ISSN: 2538-2152; ECOPERSIA 2026;14(1): 1-22.

Tarbiat Modares University

Zoning of Qanat Systems Extension using Machine Learning Models (Case Study: East and Northeast of Iran)

ARTICLEINFO

Article Type
Original Research

Authors
Javad Momeni Damaneh, *Ph.D.*[1]
Mohammad Ehteram, *Ph.D.*[2]
Fatemeh Panahi, *Ph.D.*[3*]

How to cite this article
Momeni Damaneh J., Ehteram M., Panahi F. Zoning of Qanat Systems Extension using Machine Learning Models (Case Study: East and Northeast of Iran). ECOPERSIA 2026;14(1): 1-22.
DOI:
10.48311/ecopersia.2025.104004.0

[1] Ph.D., Department of Natural Resources Engineering, Faculty of Agriculture and Natural Resources, University of Hormozgan, Bandar Abbas, Iran.
[2] Post-Ph.D., Department of Water Engineering and Hydraulic Structures, Faculty of Civil Engineering, Semnan University, Semnan.
[3] Ph.D., Faculty of Natural Resources and Earth Sciences, University of Kashan, Kashan, Iran.

*** Correspondence***
Address: Ph.D., Faculty of Natural Resources and Earth Sciences, University of Kashan, Kashan, Iran.
E-mail: alabd_fpanahi@yahoo.com

Article History
Received: May 18, 2025
Accepted: December 13, 2025
Published: December 22, 2025

ABSTRACT

Aims: As a valuable human heritage, the Qanat is of great significance to groundwater systems. This research aims to evaluate the effectiveness of environmental variables in the construction of Qanat systems in the east and northeast of Iran and to present the best machine learning model for modeling.
Materials & Methods: Using GIS and R-biomod2, 40 environmental parameters were selected as predictive variables, and GLM, GBM, CTA, SRE, FDA, MARS, RF, and ESMs models were used to determine the relationship between Qanat potential areas and environmental factors. Their Accuracy was evaluated using Kappa, Accuracy, TSS, and ROC.
Findings: Results revealed that random forest (RF) and ensemble (ESMs) models achieved the highest Accuracy in determining Qanat potential reas. SRE performed worse than the other eight models. The results also indicated that climatic factors (BIO4), physiographic factors (DEM& Topographic wetness index & slope), soil factors (Organic 60-100 cm, Cations 60-100 cm, Land Surface Temperature) and Geology has a considerable significance in geographical distribution of areas prone to Qanat existence but terrain roughness index showed the least contribution in determining the groundwater potential areas.
Conclusion: According to the results, the areas of regions with good to outstanding potential for the existence of Qanats were estimated at 13.15% and 13.31% of the total area using ESMs and RF models, respectively. In general, the use of RS in combination with DEM can reveal numerous significant correlations in groundwater research.

Keywords: Machine Learning Models; Environmental parameters; Geographical Information System; Qanat water systems.

CITATION LINKS

[1] Naghibi S.A., Pourghasemi H.R., Pourtaghi Z.S., Rezaei, ... [2] Naghibi S.A., Pourghasemi H.R., Abbaspour K. A compariso... [3] Kharrat F.M., Ragala R., Moussa M. Promoting a sustainab... [4] Yazdi A.A., Khaneiki M.L. Veins of the desert: a review ... [5] Salih, A. (006). Qanats: a unique groundwater managemen... [6] Fattahi, M. (2015). OSL dating of the Miam Qanat (Kāriz)... [7] Labbaf Khaneiki, M. (019). Qanat and territorial cooper... [8] Semsar Yazdi, A.A., Labbaf Khaneiki, M. (017). Qanat kn... [9] Beaumont, P. (971). Qanat systems in Iran. Hydrol. Sci.... [10] Abbasnejad A., Abbasnejad, B., Derakhshani, R., Sarapard... [11] IMA (2020). National Qanat geo-information system, archi... [12] Zamanirad, M., Sedghi, H., Sarraf, A., Saremi, A., Rezae... [13] Behling, R., Roessner, S., Foerster, S., Saemian, P., To... [14] Mansouri Daneshvar, M. R., Ebrahimi, M., Nejadsoleymani,... [15] Pourghasemi, H. R., Beheshtirad, M. (015). Assessment o... [16] Yu, H., Wen, X., Wu, M., Sheng, D., Wu, J., Zhao, Y. (0... [17] Masoudi, R., Mousavi, S. R., Rahimabadi, P. D., Panahi, ... [18] Pradhan, B. A. (013). A comparative study on the predic... [19] Naghibi, S. A., Pourghasemi, H. R., Dixon, B. (016). GI... [20] Chen, W., Panahi, M., Khosravi, K., Pourghasemi, H. R., ... [21] Golkarian, A., Naghibi, S. A., Kalantar, B., Pradhan, B.... [22] Momeni Damaneh, J., Ahmadi, J., Rahmanian, S., Sadeghi, ... [23] Momeni Damaneh, J., Gerami, S., and Hajibigloo, M. (025... [24] Thuiller, W., Lafourcade B., Engler R., Araújo M. B. ... [25] Fielding, A. H., Bell, J. F. (997). A review of methods... [26] Walther, G. R., Post E., Convey P., Menzel A., Parmesan ... [27] Smeeton, N.C. (985). Early history of the Kappa Statist... [28] Galton, F. (892). Finger Pints. Macmillan, London... [29] Saha, S., Roy, J., Arabameri, A., Blaschke, T., Tien Bui... [30] Yi, Y. J., Cheng X., Yang Z. F., Zhang S. (016). Maxent... [31] Swets, J. A. (988). Measuring the Accuracy of diagnosti... [32] Naghibi, S. A., Ahmadi, K., Daneshi, A. (017). Applicat... [33] Moghaddam, D. D., Rahmati, O., Panahi, M., Tiefenbacher,... [34] Moghaddam, D. D., Rahmati, O., Haghizadeh, A., Kalantari... [35] Catani, F., Lagomarsino, D., Segoni, S., Tofani, V. (01... [36] Caruana, R., Niculescu-Mizil, A. (006). An empirical co... [37] Naghibi, S. A., Dolatkordestani, M., Rezaei, A., Amouzeg... [38] Mousavi, S.M., Golkarian, A., Naghibi, S.A., Kalantar, B... [39] Kim, J. C., Jung, H. S., Lee, S. (019). Spatial mapping... [40] Prasad, P., Loveson, V. J., Kotha, M., Yadav, R. (020)....

Introduction

A qanat (pronounced "Qanat" or "Kāriz" in Persian) is an underground system that conveys groundwater from high-elevation areas to lower ones in a plain [1,2]. A. A traditional Qanat consists of various parts, including 1. Mena, which is an underground passage or tunnel dug to collect and transmit at r. 2. The mother well: The deepest well, which is located upstream and is the main source of Qanat at r. 3. Vents, vertical wells dug along the water pathway of Qanat to aid ventilation and access to the t nn l. 4. Maz'har, the place where the Qanat water exits from the subsurface to the Earth su fa e. 5. Distribution channels, a network of streams and rivers that carry water from the reservoir to farms and homes [3,4]. As a valuable part of human heritage, Qanats play a significant role in the sustainable management of groundwater resources [5]. For centuries, this historical system has provided water to people, especially in arid and semi-arid areas, across more than 34 countries [6]. Today, Qanats are widely distributed worldwide. Qanat is the main symbol of the relationship between man and water. Besides sustainable agricultural development, Qanats play a principal role in Qanat tourism and, as a result, employment generation. They have also represented a technological advancement and a major solution for water supply in arid and semi-arid regions for thousands of years, and their history can be traced back to the emergence of several civilizations. It is essential to maintain Qanats as sustainable water sources in arid regions. To manage canals, the main requirement is to prepare spatial data for them [8]. A preliminary review of the literature showed that, according to Beaumont (1971), the Qanat system has been of great importance in the landscape of Iran during the Achaemenid Empire (330-550 BC [9]). In addition, Qanats have supplied water to most settlements in Iran for centuries [10]. The mechanization of Qanat construction reflects Iran's environment and culture. In 2020, about 40000 Qanats with an average annual harvest of 4.7 bm^3 were recognized as the Qanat National Geomatic System in the Ministry of Agriculture Jihad data center [11]. Examining Qanat time-series data from 2008 to 2018 shows that the average annual water withdrawal from Qanats decreased by almost 50%. This decrease is mainly due to climate change and the overexploitation of land resources in Iran [7, 12]. Water resources are severely limited, and according to Behling et al. (2022), more than 90% of Iran is under dry or semi-arid conditions [13]. In Iran, data on climate change, including a 2.6 ∘ C increase in average temperature and a 35% decrease in precipitation in the coming decades, are very worrying [14]. Besides springs and wells, Qanats are also a crucial source of groundwater in Iran [15]. Various environmental parameters may affect water resource types. Supervised classification algorithms explore the relationships between the variables and known groundwater data. Algorithms provide accurate predictions, and the results are extrapolated to estimate groundwater potential in a specified region. Since artificial intelligence models can evaluate large amounts of data and produce accurate predictions, they have been progressively utilized in recent water management studies [16, 17]. Most machine learning groundwater potential map studies face two principal deficiencies. First, since the number of presence data available for training and testing algorithms may be small and the number of environmental variables may be large, how to reclassify the environmental variables to reduce their error is of great importance. The second issue is that machine learning study results are almost always evaluated using standard big data metrics such as precision, recall, and the area under the receiver operating characteristic curve [1-20]. Using several environmental variables and creating experimental models in Geographic Information System (GIS), researchers have particularly addressed the potential map of Qanat [1, 21]. Prior studies by Naghibi et al. (2015) and Moghaddam et al. (2020a) established that physiographic variables (elevation, slope, TWI) and soil properties (organic content, texture) significantly control groundwater potential in Qanat systems [33]. Additionally, Pourghasemi & Beheshtirad (2015) and Golkarian et al. (2018) validated the relevance of land surface temperature (LST) and drainage density in machine learning-based Qanat zoning [1, 21]. There is no effort to evaluate the efficiency of each machine learning model and variable used to construct Qanat routes and positions in Iran. In particular, no study has been carried out to demonstrate the specific correlations between variables and Qanat across different parts of Iran, nor to compare machine learning models. Therefore, the novelty of this study lies in developing multiple statistical methods using GIS to examine evidence of Qanat systems in the watersheds of East and North East Iran. The primary objective of this paper, as argued, is to evaluate the effectiveness of environmental variables in the construction of Qanats in the east and northeast of Iran and to present the best machine learning model.

