



## UWS Academic Portal

### **Comparative analysis of GIS and RS based models for delineation of groundwater potential zone mapping**

Islam, Fakhru; Tariq, Aqil ; Guluzade, Rufat; Zhao, Na; Shah, Safeer Ullah; Ullah, Matee; Hussain, Mian Luqman; Ahmad, Muhammad Nasar; Alasmari, Abdulrahman; Alzuaibr, Fahad M.; El Askary, Ahmad; Aslam, Muhammad

*Published in:*  
Geomatics, Natural Hazards and Risk

*DOI:*  
[10.1080/19475705.2023.2216852](https://doi.org/10.1080/19475705.2023.2216852)

Published: 01/06/2023

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication on the UWS Academic Portal](#)

#### *Citation for published version (APA):*

Islam, F., Tariq, A., Guluzade, R., Zhao, N., Shah, S. U., Ullah, M., Hussain, M. L., Ahmad, M. N., Alasmari, A., Alzuaibr, F. M., El Askary, A., & Aslam, M. (2023). Comparative analysis of GIS and RS based models for delineation of groundwater potential zone mapping. *Geomatics, Natural Hazards and Risk*, 14(1), [2216852]. <https://doi.org/10.1080/19475705.2023.2216852>

#### **General rights**

Copyright and moral rights for the publications made accessible in the UWS Academic Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

#### **Take down policy**

If you believe that this document breaches copyright please contact [pure@uws.ac.uk](mailto:pure@uws.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## Comparative analysis of GIS and RS based models for delineation of groundwater potential zone mapping

Fakhrul Islam, Aqil Tariq, Rufat Guluzade, Na Zhao, Safeer Ullah Shah, Matee Ullah, Mian Luqman Hussain, Muhammad Nasar Ahmad, Abdulrahman Alasmari, Fahad M. Alzuaibr, Ahmad El Askary & Muhammad Aslam

To cite this article: Fakhrul Islam, Aqil Tariq, Rufat Guluzade, Na Zhao, Safeer Ullah Shah, Matee Ullah, Mian Luqman Hussain, Muhammad Nasar Ahmad, Abdulrahman Alasmari, Fahad M. Alzuaibr, Ahmad El Askary & Muhammad Aslam (2023) Comparative analysis of GIS and RS based models for delineation of groundwater potential zone mapping, *Geomatics, Natural Hazards and Risk*, 14:1, 2216852, DOI: [10.1080/19475705.2023.2216852](https://doi.org/10.1080/19475705.2023.2216852)

To link to this article: <https://doi.org/10.1080/19475705.2023.2216852>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 01 Jun 2023.



Submit your article to this journal [↗](#)



Article views: 215



View related articles [↗](#)



View Crossmark data [↗](#)



## Comparative analysis of GIS and RS based models for delineation of groundwater potential zone mapping

Fakhrul Islam<sup>a</sup>, Aqil Tariq<sup>b,c</sup> , Rufat Guluzade<sup>d</sup>, Na Zhao<sup>e</sup>, Safeer Ullah Shah<sup>f</sup>, Matee Ullah<sup>g</sup>, Mian Luqman Hussain<sup>h</sup>, Muhammad Nasar Ahmad<sup>c</sup>, Abdulrahman Alasmari<sup>i</sup> , Fahad M. Alzuaibr<sup>j</sup>, Ahmad El Askary<sup>j</sup> and Muhammad Aslam<sup>k</sup>

<sup>a</sup>Department of Geology, Khushal Khan Khattak University, Karak, Pakistan; <sup>b</sup>Department of Wildlife, Fisheries and Aquaculture, College of Forest Resources, Mississippi State University, MS, USA; <sup>c</sup>State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China; <sup>d</sup>School of Earth Science and Engineering, majoring in Geodesy and Survey Engineering, Hohai University, Nanjing, China; <sup>e</sup>State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing, China; <sup>f</sup>Board of Revenue, Government of Pakistan, Peshawar, KPK, Pakistan; <sup>g</sup>Faculty of Earth sciences, Geography and Astronomy, University of Vienna, Vienna, Austria; <sup>h</sup>National Centre of Excellence in Geology, University of Peshawar, Peshawar, KPK, Pakistan; <sup>i</sup>Department of Biology, Faculty of Science, University of Tabuk, Tabuk, Saudi Arabia; <sup>j</sup>Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Taif University, Taif, Saudi Arabia; <sup>k</sup>School of Computing Engineering and Physical Sciences, University of West of Scotland, Paisley, UK

### ABSTRACT

Groundwater is a crucial natural resource that varies in quality and quantity across Khyber Pakhtunkhwa (KPK), Pakistan. Increased population and urbanization place enormous demands on groundwater supplies, reducing both their quality and quantity. This research aimed to delineate the groundwater potential zone in the Kohat region, Pakistan by integrating twelve thematic layers. In the current research, Groundwater Potential Zone (GWPZ) were created by implementing Weight of Evidence (WOE), Frequency Ratio (FR), and Information Value (IV) models of the Kohat region. In this study, we used Sentinel-2 satellite data were utilized to generate an inventory map of groundwater using machine learning algorithms in Google Earth Engine (GEE). Furthermore, the validation was done with a field survey and ground data. The inventory data was divided into training (80%) and testing (20%) datasets. The WOE, FR, and IV models are applied to assess the relationship between inventory data and groundwater factors to generate the GWPZ of the Kohat region. Finally, the current research results of Area Under Curve (AUC) technique for WOE, FR, and IV models were 88%, 91%, and 89%. The final GWPZ can aid in better future planning for groundwater exploration, management, and supply of water in the Kohat region.

### ARTICLE HISTORY

Received 15 February 2023  
Accepted 17 May 2023

### KEYWORDS

Ground water potential zones; GIS; RS; FR; AUC

## 1. Introduction

Groundwater is an important natural resource that makes up about 34% of the world's freshwater supply (Tariq, Siddiqui, et al. 2022). It is the primary water supply and is regarded as less contaminated than other water sources. It supplies approximately half of the freshwater that can easily be accessed and used for cleaning, drinking, and cooking regularly (Termeh et al. 2019). Groundwater meets the requirements of 97% of the world's population for freshwater and provides 50% of the world's irrigation (Tariq and Shu 2020). It can be considered the essential capital natural possessions that occurred in the sediments and fractures of soil and rock (Ahmad et al. 2020). Groundwater is commonly utilized for domestic, industrial, and farming purpose in numerous parts of the world (Mumtaz et al. 2023). The tremendous demand for groundwater is increasing rapidly, and this rising need for water frequently causes overutilization, which is striking massive stress on the inadequate groundwater source. Furthermore, freshwater matter has become a critical issue in the tropical and subtropical areas of the world due to unscientific irrigation, exploration, urbanization, and changes in climatic factors.

Furthermore, these methods only sometimes account for the several parameters that influence groundwater's existence, storage, and mobility in rocks and soil (Keesstra et al. 2012). While currently, with the advancement of computer technology, geospatial techniques have become the most effective, emerging, and innovative ways to delineate potential groundwater regions. These methods can be applied locally and regionally based using ground and satellite data. These ground and remote sensing (RS) data are processed in the geographic information system (GIS) platform to detect and categorize the influencing parameters of groundwater (Basharat et al. 2022). RS provides inexpensive data as input to the GIS platform for the regional and inaccessible regions with a temporal, spectral, and spatial resolution (Majeed et al. 2022; Sadiq Fareed et al. 2022; Tariq, Mumtaz, et al. 2023). Satellite data can compute geological information (fault, fold, fractures, lithology) and topographic, climatic, and hydrological parameters for groundwater assessment. Geospatial technology is an innovative science to collect, store, display, and analyze various ground and RS data for groundwater in the form of surface water inventory (e.g. dams, ponds, open wells, and springs), groundwater demarcation, surface water modelling, and groundwater contamination (Siddiqui et al. 2020; Zainab et al. 2021; Tariq, Yan, et al. 2022; Tariq, Jiango, Li, et al. 2023).

In the past, numerous GIS and RS based models have been applied by scientist, i.e. Evidential Belief Function (EBF) (Shah et al. 2021), Weight of Evidence (WOE) (Lee, Kim, et al. 2012), Frequency Ratio (FR) (S. Hasan AL-Zuhairy et al. 2017), Decision Tree (DT), (Lee, Song, et al. 2012), Classification and Regression Tree (CART), (Mohammadi et al. 2021), Boosted Regression Tree (BRT), (Kordestani et al. 2019), Artificial Neural Network (ANN), (Lee, Song, et al. 2012), Multivariate Adaptive Regression Splines (MARS), (Kalantar et al. 2018), Binary Logistic Regression (BLR), (Chen et al. 2018), Analytic Hierarchy Process (AHP), (Singh et al. 2018), Random Forest (RF), (Tariq et al. 2021), Fuzzy Logic (FL), (Shahid et al. 2002), Support Vector Machine (SVM), (Lee et al. 2017), Multi-criteria Decision Analysis (MCDA)

(Kaliraj et al. 2014), Quadratic Discriminant Analysis (QDA) (Baloch et al. 2021), K-Nearest Neighbour (KNN) (Naghibi et al. 2018). The SVM is the technique to predict Groundwater Potential Zone (GWPZ) (Eid et al. 2023).

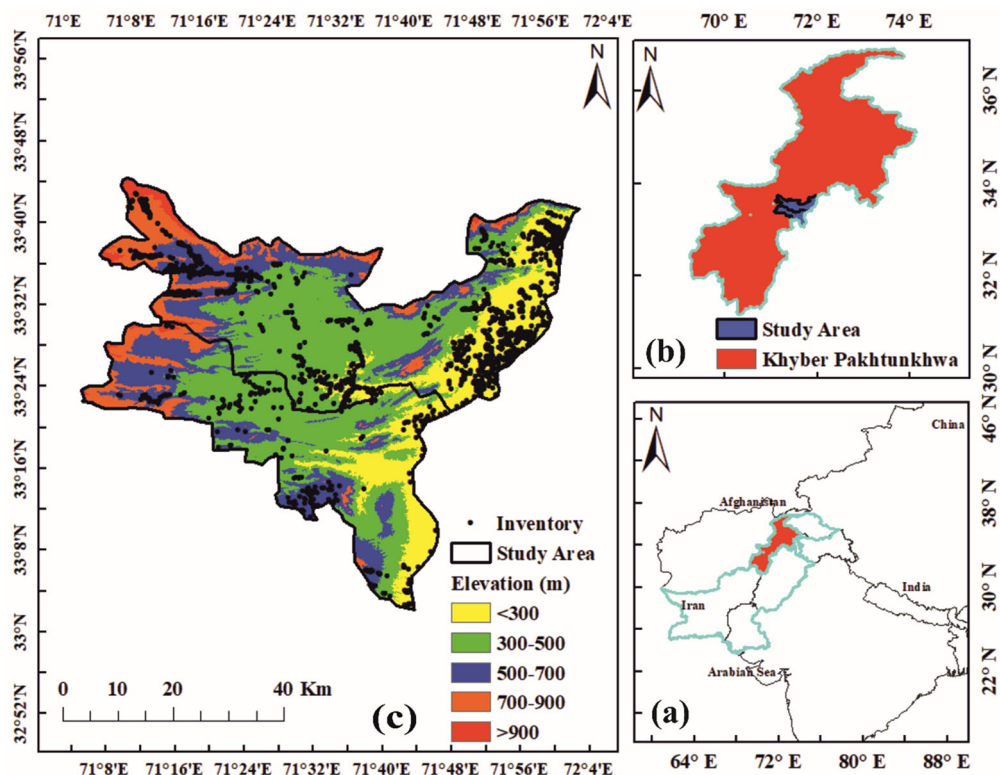
Groundwater is most vital and significant natural resource for sustainability due to its agriculture-dependent economy in Pakistan. Groundwater is a vital element in the economy, but human population, industrialization, unscientific exploration, and groundwater mismanagement have twisted a chief risk to this treasured energy source (Moazzam et al. 2022). Therefore, GWPZ is an indispensable technique for mapping and managing the precious water resources in the area of interest (Baloch et al. 2021). Numerous field survey mechanisms, i.e. geological, geophysical, and hydrological studies, have been used by researchers to demarcate potential groundwater zones (Israil et al. 2006). These methods need several human resources financial budget, and it can be most time-consuming.

