**Geospatial technology-based MCDA and machine learning algorithms: an ensemble and inter-evaluating frameworks for conceptualizing groundwater potential**

Kesyton Oyamenda Ozegin a**\***, Soliu Ademola Mudashirub, Jude Steven Ejepuc

aDepartment of Physics, Ambrose Alli University Ekpoma, Edo State, Nigeria

bDepartment of Applied Geophysics, Federal University of Technology, Akure, Nigeria

cDepartment of Geology, Federal University of Technology, Minna, Niger State, Nigeria

\*Corresponding author’s e-mail: [ozeginkess@yahoo.com](mailto:ozeginkess@yahoo.com)

**Abstract**

Humanity's usage of water cannot be halted, as water is a necessary resource for existence and is highly valued in a variety of quotidian usages and initiatives. Hence, reliable identification of groundwater potential locations in diverse geographical areas (from low to high) is critical for ensuring its management and stability. This study aimed to predict and define groundwater potential (GWP) zones by combining machine learning (ML) models, such as random forest (RF), support vector machine (SVM), adaptive boosting (AdaBoost), and eXtreme gradient boost (XGB), with analytical hierarchy process (AHP)—multi-criteria decision analysis (MCDA) that integrates anthropological (normalized difference vegetation index), hydrogeological (geology and lineament density), hydrological (rainfall distribution and proximity to surface water bodies), and topographical factors (slope, aspect, drainage density, and topographic wetness index) as groundwater-influencing parameters. Prognostication competencies of these factors were established through the application of multicollinearity (MC) evaluation. The AHP approach was used to weight and scale these influencing variables and determine their impact on groundwater prospects. Given the constrained amount of data available, 'CreateFishnet\_management' requires training the ML models with 100% borehole data. Every model's reliability was quantified using the area under the receiver-operating characteristic (AU-ROC) curve. The four (4) machine learning algorithms and AHP model divided the GWP zone into four distinct groups: high, moderate, low, and very low. The RF, AdaBoost, SVM, XGB, and AHP models show that the beneficial GWP zones encompass 41.59%, 36.42%, 45.93%, 25.30%, and 14.03% of the research region, respectively. The model dependability using measured AU-ROC curve metrics indicates that the RF model has the greatest rate of success (78.00%), preceding the AdaBoost (77.00%), SVM (73.00%), XGB (71.00%), and AHP (50.00%) models. Consequently, this multi-approach gives valuable perspectives to promote long-term groundwater governance, directing choices on the advancement of the availability of groundwater.

**Keywords:** AHP; AU-ROC; Fishnet; Groundwater potential; Machine learning; Multicollinearity

**1. Introduction**

The crucial significance of water to humanity is widely acknowledged—water resources are required for the continued existence of individuals, organisms, and ecosystems (e.g., Ozegin et al., 2023; Kaya et al., 2023; Ozegin et al., 2024a; Chatterjee et al., 2024; UNWWDR, 2025). Population growth and rising economies have created an increased need for access to water. Currently, over forty percent of mankind is facing a shortage of water, and if this trend continues, over 6.3 billion people (globally) will be affected by various levels of water-related problems by 2030 (Opoku et al., 2024). In the three preceding centuries, approximately eighty-five percent of the world's wetlands have been desiccated, while the other 15% of wetlands have deteriorated in quality (UN DESA, 2023; Opoku et al., 2024). The Sustainable Development Goals (SDGs) are built on the sustained existence as well as management of fundamental water supplies, with the goal of providing "access to clean water and sanitation (SDG 6), zero hunger (Goal 2), and good health and well-being (Goal 3) through sustainable management of terrestrial ecosystems and their services (SDG 15) and mitigating the effects of climate change (SDG 13)" (United Nations, 2015). Thereby strengthening sustainable water, food, and energy nexuses and fostering perspectives for a reliable water future. Given the ever-increasing need for water and a dearth of freshwater resources, a shortage of water presents serious problems, including inadequate access to water, farming impacts, ecological damage, and social and economic inequality. Furthermore, changes in the climate exacerbate water scarcity concerns by altering trends in precipitation, lengthening the degree, severity, and incidence of droughts.

According to Scanlon et al. (2023), over two billion people worldwide rely on groundwater (GW) for their water requirements on a daily basis. Considering the rising worldwide shortages of water, the immense potential benefits of groundwater can hardly be discounted. GW, considered an integral part of the hydrologic process, is essential for the survival of several ecosystems, both aquatic and terrestrial. Present-day extraction of groundwater accounts for roughly 26% of all freshwater removal worldwide (Van der Gun, 2012). The frequency of removal of groundwater globally has at least increased threefold and continues to rise at a pace of approximately 0.5-1.0% annually. Specifically, global extractions of freshwater rose 14% between 2000 and 2021—representing a mean yearly increase of 0.7%. As stated by the most contemporary worldwide estimates (from 2021), agriculture accounts for 72% of freshwater extractions, preceded by industry (15%) and home (or municipal) consumption (13%) (UNWWDR, 2025). Sector-driven extractions of freshwater vary greatly depending on the economic growth of a nation. Higher-income nations utilize greater amounts of water for industry, while low-income countries use ninety percent or greater for agricultural irrigation. As climate change makes water supply increasingly unpredictable, groundwater serves as a key reserve during seasons of scarcity, increasing the significance of aquifers for preserving water. GW is a critical worldwide resource that stores the most freshwater, aside from glaciers (WWQA, 2021). The overall quantity of water on the earth is projected to be 1386 trillion litres, with oceans (salt water) accounting for 97% of the water (Senanayake et al., 2016). Of the remaining 3% of freshwater, two-thirds is frozen in ice caps. Less than 1% (0.67) of the world's water is viable freshwater, and groundwater constitutes about 97 percent of total usable freshwater (UNWWDR, 2022; Ozegin et al., 2024a).

The accessibility of water is an ongoing issue for billions of people throughout the globe! Water-scarce urban populations are estimated to be 1.7–2.4 billion globally by 2050, from 930 million as of 2016 (United Nations, 2023). The lack of water affects the African continent both physically and economically. Physically, it is mostly caused by the depletion of resources as a result of climate change and human activity. While economic constraint comes when water is inaccessible due to inadequate capacity or organizational failings requiring a concerted effort, which is Africa's major issue. Nigeria possesses an abundance of water assets, but the government hasn't leveraged them to their fullest ability to give its citizenry a consistent and adequate supply of drinkable water at an affordable price (e.g., Kolawole, 2015; Ozegin et al., 2023). With a projected population of over 215 million, Nigeria is among the six largest populated nations around the world (UN DESA, 2022; Ozegin and Ilugbo, 2024), making the issue of water scarcity extremely onerous. With only 61% of Nigerians having access to safe water for consumption and an estimated 71 million individuals still requiring improved water services, the country's water industry is beset by serious problems (World Bank Group, 2017; Adewumi et al., 2020; Ozegin et al., 2024b). A large portion of the study area is underpinned with impermeable or nearly impervious materials. Improper borehole placement and considerable borehole failure probabilities are frequently caused by adverse hydrogeological settings and a lack of information about aquifers. Furthermore, accessing water is especially difficult in locations where precipitation is low or uneven, as well as in areas where the consequences of climate change tend to be more severe. According to Ozegin et al. (2023), the bulk of the aquifer networks within the research area are composed of crystalline and basaltic sediments that are distinguished by secondary porosity, lineaments, and fragmented rocks. The penetrability exhibited by these unconfined aquifers is comparatively higher, and their ability to hold water is diminutive. Based on hydrogeological settings, approximately 76% of the research region's land area comprises hard rocks possibly exhibiting low groundwater potential, with the remaining 24% consisting of sedimentary deposits that have possibly significant potential for groundwater (Ozegin et al., 2023). Sedimentary rock is only found in the southern section of the research area (Ozegin et al., 2023).