Materials & Methods

Case Study

The study has been carried out in the east and northeast of Iran, covering an area of 295,11 km^2. It is located between 55° 26' 36'' to 61° 26’ 34’’ altitude and 30° 31’ 54’’ to 38° 11’ 53’’ latitude. The studied area is situated in the Iranian-Turanian vegetation region and, due to its vast size and diverse natural conditions, has unique ecological characteristics across its various areas. Based on the De Martonne aridity index classification, the entire province falls under a cold dry climate, with specific areas classified as cold semi-arid [22]. In general, the climate of Khorasan province is arid and semi-arid. The province receives an average of 209.8 mm of rainfall, but precipitation is not uniform, and amounts decrease from north to south. The lowest precipitation is estimated at 116.2 mm, and the highest is 312.8 mm. The province has an average annual temperature of 15.6 ˚C, with minimum and maximum temperatures of 12.2 ˚C and 18.2 ˚C, respectively, during the statistical period [22]. The study included 2,482 plant species across 585 genera and 115 families. The sunflower family is the most numerous in the region, accounting for 303 species, followed by the Gramineae family with 180 and the Laminaceae family with 122 species. Among the most important plants of the area, we can mention *Astragalus sp.*, *Acantholimon sp.*, *Seidlitzia sp.*, *Onobrycis sp.*, as well as *Juniperus sp.*, *Pistacia sp.*, *Tamarix sp.*, *Berberis sp.,* and *Rosa pe sica*. Some plants of high nutritional and medicinal value, such as *Ziziphora sp.*, *Thymus sp.*, *Cuminum sp.*, and *Ferula sp.*, are becoming extinct and are being destroyed due to excessive grazing and rangeland degradation [23].

Determining the Presence Points of Qanats

The areas of active Qanats presence for which the location of Qanat debouchure (water outlet of the Qanat) was obtained from water organization of Khorasan Province (North, Razavi, South) and the areas of absence were determined randomly from the background with some 20,000 points; a total of 11,957 points for active Qanats that have already been officially registered by the regional water company using the Global Positioning System (Figure 1).

Determination of Environmental Variables

Based on a review of pertinent studies and regional information sources, 92 key variables were identified as significant for

Table 1) The list of used models from the Biomod2 package.

Abbreviation	Full Name	References
GLM	Generalized Liner Model	Austin et al. (1984)
GBM	Generalized Boosting Method	Austin et al. (1984)
CTA	Classification Tree Analysis	Hastie et al. (1994)
SRE	Surface Range Envelope	Harrell et al. (1996)
FDA	Flexible Denotative Analysis	Nix (1986)
MARS	Adaptive Regression Spline Multivariate	Hastie et al. (1994)
RF	Random Forest	Friedman (1991)
ESMs	Techniques and their ensembles	Damaneh et al. (2022)

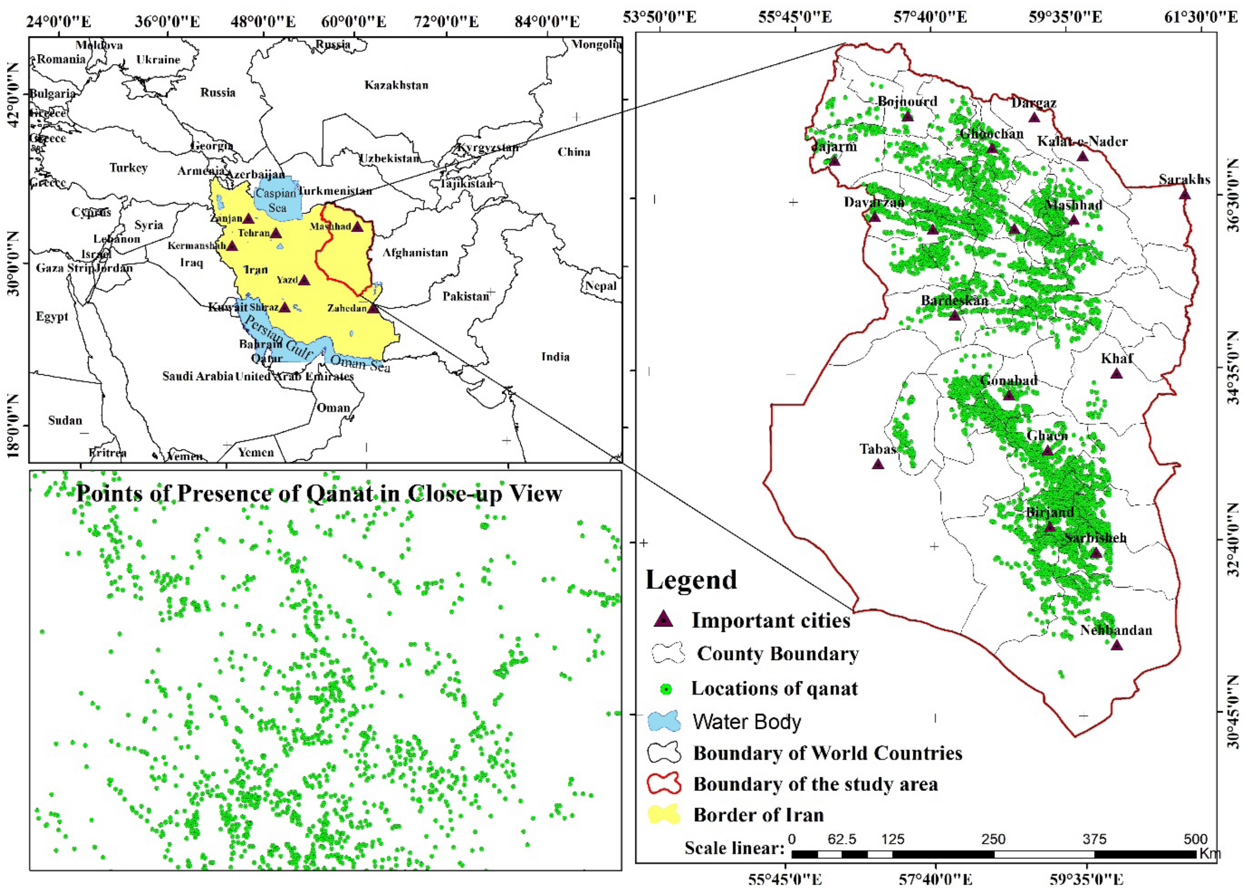

Figure 1) Location of the studied area and the distribution of Qanats' presence in the east and northeast of Iran (presented by the Regional Water Company of Iran, 2023).

mapping and identifying areas with Qanat potential. The corresponding data layers for these variables were compiled from various sources.

Using the available information, 92 environmental variables including seven physiographic variables (Iran Geological Organization, Google Earth Engine), 24 climate variables with an Accuracy of 30 seconds (https://www.worldclim.org, https://globalsolaratlas.info), 54 soil variables with an Accuracy of 250*250 km (https://soilgrids.org, Google Earth Engine), three geological variables (Iran Geological Organization), two vegetation (Google Earth Engine) and two hydrological variables (Google Earth Engine) were used to generate the odel. Since all of the input data layers of the model must have the same georeference, coordinate system and scale, the initial preparation and processing of the data layers were performed using Idrisi Selva software and preparing data layers (Figure 2) and matching the layers with a pixel size of 1000 × 1000 m was carried out in Idrisi Selva software using the Variance Inflation Factor (VIF) to verify collinearity severity between independent vari bles. Variables with VIFs less than 10 were selected, and 40 environmental parameters were ultimately chosen as predictive variables (Table 2) [42]. They were called in R using the Grid format, along with the presence points of Q nats. Then, they were predicted using GLM, GBM, CTA, SRE, FDA, MARS, RF, and ESMs models (Table 1) to determine the relationship between Qanat potential and environmental factors in the east and northeast of Iran. The Accuracy of the models was evaluated using KAPPA, ACCURACY, TSS, and ROC values, which are widely used indicators for identifying areas of equal potential [23].

Modeling the Distribution of the Equipotential Zones of the Presence of Qanats

In the present research, seven algorithms from the Biomod2 package [24] were used to model spring-prone areas and generate points of absence (Table 1). In the modeling process, 70% of the presence points were used to create the models, and 30% were used to evaluate them. Also, the number of repetitions was set to 5 to enhance modeling Accuracy. The models' Accuracy was calculated using three statistical coefficients for each model type. The first method is to check the Receiver Operating Characteristic (ROC) curve. A ROC chart is a graphical method for evaluating the model's ability to predict the presence or absence of groundwater resources (springs) based on related environmental variables [25]. Accurate Skill Statistic (TSS) can be considered evidence for interpreting real ecological phenomena. Various studies demonstrate that ROC values show a strong correlation with TSS; therefore, in studies with presence-absence maps, TSS may be a suitable alternative to ROC [26]. The third method, Cohen's Kappa, can measure the agreement between two raters who classify N items into C mutually exclusive classes. The first use of Kappa-like statistics is by Galton and Smeaton [27, 28]. The fourth method of Diagnostic Accuracy (ACCURACY) is another universal measure of "Diagnostic Accuracy" expressed as the proportion of correctly classified subjects (TP + TN) among all subjects (TP + TN + FP + FN) [29]. KAPPA, ACCURACY, TSS, and ROC values less than 0.5 are considered as inappropriate modeling performance, very poor, poor, moderate, good, and high (susceptible) for values between 0.5-0.6, 0.7-0.6, 0.7-0.8, 0.8-0.9, and more than 0.9, respectively [30, 31].

