[CrossRef](#)]
21. Barakat, A.; Meddah, R.; Afdali, M.; Touhami, F. Physicochemical and microbial assessment of spring water quality for drinking supply in Piedmont of Béni-Mellal Atlas (Morocco). *Phys. Chem. Earth Parts A/B/C* **2018**, *104*, 39–46. [[CrossRef](#)]
22. Bouchaou, L.; Chauve, P.; Mudry, J.; Mania, J.; Hsissou, Y. Structure et fonctionnement d'un hydrosystème karstique de montagne sous climat semi-aride: Cas de l'Atlas de Beni-Mellal (Maroc). *J. Afr. Earth Sci.* **1997**, *25*, 225–236. [[CrossRef](#)]
23. Bouchaou, L.; Michelot, J.L. Contribution Des Isotopes à l'étude de La Recharge Des Aquifères de La Région de Béni Mellal (Tadla, Maroc). *IAHS-AISH Publ.* **1997**, *244*, 37–44.
24. Ettazarini, S. Incidences of water-rock interaction on natural resources characters, Oum Er-Rabia Basin (Morocco). *Environ. Geol.* **2004**, *47*, 69–75. [[CrossRef](#)]
25. Ettazarini, S. Groundwater pollution risk mapping for the Eocene aquifer of the Oum Er-Rabia basin, Morocco. *Environ. Geol.* **2006**, *51*, 341–347. [[CrossRef](#)]
26. Archambault, C.; Combe, M.; Ruhard, J.P. The Phosphates Plateau. Water Resources. *Notes Mém. Sér. Géol. Rabat* **1975**, *231*, 239–258.
27. Brede, R. Structural aspects of the Middle and the High Atlas (Morocco) phenomena and causalities. *Geologische Rundschau* **1992**, *81*, 171–184. [[CrossRef](#)]

28. Boudhar, A.; Ouatiki, H.; Bouamri, H.; Lebrini, Y.; Karaoui, I.; Hssaisoune, M.; Arioua, A.; Benabdelouahab, T. Hydrological Response to Snow Cover Changes Using Remote Sensing over the Oum Er Rbia Upstream Basin, Morocco. In *Mapping and Spatial Analysis of Socio-Economic and Environmental Indicators for Sustainable Development*; Springer: Cham, Switzerland, 2020; pp. 95–102.
29. 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](#)]
30. Ouatiki, H.; Boudhar, A.; Trambly, Y.; Jarlan, L.; Benabdelouahab, T.; Hanich, L.; El Meslouhi, M.R.; Chehbouni, A. Evaluation of TRMM 3B42 V7 Rainfall Product over the Oum Er Rbia Watershed in Morocco. *Climate* **2017**, *5*, 1. [[CrossRef](#)]
31. Shaban, A.; Khawlie, M.; Abdallah, C. Use of remote sensing and GIS to determine recharge potential zones: The case of Occidental Lebanon. *Hydrogeol. J.* **2005**, *14*, 433–443. [[CrossRef](#)]
32. Yeh, H.-F.; Cheng, Y.-S.; Lin, H.-I.; Lee, C.-H. Mapping groundwater recharge potential zone using a GIS approach in Hualian River, Taiwan. *Sustain. Environ. Res.* **2016**, *26*, 33–43. [[CrossRef](#)]
33. Botzen, W.J.W.; Aerts, J.C.J.H.; Bergh, J.V.D. Individual preferences for reducing flood risk to near zero through elevation. *Mitig. Adapt. Strat. Glob. Chang.* **2013**, *18*, 229–244. [[CrossRef](#)]
34. Gokceoglu, C.; Sonmez, H.; Nefeslioglu, H.A.; Duman, T.Y.; Çan, T. The 17 March 2005 Kuzulu landslide (Sivas, Turkey) and landslide-susceptibility map of its near vicinity. *Eng. Geol.* **2005**, *81*, 65–83. [[CrossRef](#)]
35. Riley, S.J.; DeGloria, S.D.; Elliot, R. Index that quantifies topographic heterogeneity. *Intermt. J. Sci.* **1999**, *5*, 23–27.
36. Juliev, M.; Mergili, M.; Mondal, I.; Nurtaev, B.; Pulatov, A.; Hübl, J. Comparative analysis of statistical methods for landslide susceptibility mapping in the Bostanlik District, Uzbekistan. *Sci. Total Environ.* **2019**, *653*, 801–814. [[CrossRef](#)]
37. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
38. Knoll, L.; Breuer, L.; Bach, M. Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning. *Sci. Total Environ.* **2019**, *668*, 1317–1327. [[CrossRef](#)] [[PubMed](#)]
39. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013.
40. Bourenane, H.; Guettouche, M.S.; Bouhadad, Y.; Braham, M. Landslide hazard mapping in the Constantine city, Northeast Algeria using frequency ratio, weighting factor, logistic regression, weights of evidence, and analytical hierarchy process methods. *Arab. J. Geosci.* **2016**, *9*, 1–24. [[CrossRef](#)]
41. Saito, H.; Nakayama, D.; Matsuyama, H. Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaishi Mountains, Japan. *Geomorphology* **2009**, *109*, 108–121. [[CrossRef](#)]
42. Tseng, C.-J.; Lu, C.-J.; Chang, C.-C.; Chen, G.-D.; Cheewakriangkrai, C. Integration of data mining classification techniques and ensemble learning to identify risk factors and diagnose ovarian cancer recurrence. *Artif. Intell. Med.* **2017**, *78*, 47–54. [[CrossRef](#)]
43. Quinlan, R.C. *4.5: Programs for Machine Learning*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1993.
44. Kawabata, D.; Bandibas, J. Landslide susceptibility mapping using geological data, a DEM from ASTER images and an Artificial Neural Network (ANN). *Geomorphology* **2009**, *113*, 97–109. [[CrossRef](#)]
45. Song, Y.; Niu, R.; Xu, S.; Ye, R.; Peng, L.; Guo, T.; Li, S.; Chen, T. Landslide Susceptibility Mapping Based on Weighted Gradient Boosting Decision Tree in Wanzhou Section of the Three Gorges Reservoir Area (China). *ISPRS Int. J. Geo-Inf.* **2018**, *8*, 4. [[CrossRef](#)]
46. Demir, G.; Aytakin, M.; Akgun, A. Landslide susceptibility mapping by frequency ratio and logistic regression methods: An example from Niksar–Resadiye (Tokat, Turkey). *Arab. J. Geosci.* **2014**, *8*, 1801–1812. [[CrossRef](#)]
47. Yesilnacar, E.; Topal, T. Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Eng. Geol.* **2005**, *79*, 251–266. [[CrossRef](#)]
48. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
49. Avand, M.; Janizadeh, S.; Bui, D.T.; Pham, V.H.; Ngo, P.T.T.; Nhu, V.-H. A tree-based intelligence ensemble approach for spatial prediction of potential groundwater. *Int. J. Digit. Earth* **2020**, *13*, 1408–1429. [[CrossRef](#)]
50. 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](#)]
51. 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](#)]
52. Karami, G.H.; Bagheri, R.; Rahimi, F. Determining the groundwater potential recharge zone and karst springs catchment area: Saldoran region, western Iran. *Hydrogeol. J.* **2016**, *24*, 1981–1992. [[CrossRef](#)]
53. Tiwari, M.K.; Chatterjee, C. Development of an accurate and reliable hourly flood forecasting model using wavelet–bootstrap–ANN (WBANN) hybrid approach. *J. Hydrol.* **2010**, *394*, 458–470. [[CrossRef](#)]