In this research, WOE, FR, and IV models were utilized to locate the GWPZ in the KPK region of Pakistan. Although several studies have been conducted across Pakistan utilizing RS and GIS techniques to delineate the groundwater potential map, none of those studies have been conducted in this region where sustainable groundwater resources management is essential for the industrial, commercial development, and economy of the country. In the current study, we used twelve influencing parameters that are considered significant to explore the deficient, low, medium, high, and very high potential groundwater regions. These parameters were prepared in the GIS platform from various ground and RS data. In the current research, three geospatial techniques were used to compute the association of influencing parameters with groundwater inventory data and to delineate potential groundwater regions in district Kohat, Pakistan. These GIS-based groundwater mapping models have not been investigated previously in the district Kohat. The final GWPZ map can be helpful for decision-makers to assess and manage groundwater in various regions of the study area.

## 2. Material and methodology

### 2.1. Study area

The current research is conducted in district Kohat, situated in the southern part of Khyber Pakhtunkhwa (KPK), Pakistan. The Kohat region is geographically extended from 33° 35' 13" 33° 49' 73" N and 71° 52' 49" E to 71° 26' 32" (Figure 1a–c) (Hussain 2014). The study area occurred at an elevation of about 2000 m. Climatically the research region is considered a limited steppe climate region with slight precipitation throughout the year. In Kohat District, the summers are long, hot, humid, and clear, while the winters are brief, cold, and mostly clear. Both seasons have clear skies. Temperatures below  $-0.5^{\circ}\text{C}$  or above  $43.33^{\circ}\text{C}$  are extremely uncommon throughout the year. On average, the temperature ranges from  $2.2^{\circ}\text{C}$  to  $39.45^{\circ}\text{C}$  (Azra et al. 2019). Geologically the current study area is situated in Kohat Plateau. The study area includes fold and thrust belt collections which are thin-skinned structures covered by thick-skinned structures. Compressional structures subject a significant portion of the plateau; however, the strike-slip faulting is limited to the southern Kohat plateau (Hussain and Zhang 2018). The Kohat plateau is mainly



**Figure 1.** (a) Geographical location of Pakistan, (b) Provincial boundary of KPK where study area exists, and (c) Location map of study area with elevation.

occupied by lithologies of Eocene limestone, shale, evaporates, and subordinate clays, and younger clastic sedimentary rocks of the Miocene–Pliocene age (Hussain et al. 2021). The age of sedimentary rocks in the plateau is composed of Paleocene to Pliocene, which was first deposited on the northern Indian plate margin (Tariq and Qin 2023).

## 2.2. Datasets

Various datasets are applied to generate different parameters used in the current work. The datasets utilized in the current research comprise organization ground (field survey and ground data) and RS data. The ground and satellite information applied to prepare twelve influential parameters for groundwater potential were acquired from appropriate national and international research platforms. The data details and sources of information are mentioned in Table 1.

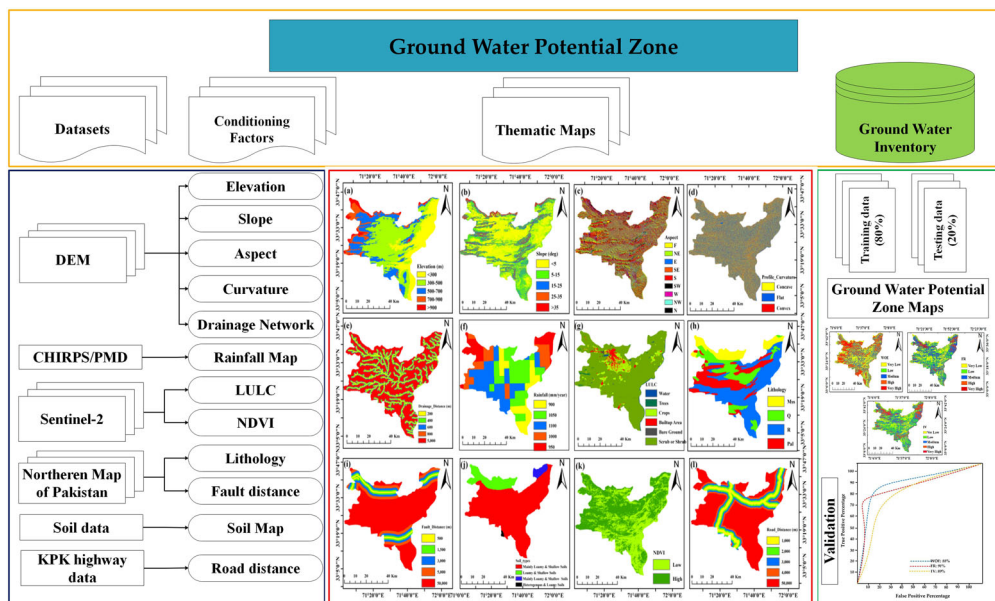
## 2.3. Methodology

The study was established in four phases: i) preparation of ground water inventory map of the study area using different geospatial, machine learning, and field survey

Table 1. Description of datasets were used in this research.

Data	Scale/resolution	Data availability	The data supporting this study's findings are openly available in [USGS] at <a href="https://earthexplorer.usgs.gov">https://earthexplorer.usgs.gov</a> .	Data availability statement/source	Maps of parameters
Sentinel –2	10 m	Freely available	The data that support the findings of this study are openly available at <a href="https://asf.alaska.edu">https://asf.alaska.edu</a>		LULC, NDVI Map and ground water inventory
ALOS DEM	12.5m	Freely available	The data supporting this study's findings are available in [UCSB] at <a href="https://www.chc.ucsb.edu">https://www.chc.ucsb.edu</a> . The spatial resolution of CHIRPS is 0.05° (5.54 Km) and daily gridded.		Elevation, Slope, Aspect, Curvature, and drainage
CHIRPS	0.05°	Freely available			Rainfall maps
Soil	1: 2000000	Freely available	Freely available at Soil type of Pakistan and FAO		Soil Texture Map
Geology	1:650000		Geological Map		Lithology and Fault map





**Figure 2.** Flowchart of present research work for the current research study.

techniques, ii) generation of twelve influential groundwater parameters, iii) generating GWPZ using three geospatial models like WOE, FR, and IV and, vi) performing validation and accuracy assessment using AUC technique. The comprehensively organized methodology for the present investigation is shown in Figure 2.

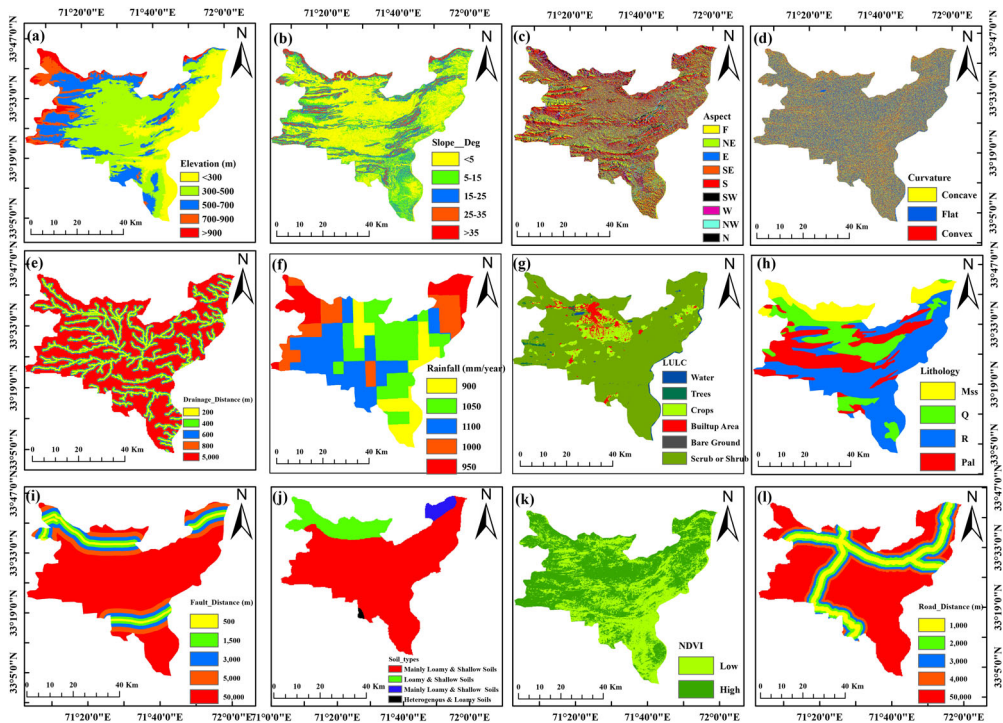
### 2.3.1. Inventory map of surface water bodies

The accurate water inventory map is the primary and essential parameter to generate GWPZ for the region of interest. The ground and RS data for the inventory map were collected from various public organizations and satellite sources. The inventory map of different water bodies was prepared from Sentinel-2 using a ML-model. The inventory map was validated and verified with ground data collected from the public department of district Kohat and various field surveys in the Kohat region. Finally, detected inventory data of current research is divided into training (80%) and testing (20%) datasets (Zhu et al. 2022).

### 2.3.2. Preparation of GWPZ conditioning parameters

Considering groundwater potential, conditioning parameters is a significant task affecting the final output map of GWPZ; hence, conditioning parameters should be cautiously designated (Bui et al. 2019). The existence and yield of groundwater in a specified aquifer are influenced by numerous parameters. In the present study, twelve influential conditioning factors like elevation, slope angle, aspect, curvature, drainage network, rainfall, LULC, soil, NDVI and the road distance are considered to evaluate the influences of mentioned parameters on groundwater potential in the study area.