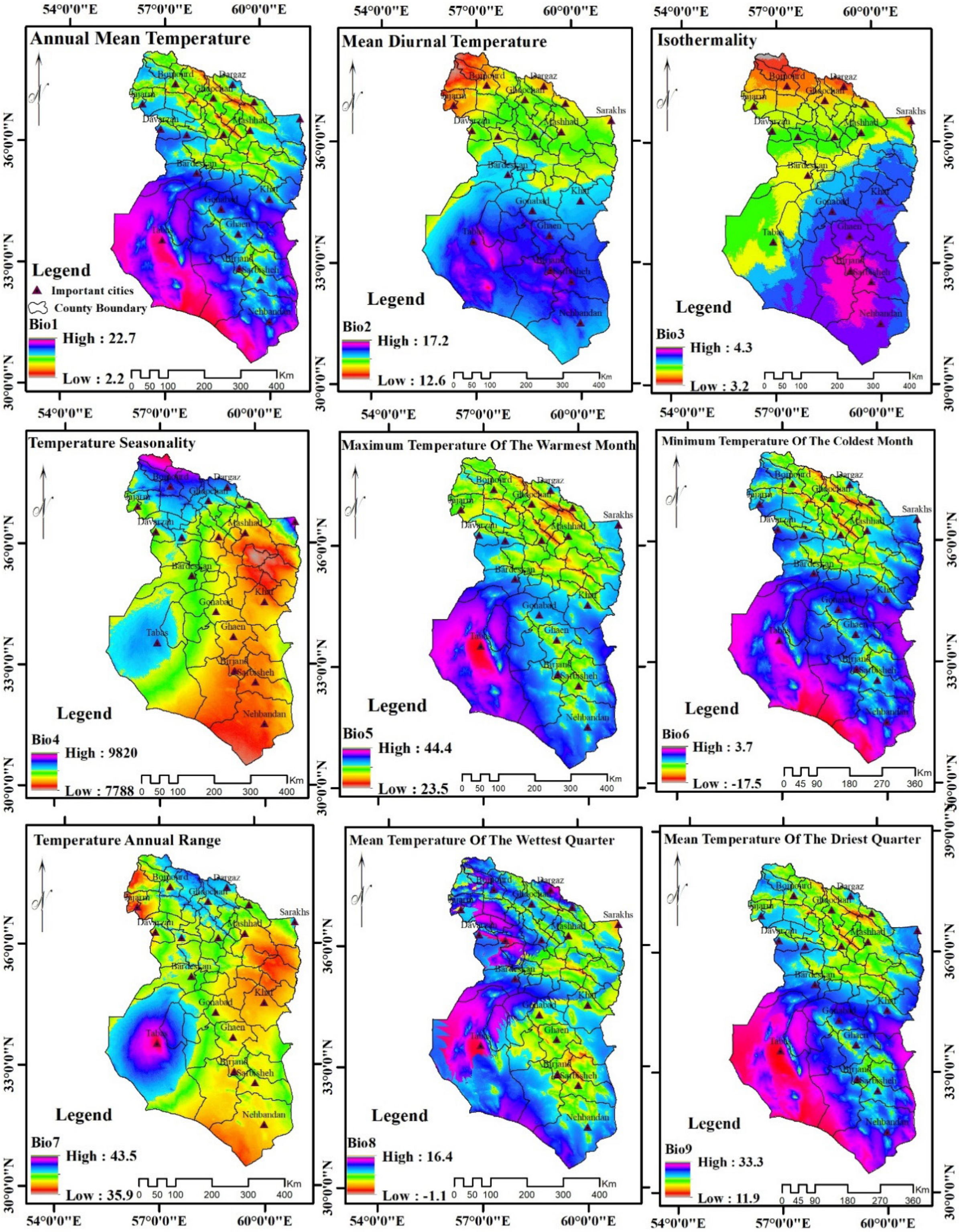

Figure 2) Data layers of the variables used.

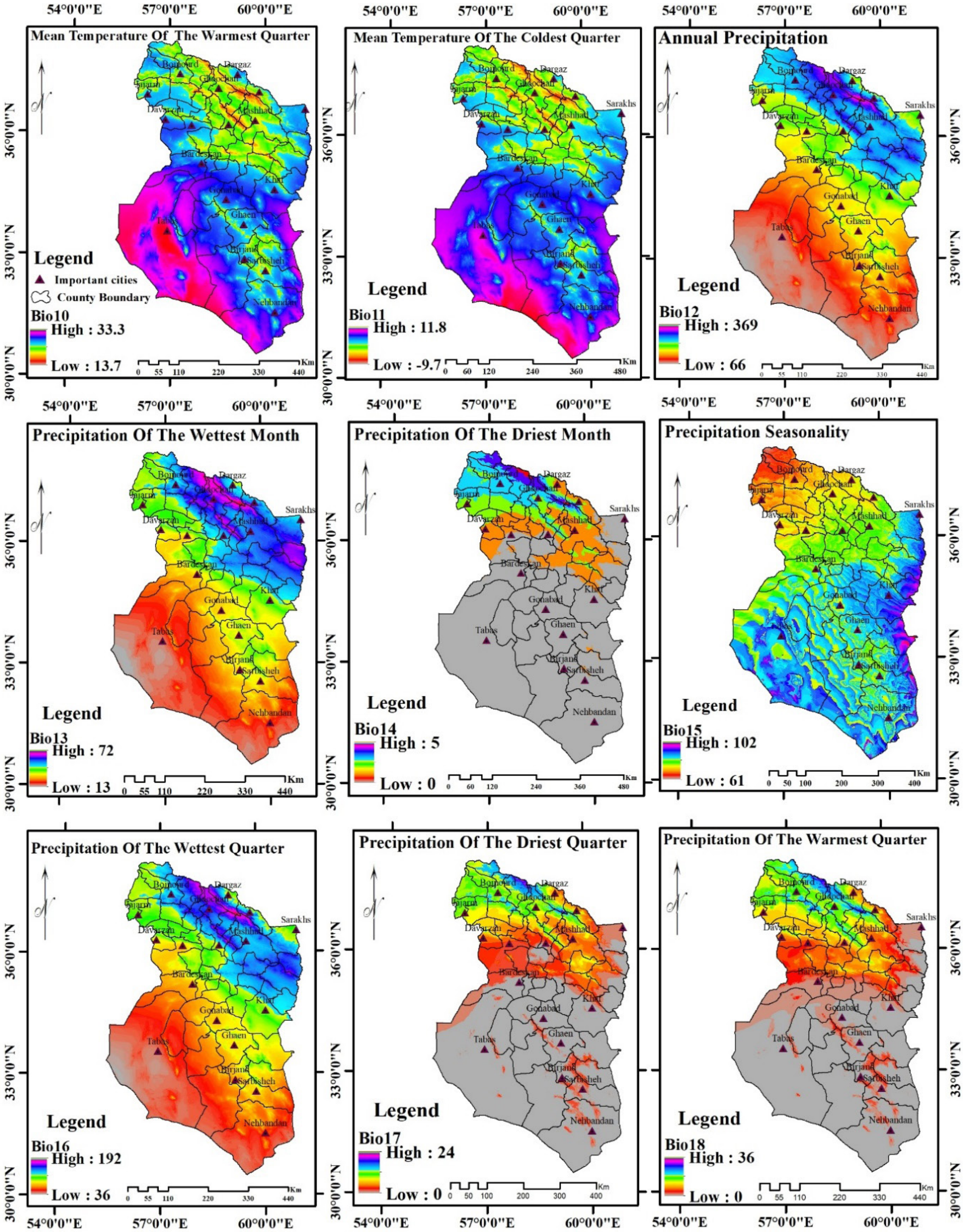

Figure 2 Continued) Data layers of the variables used.

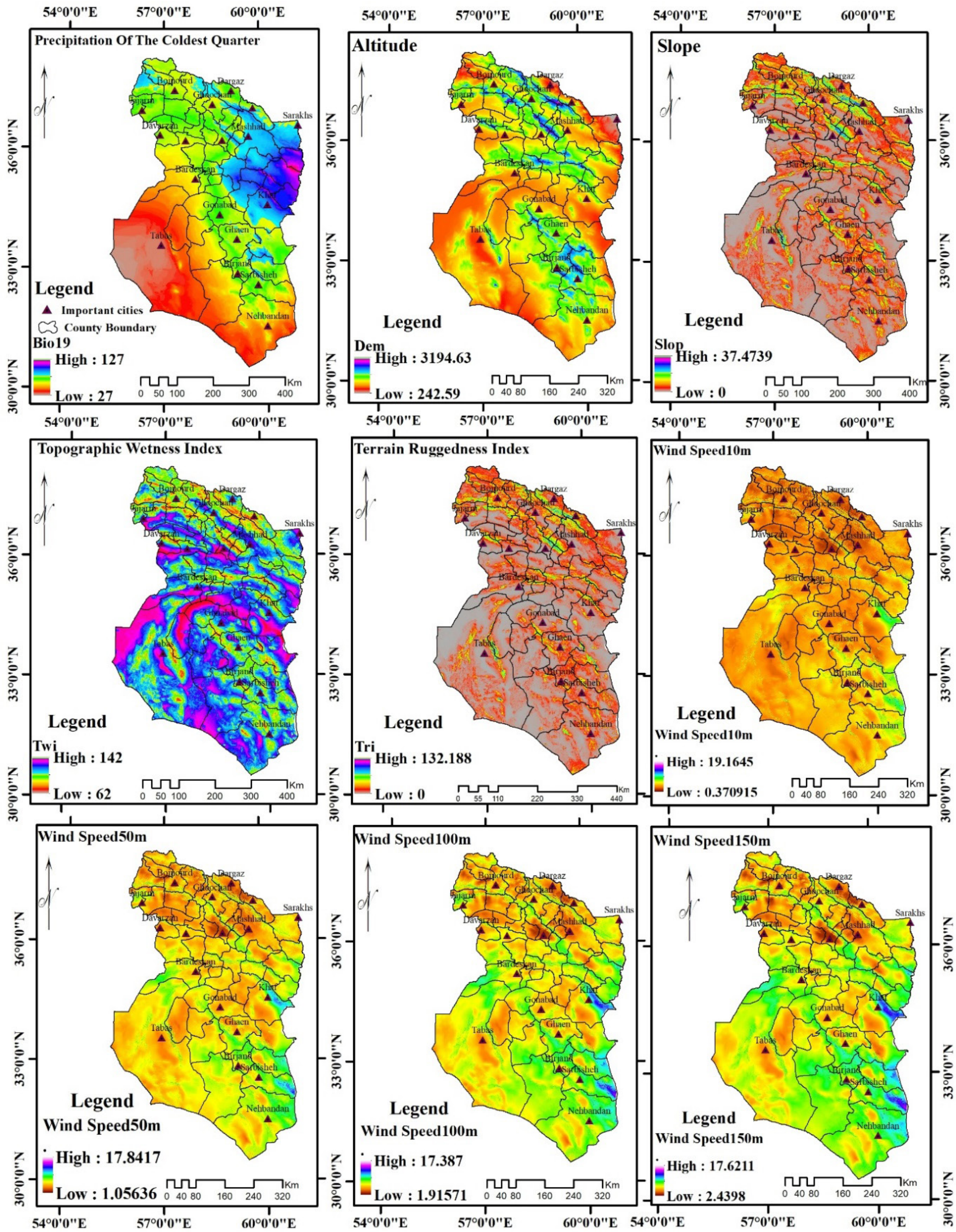

Figure 2 Continued) Data layers of the variables used.

