*Journal of Sustainable Civil Engineering and Technology*
*e-ISSN: 2948-4294 | Volume 4 Issue 2 (September 2025), 47-61*
https://doi.org/10.24191/jscet.v4i2.JSCET_ICoSCID2025_005

# Groundwater Potential Identification Using Frequency Ratio Method in Kedah

**Syarifah Raihana Syed Zabidi[1], Sharifah Norashikin Bohari[1*], Rohayu Haron Narashid[1], Rizauddin Saian[2], Suhaila Hashim[1]**

[1]Faculty of Built Environment, Surveying Science and Geomatics Studies, Universiti Teknologi MARA, Perlis Branch, Arau Campus, Arau, Malaysia
[2] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Perlis Branch, Arau Campus, Arau, Malaysia

*Corresponding author: ashikin10@uitm.edu.my

Received: 10 July 2025 / Accepted: 08 August 2025 / Published online: 30 September 2025

**Abstract**

Water demand has been rising globally due to various factors such as lack of clean water, climate change, weather conditions, and water pollution. This increasing demand for water is deeply concerning as it has become a basic need for humans, animals, and plants. Therefore, an alternative solution such as groundwater can become one of the sources to support the current needs of water supply as well as to prevent water shortages from occurring. This study aims to identify the groundwater potential area in Kedah by using the Frequency Ratio (FR) method. 12 groundwater conditioning parameters has been divided into three factors which are topographical (elevation, slope, aspect, topographical wetness index (TWI), plan curvature and geomorphology), hydro-geological (drainage density, lithology, aquifers and distance to faults) and other factors (rainfall and landuse) which are obtained through various resources, departments and also agencies. The FR method determines the relationship between dependent variables (tubewell points) and independent variables (groundwater parameters). A total of 354 tubewell points were divided into 70% (248 points) and 30% (106 points) for the training and testing datasets. Subsequently, the identified groundwater potential area was classified into five different classes, which are very high, high, medium, low, and very low. The result shows the highest groundwater area using FR was 23.99% and the lowest groundwater area was 17.09% which is located in the western and the northeast part of Kedah, respectively. The validation using Area Under Curve (AUC) shows a success rate (85.5%) and prediction rate (82.8%), which indicates a good and reliable model. The findings of this study can provide a significant reference for any related agencies and authorities in developing more effective and sustainable groundwater management strategies.

**Keywords:** Groundwater Potential Mapping, Frequency Ratio, Kedah

## 1. Introduction

The increasing water demand can be seen in many countries due to the rapid population growth, urbanisation, development, and climate change impacts. According to The United Nations Children's Fund (UNICEF), in early 2025, half of the world's population may reside in areas facing water scarcity. In Malaysia, statistics from Air Selangor indicate that 53% of water use is allocated to home and industrial sectors, while the agriculture sector utilises the remaining 47%. To address this issue, an alternative solution such as groundwater can be utilised to meet the increasing water demand (Chatterjee & Dutta, 2022; Masroor et al., 2021; Mridha et al., 2020). As supported by the report from Water Sector Transformation (WST), Malaysia aims to use groundwater resources and surface water for up to 20% of the water supply in 2040. It is due to the minimal abstraction of groundwater, which is less than 5% of the total water resources in Malaysia. Thus, proper planning and management of groundwater resources is needed to ensure long-term sustainability and availability to fulfil the needs of humans, animals, and plants.

The identification of groundwater potential resources is difficult, as groundwater lies beneath the Earth's surface. Direct monitoring is essential to effectively identify the groundwater resources. Earlier studies have utilized field surveys such as drilling and pumping tests in order to extract and identify the groundwater resources (Masroor et al., 2021; Ponnusamy & Elumalai, 2022). However, it is not suitable for large-scale mapping as it is time-consuming and costly due to the expensive instrument (Das, 2019; Thapa et al., 2017; Vafadar et al., 2023). According to Ghosh & Bera (2024), Maskooni et al. (2020), and Pawar et al. (2024), nowadays, the integration of geographic information systems (GIS) and remote sensing has been widely used to identify and predict the groundwater resources. Some of the methods that have been used were multicriteria decision making (MCDM) (Arabameri et al., 2019; Elubid et al., 2020), analytical hierarchy process (AHP) (Arunbose et al., 2021; Das, 2019; Elvis et al., 2022), weight of evidence (WOE) (Arabameri et al., 2019), as well as frequency ratio (FR) (Das, 2019; Elvis et al., 2022; Hasanuzzaman et al., 2022). The integration of these methods with various factors such as topography, hydrology, geology, and climatic changes is essential as it influences the groundwater movement and occurrence (Jari et al., 2022; Liu et al., 2022). Thus, these approaches will help to give a good and reliable result.

This study aims to utilize the method of FR in delineating the groundwater potential in Kedah by integrating various factors that influence groundwater occurrence. A total of 12 different conditioning parameters from topography, hydrogeology, and other factors have been utilized in the study. For topographical factors, the parameters consist of slope, elevation, aspect, topographical wetness index (TWI), plan curvature, and geomorphology. Meanwhile, in hydrogeological factors, it consists of drainage density, distance to fault, lithology, and aquifer. The other remaining parameters were rainfall and landuse. The validation was utilized by using ROC(AUC), which is most commonly used by past researchers for groundwater validation (Guru et al., 2017; Pawar et al., 2024; Prasad et al., 2020; Sharma et al., 2024). The findings from this study are expected to serve as a baseline to sustain and manage groundwater resources in Kedah more efficiently.

## 2. Material and Methods

### 2.1 Study Area

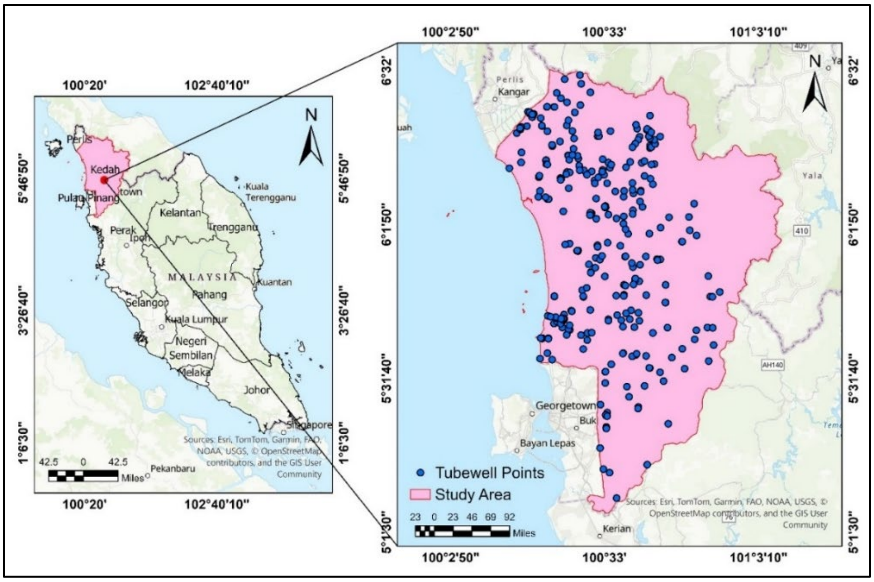


**Figure 1.** Study Area

The study area is in Kedah, which is located in the northwestern region of Peninsular Malaysia between latitudes 5˚40'N to 6˚40'N and longitudes 99˚40'E to 100˚40'E as shown in Figure 1. The study area covers approximately 9500 $km^2$ with a population of over 2 million people. It consists of several administrative districts such as Kubang Pasu, Sik, Yan, Kota Setar, Penang, Kuala Muda, Baling, Kulim, and Bandar Baharu. The higher elevation is located in the eastern and southern parts of the study area, which is covered with a forest. Meanwhile, the lower elevation is located in the western and central parts, which are covered with agriculture and urban land. For the hydrogeological setting, most of the high aquifers cover the study area, while the medium and low aquifer covered the northwestern and southeastern parts, respectively.