Considering the intrinsic variability of hard-rock aquifers, identifying optimal spots for groundwater (also known as groundwater potential or groundwater potential (GWP) zones) is difficult and necessitates an extensive knowledge of all variables that control the accessibility of groundwater in the study area. Groundwater supply and movement are influenced by topography, hydrology, ecology, geology, and the atmosphere (e.g., Oh et al. 2011; Golkarian et al. 2018; Arshad et al., 2020; Ilugbo et al., 2023a; Reza et al., 2023). Groundwater researchers are increasingly employing these parameters to establish groundwater potential (e.g., Al-Shabeeb et al., 2018; Ozegin et al., 2024a, b). To ensure the continued viability of water management strategies and to effectively evaluate the availability of groundwater in the study area, a thorough understanding of the features and dynamics of aquifers—geological structures that hold and distribute water—is required using machine learning (ML) and geoinformatics-based methods for the geohydrological assessment of groundwater-enriching sites.

In broad terms, there are two main approaches for assessing groundwater potential in any area: standard techniques and advanced methods—expert decision frameworks and machine learning (see Ozdemir, 2011; Agarwal et al., 2013; Naghibi et al., 2016; Guru et al., 2017; Abrams et al., 2018; Mohammadi-Behzad et al., 2019; Arefin, 2020). The standard approach is based on traditional field surveys, which are laborious and costly. Expert decision frameworks include the analytical hierarchy process (AHP), weighted overlay method, evidential belief function (EBF), fuzzy AHP, and stepwise weight assessment ratio analysis (SWARA) (e.g., Arefin, 2020; Abrar et al., 2021; Fajemilo and Ozegin, 2025), whereas ML algorithms (MLAs) include random forest (RF), artificial neural network (ANN), K-nearest neighbour (KNN), gradient boosting (GB), and so on. These advanced approaches provide more accurate and dependable results than conventional approaches for assessing GW potential (Lee et al., 2012; Das, 2019; Rana et al., 2025). ML models commonly rely on enormous amounts of data, which are either not readily accessible or inadequate in multiple practical realities of hydrologic situations. This difficulty has sparked interest in expert decision frameworks. The application of geospatial-based MCDA and hybrid learning models has significantly improved GWP mapping reliability. Arguably the most common approach for modelling an area with potential for groundwater is to combine a variety of ML algorithms and AHP approaches with geospatial (GIS and remote sensing (RS)) technology (Hasanuzzaman et al., 2022; Dey et al., 2023; Rana et al., 2025). Acknowledging the drawbacks of conventional techniques, this study proposes an innovative amalgamation of machine learning (ML) models, including random forest (RF), support vector machine (SVM), adaptive boosting (AdaBoost), and eXtreme gradient boost (XGB), with analytical hierarchy process (AHP) approaches within a GIS structure. By integrating these cutting-edge methods with varied datasets, this study provides a quint-perspective strategy for groundwater potential modelling that improves reliability and usability in the research area.

ML algorithms are a recently developed technique that has produced fascinating outcomes (Vafadar et al., 2023; Fajemilo and Ozegin, 2025). The models are entirely based on computer technologies and are designed to address tough problems involving exponential and intricate datasets. The fundamental advantage of algorithms is that they deal directly with unprocessed data, which substantially decreases experts’ influence. Several research studies (e.g., Lee et al., 2012; Naghibi and Pourghasemi, 2015; Rahmati et al., 2016; Sahoo et al., 2017; Golkarian et al., 2018; Lee et al., 2018; Chen et al., 2019; Sameen et al., 2019; Avand et al., 2020; Maskooni et al., 2020; Patidar et al., 2021; Karimi-Rizvandi et al., 2021; Pham et al., 2021; Hasanuzzaman et al., 2022; Vafadar et al., 2023; Shandu and Atif, 2023; Khan and Jhamnani, 2023; Sharma et al., 2024; Rana et al., 2025) have used ML algorithms to map groundwater potential zones, including random forest (RF), eXtreme gradient boost (XGB), logistic regression (LR), classification and regression tree (CART), support vector machine (SVM), adaptive boosting (AdaBoost), genetic algorithm (GA), and K-neighbour (KN). These researchers created reliable maps with high levels of accuracy and yielded convincing outcomes. The random forest (RF), support vector machine (SVM), adaptive boosting (AdaBoost), and eXtreme gradient boost (XGB) were chosen for the study because they have been reported to be more efficient than other MLAs such as genetic algorithm, boosted regression tree (BRT), artificial neural network model, classification and regression tree, and LR (Naghibi et al., 2017; Moghaddam et al., 2020; Patidar et al., 2021; Vafadar et al., 2023). While these machine learning algorithms are capable of handling various predictive features and facilitating fitting interactions between indicators (e.g., Olden et al. 2008; Khan and Jhamnani, 2023), they are additionally susceptible to overfitting data, which makes fishnet (FN) analysis crucial. The prediction power of the four chosen ML algorithms is additionally contrasted to guarantee the dependability of the resulting GWP zones map.

Remote sensing technologies are developing rapidly, and a significant amount of RS materials are freely available globe-wide. These can be used to identify global driving features and effectively calculate GW’s future potential. RS data and GIS technology play a pivotal part in locating and mapping possible groundwater sites. For evaluating GWP zones, this study espoused the expert judgment system known as the analytical hierarchical process (AHP). Spatial challenges such as exploring groundwater present a multiple-attribute choice-making task since they incorporate a spectrum of features that are evaluated based on asymmetric and competing criteria (Malczewski, 2010; Ilugbo et al., 2023b). The combination of these several elements aids in creating a realistic and dependable predictive map that reflects the location's foreseeable groundwater extraction plans. Using multi-criteria decision analysis (MCDA), e.g., AHP, can help accomplish this combination. The AHP is a multifaceted decision-making approach that creates a framework of hierarchy out of an intricate situation— splitting problems into tiers, encompassing objectives, variables, and choices, which are then assessed both intuitively and analytically (Wind and Saaty, 1980; Dar et al., 2021). This approach adequately overcomes the challenge of weight assignment for evaluating indexing; therefore, numerous studies have used the AHP in groundwater potential evaluations (e.g., Singh et al., 2016; Patra et al., 2018; Ajay Kumar et al., 2020; Muthu and Sudalaimuthu, 2021; Dar et al., 2021; Abrar et al., 2021; Zhang et al., 2021; Castillo et al., 2022; Ozegin et al., 2023; Ilugbo et al., 2023a, b; Sharma et al., 2024; Opoku et al., 2024). The proportional relevance of the various objectives is established by a pairwise contrast method that generates a prioritizing scale based on expert judgments (Saha et al., 2024; Opoku et al., 2024). By taking into account a variety of parameters, this method helps analysts establish regions with groundwater potential without being influenced by the geographic environmental spatial complexities (e.g., Dar et al., 2021; Thanh et al., 2022; Baghel et al., 2023). Danumah et al. (2016) employed this approach to identify flood-prone locations, while Akıncı et al. (2013), Mansour et al. (2019), Morales and de Vries (2021), Kalura et al. (2021), and Rao et al. (2022) utilized an integrated method that included AHP and GIS to assess the ecological functioning of various regions. In recognition of the intrinsic intricacy and unpredictability of groundwater systems, identifying GWP zones requires the incorporation of multiple data sources into a GIS. A GIS is an effective means for comprehending and inferring geographic information. It can visualize data by processing remote sensing photos, geologic charts, topographic diagrams, and various other kinds of data. This attribute allows for an accurate depiction of the geographical distribution and availability of groundwater supplies (Pande et al., 2020; Baghel et al., 2023). The use of AHP and GIS tools provides an efficient method for conducting quantitative and methodical/systematic evaluations of groundwater potential.