**Figure 3.** Parameters for GWPZ; (a) elevation, (b) slope, (c) aspect, (d) curvature (e) drainage network (f) rainfall, (g) LULC, (h) lithology, (i) fault distance, (j) soil types, (k) NDVI (l) road distance.

**2.3.2.1. Elevation.** Altitude influences the potential groundwater zone as it is conversely associated with the reservoir (Karimi-Rizvandi et al. 2021). The altitude of the present research area is computed from having 12.5 m spatial resolution and reclassified into five categories in ArcMap 10.8 as revealed in Figure 3a.

**2.3.2.2. Slope.** The slope gradient is another significant parameter for groundwater potential because the slope angle directly influences the amount of rainwater water intrusion and surface run-off in any region. A steep slope gradient negatively impacts groundwater reservoirs because a higher slope enables a rapid run-off area and reduces water infiltration. In contrast, a low slope promotes water infiltration and potential recharge area (Maskooni et al. 2020). The slope of the Kohat region reclassified into five classes, i.e.  $<5^\circ$ ,  $5-15^\circ$ ,  $15-25^\circ$ ,  $25-35^\circ$  and  $>35^\circ$  using ArcGIS 10.8 as shown in Figure 3b. The highest slope of the present research area is verified as  $78^\circ$ , while the lowest slope of the region is recorded as  $0^\circ$ .

**2.3.2.3. Aspect.** The slope aspect presents slope directions that affect the quantity of precipitation, radiation of the sun, wind speed, and LULC, which concomitantly strike the amount of water permeation to the pore spaces of sediments influencing groundwater potential in the region (Solomon and Quiel 2006). The aspect of the present area is generated and reclassified into nine classes, as revealed in Figure 3c.

**2.3.2.4. Curvature.** The curvature map shows the association with the capacity to store and hold water reserves on the area of surface. Usually, the dipped structures accumulate more water bodies (Pham et al. 2019). The curvature of the Kohat region is calculated from ALOS DEM having 12.5m resolution and reclassified into open, flat, and convex groups, as mentioned in Figure 3d.

**2.3.2.5. Drainage network.** Drainage network presents an inverse association with the percolation of water in fracture and sediments of strata because river density discourages water retention (Kordestani et al. 2019). As the river network density is high, water recharge in the area will be low and vice versa because river density favours surface runoff and decreases infiltration. The five buffers were applied to the stream network of the present research area stream, as shown in Figure 3e.

**2.3.2.6. Rainfall.** Climatic parameters are influential in controlling the water table from the seasonal perception. Precipitation is a significant climatic factor that affects groundwater recharge. Precipitation is considered a vital parameter for potential groundwater mapping because the probabilities of penetration are more in cases of high rainfall; consequently a chief source of water recharge in the area (da Silva Monteiro et al. 2022). The precipitation of the present study area is computed from Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) of 2010 to 2022 using a ML Algorithm in GEE. The concluding precipitation map was then reclassified into five classes using GIS environment as shown in Figure 3f.

**2.3.2.7. LULC.** LULC is a significant parameter influencing groundwater recharge, occurrence, and availability. LULC presents environmental parameters having a significant impression on groundwater because they affect penetration and surface runoff (Bui et al. 2019). Moreover, bare ground and built-up regions usually display low potential, while vegetation and the area near water reservoirs illustrate higher groundwater potential. LULC map of the Kohat region is generated from Sentinel-2 data in GEE using a ML algorithm. This causative factor was further categorized into six classes for evaluating these classes on groundwater, as shown in Figure 3g.

Confusion matrices were used to create classification accuracy processes, such as overall accuracy, omission and commission errors, and Cohen's kappa statistic. These classification accuracy matrices were derived using confusion matrices (Firdaus 2014). As per reference data, the commission error is the percentage of pixels that have been incorrectly assigned to classes they do not belong. On the other hand, the percentage of pixels that should have been assigned to a particular class according to the reference data but have not been assigned to that class is an omission error. It was figured out how to determine the omission and commission error for each LULC class and estimate the average for all classes.

**2.3.2.8. Lithology.** The lithology of strata controls the porosity and permeability of aquifers and influences groundwater due to its conductivity and penetration. These rock and soil properties affect groundwater's existence, accumulation, and mobility

(Muavhi et al. 2022). The lithology of the Kohat region is extracted from the Northern Geological Map of Pakistan, as shown in Figure 3h.

**2.3.2.9. Fault distance.** The fault Buffer of various distance gaps was selected for analysis because it influences the subsurface flow of fluids (Yin et al. 2018). Therefore, geological faults and fractures are critical parameters in detecting groundwater sources. The fault parameter is digitized in this research, as shown in Figure 3i. Five buffers were applied to calculate the relationship of fault with groundwater potential in the study area.

**2.3.2.10. Soil types.** Soil is the uppermost horizon of land which helps in water infiltration. Soil type is a crucial factor in investigating potential groundwater mapping as the penetration capability of an area controlled by pore spaces of soil (Tariq, Jiango, Lu, et al. 2023). Similarly, the soil is a significant conditioning parameter in the groundwater potential zone mapping. The soil's texture and structure determine its permeability, which in turn represents the soil's capacity for allowing water and other substances to penetrate it (Tariq, Mumtaz, et al. 2023). The soil type map of the research region is produced from FAO and the soil survey of Pakistan, as shown in Figure 3j.

**2.3.2.11. NDVI.** There is a secondary association between NDVI with groundwater. For example, the region increases plant density, the groundwater table decrease, and vice versa. The value of NDVI in the various depths of the water table revealed that dense vegetation occurs in shallow water regions (Islam et al. 2022). The NDVI map of the Kohat region is calculated from Sentinel-2 using machine learning techniques in GEE. The final NDVI map was reclassified into two classes in the GIS environment, as shown in Figure 3k.

**2.3.2.12. Road distance.** The road distance of District Kohat is generated using the Google Earth platform and road network map of KPK Highway Authority in ArcGIS platform as shown in Figure 3l.

### 3. Ground water potential zone mapping models

Geospatial modelling was applied in the current research to evaluate the association of groundwater conditioning parameters and groundwater inventory data to generate GWPZ of the Kohat region. The explanation of the applied three models in the present study is as follows.

#### 3.1. WOE model

This GIS-based technique employed linear logic based on Bayesian law to combine data to approximate events' non-conditional and conditional probability (Elmoulat et al. 2015). WOE models compute the spatial association of dependent variables, i.e. water bodies' location and independent variables like groundwater potential mapping conditioning parameter and compute the weight of each class of parameters. The WOE method was first considered to evaluate mineral potential mapping using GIS-based

modelling (Bonham-Carter et al. 1989). In this technique, the  $W^+$  and  $W^-$  weights should be considered as the dynamic aspects. The weight of conditioning factors (B) established on the existence or non-existence of the water bodies (C) of the study region is estimated using the following Eqs. (1)–(3) (Bonham-Carter et al. 1989).

$$W^+ = \ln \frac{P\left(\frac{B}{C}\right)}{h\left(\frac{B}{C}\right)} \quad (1)$$

$$W^- = \ln \frac{h\left(\frac{\bar{B}}{\bar{C}}\right)}{h\left(\frac{B}{C}\right)} \quad (2)$$

$$L = W^+ - W^- \quad (3)$$

In the mentioned equation,  $p$  is the likelihood and  $\ln$  is the natural logs. However,  $BB$  and  $\bar{B}\bar{B}$  are the existence and nonexistence in the causative factor, correspondingly. Similarly,  $CC$  and  $\bar{C}\bar{C}$  show the occurrence and absence of inventory, respectively (Xu et al. 2012).  $W^+$  signifies the occurrence of the conditioning parameters at the spatial positions. Its amount demonstrates the positive relationship between conditioning parameters and water bodies occurrence, respectively. While  $W^-$  represents the nonappearance of groundwater parameters and suggests the level of a contrary relationship.

### 3.2. FR model

The FR technique is the finest bivariate statistical model applied as a valuable GIS-based model for evaluating groundwater inventory and groundwater conditioning parameters (Guru et al. 2017). Currently, the FR model has been effectively utilized for GWPZ in various regions of the world. The FR value equal to or greater than one shows a strong positive correlation between different variables. The following Eq. (4) calculation is applied to compute FR for all causative factors in the present study (Ahmad et al. 2022).

$$FR = \frac{E/F}{M/L} \quad (4)$$

Where  $FR$  = Frequency Ratio for each conditioning parameter,  $E$  = number of water body pixels in each landslide's causative parameter class,  $F$  = total number of all well pixels in research region,  $M$  = number of pixels in each landslide conditioning factors,  $L$  = total number of all pixels in study area.

### 3.3. IV model

In the current research, the IV model is applied to make GWPZ of the Kohat region. IV is one of the most suitable practices for choosing significant parameters, ranking variables based on their position, and computing their association with inventory

data of the study area in the predictive model (Pardeshi et al. 2013). IV model was first improved by Shano et al. (2020). This article considers the IV for each parameter class based on the presence of groundwater inventory pixels in the given region. The computed information value supports governing the role of each parameter class for groundwater occurrence (Ali et al. 2023). The conditional probability was calculated by dividing the groundwater pixels in each parameter class into pixels of a subclass of groundwater parameter, while the prior probability was considered by dividing the total groundwater pixels in the research region by the entire pixels in the whole research region using the Eq. (5) (Pardeshi et al. 2013).

$$W = \log \frac{\frac{MQox(R_o)}{MQ_o x(R_o)}}{\sum \frac{MQ_o x(R_o)}{\sum MQ_o x(M_i)}} \quad (5)$$

W symbolize the weight of parameters for groundwater.  $Mox(R_o)$  illustrates water number of pixels within class 'o'.  $MQ_o x(M_o)$  number of all pixels within class 'o',  $MQ_o x(R_o)$  total number of water pixels  $\sum MQ_o x(M_i)$  practice for entire number of pixels in region.

### 3.4. Delineation of the GWPZs

#### 3.4.1. Delineation of groundwater using WOE

The groundwater potential index (GWPI) was calculated (Eq. (6)) and mapped based on  $\tau$  values.

$$GWPI = \tau_1 + \tau_2 \dots + \tau_n \quad (6)$$

where  $\tau$  is the final weight for the WOE model.

#### 3.4.2. Delineation of groundwater using FR

In contrast to the WOE, the weightage of each class in FR is not determined based on the characteristics of the conditioning factor; instead, it is given in the form of the spatial occurrence of the wells in each class. This contrasts with the WOE, which determines the weightage of each class based on the properties of the conditioning factor. Similarly, the FR is computed for each of the conditioning variables. The succeeding scientific Eq. (7) has been applied to produce GWPZ of the Kohat region (Guru et al. 2017).

$$GWPZM = \sum_{i=1}^n FR_{ij} \quad (7)$$

#### 3.4.3. Delineation of groundwater using IV

The GWPZ can be produced for Kohat region using the Eq. (8).

$$GWPZ = W_E + W_S + W_A + W_C + W_{LULC} + W_L + W_P + W_F + W_R + W_D \quad (8)$$

$W_E$  = Weight of Elevation,  $W_S$  = Weight of Slope,  $W_A$  = Weight of Aspect,  $W_C$  = Weight of Curvature,  $W_{LULC}$  = Weight of Landuse Landcover,  $W_L$  = Weight of Lithology,  $W_F$  = Weight of fault,  $W_R$  = Weight of Road,  $W_P$  = Weight of Rainfall,  $W_D$  = Weight of stream network.