## 2.2 Methodology

The methodology for this study is presented in Figure 2 which has been divided into four phases which are data acquisition, pre-processing, data processing and lastly, data validation.

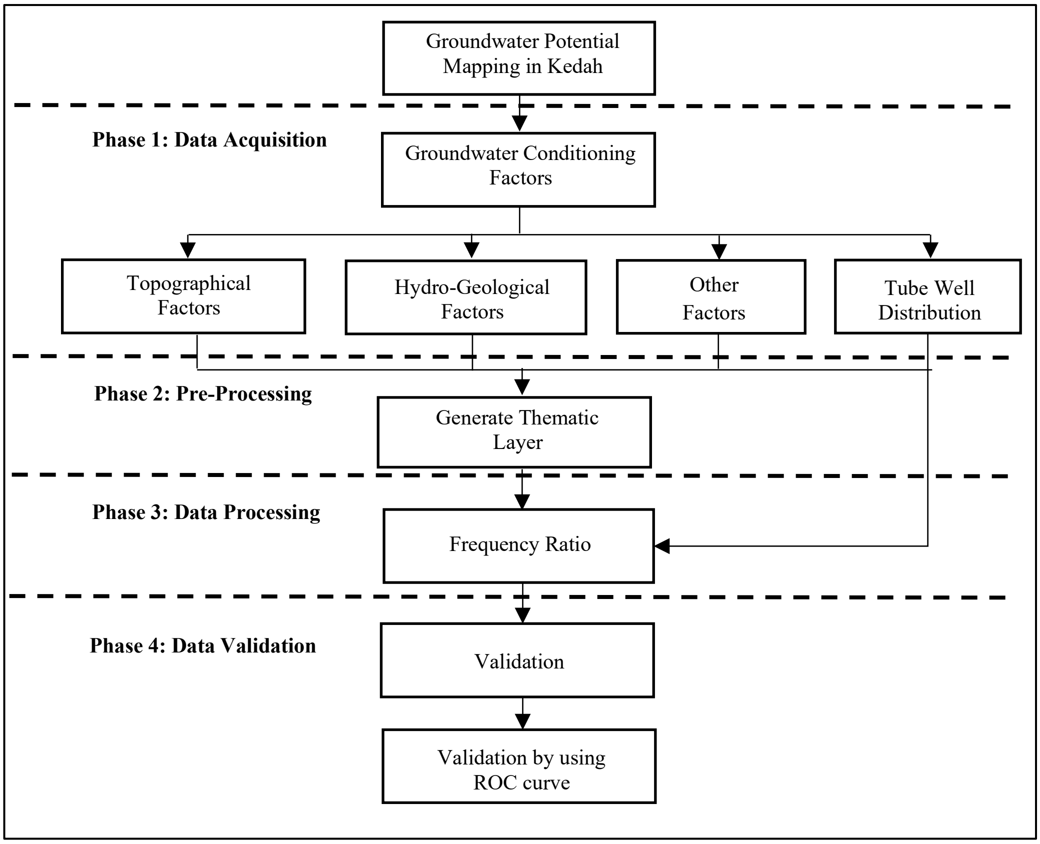


**Figure 2.** Methodology

In this study, 12 such parameters were divided into 3 categories which are topographical, hydrogeological, and other factors. These parameters include slope, elevation, aspect, TWI, plan curvature, geomorphology, drainage density, lithology, distance to fault, aquifer, rainfall, and landuse that have been obtained from various sources, including government agencies and open-source data. All of the parameters were converted into thematic maps with a uniform grid of 10 x 10 m by using ArcGIS Pro software. A total of 354 tubewell points were divided into 70% for the training dataset and 30% for the testing dataset. The validation of the result was then executed by using the ROC(AUC) method.

## 2.3 Groundwater Conditioning Parameters

### 2.3.1 Topographical Factors

Tandem-X DEM was used to produce the following topographic factors. Elevation is one of the vital indicators in predicting groundwater potential (Chen et al., 2019). Elevation influences other topographical factors such as slope, aspect, TWI, and curvature. Higher elevation indicates a lower infiltration rate, while lower elevation indicates a higher infiltration rate, resulting in a high groundwater potential capacity (Sharma et al., 2024; Thapa et al., 2017; Wei et al., 2022). Figure 3(a) shows the elevation map, which ranges between -170 m to 1856 m. A gentle slope has a higher infiltration rate, which increases the occurrence and availability of groundwater. Meanwhile, a steeper slope has a lower infiltration capacity, which leads to a lower groundwater occurrence (Guru et al., 2017; Maskooni et al., 2020). The slope for the study area, as shown in Figure 3(b) was extracted from the DEM by using the Slope tool that ranges between 1˚ and 80˚. Moreover, the aspect in Figure 3(c) was extracted using Aspect tools that classified it into 10 classes of direction. Aspect influences groundwater occurrence based on the direction and orientation of the slope (Fatah et al., 2024; Roy et al., 2024; Sharma et al., 2024). TWI evaluates the influence of topography on hydrological processes, which affects the groundwater occurrence. TWI can be calculated based on the equation below (Moore et al., 1991).

$$TWI = In(\frac{fa}{tan\beta}) \quad (1)$$

Based on the equation, fa indicates flow accumulations, and B represents the slope angle at the specified location (Moore et al., 1991). Therefore, a higher TWI value indicates a lower elevation, which results in higher groundwater occurrence. TWI in Figure 3(d) shows the range between -6 and 15 (Anh et al., 2023; Huang et al., 2024; Jari et al., 2022). Figure 3(e) shows the map of plan curvature that has been classified into convex, linear, and concave. This parameter describes the convergence and divergence of water flow (Benjmel et al., 2020). The negative values on plan curvature are described as the convergence of water flow, while the positive values are described as the divergence of water flow, which results in the groundwater occurrence (Kumar et al., 2023; Prasad et al., 2020). Geomorphology served as one of the indicators for groundwater potential. Figure 3(f) shows the geomorphology map that ranges between 0˚ and 85˚.

### 2.3.1 Hydrogeological Factors

Hydrogeological factors in this study include several parameters such as drainage density, lithology, geology, aquifer, and distance to fault. Drainage density is crucial to the availability of groundwater. According to Al-Kindi & Janizadeh (2022) and Dey et al (2023), a high value of drainage density indicates a low infiltration of water, while a low drainage density indicates a large infiltration of water. Figure 3(g) shows the drainage density map that has been classified into five classes that range from 0 to 466. Distance to fault map in Figure 3(h) was extracted using Euclidean Distance tools with the input of major and minor faulting that was obtained from JMG. The map was classified into five different classes, from very low to very high. Lithology and aquifer in Figure 3(i) and Figure 4(a) for this study have been obtained from the Department of Mineral and Geoscience (JMG), which has been classified into each of the distinct classes.