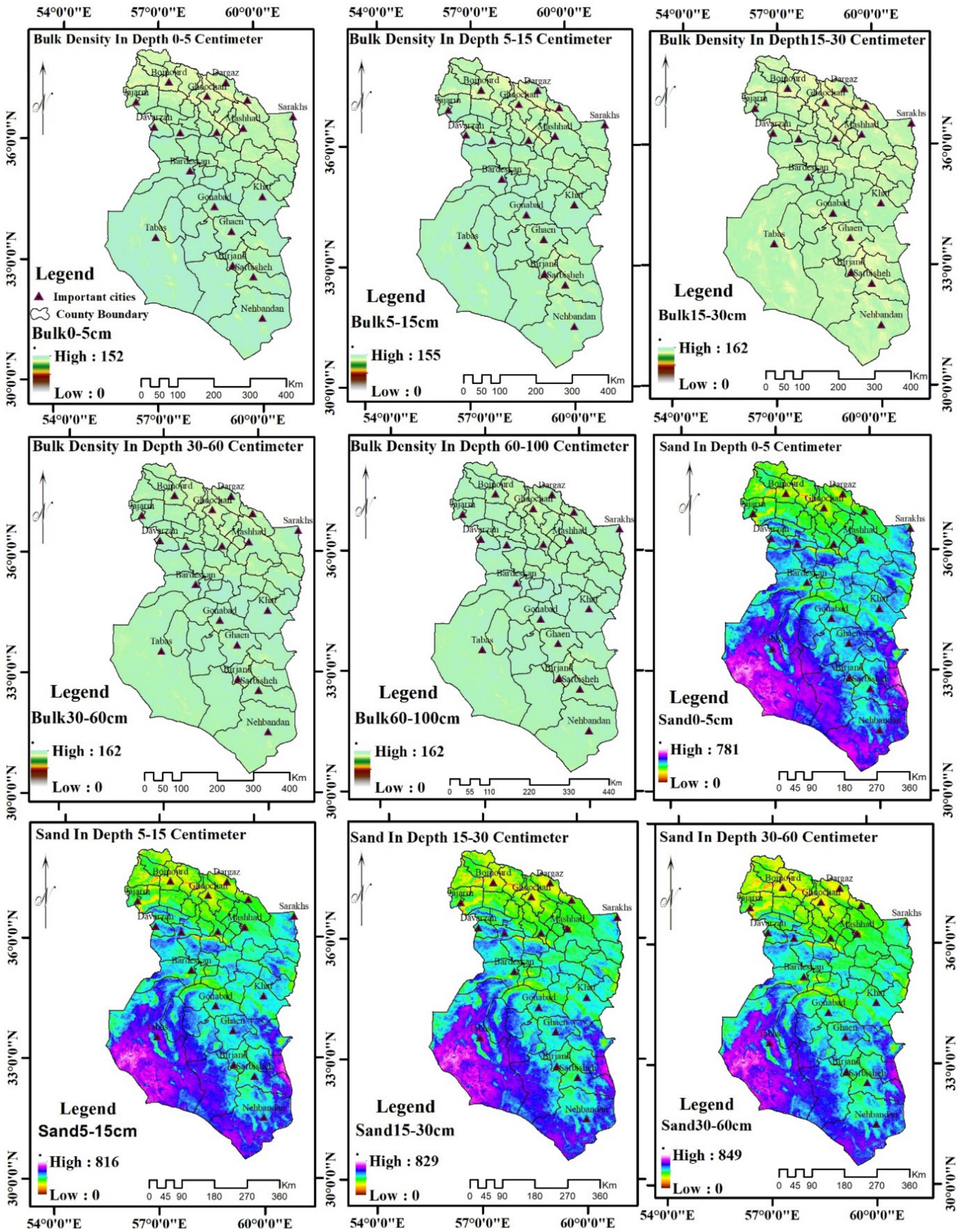

Figure 2 Continued) Data layers of the variables used.

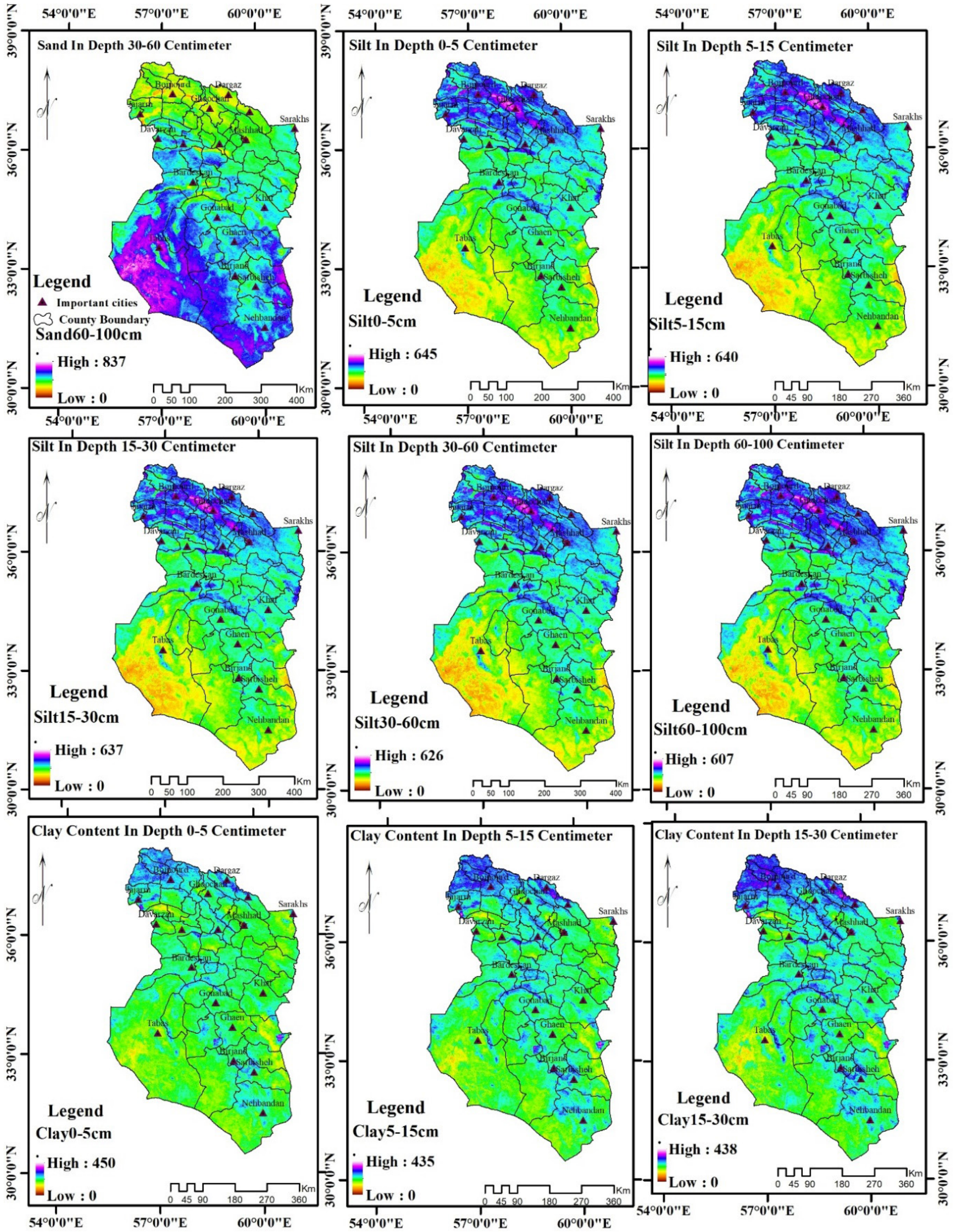

Figure 2 Continued) Data layers of the variables used.

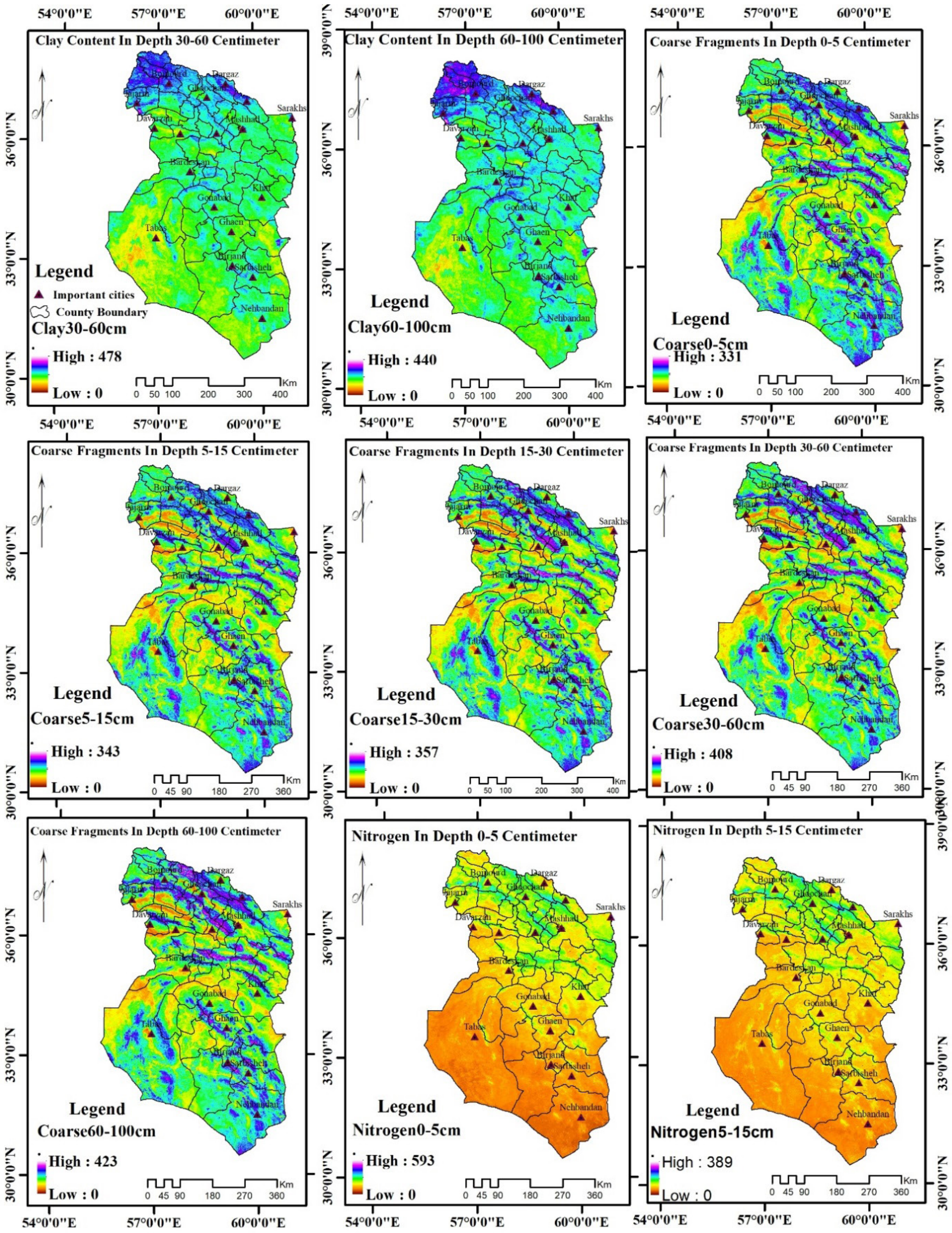

Figure 2 Continued) Data layers of the variables used.