This study accentuates the importance of incorporating all-inclusive geospatial (GIS and RS), machine learning (ML), and AHP techniques to establish the GWP zones in a geographical region with a diversity of anthropological (normalized difference vegetation index (NDVI)), hydrogeological (geology and lineament density), hydrological (rainfall distribution and proximity to surface water bodies), and topographical factors (aspect, drainage density, slope, and topographic wetness index (TWI)). The study's objectives are fourfold: (i) to discern and assess the variables that influence groundwater; (ii) to integrate the AHP and ML techniques within a GIS context in order to generate high-accuracy prospective groundwater maps of the research area; (iii) to determine the reliability of AHP and ML technology in determining GWP zones using the AU-ROC curve, guaranteeing a rational system for assessing groundwater potential; and (iv) to leverage cutting-edge geospatial, machine learning, and multi-subjective analysis approaches to offer a useful decision-support system for farming strategy and sustainable use of water resources in the research area. These will support the development of economical, efficient, and scientific approaches to sustainable management of groundwater by water administrators and regulators. The findings of this study may potentially offer sensible suggestions for managing and developing groundwater in regions with complicated geologic origins. Besides, this study advances the domain of geospatial evaluation and environmental modelling by demonstrating the ability of machine learning techniques to resolve difficult hydrogeological concerns while enhancing the management of water assets.

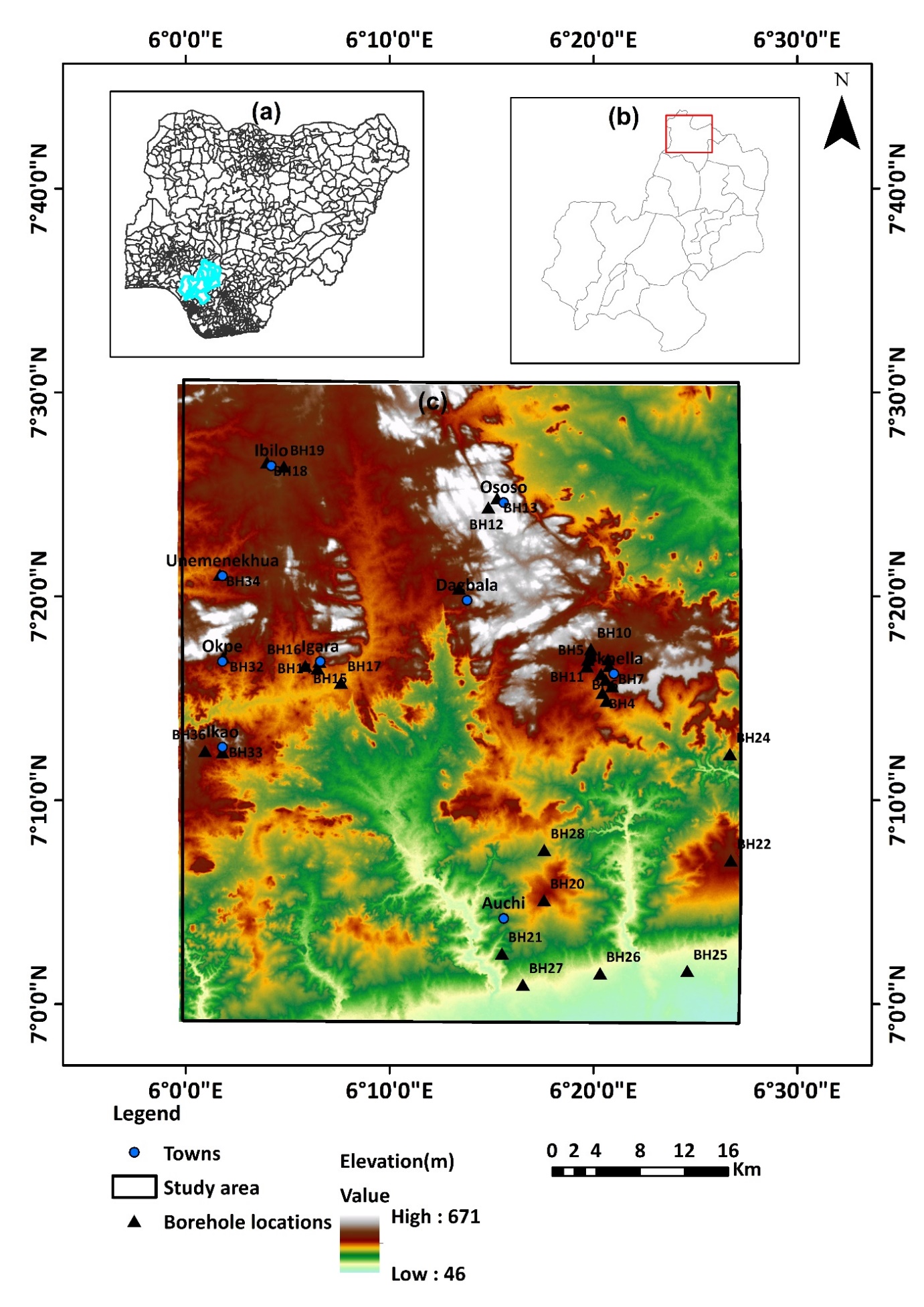
The remaining sections of the paper are structured as follows: Section 2 highlights the study area. Section 3 presents the data sources and proposed approaches. Section 4 evaluates and discusses the modelling results, which include experimental appraisal, comparative analysis, model constraints, and a critique of the SDGs' potential consequences. Finally, Section 5 presents important findings, their potential for improvement, and ideas for further investigation.

**2. Study Area**

The research area is situated in Edo North District (sheet 226), Edo State, Nigeria (**Fig. 1**). This area (6°00-6°30E, 7°00-7°30N) was selected for its distinct assembly of geographical, meteorological, and demographic factors, making it an appealing perspective for groundwater evaluation in data-scarce areas.