### 2.3.1 Other Factors

There are three parameters in the other factors group, which are rainfall, landuse, and soil types. Rainfall is most important as it serves as a primary source for the groundwater recharge, which is crucial for groundwater occurrence (Prasad et al., 2020; Seifu et al., 2023; Thapa et al., 2017). There were total of 9 stations across the study area, which are Alor Setar, Bukit Sidim, Felda Bukit Tangga, Felda Bukit Tembaga, Felda Sungai Tiang, Hospital Baling, Mardi Bukit Raya, and P. Pert. Teluk Chengai and Pusat Perhentian Batu Seketol. The rainfall map in Figure 4(b) was extracted using Inverse Distance Weighted (IDW) tools with a classification that ranges between 142 mm and 353 mm. Landuse data has been produced by using Sentinel 2A satellite images obtained from Copernicus Data Space Ecosystem. The classification of landuse has been extracted by using the

Supervised Classification tool with 5 classes of barren land, agriculture, forest, water, and urban land as shown in Figure 4(c).

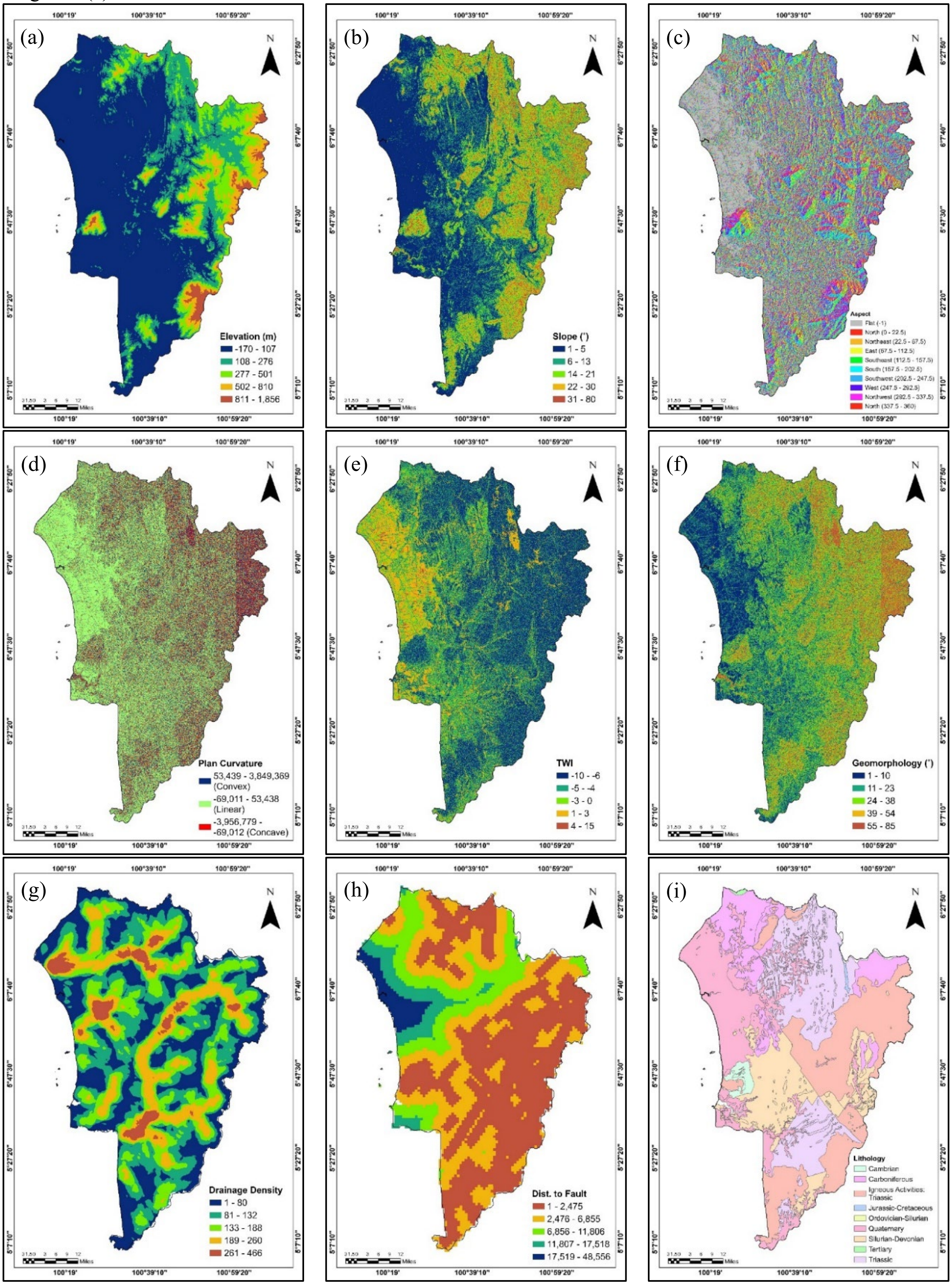

**Figure 3.** Thematic layers for groundwater conditioning parameters: (a) Elevation, (b) Slope, (c) Aspect, (d) Plan Curvature, (e) TWI, (f) Geomorphology, (g) Drainage Density, (h) Distance to Fault, (i) Lithology

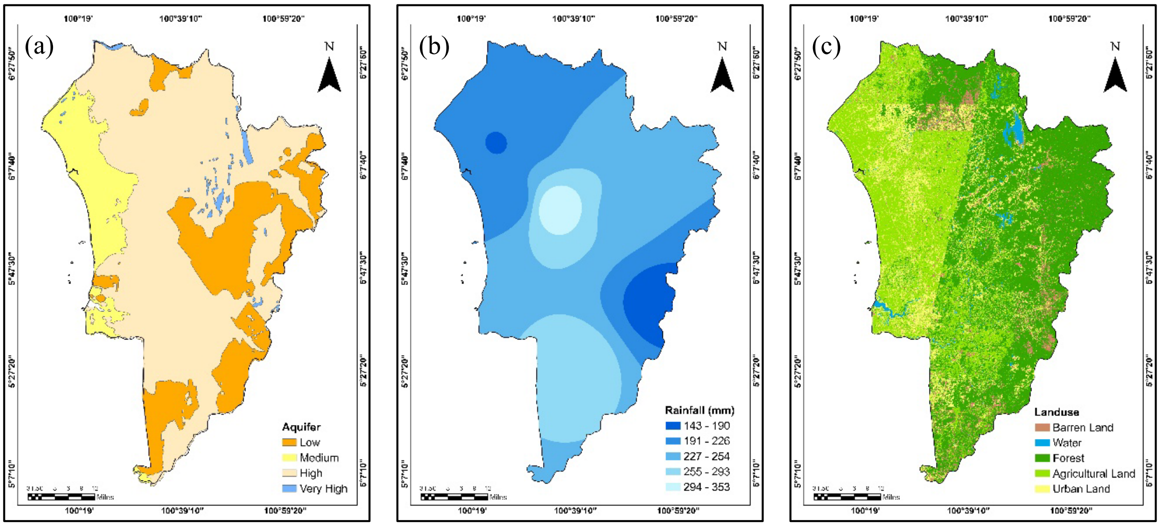


**Figure 4.** Thematic layers for groundwater conditioning parameters: (a) Aquifer, (b) Rainfall, (c) Landuse

## 2.4 Application of Frequency Ratio (FR)

FR is a statistical method used for analysing the possibility of a given attribute that occur. FR defined the relationship between dependent variables (tubewell distribution) and independent variables (groundwater conditioning parameters) (Das, 2019; Kumar et al., 2023; Prasad et al., 2020). The FR value was calculated using formula below.