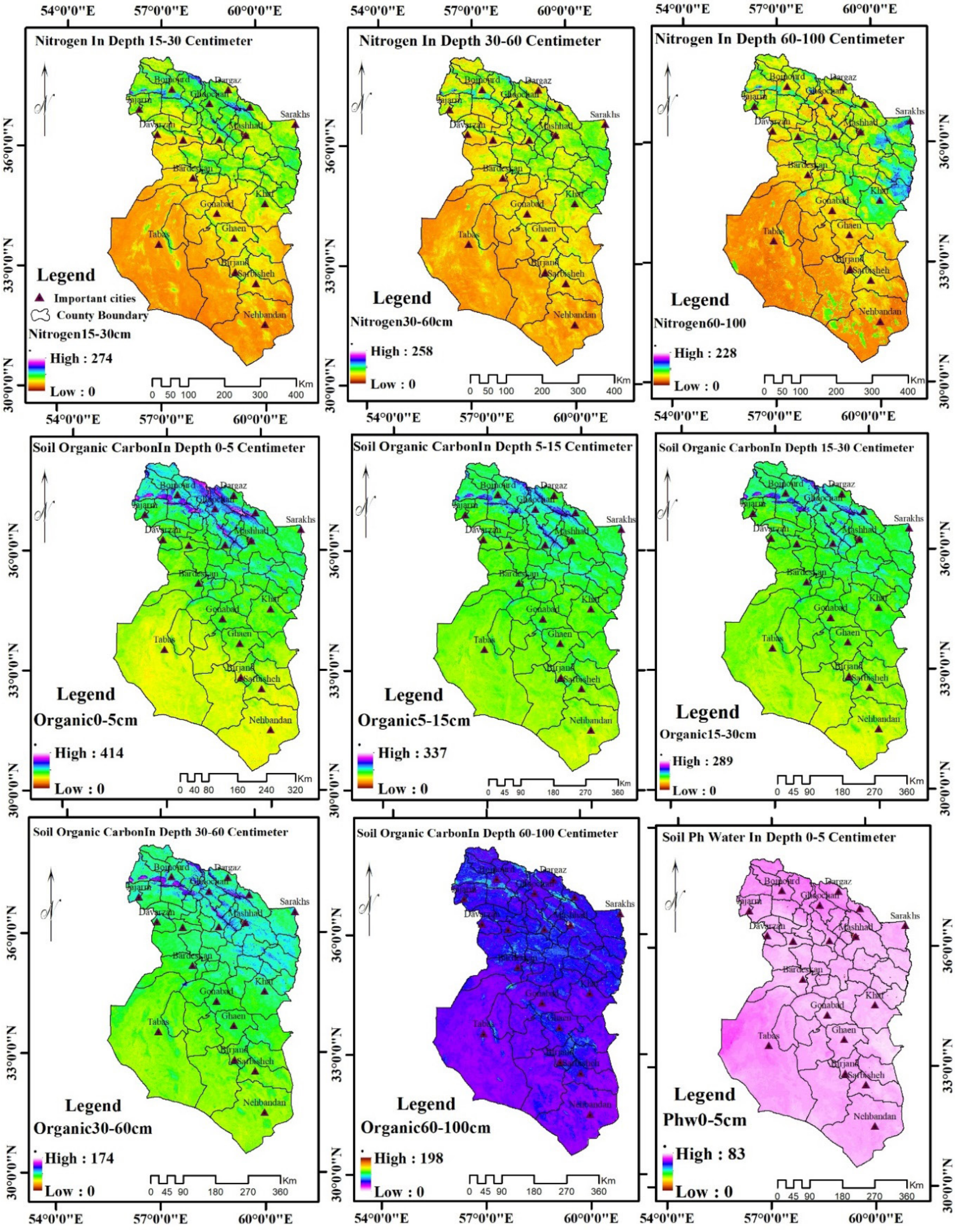

Figure 2 Continued) Data layers of the variables used.

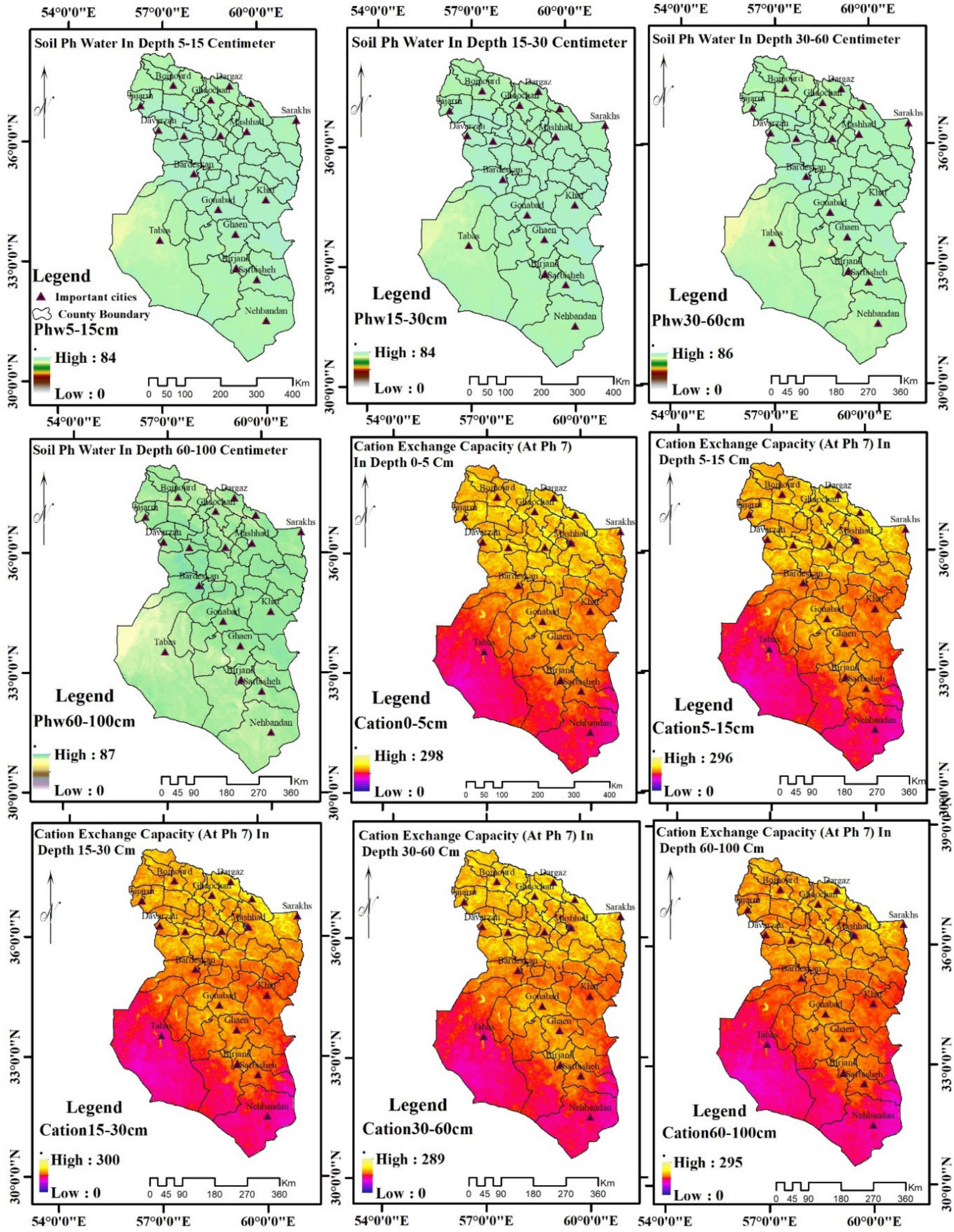

Figure 2 Continued) Data layers of the variables used.