According to Odeyemi (1981) and Obaje (2009), the research region is notably located among the Precambrian basement complex rocks of the southwest region of Nigeria, forming part of the Nigerian basement complex and the Pan-African mobile belt, which is located east of the West African Craton. The eastern and western districts are the two separate provinces that make up the basement complex of southwest Nigeria. Migmatite, Gneisses, and substantial amounts of Pan-African Granites, interposed with Mesozoic Younger Granites, are the features that define the eastern province. In contrast, the western province was primarily composed of gneisses or granites, migmatite, and poor-quality Schist Belt. The southwest section of the basement complex in southwest Nigeria constitutes where the research region is located. The rocks in this region are migmatite–gneiss complexes, which are primarily sedimentary series with modest igneous rock incursions believed to have undergone granitic, migmatitic, and metamorphic changes (see Obaje, 2009) over the course of geological history. Considering already existing basic structures were obliterated by succeeding distortion, Odeyemi (1981) proposes that nearly each of the foliations seen in the rocks of southwestern Nigeria, with the exception of intrusives, is primarily tectonic in provenance. The landscape is mostly gently sloping, with granitic mountains extending eastward. Given that the majority of the studied region is fundamentally resistant to groundwater possibility, significant faulting and weathering processes triggered by secondary porosity commonly occur in the availability of groundwater in the studied region. Accordingly, fractures, geologic interactions, shearing regions, faults, and various other discontinuities typically regulate the movement of groundwater in the hard rocks; their complex interactions regulate the process of the watershed as a whole (Ozegin et al., 2023). Groundwater is accessed in a study region through springs, hand-dug wells, and deep reservoirs.

The study region is between 46 and 671 meters above sea level. The human population in some parts of the study area is dispersed widely because of the region's rough terrain. The region's climate and vegetation make it easier to grow food crops, including cereals, yams, beans, maize, cassava, plantains, and some tree crops like avocados, mangos, oil palms, and cashews, in addition to a variety of other income crops. Since there are several agrarian community farms in the area, agriculture is the primary occupation of the people living in the studied area. Recognizing that groundwater is the inhabitants' primary supply of potable and irrigation water, it is critical to assess aquifer possibilities in this region. Furthermore, given the region’s relatively tropical climate, the development of soil and weathering of rocks are encouraged. Rain and dry conditions are the two basic ones. Low to moderate rainfall occurs throughout the rainy period, which starts in basically April and spans through October. November marks the start of the dry period, which ends in March. The geologic framework controls the relief, which is characterized by lower-lying areas to the south and rising older granites in the uneven hilly higher elevations, especially across the northern margins. In the context of pedology, the research area exhibits saturated loamy and stony sandstone varieties of soil in the western portion, alternating with clayey soil in the eastern portion, which primarily covers the middle and eastern parts.



**Fig. 1.** The map depicts the region of focus

**3. Data sources and Methodology**

In order to determine the GWP zones in a geographical area located in the Edo North Sub-basin of Nigeria, this study emphasizes the significance of combining comprehensive geospatial data, ML algorithms, and AHP techniques. Numerous influencing criteria were used in this evaluation. However, a multicollinearity (MC) analysis helped narrow down the influencing factors to the nine features subsequently used: normalized difference vegetation index (NDVI), geology (GY), lineament density (LD), rainfall distribution (RD), proximity to surface water bodies (PSW), aspect (AP), drainage density (DD), slope (SP), and topographic wetness index (TWI). GWP zone evaluation and spatial modelling are conducted using data from multiple databases.

**Table 1** outlines the geographical data used in the study, indicating the sources that were used. Nine theme layers were created by combining data from various sources. In the scientific field, thematic layers are an effective method for establishing potential GW locations. The Shuttle Radar Topography Mission (SRTM) and digital elevation model (DEM) information were combined to generate distinct geographic features known as the conceptual strata. These layers comprised lineament density (LD), aspect (AP), drainage density (DD), slope (SP), and topographic wetness index (TWI). PERSIANN rainfall data from 2021-2024 was leveraged to generate the rainfall distribution data because the Climatic Research Unit (CRU) has a significantly broader resolution and consequently provides no detailed resolution for our study area; a geologic map was produced by digitizing the previous EDO state map (Ozegin and Ilugbo, 2025); and an NDVI map was generated using the MODIS13Q1 dataset retrieved for this date (12-07-2022) and explored from the NASA LAAD DAACS site. PSW was produced using the topographic map from ArcGIS online digitized. To create maps indicating the potential of groundwater, a thorough evaluation has been carried out using knowledge-centred weighted analysis and AHP models. This is a popular method for choosing a decision that incorporates a number of factors (Saaty, 1990). Once the weights of each class were established, an overlay analysis was performed using ArcGIS's raster calculator. Weights in AHP were normalized using the geometrical average algorithm. In order to get the geometric mean, the variables were scored using a predetermined scale (1–9). The geometric average is calculated by dividing the overall scale weight—the sum of the scales of all the parameters—by a total quantity of variables. Also, machine learning algorithms, including RF, SVM, AdaBoost, and XGB, were incorporated. The illustration graphic (**Fig. 2**) provides a detailed representation of the entirety of the processes used in the study.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Factors** | **The categories of data and formats.** | **Description** | **Date last accessed** | **Source** |
| Geology | Image | Georeferenced | April, 28, 2025 | NGSA |
| Lineament density | SRTM DEM 30m – Raster | Derived from DEM | April, 25, 2025 | <https://opentopography.org/> |
| Drainage density | SRTM DEM 30m – Raster | Derived from DEM | April, 25, 2025 | <https://opentopography.org/> |
| Slope | SRTM DEM 30m – Raster | Derived from DEM | April, 25, 2025 | <https://opentopography.org/> |
| Aspect | SRTM DEM 30m – Raster | Derived from DEM | April, 25, 2025 | <https://opentopography.org/> |
| Topographic wetness index | SRTM DEM 30m - Raster | Derived from DEM | April, 25, 2025 | <https://opentopography.org/> |
| NDVI | MODIS 13Q1- Raster | Downloaded | April, 28, 2025 | <https://ladsweb.modaps.eosdis.nasa.gov/search/> |
| Proximity to surface water bodies | Topographic map – ArcGIS Online | Digitized | April, 29, 2025 | ArcGIS |
| Rainfall | PERSIANN – Raster | Downloaded | 2021- 2024 | <https://chrsdata.eng.uci.edu/> |
| Borehole data | Vector | Derived from detailed fieldwork | April, 19, 2025 | Research fieldwork survey |

**Table 1**

The dataset utilized in the research for GWP categorization modelling.