$$FR = (P_{TW}/T_{TW})/(P_T/T_T) = (a/b) \tag{2}$$

whereas:
$P_{TW}$: Number of tubewell points in a class
$T_{TW}$: Total number of tubewell points
$P_T$: Area of the class
$T_T$: Total area
$a$: Percentages of tubewell points in a class
$b$: Percentages of area in a class

The processing of FR started by classifying each of the groundwater parameters into several classes using Natural Breaks by Jenks tools in ArcGIS software. The $P_{TW}$ value was obtained by overlaying the tubewell points data with each classified parameter layer using the Tabulate Area tools. This value represents the number of tubewell points that fall within each class. Meanwhile, the $T_{TW}$ value is the total number of tubewell points. The $P_T$ value was obtained from the pixel value for each classified parameter layer, while the value of $T_T$ was obtained from the total pixel area for each groundwater parameter. Then, the FR was obtained by dividing the percentages of tubewell points in a class $(a)$ with the percentages of area in a class $(b)$. Thus, FR value indicates that a value greater than 1 indicates high groundwater potential, a value less than 1 indicates low groundwater potential, and a value equal to 1 represents the average of the model (Kumar et al., 2023; Manap et al., 2014). The final result of groundwater potential mapping using FR was obtained by summing all of the newly classified parameters using Raster Calculator tools. The equation is shown below.

$$GWP = SLP_{FR} + ELV_{FR} + ASP_{FR} + PLC_{FR} + TWI_{FR} + GMP_{FR} + DD_{FR} + LTO_{FR} + DTF_{FR} + AQF_{FR} + RAIN_{FR} + LU_{FR} \tag{3}$$

After the result has been obtained, the map classified into five different classes of very high, high, medium, low and very low by using Natural Breaks by Jenks in ArcGIS Pro.

### 2.5 Validation Using ROC(AUC) Method

The validation of this study was made by using the Receiver Operating Characteristics (ROC) curve with the Area Under Curve (AUC), which is one of the most commonly used methods for validation purposes. ROC curve is a graphical plot of the diagnostic test for the performance models (Golkarian et al., 2018). The curve plots two parameters, which are the True Positive Rate (TPR) on the Y-axis and False Positive Rate (FPR) on the X-axis. Meanwhile, AUC indicates an area under the curve that summarizes overall performance.

**Table 1.** AUC Values

| AUC Values | Description |
|---|---|
| 0.9 – 1.0 | Excellent |
| 0.8 – 0.9 | Good |
| 0.7 – 0.8 | Fair |
| 0.6 – 0.7 | Poor |
| 0.5 – 0.6 | Fail |

Table 1 shows the range of AUC value that ranges between 0 and 1, whereas the value that is closer to 1 represents an excellent and good result; meanwhile, the value around 0.5 and below shows low performance of the result (Prasad et al., 2020).

## 3. Results and Discussion

### 3.1 Spatial Relationship Between Groundwater Tubewell and Groundwater Conditioning Parameters

Table 2 shows the relationship between tubewell and each of the groundwater conditioning parameters using FR method.

**Table 2.** FR Value for Each Class of Groundwater Conditioning Parameters

| Factors | Factor Classes | Tubewell Points | % Tubewell Points in a Class (a) | % Area in a Class (b) | FR (a/b) |
|---|---|---|---|---|---|
| Elevation (m) | -170 – 107 | 221 | 97.36 | 76.56 | 1.27 |
| | 108 – 276 | 6 | 2.64 | 23.44 | 0.11 |
| | 277 – 501 | 0 | 0 | 0 | 0 |
| | 502 – 810 | 0 | 0 | 0 | 0 |
| | 811 – 1858 | 0 | 0 | 0 | 0 |
| Slope (˚) | 1 – 6 | 162 | 71.37 | 43.29 | 1.65 |
| | 7 – 14 | 46 | 20.26 | 22.66 | 0.89 |
| | 15 – 23 | 12 | 5.29 | 16.63 | 0.32 |
| | 24 – 33 | 6 | 2.64 | 12.00 | 0.22 |
| | 34 – 28 | 1 | 0.44 | 5.42 | 0.08 |
| Aspect | North | 52 | 22.91 | 20.45 | 1.12 |
| | Northeast | 19 | 8.37 | 9.19 | 0.91 |
| | East | 33 | 14.54 | 10.83 | 1.34 |
| | Southeast | 21 | 9.25 | 9.83 | 0.94 |
| | South | 24 | 10.57 | 11.99 | 0.88 |

| Factors | Factor Classes | Tubewell Points | % Tubewell Points in a Class (a) | % Area in a Class (b) | FR (a/b) |
|---|---|---|---|---|---|
| | Southwest | 23 | 10.13 | 10.52 | 0.96 |
| | West | 21 | 9.25 | 11.98 | 0.77 |
| | Northwest | 25 | 11.01 | 10.63 | 1.04 |
| | North | 9 | 3.96 | 4.57 | 0.87 |
| Plan Curvature | Convex | 2 | 0.88 | 7.92 | 0.11 |
| | Linear | 183 | 80.62 | 67.98 | 1.19 |
| | Concave | 42 | 18.50 | 24.10 | 0.77 |
| TWI | -10 – -6 | 54 | 23.79 | 39.84 | 0.60 |
| | -5 – -4 | 71 | 31.28 | 30.51 | 1.03 |
| | -3 – 0 | 42 | 18.50 | 12.75 | 1.45 |
| | 1 – 3 | 50 | 22.03 | 14.83 | 1.49 |
| | 4 – 15 | 10 | 4.41 | 2.07 | 2.12 |
| Geomorphology | 1 – 10 | 56 | 24.67 | 22.90 | 1.08 |
| | 11 – 23 | 93 | 40.97 | 28.46 | 1.44 |
| | 24 – 38 | 56 | 24.67 | 21.54 | 1.15 |
| | 39 – 54 | 15 | 6.61 | 16.56 | 0.40 |
| | 55 – 85 | 7 | 3.08 | 10.54 | 0.29 |
| Drainage Density | 0.01 – 80.47 | 56 | 24.67 | 28.64 | 0.86 |
| | 80.48 – 131.67 | 53 | 23.35 | 29.76 | 0.78 |
| | 131.68 – 188.37 | 53 | 23.35 | 21.81 | 1.07 |
| | 188.38 – 259.69 | 47 | 20.70 | 15.53 | 1.33 |
| | 259.7 – 466.35 | 18 | 7.93 | 4.26 | 1.86 |
| Distance to Fault (m) | 1 – 2475 | 81 | 35.68 | 46.74 | 0.76 |
| | 2476 – 6665 | 78 | 34.36 | 28.79 | 1.19 |
| | 6666 – 11615 | 44 | 19.38 | 14.66 | 1.32 |
| | 11616 – 17328 | 21 | 9.25 | 6.17 | 1.50 |
| | 17329 – 48556 | 3 | 1.32 | 3.65 | 0.36 |
| Lithology | Quaternary | 60 | 26.55 | 21.04 | 1.26 |
| | Silurian-Devonian | 43 | 19.03 | 15.26 | 1.25 |
| | Cambrian | 13 | 5.75 | 1.30 | 4.41 |
| | Carboniferous | 32 | 14.16 | 11.86 | 1.19 |
| | Triassic | 61 | 26.99 | 24.91 | 1.08 |
| | Jurassic – Cretaceous | 0 | 0 | 0 | 0 |
| | Ordovician – Silurian | 0 | 0 | 0 | 0 |
| | Igneous Activities: Triassic | 17 | 7.52 | 25.62 | 0.29 |
| | Tertiary | 0 | 0 | 0 | 0 |
| Aquifer | Very High | 2 | 0.88 | 1.09 | 0.81 |
| | Medium | 30 | 13.22 | 11.95 | 1.11 |
| | Low | 19 | 8.37 | 23.56 | 0.36 |
| | High | 176 | 77.53 | 63.39 | 1.22 |
| Rainfall (mm) | 142.11 – 190.09 | 5 | 2.20 | 4.76 | 0.46 |
| | 190.1 – 225.67 | 91 | 40.09 | 31.80 | 1.26 |
| | 225.68 – 253.8 | 83 | 36.56 | 41.84 | 0.87 |
| | 253.81 – 292.69 | 43 | 18.94 | 19.16 | 0.99 |
| | 292.7 – 353.09 | 5 | 2.20 | 2.44 | 0.90 |
| Landuse | Barren Land | 11 | 4.85 | 5.44 | 0.89 |
| | Water | 3 | 1.32 | 1.30 | 1.01 |