### 3.5. Validation of the GWPM

Evaluation of generated GWPZ is crucial because models without validation have no empirical value. Rather than using the hydraulic parameter of specific capacity, as previous studies did, an indirect indicator of groundwater yield measurement was used in the present research (Jha et al. 2010). From a groundwater sustainability point of view, groundwater yield measurement has been used widely by several researchers such as (Qureshi et al. 2010; Pardeshi et al. 2013; Ji et al. 2015; Fayez et al. 2018; Arabameri et al. 2019) for validation of GWPZ. For many investigations, the receiver operating characteristics (ROC) curve has been the gold standard for evaluating the precision of the GWPZ (Shirazi et al. 2012). The area under the AUC measures the accuracy with which a prediction system can determine whether or not an incident will occur (Shah et al. 2022). In order to validate the WOE, RF and IV-generated GWPZ, the healthy dataset (20%) was used for testing. Areas under the ROC curve were used to evaluate the GWPZ, spatial efficacy (AUC). The rate explains the accuracy with which the model and influencing variables predict the potential. AUC determines which model is superior, and the one with the most outstanding value wins (Rahman 2008).

## 4. Results

In this article, we developed an inventory of groundwater bodies from Sentinel-2, imageries using various advanced JavaScript algorithms, Google Earth Pro and Google Earth images. The spatial location of surface water bodies like well, ponds, and springs is mentioned in Figure 1. In the present research, we accomplished three bivariate models to generate GWPZ for the Kohat area.

### 4.1. WOE model

The contrast value can be computed from the calculation of both mentioned weights and calculate the association of both dependent and independent variables. The concluding LSM of the Kohat region is mentioned in Figure 4. Table 2 shows the analytical results of GIS-based models. The two variables' low correlation shows groundwater's low potential zone, and the high value illustrates the high groundwater potential zone in the research region. Based on the results of the WOE model in the elevation parameter, an altitude less than 500 m shows a strong association with groundwater. However, more than 800 m elevation class shows the slightest relationship with groundwater.



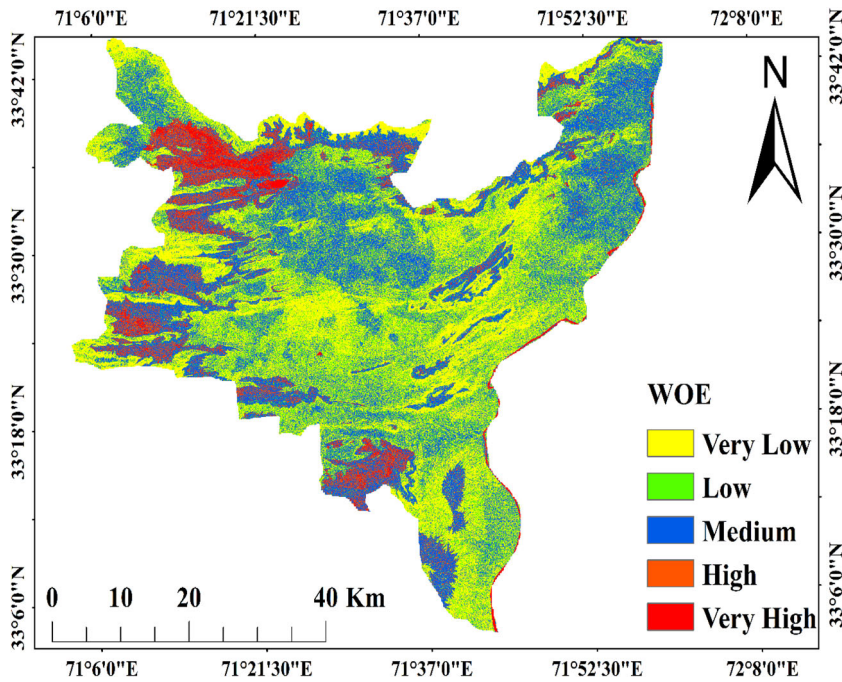


Figure 4. The WOE model for groundwater potential zone.

#### 4.2. FR model

The ultimate output map by FR is mentioned in Figure 5. Estimating the GWPZ with the FR model cannot be overstated. FR model carried out the GWPZ by correlating the various variables that conditioned the water with the specific locations of bore wells. In addition, a more excellent correlation value suggests a more significant groundwater potential and vice versa. Finally, LULC classes have a significant bearing on the effect of industrialization on the potential of groundwater. According to the findings of this research, the water body was a factor in the highly prospective ability of 9.923. The mining/industrial region and the vegetation cover area both discovered insignificant FR values when the FR was analyzed by the conditioning factor and the bore wells data. This is the case because FR is analyzed. The result suggested a low FR value because more data is needed from these classes' bore wells. In contrast, the vegetation cover was always found to influence the infiltration rate significantly.

The FR model's elevation class of less than 500 m, as shown in Table 2 in the present article, illustrates a strong correlation with the groundwater. In contrast, the less correlated elevation class with groundwater is more than 800 m elevation. In the results of the IV model, as shown in Table 2, the most significant elevation for groundwater is <500 m altitude, while the less critical class is more than 800 m. The results revealed that the correlation value of the WOE, FR, and IV model for elevation classes less than 500 m are 1.62, 3.39, and 1.22, respectively. However, the variables association results show that the correlation value of altitude class > 800 m for WOE, FR, and IV model are -1.84, 0.18, and -1.72, respectively. The results of three



**Table 2.** Statistical analysis for GWPZ of District Kohat, Pakistan.

Parameters	Class	No of pixels in class	No of Landslide pixels in a class	W <sup>+</sup>	W <sup>−</sup>	WoE	% Pixels in Class	% LS pixels in Class	(FR)	IV = log (A/B)
Elevation	< 500	8360	149	0.58	−0.25	0.83	22.16	39.21	1.77	0.57
	500–600	16236	138	−0.17	0.11	−0.28	43.04	36.32	0.84	−0.17
	600–700	8448	63	1.49	−5.51	7.00	22.40	16.58	0.74	0.55
	700–800	3234	22	−0.40	0.03	−0.43	8.57	5.79	0.68	−0.39
	> 800	1443	8	−0.60	0.02	−0.62	3.83	20	3.41	−0.60
Slope	< 5°	2211	76	1.11	0.00	0.73	5.87	20	3.41	1.09
	10–20°	1186	130	2.35	0.00	2.68	3.15	34.21	10.87	2.25
	15–25°	5998	70	−0.61	0.00	2.36	15.92	18.42	1.16	−0.60
	25–35°	11056	61	−0.13	0.00	1.22	29.35	16.05	0.55	−0.13
	> 35°	17223	43	−1.54	0.00	1.53	45.72	11.32	0.25	−1.53
Aspect	F	2779	115	1.44	−0.29	1.73	1.44	7.37	30.26	4.11
	NE	2970	31	0.03	0.00	0.04	0.03	7.87	8.16	1.04
	E	5471	42	−0.28	0.04	−0.32	−0.28	14.50	11.05	0.76
	SE	5673	49	−0.16	0.03	−0.18	−0.16	15.04	12.89	0.86
	S	5416	27	−0.71	0.08	−0.79	−0.71	14.36	7.11	0.49
	SW	3920	41	0.04	0.00	0.04	0.04	10.39	10.79	1.04
	W	4073	21	−0.68	0.06	−0.73	−0.68	10.80	5.53	0.51
	NW	3762	27	−0.34	0.03	−0.38	−0.34	9.97	7.11	0.71
Curvature	N	3610	27	−0.30	0.03	−0.33	−0.30	9.57	7.11	0.74
	Concave	14057	213	0.4	−0.4	0.77	37.27	56.05	1.50	0.41
	Flat	6955	117	0.5	−0.2	0.69	18.44	30.79	1.67	0.51
Distance to Stream	Convex	16709	50	−1.2	0.4	−1.67	44.30	13.16	0.30	−1.21
	< 200	5359	206	1.36	−0.63	2.00	14.27	54.21	3.80	1.33
	200–400	6009	66	0.08	−0.02	0.10	16.01	17.37	1.09	0.08
	400–600	6951	45	−0.45	0.08	−0.53	18.51	11.84	0.64	−0.45
	600–800	9656	33	−1.09	0.21	−1.30	25.72	8.68	0.34	−1.09
Precipitation (mm/year)	> 800	9069	30	−1.13	0.20	−1.32	24.16	7.89	0.33	−1.12
	<900	5829	30	−0.68	0.09	−0.77	15.49	7.89	0.51	−0.67
	900–950	10341	62	−0.53	0.14	−0.67	27.47	16.32	0.59	−0.52
	950–1000	11035	113	0.01	−0.01	0.02	29.32	29.74	1.01	0.01
	1000–1050	5823	115	0.68	−0.19	0.87	15.47	30.26	1.96	0.67
LULC	>1050	4612	60	0.26	−0.04	0.30	12.25	15.79	1.29	0.25
	Water	200	150	5.69	−0.50	6.19	0.53	39.47	74.41	4.31
	Trees	337	10	1.10	−0.02	1.12	0.89	2.63	2.94	1.08
	Crops	3581	52	0.37	−0.05	0.42	9.50	13.68	1.44	0.37
	Builtup Area	1650	6	−1.03	0.03	−1.06	4.38	1.58	0.36	−1.02
	Bare Ground	234	2	−0.17	0.00	−0.17	0.62	0.53	0.85	−0.16
	Scrub/Shrub	31700	160	−0.70	1.32	−2.01	84.08	42.11	0.50	−0.69
	Mss	4802	36	−0.30	0.04	−0.34	12.73	9.47	0.74	−0.30
Lithology	Q	7565	131	0.55	−0.20	0.75	20.06	34.47	1.72	0.54
	R	16554	175	0.05	−0.04	0.09	43.89	46.05	1.05	0.05
	Pal	8797	38	−0.85	0.16	−1.01	23.32	10	0.43	−0.85
Fault Buffer	<500	1300	40	1.14	−0.08	1.21	3.45	10.53	3.05	1.12
	1500	2381	62	0.97	−0.11	1.08	6.31	16.32	2.58	0.95
	3000	3420	40	0.15	−0.02	0.17	9.07	10.53	1.16	0.15
	5000	4341	45	0.03	0.00	0.03	11.51	11.84	1.03	0.03
	>5000	26285	193	−0.32	0.49	−0.81	69.68	50.79	0.73	−0.32
Soil	Loamy and shallow soil	7510	54	−0.67	0.64	−1.30	53	21	0.71	0.66
	Mainly loamy soil	30035	326	0.64	−0.67	1.55	27	57	2.69	1.23
NDVI	Low	15406	117	−0.29	0.16	−0.45	40.91	40.91	−0.01	−0.28
	High	22252	263	0.16	−0.29	0.45	59.09	59.09	0.01	0.16
Distance to Road	< 1000	3145	22	−0.37	0.03	−0.40	8.34	5.79	0.69	−0.36
	1000–2000	3993	40	−0.01	0.00	−0.01	10.58	10.53	0.99	−0.01
	2000–3000	4328	39	−0.11	0.01	−0.13	11.47	10.26	0.89	−0.11
	3000–4000	3569	70	0.68	−0.11	0.78	9.46	18.42	1.95	0.67
	> 4000	22693	209	−0.09	0.12	−0.21	60.15	55	0.91	−0.09