**Groundwater potential conditioning factors**

**TWI**

**Rainfall**

**Proximity to surface water bodies**

**Lineament density**

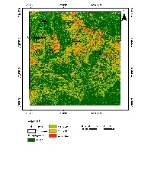
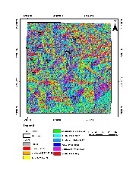
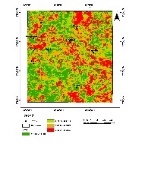
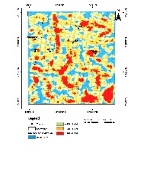
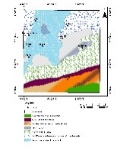
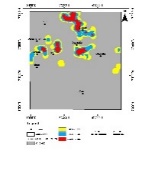
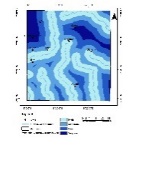
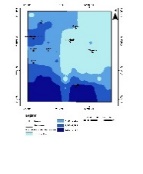
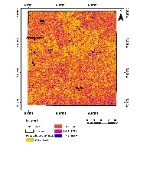
**Geology**

**Drainage density**

**NDVI**

**Aspect**

**Slope**



**Database**

**Borehole (BH) yield data map**

**Raster value extracted to BH points**

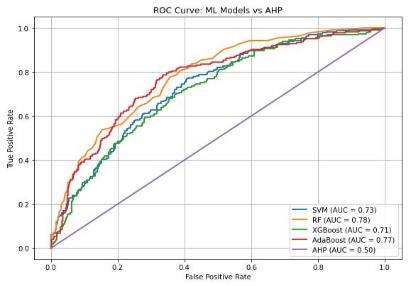
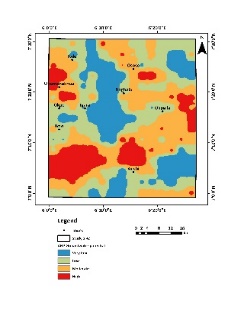
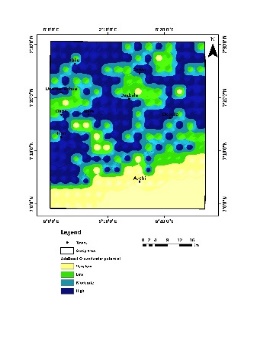
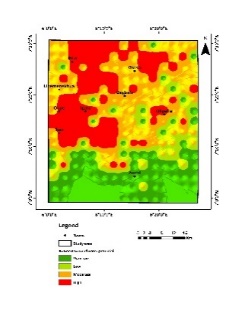
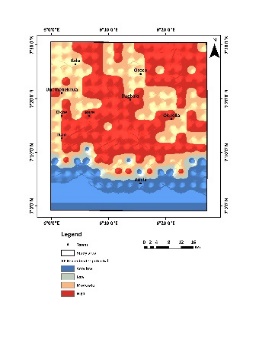
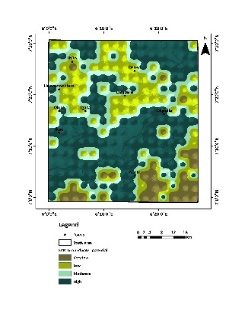
**Raster value extracted to fishnet points**

**Training data**

**Testing data**

**Groundwater potential maps**

**SVM, RF, XGBoost, AdaBoost and AHP Models**



**Validation**

**ROC curve**

**Multicollinearity test**

**Fig. 2.** A flow diagram describing the research's methodology

3.1. Collinearity statistics

To verify the plurality of affecting parameters, collinearity statistical processes were established prior to overlay analysis implementation. A significant correlation between the components can be detected during the multiple regression algorithm analysis; this is known as multicollinearity (MC) (e.g., Venkatesh and Parimalarenganayaki, 2023; Ozegin et al., 2024b). It should be noted that MC evaluations are usually conducted prior to regression studies in groundwater-conceivable models in order to guarantee the accuracy and dependability of the model. Nine (9) groundwater-predicting indicators with a significant effect on freshwater were selected for the study geographic location, and tolerance (T) and variance inflation factor (VIF) served as means to assess the degree of multiple correlations. To avoid distortion in groundwater models, these freshwater-influencing features were to be assessed for MC prior to proceeding with geospatial and ML modelling. Arguably, the leading application of regression in the sciences of hydrology is the development of spatial algorithms that link hydrological parameters, such as estimates of the lowest flow, to water attributes (Kroll and Song, 2013; Ozegin et al., 2024b). Tolerance (T) and variance inflation factor (VIF) were measured using a method known as linear regression in this study to perform collinearity statistics. This approach uses one input characteristic as a variable that depends on the remaining elements as distinct variables in order to determine the coefficient of determination (R²). The values of T and VIF are usually determined (e.g., Saha, 2017; Maity and Mandal, 2019; Venkatesh and Parimalarenganayaki, 2023; Ozegin et al., 2024b) using **Eqs.** (1) and (2).

 (1)

 (2)

The procedure stated above is repeated by modifying the dependent components until the T and VIF are determined for all factors. T limits are > 0.10, and VIF is ≤ 10. A value for each factor that deviates from the standard values (i.e., T < 0.10 and VIF values > 10) shows multicollinearity issues and should be deleted when evaluating prospective maps (Saha, 2017; Venkatesh and Parimalarenganayaki, 2023; Ozegin et al., 2024b).

The ArcMap extraction program was used to determine the values of 500 (N = 500) independently selected spots within the theme maps for the determinants' convergence statistics. T and VIF values were computed in **Table 1** after a linear regression examination was carried out using the statistical package SPSS on the obtained parameters.

3.2. MCDA-AHP-based model and GIS techniques

*3.2.1. AHP for GWP (Comparative matrix development and variable weighting)*

The AHP framework, among the MCDA-based frameworks employed to provide outcomes for difficult decision-making scenarios, was initially developed by Saaty (1990). One widespread MCDA process for addressing natural assets and the surrounding ecosystem is Saaty's (1990) AHP. Applying the widely utilized AHP approach, a standardized weight is assigned to each theme layer of the groundwater exploring feature. The principal eigenvalue of the generated matrix was used to determine the final weight for every theme layer. The relative weights and correlations of every item are examined in the first step of AHP computations using an n × n matrix (A) with diagonal parts equal to 1. As the comparison of pairs demonstrates, the value is additionally adjusted to get a weighted value (W). The foundation of the concept of AHP is a set of paired assessing matrix equations (**Eq. 3**) that compare encompassing factors in order to assess the weighting of every parameter (see Saaty, 1990; Das, 2019; Abrar et al., 2021; Ozegin et al., 2023; Ilugbo et al., 2023a; Baghel et al., 2023; Sharma et al., 2024; Zewdie et al., 2024). The most prevalent method for identifying the prioritizing vector is the Saaty technique, which asserts that the most important vector should be the dominating eigenvector of A. Given the homonymic idea, it is frequently called the Perron-Frobenius eigenvector in an algebraic linear relationship (Horn and Johnson, 1985; Brunelli, 2015; Ozegin et al., 2024a, b, c). The following realization forms the foundation of the process: **Eq. (3)** can be obtained by multiplying a matrix A by W along with its entries, which are exactly expressed as weighted proportions:

** (3)

A comparison is made between the row matrices and column matrices components. Skilled competence, actual fieldwork, and the possible influence of groundwater were used to determine the proportions of all of the distinct strata. In this research, the pairwise evaluation phase of the AHP technique was conducted using a 9 × 9 matrix.