| Factors | Factor Classes | Tubewell Points | % Tubewell Points in a Class (a) | % Area in a Class (b) | FR (a/b) |
|---|---|---|---|---|---|
| | Forest | 51 | 22.47 | 47.35 | 0.47 |
| | Agricultural Land | 99 | 43.61 | 30.37 | 1.44 |
| | Urban Land | 63 | 27.75 | 15.53 | 1.79 |

Based on Table 2, the FR values range from 0 to 4.41 in each class of the parameters. The FR analysis for elevation shows that lower elevation areas (-170 m to 107 m) are highly correlated with tubewell distribution (FR=1.27), which indicates higher potential towards groundwater occurrence. For slope, areas with gentle slope (1˚ to 6˚) exhibit the highest relationship with tubewell (FR=1.65). Meanwhile, the east-facing aspect indicates a strong association with the tubewell (FR=1.34). Linear areas in plan curvature indicate a high potential of groundwater with the highest FR value of 1.19. In TWI classes, the range of 4 to 15 has the highest FR value (2.12) compared to the lower TWI ranges. Geomorphology (11 to 23) indicates a strong relationship with tubewell distribution (FR=1.44). Regarding drainage density, most of the tubewell points occur in the areas with high drainage density (259.7 to 466.35), while fewer tubewells are found in areas with low drainage density values (80.48 to 131.67). For distance to fault, the areas within 11616 m to 17328 m from faults indicate a high potential of groundwater occurrence due to the strong relationship with tubewell location (FR=1.50). For other remaining hydrogeology factors, a higher relationship with tubewell for lithology and aquifer has been found in the Cambrian class (FR=4.41) and the high aquifer zones (FR=1.22). Meanwhile, for rainfall, the class (190.21 mm to 225.67 mm) exhibits a strong relationship with the tubewell, with the FR value (1.26). Lastly, for landuse, urban land (FR=1.79) and agricultural land (FR=1.44) indicate a strong relationship with the tubewell, while forest areas (FR=0.47) exhibit a low relationship with the tubewell, indicating a low occurrence of groundwater.

## 3.2 Weight of Individual Factors

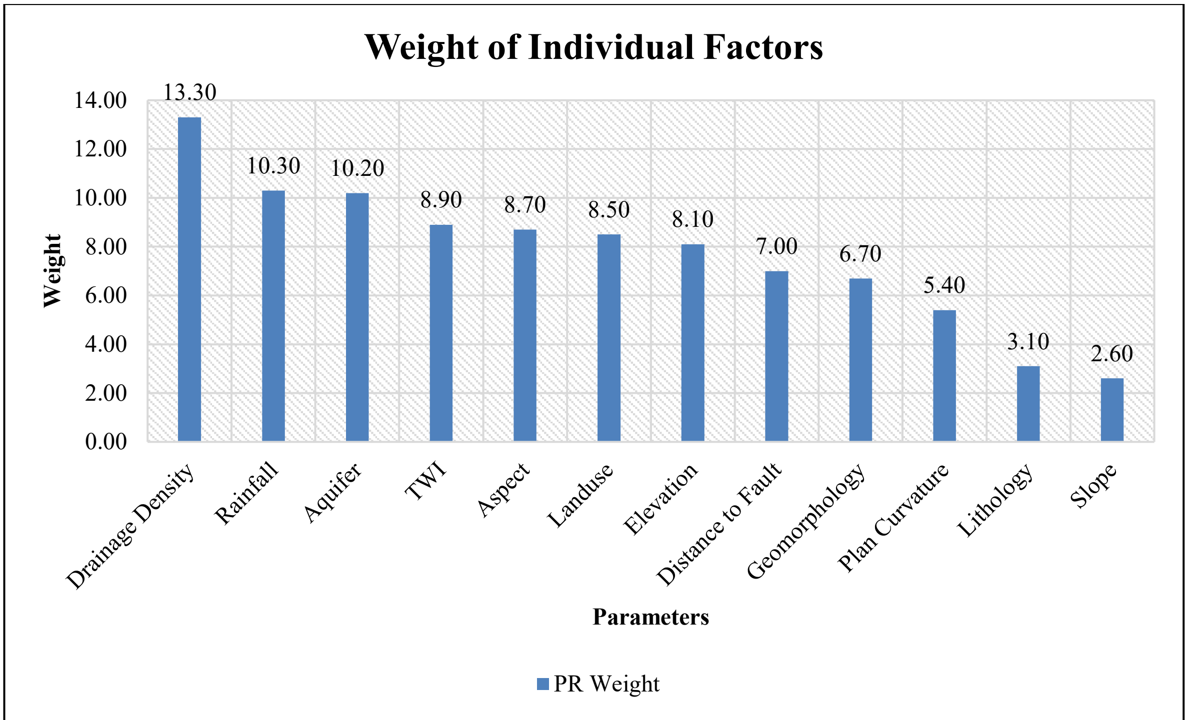


**Figure 5.** Weight of Individual Factors for Each of the Groundwater Conditioning Parameters

Figure 5 shows the weight of individual factors using the FR method. Drainage density (13.30%), rainfall (10.30%), and aquifer (10.20%) show the highest prediction weight (PR), which indicates as the most critical factors for groundwater potential. Meanwhile, plan curvature (5.40%), lithology (3.10%), and slope (2.60%) have the lowest PR weight, suggesting a relatively minor influence on groundwater potential in Kedah.