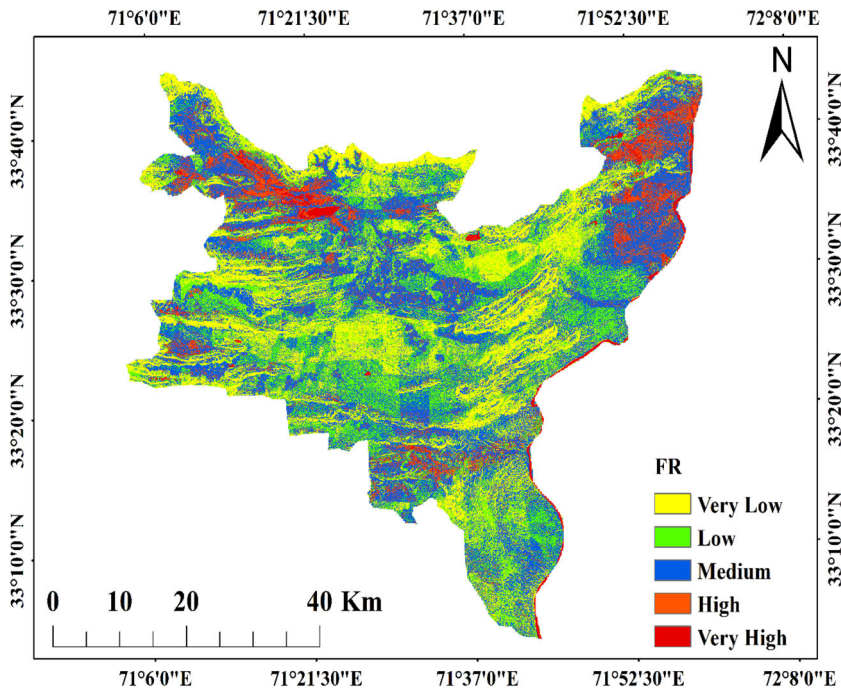


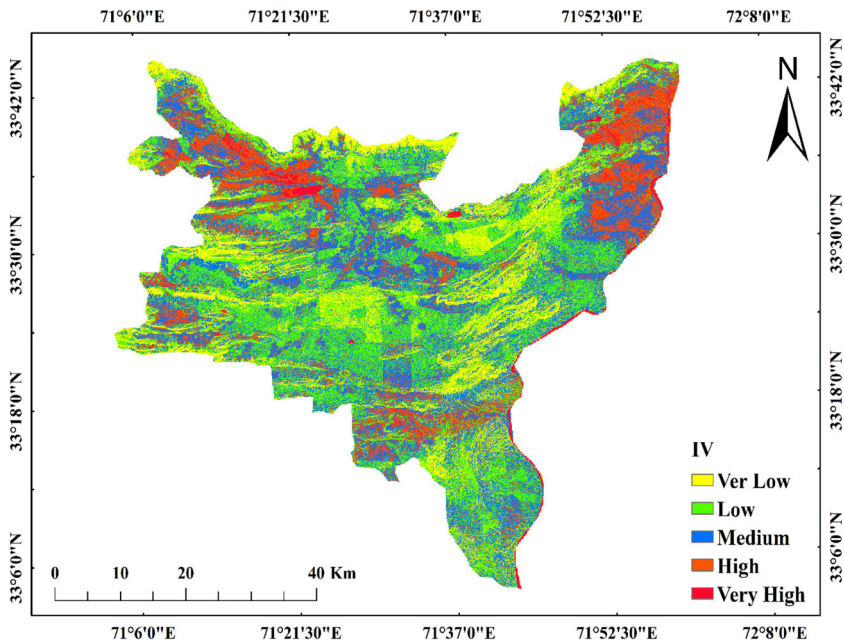
Figure 5. The FR model for groundwater potential.

bivariate models for elevation illustrate that low elevated area is more permeable and suitable for groundwater potential as compared to high elevated zones of Kohat district. The slope gradient parameter is considered a very significant parameter for GWPZ.

The bivariate results achieved from the association of groundwater conditioning parameters and groundwater pixels in the research region, as mentioned in Table 2, revealed that the most critical class for the groundwater potential is  $5^\circ$  followed by  $10^\circ$ – $20^\circ$ . The results also show that the most significant slope class for the current research area is  $>35^\circ$  slope followed by  $25^\circ$ – $35^\circ$ . Based on the slope angle parameter, the slope class with less than  $5^\circ$  has the maximum weight. The association rank for both variables in the present work for  $5^\circ$  by WOE, FR, and IV are 2.68, 3.5, and 2.37, respectively. The correlation between water inventory and groundwater conditioning parameters for slopes over  $35^\circ$  are  $-1.45$ ,  $0.25$ , and  $-1.53$  for WOE, FR, and IV models, respectively.

#### 4.3. IV model

The analytical values, as shown in Table 2 for the current research analysis between dependent and independent variables. It illustrated that a lower slope has the greatest likelihood of groundwater potential. In contrast, a steep slope has adverse impacts on the occurrence of groundwater due to high runoff in high slopes region. The analytical results of bivariate models explained that F is the most vital class of aspect, followed by NE of the Kohat area. The correlation results of the WOE, FR, and IV model for the F direction are



**Figure 6.** The IV model for groundwater potential.

1.73, 4.11, and 1.47, respectively. The NE and SW classes of aspects follow the F class of aspects. The less significant class of aspect is the S direction having  $-0.71$ ,  $0.65$ , and  $-0.70$  for WOE, FR, and IV model. According to the analysis for the association between groundwater data and curvature shape, the concave structure has the highest correlation value, i.e.  $0.99$ ,  $1.65$ , and  $0.50$  for WOE, FR, and IV, respectively. The concave structure is the most significant class of curvature for groundwater potential zone mapping. The results revealed that the spatial association of groundwater and conditioning parameters for WOE, FR, and IV models are  $-1.67$ ,  $0.29$ , and  $-1.21$ , respectively. As shown in Table 2, the results revealed that the most significant class of curvature is a concave structure, followed by Flat and convex structures. As shown in Table 2, the results revealed that groundwater is more likely to occur in a sense stream. There is a maximum likelihood of distance from the river of less than  $200$  m. The correlation value of less than  $200$  m class of stream is  $2.0$ ,  $3.80$ , and  $1.33$  for WOE, FR, and IV model, respectively. These results illustrate that fewer distances to rivers have had a more significant influence on groundwater potential (Figure 6).

However, most of the less significant class of stream parameters are greater than  $800$  m, followed by a  $600$  m– $800$  m range. The association between WOE, FR, and IV variables are  $-1.32$ ,  $0.33$ , and  $-1.12$ , respectively, for more significant than  $800$  m class of stream. In the current research, the precipitation map was formed from CHIRPS satellite data, followed by the reclassification into five categories to assess the relationship of rainwater factor with groundwater bodies. The outcomes, as revealed in Table 2 for rainfall, supported that rainfall is a significant aspect of groundwater potential. The results show that the  $1000$ – $1050$  mm/year precipitation class is the most significant for groundwater potential, followed by  $>1050$  mm/year. The  $1000$ – $1050$  class correlation

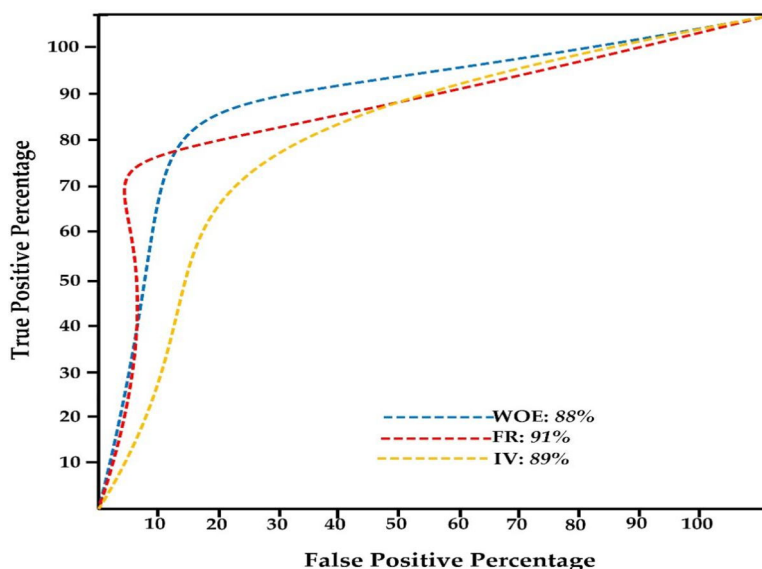
between rainfall and groundwater is 1.32, 2.73, and 1.0 for WOE, FR, and IV models. The precipitation class  $< 900$  mm/year has no significant impact on groundwater potential. The bivariate analysis for WOE, FR, and IV are  $-0.90$ ,  $0.45$ , and  $-0.79$ , respectively. Considering the above-mentioned statistical facts, it can be concluded that high precipitation classes show more groundwater occurrence and vice versa. In the current research results, the cropland area is the most important and influential class of LULC parameter in the study area. The analytical results of groundwater and conditioning parameters for WOE, FR, and IV are  $0.63$ ,  $1.70$ , and  $0.53$ , respectively. The agricultural land shows high potential results for groundwater because the agriculture region is recharged from the irrigation system of the study area. The scrub/shrub, forest, and urban class of LULC follow the agricultural land.