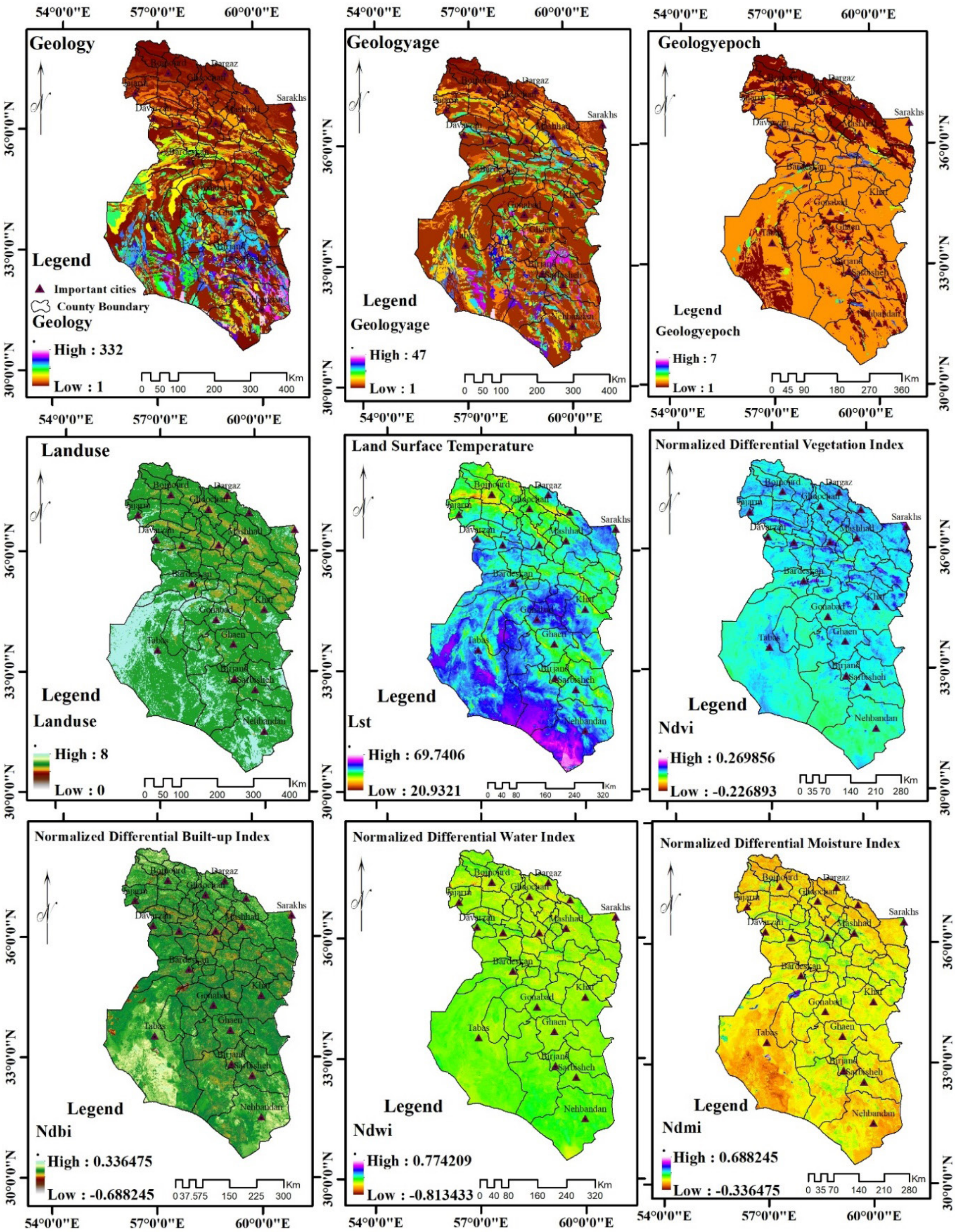

Figure 2 Continued) Data layers of the variables used.

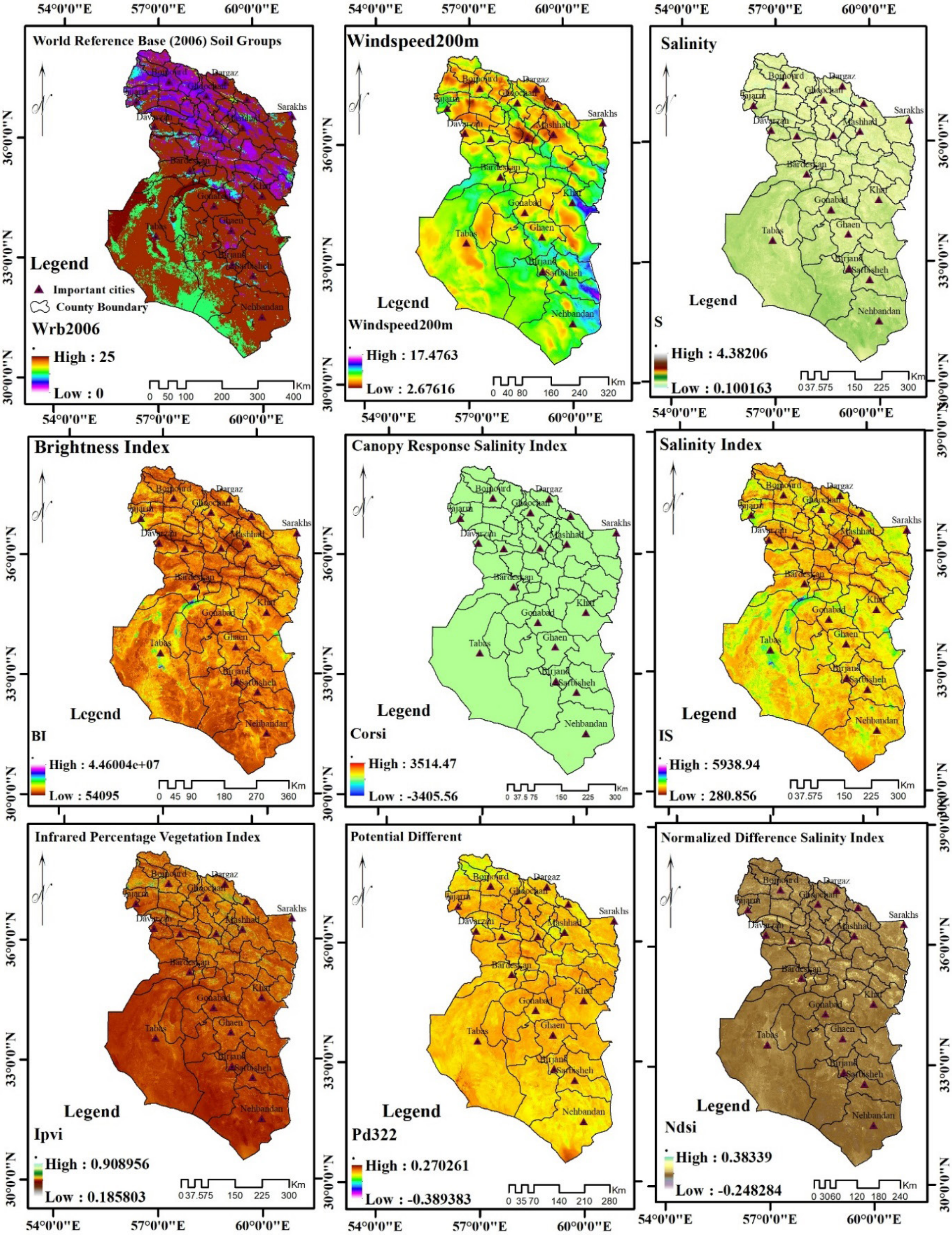

Figure 2 Continued) Data layers of the variables used.

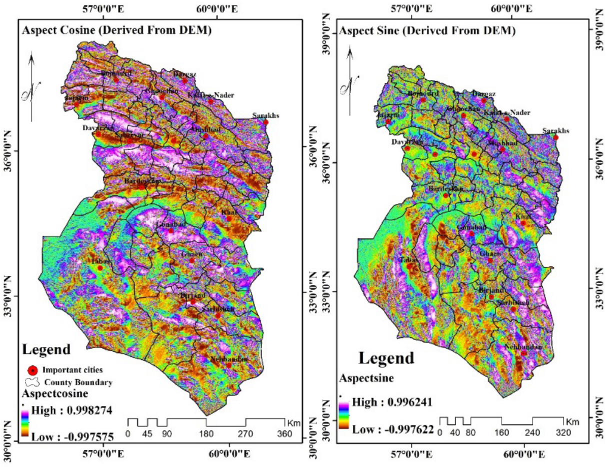

Figure 2 Continued) Data layers of the variables used.

In the following, to capture the geographical distribution of areas with favorable climatic and environmental conditions for Qanat presence in the study area, maps of areas prone to Qanat have been depicted (Figure 5). The maps of prone areas obtained by vector machine models are expressed as values from 0 to 1000. Zero is the lowest probability, and 1000 is the highest probability. To better understand the distribution, the map was classified into four classes, including areas with no potential for Qanats presence between 0 and 250, areas with a low proportion of Qanats presence between 250 and 500, areas with the average proportion of the presence of Qanats between 500 and 750 and the prone Qanats regions presence between 750 and 1000 using ArcGIS10.8 software, the Natural Breaks method or the Jenks algorithm (Table 3, Figure 3).

Findings

Table 2 presents the results of a collinearity analysis using VIF (<10) to assess the severity of collinearity among independent variables. In total, 40 of 92 environmental factors had VIFs less than 10 and were used in the calculations and modeling.

Studying the potential of Qanat using machine learning algorithms and natural factors derived from remote sensing (RS) has been the subject of various studies in this field. However, the conditional parameters based on topography and other environmental factors differ across regions. KAPPA, ACCURACY, TSS, and ROC values, which are widely used to identify and determine co-potential areas, are presented in Figure 3. The modeling process in most of the models used in this research shows a good to moderate fit. Studying the values of all four Accuracy evaluation indicators showed that the random forest (RF) and ensemble (ESMs) models achieved the highest Accuracy in determining Qanat potential areas.

Among the eight models, the ESM model achieved the highest Accuracy and ROC area,

Table 2) List of environmental factors used for modeling.

Parameter	VIF	Parameter	VIF	Parameter	VIF	Parameter	VIF
Geologyepoch	1.18	Bio8	2.33	Bio14	3.88	Clay60100	6.91
Land-Use	1.26	DEM	2.63	Cation60100	4.12	Aspect Sine	7.36
Wrb2006	1.37	Organic60100	2.96	Nitrogen60100	4.26	Slop	7.46
Coarse60100	1.97	Bio4	3.04	Clay05	5.05	Organic3060	7.80
Geologyage	1.99	Corsi	3.09	Silt60100	5.21	Organic515	8.18
TWI	2.00	Lst	3.16	Nitrogen3060	5.45	Organic1530	9.10
Eastness	7.22	Geomfootslope	1.05	Northness	6.88	Roughness	7.55
Pd322	2.05	Bio15	3.36	Nitrogen515	5.46	Bi	2.21
Is	2.12	Ndsi	3.39	Bulk1530	5.85	Geology	2.31
NDVI	3.69	PHW60100	6.52	IPVI	3.76	Aspectcosine	6.89

while RF ranked second, and SRE performed worse than the others. Measuring inter- and intra-water reliability using KAPPA values, all eight models demonstrated significant relationships. This suggests that the models are more consistent when measuring a fixed pheno enon. Therefore, according to the results and values of Figure 3, random forest (RF) and ensemble (ESMs) models will be the basis of the next calculations. The highest values of Accuracy and the influencing variables in the distribution of Qanat equipotential zones are shown in Fig. 3. The ensemble and selected models will present the results.