### 3.3 Groundwater Potential using FR

Figure 6 shows the result of the groundwater potential area in Kedah using the FR method. The map has been classified into five different classes, which are very high, high, medium, low, and very low. The green colour represents high to very high groundwater areas, yellow represents medium or moderate, orange represents low groundwater, and lastly, red represents very low areas. Based on the map, it can be seen that a very high to high groundwater potential area is located in the western and northwestern part of Kedah. The high groundwater area indicates a favourable condition for groundwater occurrence due to a gentle slope and lower elevation, which indicates a higher infiltration rate that can lead to an increase of groundwater potential (Sharma et al., 2024; Maskooni et al., 2020). Meanwhile, low to very low potential is primarily found in the eastern and southeastern parts of Kedah, which shows the higher slope and elevation that reduces the potential of groundwater occurrence.

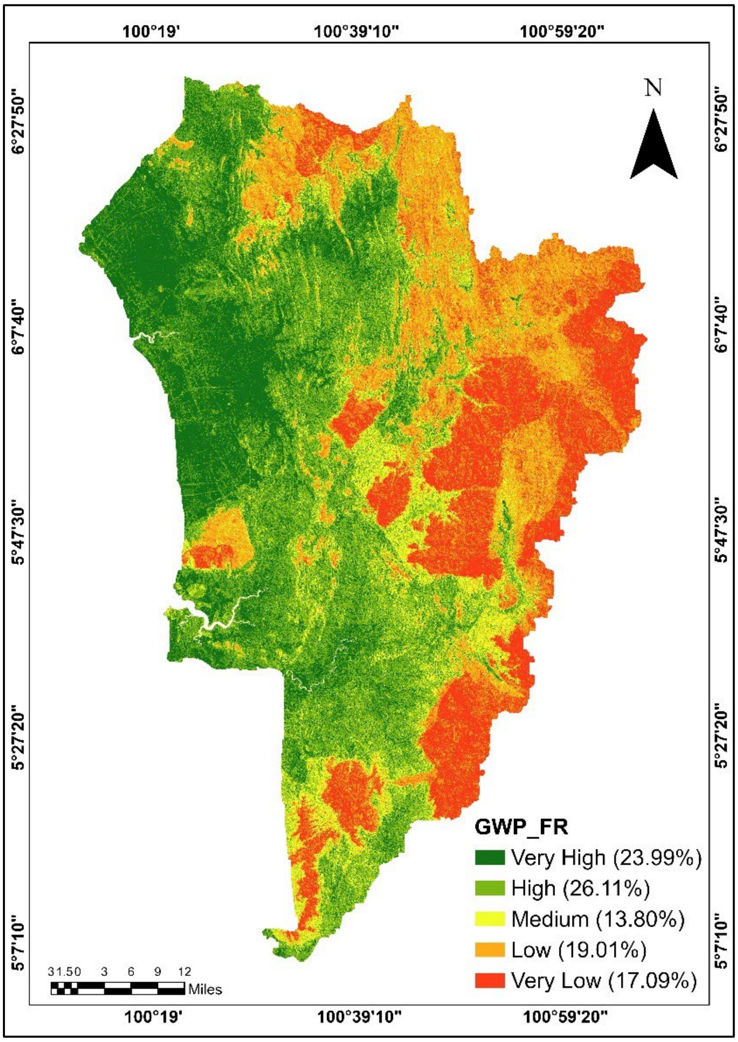


**Figure 6.** GWP Using FR

As shown in Table 3, the classification result indicates that 23.99% falls under the very high area for groundwater potential, 26.11% under high, 13.80% under medium, 19.01% under low, and 17.09% under very low area. The highest percentage area for groundwater potential in Kedah is designated as high area (2312.38 $km^2$), while the smallest percentage is classified as very low area (1513.08 $km^2$).

**Table 3.** Percentages of GWP Area

| **Classes** | **Area GWPM ($km^2$)** | **Percentage (%)** |
|---|---|---|
| Very High | 2124.71 | 23.99 |
| High | 2312.38 | 26.11 |
| Medium | 1222.18 | 13.80 |
| Low | 1683.07 | 19.01 |
| Very Low | 1513.08 | 17.09 |

3.4 Validation using ROC(AUC)

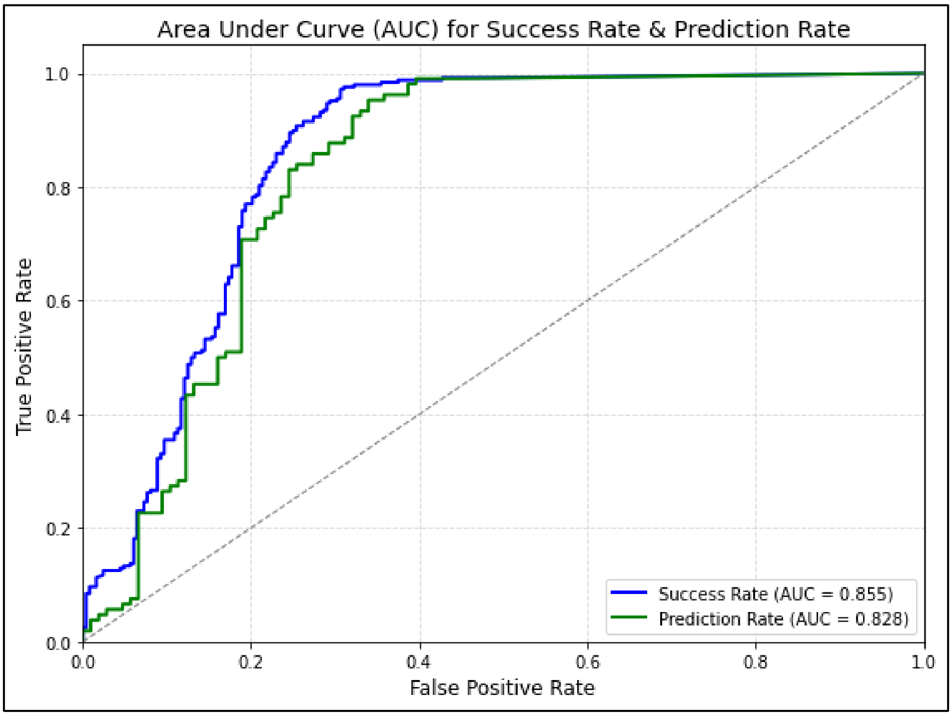


**Figure 7.** AUC Values for Success Rate and Prediction Rate

Figure 7 shows the ROC(AUC) result for the validation of groundwater potential mapping. The tubewells were divided into 70% of the training dataset (248 tubewell points) and the remaining 30% of the testing dataset (106 tubewell points). The ROC(AUC) value shows a success rate of 0.855 and a prediction rate of 0.828, which indicates a highly accurate and reliable model for groundwater potential. The study's findings are consistent with the studies from Al-Abadi et al. (2016), Guru et al. (2017), and Hasanuzzaman et al. (2022), which show an AUC value above 0.8, indicating a good and excellent result.