**Table 2**'s comparative ratings, which are determined on Saaty's 1 to 9 scale, were developed using knowledgeable views and assessments of pertinent literature to determine the relative value of each influencing feature for explorative characteristics. According to the Saaty scale, utmost significance is represented by a number of 9, high significance by a number of 7, great by a value of 5, considerable by a number of 3, equally significant by a value of 1, and intermediary significance by a number of 2, 4, 6, and 8. According to the categorization, themes are given weights according to their importance and ability to retain water. A pairwise evaluation matrix was used for contrasting the influencing variables (**Table 3**). The comparative comparison's outcome is adjusted to establish each parameter's score. To generate a diagonally oriented matrix, Saaty (1990) recommends aligning the numbers in the matrix's top triangle. The inverse values of the originating array are then inserted into the base triangular array. The proportional weights produced by this procedure are then established (**Table 4**). An estimate of the proportionate relevance of the indicators being compared is given by these averages. The final indicator priority is represented by the normalized quantities of the eigenvectors, which are linked to the optimum eigenvalues of the proportionate (ratio) matrix. The process outlined here constitutes the most effective way to lessen the impacts of ratio inequalities. Corresponding values of the criteria are shown in **Table 5**. The dominant geological characteristics and intricate recharge mechanisms describe this research locale.

The principal eigenvalue, which evaluates the uniformity of the approach to decision-making, is used to compute the consistency ratio (CR). The CR aids in assessing the consistency and dependability of the decisions reached when comparing the criteria by pairing.

**Table 2:**

Scale for value in relative terms.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Strength | 1 | 3 | 5 | 7 | 9 | 2,4,6,8 |
| Description | equally significant | considerable | great | highly significant | utmost significance | Intermediary significance |

**Table 3.**

Comparative matrix between pairings with the primary eigenvector normalized.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Matrix | **TWI** | **RD** | **PSW** | **LD** | **GY** | **DD** | **NDVI** | **AP** | **SP** | Primary eigenvector normalized |
| **TWI** | 1.000 | 0.500 | 2.000 | 0.500 | 0.333 | 1.000 | 3.000 | 5.000 | 5.000 | 10.5554% |
| **RD** | 2.000 | 1.000 | 3.000 | 1.000 | 0.500 | 2.000 | 5.000 | 7.000 | 7.000 | 17.9395% |
| **PSW** | 0.500 | 0.333 | 1.000 | 0.333 | 0.250 | 0.500 | 2.000 | 4.000 | 4.000 | 6.8214% |
| **LD** | 2.000 | 1.000 | 3.000 | 1.000 | 0.500 | 2.000 | 5.000 | 7.000 | 7.000 | 17.9395% |
| **GY** | 3.000 | 2.000 | 4.000 | 2.000 | 1.000 | 3.000 | 6.000 | 8.000 | 8.000 | 27.3092% |
| **DD** | 1.000 | 0.500 | 2.000 | 0.500 | 0.333 | 1.000 | 3.000 | 5.000 | 5.000 | 10.5554% |
| **NDVI** | 0.333 | 0.200 | 0.500 | 0.200 | 0.167 | 0.333 | 1.000 | 3.000 | 3.000 | 4.3478% |
| **AP** | 0.200 | 0.143 | 0.250 | 0.143 | 0.125 | 0.200 | 0.333 | 1.000 | 2.000 | 2.4456% |
| **SP** | 0.200 | 0.143 | 0.250 | 0.143 | 0.125 | 0.200 | 0.333 | 0.500 | 1.000 | 2.0862% |
| **Summation** |  |  |  |  |  |  |  |  |  | 100% |

Normalized difference vegetation index (NDVI), geology (GY), lineament density (LD), rainfall distribution (RD), proximity to surface water bodies (PSW), slope (SP), drainage density (DD), topographic wetness index (TWI) and aspect (AP).

**Table 4.**

Normalized matrix

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **TWI** | **RD** | **PSW** | **LD** | **GY** | **DD** | **NDVI** | **AP** | **SP** |
| **TWI** | 0.103678 | 0.10556 | 0.103379 | 0.10556 | 0.108488 | 0.103678 | 0.103807 | 0.106946 | 0.108891 |
| **RD** | 0.178193 | 0.1772 | 0.180328 | 0.1772 | 0.178413 | 0.178193 | 0.180942 | 0.182827 | 0.181259 |
| **PSW** | 0.067383 | 0.06975 | 0.064904 | 0.06975 | 0.072326 | 0.067383 | 0.064497 | 0.066985 | 0.070948 |
| **LD** | 0.178193 | 0.1772 | 0.180328 | 0.1772 | 0.178413 | 0.178193 | 0.180942 | 0.182827 | 0.181259 |
| **GY** | 0.278818 | 0.27207 | 0.281312 | 0.27207 | 0.260372 | 0.278818 | 0.281803 | 0.271214 | 0.261351 |
| **DD** | 0.103678 | 0.10556 | 0.103379 | 0.10556 | 0.108488 | 0.103678 | 0.103807 | 0.106946 | 0.108891 |
| **NDVI** | 0.043209 | 0.045112 | 0.041115 | 0.045112 | 0.047179 | 0.043209 | 0.04004 | 0.041084 | 0.045243 |
| **AP** | 0.025152 | 0.025885 | 0.02398 | 0.025885 | 0.025873 | 0.025152 | 0.023193 | 0.021796 | 0.023186 |
| **SP** | 0.021695 | 0.021662 | 0.021275 | 0.021662 | 0.020448 | 0.021695 | 0.02097 | 0.019375 | 0.018971 |

**Table 5.**

Evaluation of input parameter weights and classification for GWP zoning

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Criteria** | **Prospect for groundwater potential** | **Class** | **Area covered (Km2)** | **% of the study area covered** | **Rating** | **Normalized AHP Weight** |
| **TWI** | Maximum value preferred | -0.782 – 3.371  3.372-5.41  5.411-8.733  8.734 – 18.475  **Tota**l | 905  970  385  130  **2390** | 38  41  16  5  **100** | 1  2  3  4 | 0.10555 |
| **RD (**mm/year**)** | Maximum value preferred | 2104 – 2,339  2340 - 2446  2447 – 2579  2580 – 2767  **Tota**l | 808  767  418  397  **2390** | 34  32  17  17  **100** | 1  2  3  4 | 0.17940 |
| PSW (km) | Minimum value preferred | High  Moderate  Low  Very low  **Tota**l | 888  872  330  300  **2390** | 37  36  14  13  **100** | 1  2  3  4 | 0.06821 |
| **LD (**km-1**)** | Maximum value preferred | 0 – 0.090  0.091 – 0.270  0.271 – 0.500  0.501 – 0.860  **Tota**l | 1988  192  130  80  **2390** | 83  8  5  4  **100** | 1  2  3  4 | 0.17940 |
| GY | Type descriptive | Clays, shales with limestone  Coal, sandstone and shale  Falsebedded sandstone, coal and shale  Older granite  Shale and mudstones  Undifferentiated basement complex with pebbles  Undifferentiated meta sediments  **Tota**l | 590  419  505  352  265  149  110  **2390** | 25  18  21  15  11  6  4  **100** | 3  2  1  3  4  3  2 | 0.27309 |
| **DD (**km-1**)** | Minimum value preferred | 0.030 – 0.750  0.751 – 1.060  1.061 – 1.380  1.381 – 2.330  **Tota**l | 515  746  703  426  **2390** | 22  31  29  18  **100** | 4  3  2  1 | 0.10555 |
| **NDVI** (o) | Maximum value preferred | 0.0184 – 0.2684  0.2685 – 0.4723  0.4724 - 0.6993  0.6994 – 0.9993  **Tota**l | 774  604  601  411  **2390** | 32  25  25  18  **100** | 1  2  3  4 | 0.04348 |
| **AP** (o) | Type descriptive | Flat (-1 - 0)  North (0 – 22.5)  Northeast (22.6 – 67.5)  East (67.6 – 112.5)  Southeast (112.6 – 157.5)  South (157.6 – 202.5)  Southwest (202.6 – 247.5)  West (247.6 – 292.5)  Northwest (292.6 – 337.5)  North (337.6 - 360)  **Tota**l | 211  213  221  261  287  294  255  224  214  210  **2390** | 9  9  9  11  12  12  11  9  9  9  **100** | 4  4  3  2  1  1  1  2  3  4 | 0.02446 |
| **SP** (o) | Minimum value preferred | 0 - 4.0  4.1 – 10.0  10.1 – 19.0  19.1 – 65.0  **Tota**l | 590  100  700  1000  **2390** | 25  4  29  42  **100** | 4  3  2  1 | 0.02086 |