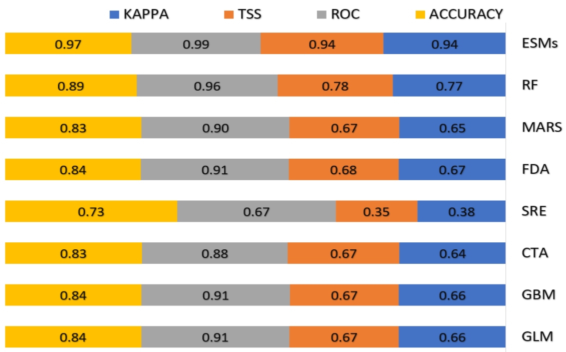

Figure 3) Evaluation of advanced machine learning algorithms (MLA) by different indicators.

Examining the relative importance of all environmental factors in the selected models indicated that climatic factors (BIO4), physiographic factors (DEM& Topographic wetness index & slope), soil factors (Organic 60-100 cm, Cations 60-100 cm, Land Surface Temperature) and Geology has a considerable significance in geographical distribution of areas prone to Qanats existence in the east and northeast of Iran (Figure 4).

According to Table 3, the results of this section demonstrated that in the selected models of ESMs and RF, the area between 38,427 and 38,884 km^2, equivalent to 13.15 to 13.31 percent of the regions studied, has good to outstanding potential for the existence or reclamation of ground water resources (Qanats) and showed the highest geographical distribution of the probability of the success of construction or existence of Qanats (Table 3 and Figure 5).

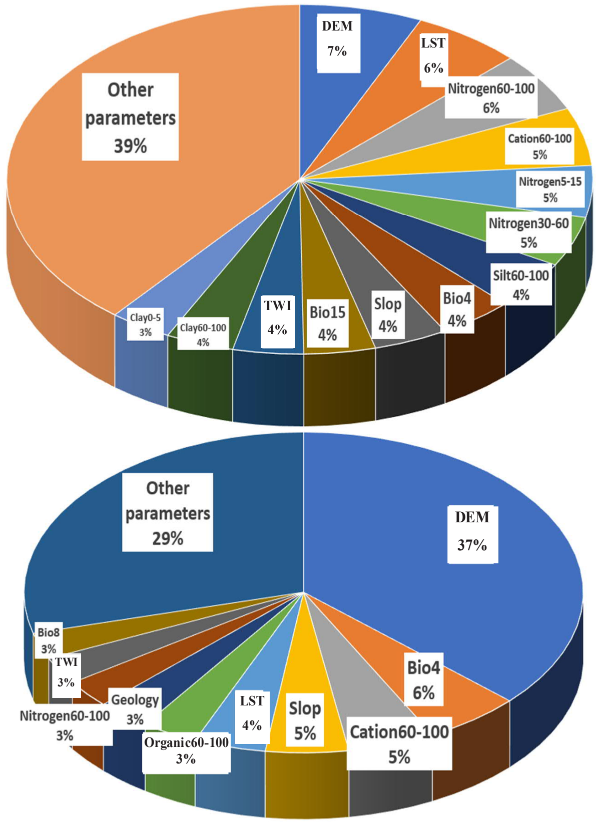

Figure 4) Relative importance of environmental parameters affecting the determination of groundwater equipotential zones using Random Forest model (above) and Ensemble model (bottom).

Table 3) The area and percentage of areas with similar potentials for the presence of groundwater Qanat in the modeling of the east and northeast of Iran.

Class	ESMs		RF	
	Area (Km^2)	Area (%)	Area (Km^2)	Area (%)
Low	237519	81.28	237795	81.38
Moderate	16270	5.57	15540	5.32
High	12368	4.23	12069	4.13
Very High	26059	8.92	26815	9.18

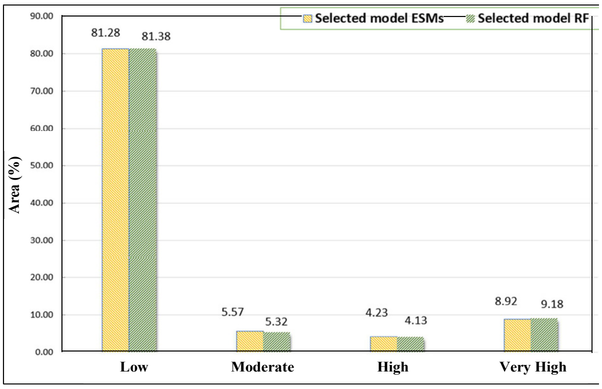

Figure 5) Percentage of each class of groundwater Qanat potential maps in the modeling of the east and .northeast of Iran

In general, as shown in Figure 6, without proper and serious planning, we will witness a reduction and loss of water resources in the area, which would increase immigration, unemployment, and the destruction of industries and agriculture in the study area (Statistics Organization of Iran).

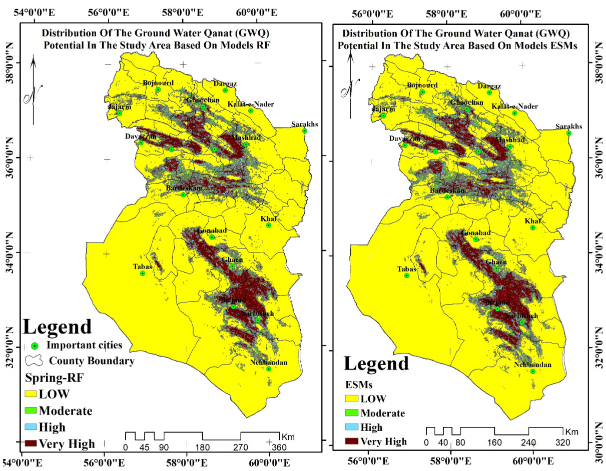

Figure 6) Distribution of the groundwater Qanat (GWQ) potential in the study area based on selected random forest and ensemble models.

Discussion

Various studies have demonstrated that Random Forest (RF) and other Ensemble Models (ESMs) are highly effective tools for generating groundwater potential maps. This finding is consistent across multiple research efforts. For instance, Golkarian et al. (2011) reported that the RF model demonstrated highly acceptable performance with an AUC-ROC of 79.7% in the Mashhad Plain, while the MARS model performed best [21]. Similarly, Naghibi et al. (2011) concluded that RF and its optimized version, RFGA, outperformed the LSVM model, with AUC values of 84.6% and 85.6%, respectively [32]. In a separate study, Moghaddam et al. (2011) confirmed the RF model's superior performance in a mountainous area, achieving a mean AUC of 0.995 compared to the GARP and QUEST models [33, 34]. Additionally, Momeni Damaneh et al. (2025) employed a range of machine learning models for groundwater potential assessment, emphasizing the considerable potential of these models, particularly ESMs, in groundwater research [23]. These collective results highlight the high Accuracy and efficiency of Random Forests and ESMs in predicting groundwater potential across diverse geographical environments. There are multiple decision trees, and the data is not over-fitted, which represents the best fit. In addition, nonlinear and determining variables in the RF model can interact, thereby increasing the model's effectiveness [35]. Also, the high Accuracy of some models, like random forests, may be due to their low tendency to overfit and their ability to handle small-scale datasets [36, 37].

Previous studies confirm the effectiveness of machine learning models, particularly Random Forest (RF), for groundwater potential assessment, findings that align with this research. For instance, a study by Naghibi et al. (2019) using an RF model achieved an impressive AUC-ROC of 90.1%, demonstrating its high effectiveness [37 38]. The ensemble approach used in this research also indicates improved performance of the base models. Similarly, in a study by Kim et al. (2019) in South Korea, BRT and RF models were identified as valuable tools for groundwater resource development, showing an acceptable Accuracy of over 78% [39]. Furthermore, research by Momeni et al. (2025) in Iran found RF to be the most accurate among several algorithms [23]. This consistency in results suggests that the models used in your study perform comparably to prominent models in other studies, reinforcing the validity of machine learning methods for this type of analysis. This review indicates that the findings on the significance of environmental factors differ from those of other researchers, underscoring the need for regional approaches in hydrological studies. For instance, studies by Naghibi et al. (2016) and Chen et al. (2019) identify drainage density and NDVI as important factors influencing groundwater potential [19, 20]; however, in recent research, these factors were not highly significant.

Furthermore, Prasad et al. (2020) considered NDVI a key factor in their study in India, which contradicts our findings [23]. These differences highlight the reality that the set of influential factors in each region varies due to specific geological, climatic, and topographical characteristics. This inconsistency in results is a strong point in the discussion section, demonstrating the importance of using indigenous models and data in groundwater resource management. A review of similar research indicates that our findings regarding the low significance of the Topographic Roughness Index (TRI) are entirely consistent with the approaches and results of other scholars. In the study by Naghibi et al. (2016), despite a focus on topographic factors, TRI was not mentioned, implicitly underscoring its lower importance relative to other factors, such as slope and elevation. Similarly, in papers by Chen et al. (2019) and Prasad et al. (2020), which examined multiple factors including topographic ones, TRI was not considered a key determinant [20, 40]. This consistency in results suggests a general pattern in groundwater potential modeling studies. In these studies, larger-scale topographic parameters typically have a greater influence on water flow and infiltration. Therefore, our findings on the limited contribution of TRI are not only valid but also align with the broader scientific literature in this field. Findings of this study regarding the significance of climatic factors (BIO4), physiographic factors (DEM, TWI, and slope), soil factors (Organic 60-100 cm, Cations 60-100 cm, and Land Surface Temperature), and geology in Qanats distribution are in complete agreement with the results of other research on groundwater potential. This broad consistency highlights the validity and importance of these factors. For instance, in the study by Naghibi et al. (2016) in Iran, slope, elevation, and TWI were identified as key physiographic parameters, which align directly with the key factors identified in our research [19]. This alignment confirms the vital role of topographic factors in controlling groundwater flow and the formation of springs and Q nats.

Furthermore, similar factors, such as slope, elevation, and TWI, were used in the research by Chen et al. (2019) in China and Prasad et al. (2020) in India [20, 40]. Prasad et al. (2020) also highlighted the importance of geological and soil factors [40], findings consistent with ours. This convergence in results across various studies in different geographical regions strengthens the validity of our findings. It demonstrates that physiographic, geological, and soil factors are universal and crucial for assessing groundwater potential.