The analysis of both variables treasures that Q is the most influential geological formation of lithology parameters for groundwater potential in the Kohat region. The correlation of both variables is clear in WOE, FR, and IV model. The results for groundwater potential zone mapping of the present study show that the lithological parameter is the least significant class for groundwater. A fault is a significant parameter for groundwater percolation. Geological faults strongly influence groundwater mobility because they enhance the strata's mobility mechanism for groundwater. The maximum likelihood of groundwater potential in distances to a fault is  $< 500$  m buffer region. The correlation value of both variables of class  $< 500$  m for WOE, FR, and IV are  $0.60$ ,  $1.73$ , and  $0.55$ , respectively. The geological fault's  $> 5000$  m fault buffer has no significant impact on groundwater. The results of  $> 5000$  m buffer for the WOE, FR, and IV model are  $-0.45$ ,  $0.82$ , and  $-0.20$ , respectively. The results concluded that fault is the influential parameter for groundwater potential for the present research area. As shown in Table 2, the results explained that mainly loamy soil is the suitable class for groundwater in the Kohat area. The correlation of both variables for WOE, FR, and IV are  $1.55$ ,  $2.69$ , and  $1.23$ , respectively. As shown in Table 2, the present research results illustrate that NDVI is a crucial parameter for groundwater potential. Both variables' association ranks are  $0.57$ ,  $1.83$ , and  $0.27$  for WOE, FR, and IV model, respectively. The results of the present study between NDVI and GWPZ revealed that NDVI and groundwater Table depth are inversely related, i.e. high NDVI will have low water table depth and vice versa for the current investigation. The current study considered the road to compute the association between the road network in the Kohat region and the groundwater potential. The results explained that the most influential class of road network is a  $3000$ – $4000$  m buffer followed by  $> 4000$  m and  $2000$ – $3000$  m buffer. The correlation value for WOE, FR, and IV model are  $0.78$ ,  $1.95$ , and  $0.67$ , respectively. Table 2 revealed that the  $< 1000$  m class correlation values between the road and groundwater are  $-0.40$ ,  $0.69$ , and  $-0.36$  for WOE, FR, and IV model, respectively. The results of all road network buffers revealed that the road network has adverse impacts on groundwater in the study area.

## 4.4. Validation

### 4.4.1. Validation of models

In the modelling technique, validation of the model is a significant phase to accomplish the reliable scientific worth of the research project (Barakat et al. 2023). In



**Figure 7.** ROC curves for WOE, FR, and IV methods.

numerous research, the AUC technique was used to evaluate GWPZ. This ROC curve is considered a standard index for accuracy assessment. This technique has been extensively utilized for assessing techniques applied in various water research investigations. The receiver operating characteristics (ROC) graph validated the WOE, FR, and IV models. The region indicates the precision of the prediction or classification under the receiver operating characteristic curve (AUC) (Pourghasemi and Rossi 2017). In this investigation, we have tested and verified three different models derived from the GWPZ's final categorization. The AUC values range from 0 to 1. If the number is less than 0.5, the model's classification was inappropriate, and it should be redone. On the other hand, if the value is close to 1, it suggests that the result is clearly defined (Pourghasemi and Rossi 2017).

To verify WOE, FR, and IV models, the ROC curves of the GWPZ maps were constructed (Figure 7). The finding demonstrates that the outcome predicted by the FR model for GWPZ (AUC = 91%) is successfully achieved when compared to both the WOE model (AUC = 88%) and the IV model (AUC = 89%). Nevertheless, all the obtained findings were checked for validity and clearly defined (Pourghasemi and Rossi 2017). However, the FR is a better representative for this study area to indicate the spatial distribution of the GWPZ compared to the WOE and FR models. This is because the GWPZ is more likely to be found in areas where the FR model is more accurate. As a result, this validation method is highly recommended for research into potential groundwater evaluation. The generated validation graphs of the applied models in the present study, as mentioned in Figure 7, utilized twenty percent of the inventory data of water. The highest value of the AUC value showed the most reliable results of the model and while the lowest value showed unreliable results. The findings by mentioned validation technique for WOE, FR, and IV clearly explained that all applied models are consistent and trustworthy methods to produce GWPZ for the

**Table 3.** Accuracy assessment of land use and land cover.

S. No	Classes	UA	PA	OA	K
1	Water	90.00	91.84	87.00	0.81
2	Trees	88.00	84.62		
3	Crops	86.00	87.76		
4	Built up area	92.00	86.79	91.49	
5	Bare ground	86.00	91.49		
6	Shrubs	88.00	88.00		

Notes: UA = User's Accuracy, PA = Producer's Accuracy, OA = Overall Accuracy, K = Kappa Coefficient.

Kohat District of Pakistan. The validation outcome specifies that the FR technique is the most reliable method for GWPZ in the study area.

#### 4.4.2. LULC accuracy assessment

Accuracy evaluations were performed after obtaining land use/land cover categorization outcomes. In order to do so, we used the user accuracy matrix, the producer accuracy matrix, and the total accuracy matrix to measure precision. They calculated the users' accuracy by taking the ratio of adequately classified cells to the total number of reference points. Google Earth was used as a reference tool for this research. The overall accuracy was calculated by dividing properly classified cells by all pixels. In contrast, producer accuracy was calculated by dividing the number of cells with correct land use/land cover classification by the number of ground truth pixels as explained in Table 3.

## 5. Discussion

Due to the increased demand for water availability for urbanization, industrialization, and irrigation purposes, there has been an increase in the research investigation on the groundwater scenario. This is especially true in arid to semi-arid regions worldwide, where the need for groundwater is even more critical. There have been numerous research to understand the science behind water recharge and prepare GWPZ for the scientific exploration and management of groundwater (Arabameri et al. 2019). Therefore, proper groundwork and methods should be implemented for GWPZ to manage the groundwater because the execution method for GWPM is still an argued subject (Nampak et al. 2014; Park et al. 2014).

This article has emphasized the appreciation of the groundwater potential of the Kohat region of Pakistan has been evaluated using WOE, FR, and IV models. These models were applied to compute the correlation between water body pixels and conditioning parameters for groundwater. The lower correlation value represents low potential zones, while the high correlation shows high potential groundwater regions. The probability of groundwater potential generally diminutions with increasing elevation (S. Hasan AL-Zuhairy et al. 2017). In the present study, the spatial analysis disclosed that the elevation class of < 500 m has a higher correlation value between both variables; however, the > 800 m elevation class revealed no significant association for groundwater potential. The slope gradient is an influential parameter for groundwater potential because steep slopes are the significant parameter in GWPZ. Moreover, the slope gradient is another significant parameter for groundwater



potential. A steep slope gradient adversely impacts on groundwater because it increases the surface run-off and affects the intrusion of precipitation into the ground (Jaiswal et al. 2003). If the slope angle is greater than  $35^\circ$ , groundwater potential is reduced because it restricts the aquifer's recharge (Madrucchi et al. 2008). The current study considers slope angle a critical factor for GWPZ. The association of dependent and independent variables for slope up to  $20^\circ$  is very suitable for the high potential zone of groundwater. In contrast, the slope angle  $> 35^\circ$  has an inverse relationship with the groundwater pixel and revealed low groundwater potential, as shown in Table 2 of the results. The flat surface of the aspect is more appropriate for groundwater amount (Manap et al. 2013).

In the present study, we observed that the flat surface of the aspect has a strong association with groundwater inventory data. The flat surface correlates with 0.86, 1.97, and 0.68 using WOE, FR, and IV models. The FR and EBF model revealed that concave and convex structures are less associated with groundwater potential than flat regions. Water reservoir and aquifer recharge mainly occurred in the flat region; however, the convex and concave structures did not support the water storage and infiltration (Arabameri et al. 2020). In this study, our results concluded that the Flat class of curvature strongly correlates with groundwater, followed by Concave. At the same time, convex adversely impacts groundwater potential, as shown in Table 2. The most developed likelihood of groundwater is perceived in denser drainage networks. In the current investigation, the  $< 200$  m class of drainage network shows the most influential association with groundwater potential using WOE, FR, and IV technique, followed by 200–400 m and 400–600 m. The relationship ranks of the  $> 800$  m class of drainage revealed that this class has no impact on groundwater potential. The rainfall strongly correlates positively with aquifer recharge (Wu et al. 2020). In this study, the precipitation class 1000–1050 mm/year strongly correlates with groundwater potential having a positive correlation  $\gamma$  followed by  $> 1050$  mm/year. The low precipitated area has no inverse relationship with the groundwater potential of the study area. The precipitation class  $< 900$  mm/year has minor importance for groundwater potential in the current study area and is followed by 900–950 mm/year Crops and a garden class of the LULC parameter are significantly associated with groundwater having correlation values of 2.06 and 1.25, respectively, demonstrating these classes' high potential water zones (Falah et al. 2017).

In the context of binary classification, the Receiver Operating Characteristic (ROC) curve is a popular method to evaluate and compare the performance of different models. The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) for different thresholds of a model's predicted probability (Li et al. 2021). A model with a higher AUC (Area Under the ROC Curve) is considered better. In this study we used WOE, FR and IV and compare their performance using the ROC curve. Model WOE has an AUC of 88%. This means that it has a good balance between TPR and FPR, with relatively few false positives and false negatives. Model WOE is likely to be a good choice for classification tasks where both precision and recall are important (Wahla et al. 2022).

Model FR has an AUC of 0.91. This means that it has a high TPR and low FPR, making it suitable for applications where identifying true positives is crucial, and false



positives are less of a concern. However, Model IV may be too aggressive in classifying examples as positive, leading to a high false-negative rate. Model IV has an AUC of 91%. This means that it has a higher FPR and lower TPR compared to Model A, but still performs better than random guessing. Model IV may be useful in cases where minimizing false positives is critical, but it may not perform as well in cases where false negatives are costly.

In summary, each model has its strengths and weaknesses, and the choice of the appropriate model depends on the specific requirements of the task at hand. Model FR strikes a good balance between TPR and FPR, Model IV is useful when minimizing false positives is crucial, and Model WOE is suitable for identifying true positives at the expense of false negatives.

Our research results in the Kohat region of Pakistan showed that cropland is the most influential factor for groundwater potential. The correlation value for cropland in the current research are 0.63, 1.70, and 0.53 for WO, FR, and IV, respectively. Concerning the geological fault buffer, it was hypothesized that the association between both variables for groundwater would weaken the further away from the fault one got. Their relationship increases when the distance from the fault decreases (Falah et al. 2017). Our present study results in the Kohat area presented that fault favours water infiltration and supports the aquifer recharge in the current area. The most effective fault buffer is <500 m because this class shows a strong positive correlation of 0.60, 1.73, and 0.55 applying the WOE, FR, and IV model, followed by 1500 m and 3000 m buffers. However, the buffer of >5000 class has no significant role in groundwater potential and recharge of water. The NDVI is a vital parameter for groundwater potential. NDVI and water Table have an inverse relationship, i.e. when the NDVI increases, the water table rise and vice versa (Seeyan et al. 2014). The same scenario we observed in our current research region. The high NDVI zone strongly correlates with groundwater, while the low NDVI region adversely impacts the present area. As shown in Table 2, the results justified the above statement for NDVI association with groundwater.

According to the analytical results in Table 2, drainage network, slope, elevation, and rainfall are the most significant parameters for GWPZ in the present research area. According to GIS-based statistical models, the FR is the best technique for GWPZ in the current research project. Final GWPZ was also produced using GIS-based models and then was classified into five classes of very low, low, moderate, high, and very high groundwater potential zones. The final GWPZ can be helpful for various research organizations like agriculture and energy-related sectors to manage the groundwater in the present study area.