*3.2.2. Matrix consistency evaluation*

Considering the ratings were based on opinionated or individualized appraisals, they might be somewhat skewed and unpredictable. Therefore, using **Eq. 4**, the consistency ratio (CR) is typically employed to verify the choice to compare thematic layers and subcategories among conceptual layers by pairing (Saaty, 1990). For the analytical process to proceed, a CR of 0.10 or less is deemed suitable. It is necessary to reevaluate the evaluation considering the CR value exceeds 0.10 in order to identify any sources of inconsistencies and make the necessary corrections. On the other hand, when the CR value is zero, it indicates that the pairwise analysis is perfectly consistent.

Moreover, the proportions of weight acquired from this approach were normalized using Table 4. The basic eigenvalue generated in this process is used to evaluate the cohesiveness of conceptions using the CR. Brunelli (2015) defines the major eigenvalue as the measure of the matrix's deviation from regularity. The CR is calculated by applying **Eq. (5)** with the CI.

(4)

where RI defines random index and CI for consistency index.

(5)

where n is the variety of features used and λmax is the principal eigenvalue.

In the context of **Eq. 5**, λmax denotes the most substantial absolute eigenvalue of the contrasting matrix pairings determined by **Eq. (6).** (e.g., Ikrri et al., 2023).

 (6)

In **Eq. (3),** W is the appropriate eigenvector for λmax, and AWi (i = 1, 2, 3, ..., n) is the weighting quantity for every variable, which is conveniently derived from the matrix (Mandal et al., 2021; Ikrri et al., 2023).

Consequently, the CI for the present analysis is

CI = (9.22 – 9)/ (9 − 1) = 0.0275

The dependability of the outputs was evaluated by applying the computed CI and RI values derived from **Table 6** data equivalent to nine (9) in number (Saaty, 1990).

Thus,  
CR = = 0.0190 (1.90%)

The CR, estimated as 0.019 and established to be less than 0.1, showed the resulting weights acquired were uniform (**Table 7**). As a result, the weights established in **Table 5** can potentially be utilized.

**Table 6.**

RI estimates distinct matrix sizes (Alonso and Lamata, 2006; Ozegin et al., 2024a).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **n (matrix size)** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
| **RI** | 0.00 | 0.00 | 0.53 | 0.88 | 1.11 | 1.25 | 1.34 | 1.41 | 1.45 | 1.49 | 1.51 | 1.54 | 1.56 | 1.57 | 1.58 |

**Table 7.**

Consistency assessment

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **n** | **RI** | **CL** | **CR** | **Reliability** |
| **9.22** | 9 | 1.45 | 0.0275 | 0.0190 | Less than 0.1 |

*3..2.3. The integration of thematic layers to identify groundwater potential zones*

The GWPZ, a devoid-of-dimensions measure that represents possible groundwater zones within a given geographical area, was calculated using a weighted linear framework approach, as described by Venkatramanan et al. (2019) and Ozegin et al. (2023). The GWP zones in the study area were determined by a rigorous approach. This procedure incorporated details from nine various themes and used an overlay analysis mechanism within the GIS platform, as well as hybrid weights produced using AHP, as shown in **Eq. 7.**

GWPZ = 

i.e.,



GWPZ: groundwater potential zones,where the subscripts w and r reflect the weights and rank value of every variable, respectively. Every theme layer has a weight (Wi), and the rank of its subcategories is represented by Ri. The locations were divided into five categories predicated on their probable groundwater zones: very low, low, moderate, high, and very high.

*3.2.4.* *Sensitivity analysis (SA)*

Sensitivity analysis is the activity of analyzing the influence of alterations to input data or model variables on spatial outputs in GIS. SA is an important GIS tool for assessing the integrity and robustness of geographical frameworks since it can identify areas of ambiguity and potential shortcomings in analysis (Mukherjee and Singh, 2020; Ozegin and Ilugbo,2025). SA is usually carried out on indicator weights to assess the strength and reliability of a decision resolution after modifying the weights for a preset set of parameters among options and reconsidering each of the possible ranks. This enables an organized evaluation of the influence of alterations to criteria weights on different rank orders. The SA is applied in this study to assess the uniformity of the effects, and it serves as the foundation for a suitable appraisal of the GWP zone map. It also helps to comprehend the consequences of all of the parameters employed in GIS-AHP.

Map removal SA is an effective tool for determining the significance of various maps or sets of maps in spatial evaluation. The comparative significance of distinct components might be analyzed, and the quality of their geographical analysis improved by systematically removing different maps or groups of maps while contrasting the findings. The SA study was designed to determine the exactness of the prospective groundwater outcomes and to evaluate effectively all nine (9) criteria that quantify groundwater potential appropriateness. **Table 8** shows the formula adopted to determine the exactitude of the groundwater potential map.