Conclusion

Accurate assessment of groundwater resources is essential to achieving the Sustainable Development Goals (SDGs). In developing countries, the lack of accurate hydrogeological data for aquifers is evident. RS-based data products can provide rich information for data-poor regions. This study applied eight widely used machine learning algorithms and compared them using four criteria, namely KAPPA, ACCURACY, TSS, and ROC. Groundwater resources, especially Qanats, are vital and valuable resources in the eastern and northeastern regions of Iran. The reduction and destruction of groundwater sources (Qanats) due to inadequate planning can have negative consequences, including increased immigration, unemployment, and the destruction of industries and agriculture in the study area. The results of this research can be used for planning and managing groundwater resources, especially in areas with low- to medium-potential Qanats. Maps of equipotential areas for the presence of Qanats can serve as a management tool for locating and constructing structures in watersheds and aquifers. The findings of this study can help improve public knowledge and awareness of the importance of preserving and restoring groundwater resources, especially Q nats. The present study investigates the distribution and identification of equipotential areas of Qanats in the east and northeast of Iran using machine learning-based modeling of 92 environmental variables. RF and ESMs produced maps that were similar to and as accurate as those of other models. This emphasizes that machine learning can be used in groundwater research to identify areas of high groundwater exploitation potential. According to the results, ESMs with 94% Accuracy were chosen as the best model for predicting the equipotential zones of Q nats. Also, the area of regions with good to outstanding potential for the existence or reclamation of Qanats was estimated to be 38427-38884 km^2 (13.15%-13.31% of the total area) using ESMs and RF models, respectively. Another goal of the present study was to carefully examine the significance of environmental factors in assessing ground water potential, which showed that climatic factors (BIO4), physiographic factors (DEM& topographic wetness index & slope) and soil factors (Organic 60-100 cm, Cations 60-100 cm), Land Surface Temperature) and geological factor cannot be negl cted. This confirms the significant role of RS in modeling. RS data may be a viable substitute for regions with no data or those that are hard to access. In general, combining RS with DEM data reveals a substantial number of significant relationships in groundwater research.

Regarding transferability, algorithms are free in the R software and can be a valuable asset for land resource managers. In addition, the ASTER used in this research is free and accessible, even though the locations and lithology of Qanats may not be accessible in some regions, which may limit the method's applicability. As for presence points, they may be replaced with data obtained from well discharge. If there is no access to the location of the Qanats, researchers can use expert-judgment-based approaches, such as hierarchical analysis, to define the weights and reach the maps needed. Considering these limitations, it is recommended to use the current method, with modifications to other factors, to address data inadequacy.

Acknowledgements

Authors would like to express their sincere gratitude to the University of Kashan.

Ethical Permission

Ethical approval was not required for this study as it did not involve human participants or animals.

Authors' Contribution

J Momeni Damaneh: Methodology, formal analysis, and writing, **M Ehteram**: writing, review, editing, and formal analysis, and **F Panahi**: formal analysis and writing.

Conflict of Interest

The authors declare that they have no conflicts of interest.

Funding/Support

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

AI Use Declaration

The authors have employed AI (ChatGPT-4, 2025) for language editing to improve the readability of the work. The authors have reviewed and edited the output and take full responsibility for the content of the publication.

References

1. Naghibi S.A., Pourghasemi H.R., Pourtaghi Z.S., Rezaei, A. Groundwater qanat potential mapping using frequency ratio and Shannon's entropy models in the Moghan watershed, Iran. Earth Sci. Inform. 2015; 8(1): 171–186.
2. 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. Climatol. 2018; 131:967–984.
3. Kharrat F.M., Ragala R., Moussa M. Promoting a sustainable traditional technique of aquifers' water acquisition common to arid lands: the case of Ghassem Abad Qanat in Yazd Province (Iran). Water Supply 2019; 19(2): 527–535.
4. Yazdi A.A., Khaneiki M.L. Veins of the desert: a review on Qanat/Falaj/Karez. River Publishers 2019; 239pp.
5. Salih A. (006). Qanats: a unique groundwater management tool in arid regions: the case of Bam region in Iran. In: Proceedings of the international symposium on groundwater sustainability (ISGWAS), Alicante, Spain.
6. Fattahi M. OSL dating of the Miam Qanat (Kāriz) system in NE Iran. J. Arcaeol. Sci. (2015) 59(1): 54–63.
7. Labbaf Khaneiki M. Qanat and territorial cooperation in Iran. Water Hist. 2019; 11(3):185 –206.
8. Semsar Yazdi A.A., Labbaf Khaneiki M. (017). Qanat knowledge: Construction and maintenance. Springer. 2018:179p.
9. Beaumont P. (971). Qanat systems in Iran. Hydrol. Sci. J. 1971; 16(1): 9–50.
10. Abbasnejad A., Abbasnejad B., Derakhshani R., Sarapardeh A.H. Qanat hazards in Iranian urban areas: explanation and remedies. Environ. Earth. Sci. 2016; 75(19): 1306.
11. IMA (2020). National Qanat geo-information system, archived by the Iranian Ministry of Agriculture.
12. Zamanirad M., Sedghi H., Sarraf A., Saremi A., Rezaee, P. Potential impacts of climate change on groundwater levels in the Kerdi- Shirazi plain, Iran. Environ. Earth. Sci. 2018; 77(11): 415.
13. Behling R., Roessner S., Foerster S., Saemian, P., Tourian M.J., Portele T.C., Lorenz C. Interrelations of vegetation growth and water scarcity in Iran revealed by satellite time series Sci. Rep. 2022;12(1): 0784.
14. Mansouri Daneshvar M. R., Ebrahimi M., Nejadsoleymani, H. An overview of climate change in Iran: facts and statistics. Environ. Syst. Res. 2019;8(1):1-10.
15. Pourghasemi H. R., Beheshtirad M. Assessment of a data-driven evidential belief function model and GIS for groundwater potential mapping in the Koohrang Watershed, Iran. Geocarto. Int. 2015;30(6):662-685.
16. Yu H., Wen X., Wu, M., Sheng D., Wu, J., Zhao, Y. Data-based groundwater quality estimation and uncertainty analysis for irrigation agriculture Agric. Water Manage. 2022;262(1):107423..
17. Masoudi R., Mousavi S. R., Rahimabadi P. D., Panahi M., Rahmani A. Assessing data mining algorithms to predict the quality of groundwater resources for determining irrigation hazard. Environ. Monit. Assess. 2023;195(2): 319.
18. Pradhan B. A. A comparative study on the predictive ability of the decision tree, support vector machine, and neuro-fuzzy models in landslide susceptibility mapping using GIS. Comput. Geosci. 2013;51:350–365. .
19. Naghibi S. A., Pourghasemi H. R., Dixon B. (016). GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. Environ. Monit. Assess. 2016;188(1):44.
20. Chen W., Panahi M., Khosravi K., Pourghasemi H. R., Rezaie, F., Parvinnezhad, D. (019). Spatial

prediction of groundwater potentiality using ANFIS ensembled with teaching-learning-based and biogeography-based optimization, J. Hydrol. 572(1): 435–448.
21. Golkarian A., Naghibi, S. A., Kalantar B., Pradhan B. Groundwater potential mapping using C5.0, random forests, and multivariate adaptive regression splines in GIS. Environ. Monit. Assess. 2018;190(3):149.
22. Momeni Damaneh J., Ahmadi J., Rahmanian, S., Sadeghi S. M. M., Nasiri V., Borz, S. A. Prediction of wild pistachio ecological niche using machine learning models. Ecol. Inform. 2022; 72(8):101907.
23. Momeni Damaneh J., Gerami S., Hajibigloo M. Remote Sensing and Machine Learning for Groundwater Spring Potential Assessment in Eastern and Northeastern Iran. Water. Harvest. Res. 2025;8(1):105-122.
24. Thuiller W., Lafourcade B., Engler R., Araújo M. B. Biomod - a platform for ensemble forecasting of species distributions. Ecograph. 2009;32 (3):369-373.
25. Fielding A. H., Bell J. F. (997). A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 1997;24(1):38-49.
26. Walther G. R., Post E., Convey P., Menzel A., Parmesan C., Beebee T. J., Bairlein F. Ecological responses to recent climate change. Nature. 2002; 416(6879):38-395.
27. Smeeton N.C. Early history of the Kappa Statistic. Biometrics. 1985;41 (3):7-95.
28. Galton F. Finger Pints. Macmillan, London. 1892;246p.
29. Saha S., Roy J., Arabameri A., Blaschke T., Tien Bui D. . Machine learning-based gully erosion susceptibility mapping: A case study of eastern India. Sensor. 2020;20(5):1313.
30. Yi Y. J., Cheng X., Yang Z. F., Zhang S. Maxent modeling for predicting the potential distribution of the endangered medicinal plant (H. riparia Lour.) in Yunnan, China. Ecol. Eng. 2016;92(1):260-269.
31. Swets J. A. Measuring the Accuracy of diagnostic systems. Science. 1988; 240 (4857):1285-1293.
32. Naghibi S. A., Ahmadi K., Daneshi, A. Application of support vector machine, random forest, and genetic algorithm-optimized random forest models in groundwater potential mapping. Water Resour. Manag. 2017;31 (9):2761-2775.
33. Moghaddam D. D., Rahmati O., Panahi M., Tiefenbacher J., Darabi, H., Haghizadeh A., Torabi Haghighi A., Nalivan O. A., Tien Bui D. The effect of sample size on different machine learning models for groundwater potential mapping in mountain bedrock aquifers. Catena. 187(1):104421.
34. 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(3):679.
35. Catani F., Lagomarsino D., Segoni S., Tofani V. Landslide susceptibility estimation using the random forests technique: sensitivity and scaling issues. Nat. Hazard. Earth. Syst. 2013; 13(11):2815-2831.
36. Caruana R., Niculescu-Mizil A. An empirical comparison of supervised learning algorithms. In Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, Pennsylvania, 2006; 161–168.
37. Naghibi S. A., Dolatkordestani M., Rezaei A., Amouzegari P., Heravi M.T., Kalantar B., Pradhan B. Application of rotation forest with decision trees as base classifier and a novel ensemble model in spatial modeling of groundwater potential. Envirron. Monit. Assess. 2019;191(4):248.
38. Mousavi S.M., Golkarian A., Naghibi S.A., Kalantar B., Pradhan B. GIS-based groundwater spring potential mapping using boosted regression trees and probabilistic frequency ratio models in Iran. Aim. Geosci.2017;3(1):91-115.
39. 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(19):2285.
40. 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; 57 (6):735–752.