## 4. Conclusion

Kedah is one of the states in Malaysia that has been recognized for its contribution to the agriculture sector, especially for its paddy fields. Sufficient sources of water are essential to ensure the sustainability of agricultural activities since they influence crop activities and the livelihood of the local communities. Therefore, groundwater potential mapping served as a vital approach to address the issues of the increasing water demand, particularly in a region facing water scarcity. In this study, the FR method was utilized to assess the groundwater potential mapping in Kedah. A total of 354 tubewell points were used as dependent variables to implement the FR method effectively. The result achieved an AUC value of 0.855 for success rate and 0.828 for prediction rate, which indicates good predictive accuracy and reliability. The groundwater potential areas in Kedah were classified into very high (23.99%), high (26.11%), medium (13.80%), low (19.01%), and very low (17.09%). A very high to high potential area for groundwater is located in the western and northwestern part of Kedah, while the eastern and southeastern region showed the lowest potential area for groundwater occurrence. There were various factors, such as topography, hydrogeology, and climate, that were incorporated to identify the groundwater potential area. Drainage density, rainfall, and aquifer were nominated as the highest contributions, while plan curvature, lithology, and slope indicate the lowest contribution towards groundwater occurrence. Thus, the implementation of groundwater studies is crucial as it is the first step to be taken to address water scarcity. It is aligned with the sixth Sustainable Development Goal (SDG), which aims to ensure the availability and sustainability of water and sanitation for all. By conducting this study, the authorities and policy makers will gain the essential insights into the groundwater potential zones that enable them to plan and manage groundwater efficiently. A good groundwater management system must be developed in order to strengthen the water security, sustain the agricultural activities, and support the livelihood of the residents in Kedah.

## Acknowledgments

We would like to thank Department of Mineral and Geoscience (JMG) Malaysia, Malaysia Meteorological Department (METMalaysia) and Department of Drainage and Irrigation for providing the hydrogeological and rainfall data. We are also grateful to The Ministry of Higher Education for the funding provided under grant number (Grant no. FRGS/1/2023/WAB07/UITM/02/3) and the Universiti Teknologi MARA.

## Declaration of Conflicting Interests

All authors declare that they have no conflicts of interest.

## Author Contributions

Conceptualization, Formal analysis, Investigation, Methodology, Project Administration, Software, Validation, Visualization, Writing – Original Draft, Writing – Review & Editing, Syarifah Raihana Syed Zabidi; Conceptualization, Funding Acquisition, Resources, Supervision, Writing – Review & Editing, Sharifah Norashikin Bohari; Conceptualization, Funding Acquisition, Supervision, Writing – Review & Editing, Rohayu Haron Narashid; Conceptualization, Methodology, Software, Funding Acquisition, Supervision, Writing – Review & Editing, Rizauddin Saian; Conceptualization, Funding Acquisition, Supervision, Writing – Review & Editing, Suhaila Hashim. All authors have reviewed and approved the final version of the manuscript for publication.

## References

Al-Abadi, A. M., Al-Temmeme, A. A., & Al-Ghanimy, M. A. (2016). A GIS-based combining of frequency ratio and index of entropy approaches for mapping groundwater availability zones at Badra–Al Al-Gharbi–Teeb areas, Iraq. *Sustainable Water Resources Management*, *2*(3), 265–283. https://doi.org/10.1007/s40899-016-0056-5

Al-Kindi, K. M., & Janizadeh, S. (2022). Machine Learning and Hyperparameters Algorithms for Identifying Groundwater Aflaj Potential Mapping in Semi-Arid Ecosystems Using LiDAR, Sentinel-2, GIS Data, and Analysis. *Remote Sensing*, *14*(21). https://doi.org/10.3390/rs14215425

Anh, D. T., Pandey, M., Mishra, V. N., Singh, K. K., Ahmadi, K., Janizadeh, S., Tran, T. T., Linh, N. T. T., & Dang, N. M. (2023). Assessment of groundwater potential modeling using support vector machine optimization based on Bayesian multi-objective hyperparameter algorithm. *Applied Soft Computing*, *132*. https://doi.org/10.1016/j.asoc.2022.109848

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. *Science of the Total Environment*, *658*, 160–177. https://doi.org/10.1016/j.scitotenv.2018.12.115

Arunbose, S., Srinivas, Y., Rajkumar, S., Nair, N. C., & Kaliraj, S. (2021). Remote sensing, GIS and AHP techniques-based investigation of groundwater potential zones in the Karumeniyar river basin, Tamil Nadu, southern India. *Groundwater for Sustainable Development*, *14*. https://doi.org/10.1016/j.gsd.2021.100586

Benjmel, K., Amraoui, F., Boutaleb, S., Ouchchen, M., Tahiri, A., & Touab, A. (2020). Mapping of groundwater potential zones in crystalline terrain using remote sensing, GIS techniques, and multicriteria data analysis (Case of the ighrem region, Western Anti-Atlas, Morocco). *Water (Switzerland)*, *12*(2). https://doi.org/10.3390/w12020471

Chatterjee, S., & Dutta, S. (2022). Assessment of groundwater potential zone for sustainable water resource management in south-western part of Birbhum District, West Bengal. *Applied Water Science*, *12*(3). https://doi.org/10.1007/s13201-021-01549-4

Chen, W., Panahi, M., Khosravi, K., Pourghasemi, H. R., Rezaie, F., & Parvinnezhad, D. (2019). Spatial prediction of groundwater potentiality using ANFIS ensembled with teaching-learning-based and biogeography-based optimization. *Journal of Hydrology*, *572*, 435–448. https://doi.org/10.1016/j.jhydrol.2019.03.013

Das, S. (2019). Comparison among influencing factor, frequency ratio, and analytical hierarchy process techniques for groundwater potential zonation in Vaitarna basin, Maharashtra, India. *Groundwater for Sustainable Development*, *8*, 617–629. https://doi.org/10.1016/j.gsd.2019.03.003

Dey, B., Abir, K. A. M., Ahmed, R., Salam, M. A., Redowan, M., Miah, M. D., & Iqbal, M. A. (2023). Monitoring groundwater potential dynamics of north-eastern Bengal Basin in Bangladesh using AHP-Machine learning approaches. *Ecological Indicators*, *154*. https://doi.org/10.1016/j.ecolind.2023.110886

Elubid, B. A., Huang, T., Peng, D. P., Ahmed, E. H., & Babiker, M. M. (2020). Delineation of groundwater potential zones using integrated remote sensing, gis and multi-criteria decision making (Mcdm). *Desalination and Water Treatment*, *192*, 248–258. https://doi.org/10.5004/dwt.2020.25761

Elvis, B. W. W., Arsène, M., Théophile, N. M., Bruno, K. M. E., & Olivier, O. A. (2022). Integration of shannon entropy (SE), frequency ratio (FR) and analytical hierarchy process (AHP) in GIS for suitable groundwater potential zones targeting in the Yoyo river basin, Méiganga area, Adamawa Cameroon. *Journal of Hydrology: Regional Studies*, *39*. https://doi.org/10.1016/j.ejrh.2022.100997

Fatah, K. K., Mustafa, Y. T., & Hassan, I. O. (2024). Groundwater potential mapping in arid and semi-arid regions of kurdistan region of Iraq: A geoinformatics-based machine learning approach. *Groundwater for Sustainable Development*, *27*. https://doi.org/10.1016/j.gsd.2024.101337