## 6. Conclusions

This article describes a study that aims to investigate potential groundwater zones in the Kohat District of Pakistan using three different GIS-based models: Weight of Evidence (WOE), Frequency Ratio (FR), and Information Value (IV). The study uses various data sources, including satellite imagery, ground surveys, and public health department data, to develop an inventory map of groundwater and twelve groundwater conditioning parameters. The study then applies the three GIS-based models to

generate GWPZ maps and categorizes them into five categories based on their potential for groundwater availability. The study finds that stream, slope angle, elevation, and rainfall are the most significant parameters for GWPZ. The study uses ROC curves to assess the accuracy of the models and finds that FR is the most reliable model for the study. The study concludes that the GWPZ maps generated by the WOE, FR, and IV techniques can be useful for research and development agencies to improve groundwater exploration and development planning in the future.

## Acknowledgements

The authors would like to thank the university authority for financial support. The authors thanks to (TURSP-2020/82), Taif University, Taif, Saudi Arabia.

## Author contributions

**Fakhrul Islam:** methodology, software, formal analysis, visualization, data curation, writing—original draft, investigation, validation, writing—review and editing, **Aqil Tariq:** formal analysis, visualization, data curation, writing—review and editing, Supervision. **Rufat Guluzade:** writing—review and editing. **Na Zhao:** Funding, writing review and editing, **Safeer Ullah Shah:** data curation, writing—original draft, investigation, validation, writing—review and editing, **Matee Ullah:** writing—review and editing. **Mian Luqman Hussain:** writing—review and editing. **Muhammad Nasar Ahmad:** writing—review and editing, **Abdulrahman Alasmari:** writing—review and editing, **Fahad M. Alzuaibr:** writing—review and editing, **Ahmad El Askary:** writing—review and editing, **Muhammad Aslam:** writing—review and editing. All authors have read and agreed to the published version of the manuscript.

## Funding

The Key Project of Innovation LREIS (KPI001). The authors would like to thank the university authority for financial support. The authors thanks to (TURSP-2020/82), Taif University, Taif, Saudi Arabia.

## ORCID

Aqil Tariq  <http://orcid.org/0000-0003-1196-1248>  
Abdulrahman Alasmari  <http://orcid.org/0000-0003-1212-8581>

## Data availability statement

Data available on the reasonable request from the 1<sup>st</sup> author of this article.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## References

- Ahmad I, Dar MA, Teka AH, Teshome M, Andualem TG, Teshome A, Shafi T. [2020](#). GIS and fuzzy logic techniques-based demarcation of groundwater potential zones: a case study from Jemma River basin, Ethiopia. *J African Earth Sci.* 169:103860.
- Ahmad MN, Shao Z, Aslam RW, Ahmad I, Liao M, Li X, Song Y. [2022](#). Landslide hazard, susceptibility and risk assessment (HSRA) based on remote sensing and GIS data models: a case study of Muzaffarabad Pakistan. *Stoch Environ Res Risk Assess.* 36(12):4041–4056.
- Ali S, Khorrami B, Jehanzaib M, Tariq A, Ajmal M, Arshad A, Shafeeque M, Dilawar A, Basit I, Zhang L, et al. [2023](#). Spatial downscaling of GRACE data based on XGBoost model for improved understanding of hydrological droughts in the Indus Basin Irrigation System (IBIS). *Remote Sens. [Internet]*.15(4):873.
- Arabameri A, Lee S, Tiefenbacher JP, Ngo PTT. [2020](#). Novel ensemble of MCDM-artificial intelligence techniques for groundwater-potential mapping in arid and semi-arid regions (Iran). *Remote Sens.* 12(3):490.
- Arabameri A, Rezaei K, Cerda A, Lombardo L, Rodrigo-Comino J. [2019](#). GIS-based groundwater potential mapping in Shahroud plain, Iran. A comparison among statistical (bivariate and multivariate), data mining and MCDM approaches. *Sci Total Environ.* 658:160–177.
- Azra A, Kaleemullah M, Khattak B, Asma N, Safi AUR, Qaiser J, Afzal M, Tahir U, Sindhu ZUD, Farhan Y. [2019](#). Comparative efficacy of domestic garlic (*Allium sativum*) and neem (*Azadirachta indica*) against *Haemonchus contortus* in small ruminants. *Appl Ecol Env Res.* 17(5):10389–10397.
- Baloch MYJ, Zhang W, Chai J, Li S, Alqurashi M, Rehman G, Tariq A, Talpur SA, Iqbal J, Munir M, et al. [2021](#). Shallow groundwater quality assessment and its suitability analysis for drinking and irrigation purposes. *Water (Switzerland).* 13(23):3361.
- Barakat A, Rafai M, Mosaid H, Islam MS, Saeed S. [2023](#). Mapping of water-induced soil erosion using machine learning models: a case study of Oum Er Rbia Basin (Morocco). *Earth Syst Environ.* 7(1):151–170.
- Basharat M u, Khan JA, Khalil U, Tariq A, Aslam B, Li Q. [2022](#). Ensuring earthquake-proof development in a swiftly developing region through neural network modelling of earthquakes using nonlinear spatial variables. *Buildings.* 12(10):1713.
- Bonham-Carter GF, Agterberg FP, Wright DF. [1989](#). Integration of geological datasets for gold exploration in Nova Scotia. *Introd Readings Geogr Inf Syst.* 1:170–182.
- Bui DT, Shirzadi A, Chapi K, Shahabi H, Pradhan B, Pham BT, Singh VP, Chen W, Khosravi K, Ahmad BB, et al. [2019](#). A hybrid computational intelligence approach to groundwater spring potential mapping. *Water.* 11(10):2013.
- Chen W, Li H, Hou E, Wang S, Wang G, Panahi M, Li T, Peng T, Guo C, Niu C, et al. [2018](#). GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. *Sci Total Environ.* 634:853–867.
- da Silva Monteiro L, de Oliveira-Júnior JF, Ghaffar B, Tariq A, Qin S, Mumtaz F, Correia Filho WLF, Shah M, da Rosa Ferraz Jardim AM, da Silva MV, et al. [2022](#). Rainfall in the urban area and its impact on climatology and population growth. *Atmosphere (Basel).* 13(10):1610.
- Eid MH, Elbagory M, Tamma AA, Gad M, Elsayed S, Hussein H, Moghanm FS, Omara AED, Kovács A, Péter S. [2023](#). Evaluation of groundwater quality for irrigation in deep aquifers using multiple graphical and indexing approaches supported with machine learning models and GIS techniques, Souf Valley, Algeria. *Water.* 15(1):182.
- Elmoulat M, Ait Brahim L, Mastere M, Ilham Jemmah A. [2015](#). Mapping of mass movements susceptibility in the zoumi region using satellite image and GIS technology (Moroccan Rif). *Int J Sci Eng Res.* 6(2):210–217.
- Falah F, Ghorbani Nejad S, Rahmati O, Daneshfar M, Zeinivand H. [2017](#). Applicability of generalized additive model in groundwater potential modelling and comparison its performance by bivariate statistical methods. *Geocarto Int.* 32(10):1069–1089.

- Fayez L, Pazhman D, Pham BT, Dholakia MB, Solanki HA, Khalid M, Prakash I. 2018. Application of frequency ratio model for the development of landslide susceptibility mapping at part of Uttarakhand State, India. *Int J Appl Eng Res.* 13(9):6846–6854.
- Firdaus R. 2014. Doctoral dissertation assessing land use and land cover change toward sustainability in humid tropical watersheds, indonesia assessing land use and land cover change toward sustainability in humid tropical watersheds, Indonesia. (March):0–1.
- Guru B, Seshan K, Bera S. 2017. Frequency ratio model for groundwater potential mapping and its sustainable management in cold desert, India. *J King Saud Univ – Sci.* 29(3):333–347.
- Hussain MA. 2014. Seroprevalence of Brucellosis in sheep and humans in District Kohat, Pakistan. *Adv Anim Vet Sci.* 2(9):516–523.
- Hussain SA, Han FQ, Ma Z, Hussain A, Mughal MS, Han J, Alhassan A, Widory D. 2021. Origin and evolution of eocene rock salts in Pakistan and implications for paleoclimate studies: insights from chemistry and Cl stable isotopes. *Front Earth Sci.* 9(April):1–13.
- Hussain H, Zhang S. 2018. Structural evolution of the Kohat Fold and Thrust Belt in the Shakardarra Area (South Eastern Kohat, Pakistan). *Geosci.* 8(9):311.
- Islam F, Riaz S, Ghaffar B, Tariq A, Shah SU, Nawaz M, Hussain ML, Amin NU, Li Q, Lu L, et al. 2022. Landslide susceptibility mapping (LSM) of Swat District, Hindu Kush Himalayan region of Pakistan, using GIS-based bivariate modelling. *Front Environ Sci.* 10(October):1–18.
- Israil M, Al-Hadithi M, Singhal DC. 2006. Application of a resistivity survey and geographical information system (GIS) analysis for hydrogeological zoning of a piedmont area, Himalayan foothill region, India. *Hydrogeol J.* 14(5):753–759.
- Jaiswal RK, Mukherjee S, Krishnamurthy J, Saxena R. 2003. Role of remote sensing and GIS techniques for generation of groundwater prospect zones towards rural development – an approach. *Int J Remote Sens.* 24(5):993–1008.
- Jha MK, Chowdary VM, Chowdhury A. 2010. Groundwater assessment in Salboni Block, West Bengal (India) using remote sensing, geographical information system and multi-criteria decision analysis techniques. *Hydrogeol J.* 18(7):1713–1728.
- Ji Y, Dong C, Kong D, Lu J, Zhou Q. 2015. Heat-activated persulfate oxidation of atrazine: implications for remediation of groundwater contaminated by herbicides. *Chem Eng J.* 263: 45–54.
- Kalantar B, Pradhan B, Amir Naghibi S, Motevalli A, Mansor S. 2018. Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). *Geomat Nat Hazards Risk.* 9(1):49–69.
- Kaliraj S, Chandrasekar N, Magesh NS. 2014. Identification of potential groundwater recharge zones in Vaigai upper basin, Tamil Nadu, using GIS-based analytical hierarchical process (AHP) technique. *Arab J Geosci.* 7(4):1385–1401.
- Karimi-Rizvandi S, Goodarzi HV, Afkoeieh JH, Chung IM, Kisi O, Kim S, Linh NTT. 2021. Groundwater-potential mapping using a self-learning bayesian network model: a comparison among metaheuristic algorithms. *Water.* 13(5):658.
- Keesstra SD, Geissen V, Mosse K, Piirainen S, Scudiero E, Leistra M, van Schaik L. 2012. Soil as a filter for groundwater quality. *Curr Opin Environ Sustain.* 4(5):507–516.
- Kordestani MD, Naghibi SA, Hashemi H, Ahmadi K, Kalantar B, Pradhan B. 2019. Groundwater potential mapping using a novel data-mining ensemble model. *Hydrogeol J.* 27(1):211–224.
- Lee S, Kim JC, Jung HS, Lee MJ, Lee S. 2017. Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. *Geomatics Nat Hazards Risk.* 8(2):1185–1203. <http://www.tandfonline.com/action/journalInformation?show=aimsScope&journalCode=tgnh20#VsXodSCLRhE>.
- Lee S, Kim YS, Oh HJ. 2012. Application of a weights-of-evidence method and GIS to regional groundwater productivity potential mapping. *J Environ Manage.* 96(1):91–105.