Ghosh, A., & Bera, B. (2024). Potentialities and development of groundwater resources applying machine learning models in the extended section of Manbhum-Singhbhum Plateau, India. *HydroResearch*, *7*, 1–14. https://doi.org/10.1016/j.hydres.2023.11.002

Golkarian, A., Naghibi, S. A., Kalantar, B., & Pradhan, B. (2018). Groundwater potential mapping using C5.0, random forest, and multivariate adaptive regression spline models in GIS. *Environmental Monitoring and Assessment*, *190*(3). https://doi.org/10.1007/s10661-018-6507-8

Guru, B., Seshan, K., & Bera, S. (2017). Frequency ratio model for groundwater potential mapping and its sustainable management in cold desert, India. In *Journal of King Saud University - Science* (Vol. 29, Issue 3, pp. 333–347). Elsevier B.V. https://doi.org/10.1016/j.jksus.2016.08.003

Hasanuzzaman, M., Mandal, M. H., Hasnine, M., & Shit, P. K. (2022). Groundwater potential mapping using multi-criteria decision, bivariate statistic and machine learning algorithms: evidence from Chota Nagpur Plateau, India. *Applied Water Science*, *12*(4). https://doi.org/10.1007/s13201-022-01584-9

Huang, P., Hou, M., Sun, T., Xu, H., Ma, C., & Zhou, A. (2024). Sustainable groundwater management in coastal cities: Insights from groundwater potential and vulnerability using ensemble learning and knowledge-driven models. *Journal of Cleaner Production*, *442*. https://doi.org/10.1016/j.jclepro.2024.141152

Jari, A., Bachaoui, E. M., Jellouli, A., Harti, A. El, Khaddari, A., & Jazouli, A. El. (2022). Use of GIS, Remote Sensing and Analytical Hierarchy Process for Groundwater Potential Assessment in an Arid Region – A Case Study. *Ecological Engineering and Environmental Technology*, *23*(5), 234–255. https://doi.org/10.12912/27197050/152141

Kumar, M., Singh, P., & Singh, P. (2023). Machine learning and GIS-RS-based algorithms for mapping the groundwater potentiality in the Bundelkhand region, India. *Ecological Informatics*, *74*. https://doi.org/10.1016/j.ecoinf.2023.101980

Liu, R., Li, G., Wei, L., Xu, Y., Gou, X., Luo, S., & Yang, X. (2022). Spatial prediction of groundwater potentiality using machine learning methods with Grey Wolf and Sparrow Search Algorithms. *Journal of Hydrology*, *610*. https://doi.org/10.1016/j.jhydrol.2022.127977

Manap, M. A., Nampak, H., Pradhan, B., Lee, S., Sulaiman, W. N. A., & Ramli, M. F. (2014). Application of probabilistic-based frequency ratio model in groundwater potential mapping using remote sensing data and GIS. *Arabian Journal of Geosciences*, *7*(2), 711–724. https://doi.org/10.1007/s12517-012-0795-z

Maskooni, E. K., Naghibi, S. A., Hashemi, H., & Berndtsson, R. (2020). Application of advanced machine learning algorithms to assess groundwater potential using remote sensing-derived data. *Remote Sensing*, *12*(17). https://doi.org/10.3390/RS12172742

Masroor, M., Rehman, S., Sajjad, H., Rahaman, M. H., Sahana, M., Ahmed, R., & Singh, R. (2021). Assessing the impact of drought conditions on groundwater potential in Godavari Middle Sub-Basin, India using analytical hierarchy process and random forest machine learning algorithm. *Groundwater for Sustainable Development*, *13*. https://doi.org/10.1016/j.gsd.2021.100554

Moore, I. D., Grayson, R. B., & Ladson, A. R. (1991). DIGITAL TERRAIN MODELLING: A REVIEW zy OF HYDROLOGICAL, GEOMORPHOLOGICAL, AND BIOLOGICAL APPLICATIONS. In *HYDROLOGICAL PROCESSES* (Vol. 5).

Mridha, G. C., Hossain, M. M., Uddin, M. S., & Masud, M. S. (2020). Study on availability of groundwater resources in selangor state of malaysia for an efficient planning and management of water resources. *Journal of Water and Climate Change*, *11*(4), 1050–1066. https://doi.org/10.2166/wcc.2019.043

Pawar, U., Suppawimut, W., & Rathnayake, U. (2024). Mapping of groundwater potential zones in a drought prone Marathwada Region using frequency ratio and statistical index methods, India. *Results in Engineering*, *22*. https://doi.org/10.1016/j.rineng.2024.101994

Ponnusamy, D., & Elumalai, V. (2022). Determination of potential recharge zones and its validation against groundwater quality parameters through the application of GIS and remote sensing techniques in uMhlathuze catchment, KwaZulu-Natal, South Africa. *Chemosphere*, *307*. https://doi.org/10.1016/j.chemosphere.2022.136121

Prasad, P., Loveson, V. J., Kotha, M., & Yadav, R. (2020). Application of machine learning techniques in groundwater potential mapping along the west coast of India. *GIScience and Remote Sensing*, 735–752. https://doi.org/10.1080/15481603.2020.1794104

Roy, S. K., Hasan, M. M., Mondal, I., Akhter, J., Roy, S. K., Talukder, S., Islam, A. K. M. S., Rahman, A., & Karuppannan, S. (2024). Empowered machine learning algorithm to identify sustainable groundwater potential zone map in Jashore District, Bangladesh. *Groundwater for Sustainable Development*, *25*. https://doi.org/10.1016/j.gsd.2024.101168

Seifu, T. K., Eshetu, K. D., Woldesenbet, T. A., Alemayehu, T., & Ayenew, T. (2023). Application of advanced machine learning algorithms and geospatial techniques for groundwater potential zone mapping in Gambela Plain, Ethiopia. *Hydrology Research*, *54*(10), 1246–1266. https://doi.org/10.2166/nh.2023.083

Sharma, Y., Ahmed, R., Saha, T. K., Bhuyan, N., Kumari, G., Roshani, Pal, S., & Sajjad, H. (2024). Assessment of groundwater potential and determination of influencing factors using remote sensing and machine learning algorithms: A study of Nainital district of Uttarakhand state, India. *Groundwater for Sustainable Development*, *25*. https://doi.org/10.1016/j.gsd.2024.101094

Thapa, R., Gupta, S., Guin, S., & Kaur, H. (2017). Assessment of groundwater potential zones using multiinfluencing factor (MIF) and GIS: a case study from Birbhum district, West Bengal. *Applied Water Science*, *7*(7), 4117–4131. https://doi.org/10.1007/s13201-017-0571-z

Vafadar, S., Rahimzadegan, M., & Asadi, R. (2023). Evaluating the performance of machine learning methods and Geographic Information System (GIS) in identifying groundwater potential zones in Tehran-Karaj plain, Iran. *Journal of Hydrology*, *624*. https://doi.org/10.1016/j.jhydrol.2023.129952

Wei, A., Li, D., Bai, X., Wang, R., Fu, X., & Yu, J. (2022). Application of machine learning to groundwater spring potential mapping using averaging, bagging, and boosting techniques. *Water Supply*, *22*(8), 6882–6894. https://doi.org/10.2166/ws.2022.283