- Lee S, Song KY, Kim Y, Park I. **2012**. Regional groundwater productivity potential mapping using a geographic information system (GIS) based artificial neural network model. *Hydrogeol J.* 20(8):1511–1527.
- Li B, Wang N, Chen J. **2021**. GIS-based landslide susceptibility mapping using information, frequency ratio, and artificial neural network methods in Qinghai Province, Northwestern China. *Adv Civ Eng.* 2021:1–14.
- Madrucci V, Taioli F, de Araújo CC. **2008**. Groundwater favorability map using GIS multicriteria data analysis on crystalline terrain, São Paulo State, Brazil. *J Hydrol.* 357(3-4):153–173.
- Majeed M, Lu L, Haq SM, Waheed M, Sahito HA, Fatima S, Aziz R, Bussmann RW, Tariq A, Ullah I, et al. **2022**. Spatiotemporal distribution patterns of climbers along an abiotic gradient in Jhelum District, Punjab, Pakistan. *Forests.* 13(8):1244.
- Manap MA, Sulaiman WNA, Ramli MF, Pradhan B, Surip N. **2013**. A knowledge-driven GIS modelling technique for groundwater potential mapping at the Upper Langat Basin, Malaysia. *Arab J Geosci.* 6(5):1621–1637.
- Maskooni EK, Naghibi SA, Hashemi H, Berndtsson R. **2020**. Application of advanced machine learning algorithms to assess groundwater potential using remote sensing-derived data. *Remote Sens.* 12(17):2742.
- Moazzam MFU, Rahman G, Munawar S, Tariq A, Safdar Q, Lee B. **2022**. Trends of rainfall variability and drought monitoring using standardized precipitation index in a Scarcely Gauged Basin of Northern Pakistan. *Water.* 14(7):1132.
- Mohammadi M, Sharifi A, Hosseingholizadeh M, Tariq A. **2021**. Detection of oil pollution using sar and optical remote sensing imagery: a case study of the Persian Gulf. *J Indian Soc Remote Sens.* 49(10):2377–2385.
- Muavhi N, Thamaga KH, Mutoti MI. **2022**. Mapping groundwater potential zones using relative frequency ratio, analytic hierarchy process and their hybrid models: case of Nzhelele-Makhado area in South Africa. *Geocarto Int.* 37(21):6311–6330. <https://doi.org/10.1080/10106049.2021.1936212>.
- Mumtaz F, Li J, Liu Q, Tariq A, Arshad A, Dong Y, Zhao J, Bashir B, Zhang H, Gu C, et al. **2023**. Impacts of green fraction changes on surface temperature and carbon emissions: comparison under forestation and urbanization reshaping scenarios. *Remote Sens.* 15(3):859.
- Naghibi SA, Pourghasemi HR, Abbaspour K, Naghibi SA, Pourghasemi HR, Abbaspour K. **2018**. A comparison between ten advanced and soft computing models for groundwater qanat potential assessment in Iran using R and GIS. *Theor Appl Climatol.* 131(3-4):967–984.
- Nampak H, Pradhan B, Manap MA. **2014**. Application of GIS based data driven evidential belief function model to predict groundwater potential zonation. *J Hydrol.* 513:283–300.
- Pardeshi SD, Autade SE, Pardeshi SS. **2013**. Landslide hazard assessment: recent trends and techniques. *Springerplus.* 2(1):1–11.
- Park I, Kim Y, Lee S. **2014**. Groundwater productivity potential mapping using evidential belief function. *Groundwater.* 52(S1):201–207.
- Pham BT, Jaafari A, Prakash I, Singh SK, Quoc NK, Bui DT. **2019**. Hybrid computational intelligence models for groundwater potential mapping. *Catena.* 182:104101.
- Pourghasemi HR, Rossi M. **2017**. Landslide susceptibility modeling in a landslide prone area in Mazandarn Province, north of Iran: a comparison between GLM, GAM, MARS, and M-AHP methods. *Theor Appl Climatol.* 130(1-2):609–633.
- Qureshi AS, McCornick PG, Sarwar A, Sharma BR. **2010**. Challenges and prospects of sustainable groundwater management in the Indus Basin, Pakistan. *Water Resour Manage.* 24(8):1551–1569.
- Rahman A. **2008**. A GIS based DRASTIC model for assessing groundwater vulnerability in shallow aquifer in Aligarh, India. *Appl Geogr.* 28(1):32–53.
- S, Hasan Al-Zuhairy M, Alauldeen Abdulrahman Hasan A, Mezher Shnewer F. **2017**. GIS-based frequency ratio model for mapping the potential zoning of groundwater in the Western Desert of Iraq. *Int J Sci Eng Res.* 8(7):52–65.

- Sadiq Fareed MM, Raza A, Zhao N, Tariq A, Younas F, Ahmed G, Ullah S, Jillani SF, Abbas I, Aslam M. 2022. Predicting divorce prospect using ensemble learning: support vector machine, linear model, and neural network. *Comput Intell Neurosci*. 2022:3687598.
- Seeyan S, Merkel B, Abo R. 2014. Investigation of the relationship between groundwater level fluctuation and vegetation cover by using NDVI for Shaqlawa Basin, Kurdistan Region – Iraq. *JGG*. 6(3):p187–p187.
- Shahid S, Nath SK, Maksud Kamal ASM. 2002. GIS integration of remote sensing and topographic data using fuzzy logic for ground water assessment in midnapur district, India. *Geocarto Int*. 17(3):69–74.
- Shah SHIA, Jianguo Y, Jahangir Z, Tariq A, Aslam B. 2022. Integrated geophysical technique for groundwater salinity delineation, an approach to agriculture sustainability for Nankana Sahib Area, Pakistan. *Geomatics, Nat Hazards Risk*. 13(1):1043–1064.
- Shah SHIA, Yan J, Ullah I, Aslam B, Tariq A, Zhang L, Mumtaz F. 2021. Classification of aquifer vulnerability by using the DRASTIC index and geo-electrical techniques. *Water*. 13(16):2144.
- Shano L, Raghuvanshi TK, Meten M. 2020. Landslide susceptibility evaluation and hazard zonation techniques – a review. *Geoenviron Disaster*. 7(1):1–19.
- Shirazi SM, Imran HM, Akib S. 2012. GIS-based DRASTIC method for groundwater vulnerability assessment: a review. *J Risk Res*. 15(8):991–1011.
- Siddiqui S, Safi MWA, Tariq A, Rehman NU, Haider SW. 2020. GIS based universal soil erosion estimation in district Chakwal Punjab, Pakistan. *IJEEG*. 11(2):30–36.
- Singh LK, Jha MK, Chowdary VM. 2018. Assessing the accuracy of GIS-based multi-criteria decision analysis approaches for mapping groundwater potential. *Ecol Indic*. 91(March):24–37.
- Solomon S, Quiel F. 2006. Erratum: groundwater study using remote sensing and geographic information systems (GIS) in the central highlands of Eritrea. *Hydrogeol J*. 14(6):1029–1041.
- Tariq A, Jiango Y, Li Q, Gao J, Lu L, Soufan W, Almutairi KF, Habib-Ur-Rahman M. 2023. Modelling, mapping and monitoring of forest cover changes, using support vector machine, kernel logistic regression and naive bayes tree models with optical remote sensing data. *Heliyon*. 9(2):e13212.
- Tariq A, Jiango Y, Lu L, Jamil A, Al-Ashkar I, Kamran M, Sabagh, A, El. 2023. Integrated use of Sentinel-1 and Sentinel-2 data and open-source machine learning algorithms for burnt and unburnt scars. *Geomat. Nat Hazards Risk*. 14(1):28.
- Tariq A, Mumtaz F, Majeed M, Zeng X. 2023. Spatio-temporal assessment of land use land cover based on trajectories and cellular automata Markov modelling and its impact on land surface temperature of Lahore district Pakistan. *Environ Monit Assess*. 195(1):114.
- Tariq A, Qin S. 2023. Spatio-temporal variation in surface water in Punjab, Pakistan from 1985 to 2020 using machine-learning methods with time-series remote sensing data and driving factors. *Agric Water Manag*. 280(February):108228.
- Tariq A, Shu H. 2020. CA-Markov chain analysis of seasonal land surface temperature and land use landcover change using optical multi-temporal satellite data of Faisalabad, Pakistan. *Remote Sens*. 12(20):3402.
- Tariq A, Shu H, Kuriqi A, Siddiqui S, Gagnon AS, Lu L, Linh NTT, Pham QB. 2021. Characterization of the 2014 indus river flood using hydraulic simulations and satellite images. *Remote Sens*. 13(11):2053.
- Tariq A, Siddiqui S, Sharifi A, Shah SHIA. 2022. Impact of spatio-temporal land surface temperature on cropping pattern and land use and land cover changes using satellite imagery, Hafizabad District, Punjab, Province of Pakistan. *Arab J Geosci*. 15(11):1045.
- Tariq A, Yan J, Gagnon AS, Riaz Khan M, Mumtaz F. 2022. Mapping of cropland, cropping patterns and crop types by combining optical remote sensing images with decision tree classifier and random forest. *Geo-Spatial Inf Sci*. 00(00):1–19.
- Termeh SVR, Khosravi K, Sartaj M, Keesstra SD, Tsai FT-C, Dijksma R, Pham BT. 2019. Optimization of an adaptive neuro-fuzzy inference system for groundwater potential mapping. *Hydrogeol J*. 27(7):2511–2534.

- Wahla SS, Kazmi JH, Sharifi A, Shirazi SA, Tariq A, Joyell Smith H. 2022. Assessing spatio-temporal mapping and monitoring of climatic variability using SPEI and RF machine learning models. *Geocarto Int.* 37(27):14963–14982.
- Wu WY, Lo MH, Wada Y, Famiglietti JS, Reager JT, Yeh PJF, Ducharne A, Yang ZL. 2020. Divergent effects of climate change on future groundwater availability in key mid-latitude aquifers. *Nat Commun.* 11(1):1–9.
- Xu C, Xu X, Dai F, Xiao J, Tan X, Yuan R. 2012. Landslide hazard mapping using GIS and weight of evidence model in Qingshui River watershed of 2008 Wenchuan earthquake struck region. *J Earth Sci.* 23(1):97–120.
- Yin H, Shi Y, Niu H, Xie D, Wei J, Lefticariu L, Xu S, Yin H, Shi Y, Niu H, et al. 2018. A GIS-based model of potential groundwater yield zonation for a sandstone aquifer in the Juye Coalfield, Shangdong, China. *JHyd.* 557:434–447.
- Zainab N, Tariq A, Siddiqui S. 2021. Development of Web-Based GIS Alert System for Informing Environmental Risk of Dengue Infections in Major Cities of Pakistan. *Geos Ind.* 6(1):77.
- Zhu Z, Wu Y, Liang Z. 2022. Mining-induced stress and ground pressure behavior characteristics in mining a thick coal seam with hard roofs. *Front Earth Sci.* 10(March):1–12.