**Table 8**

Sensitivity analysis equation

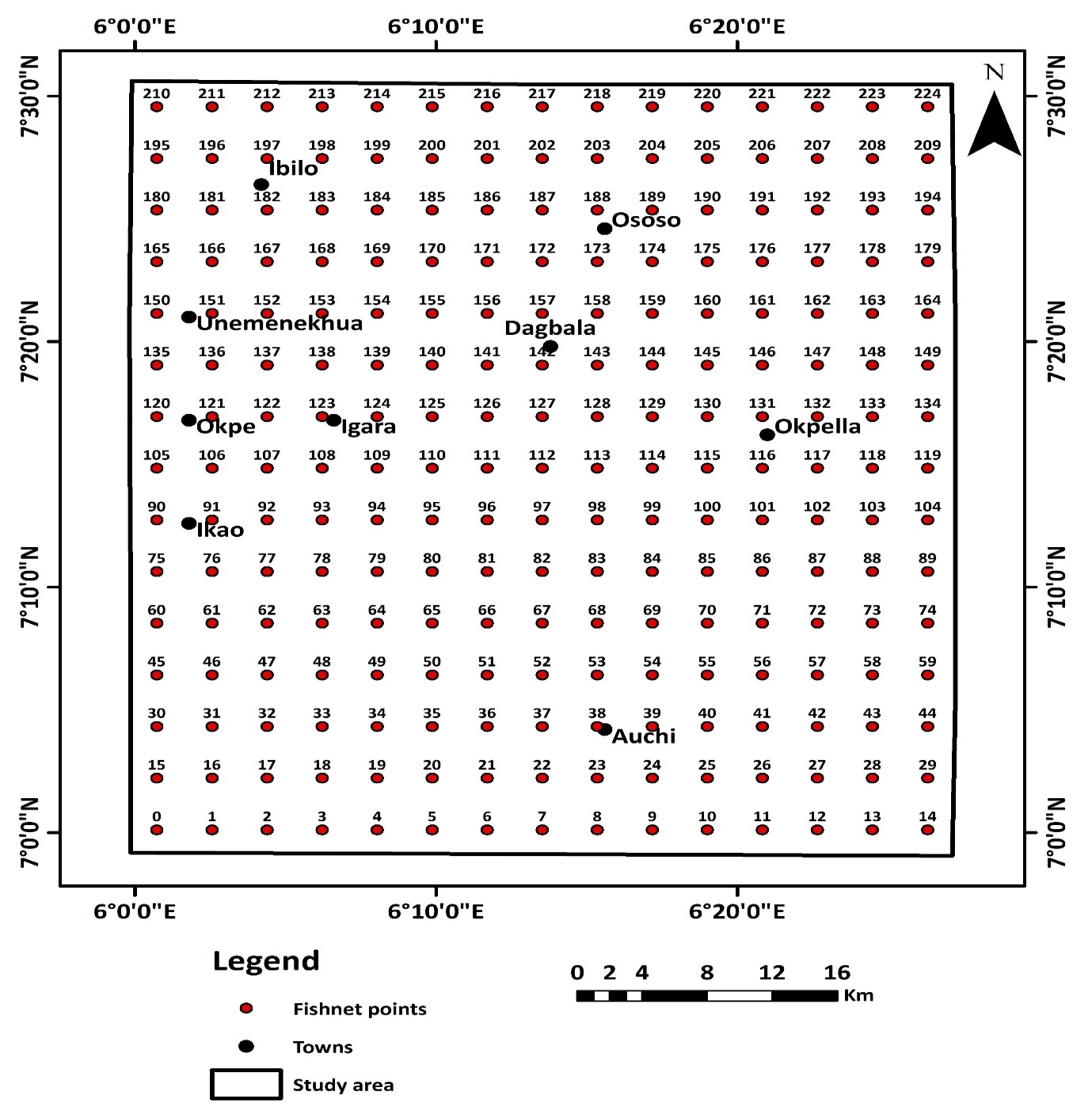
|  |  |  |
| --- | --- | --- |
| Sensitivity analysis | Formula | Description |
| Map removal |  | where   * N and n are the number of components considered in the creation of the GWPZ and GWPZʹ maps, respectively. * GWPZ denotes the groundwater potential zone created by combining all thematic layers. * GWPZʹ denotes the groundwater potential zones created by deleting one thematic layer. * SI represents sensitivity index. |

3.3. Data preparation for ML models

3.3.1. *Transforming parameterized values to points*

The number of training and testing/prediction datasets must be considered before separating the datasets for the training and testing task so as to use the ML models for the purpose of groundwater potential modelling (Lee et al., 2020). The main limitation in applying ML is the dearth of sufficient data for training (Hutchinson et al. 2017; Khan et al. 2021). This difficulty can be overcome by using the 'CreateFishnet\_management' before applying the 'ExtractValueToPoint' syntax in ArcGIS (e.g., Khan et al. 2021). A 3 × 3-meter grid (**Fig. 3**) was formed on all of the nine water raster layers (based on the nine-layer theme utilized in the study), and 224 grid points' raster values were retrieved. These grids are rectangular or square cells, resembling a fishnet (FN), hence the name. Fishnets are a type of grid that we use to overlay a feature on a map. It is used for organizing spatial data into manageable, comparable units. This segmentation allows users to aggregate data within each cell. To train the ML models, a new technique was used because the study area contained minimal borehole (BH) data (i.e., 36). This method involves training the models with 100% BH data (i.e., target indicator) and then utilizing the trained model for predicting 100% of the variables serving as predictors, which were designed to be completely hidden from the models during the training process. To implement this technique, raster values of the estimating variables were taken to the BH location points. The multiple classes designated BH data (i.e., low (0), moderate (1), and high (3)) were utilized for model training. Also, the raster values of indicators extracted to the 224 fishnet locations (**Fig. 3**) created using the fishnet development tool in ArcGIS 10.7 software were applied as the forecasting parameters. This approach provides a simple yet reliable mechanism for applying ML models to geospatially scarce target variable data locations.

A borehole yield map will be created to evaluate the classification evaluation criteria specified, and fishnet points will be retrieved from the drill yield data. This categorization was subsequently contrasted with the kind of classification projected by all of the ML models, yielding the AUC-ROC.



**Fig. 3.** The Fishnet developed for the research region to extract values from a raster layer.

3.3.2. Random forest (RF)

RF is a controlled ML algorithm that builds a collection of categorization trees derived by the randomized sampling of a cluster of features from the variable space and a bootstrap process that repeatedly picks a fraction of the input platform to adapt the model. It is an additional powerful and effective ML algorithm that was designed by Breiman (2001) as a further development of the “classification and regression trees” to increase its capacity for predicting competency (Razavi-Termeh et al., 2019; Maskooni et al., 2020). Diverse decision trees are generated using randomized bootstrapped calibrated sets (Breiman, 2001). The model then incorporates the mean results from every tree (Moghaddam et al., 2020). The individual performing the operation is required to choose two parameters, namely "the number of variables at each split" and "the number of trees" (Razavi-Termeh et al., 2019). This model does not employ the entire data set to create the tree; it only uses 66.66% of the bootstrap data. Subsequently, 33.33% of the other information is employed to assess the suited tree. In the present study, the framework was carried out in R using the 'randomForest' program (Liaw and Wiener, 2024). As a result of its dependable and efficient productivity, it is commonly used as a tree-based ML algorithm to study the complex link between groundwater-influencing features (Hasanuzzaman et al., 2022; Shandu and Atif, 2023). The model uses an identical generating approach as classification and regression trees (CART), but it creates multiple trees, resulting in a "forest" (Wiesmeier et al., 2011). It is capable of handling data from a variety of measurements without making mathematical suppositions (Rahmati et al., 2016). The RF model is famous because it is capable of addressing what is termed as the "black box," which frequently serves as a shortcoming of ML approaches, including ANN, and it is sturdy in dealing with outliers (e.g., Shandu and Atif, 2023). It provides explanations for connection and nonlinear correlations among components (e.g., Naghibi et al., 2016).

The covariance and variance that exist across grid cells can be determined by utilizing out-of-bag (OOB) error data (McKay and Harris, 2016; Naghibi et al., 2016). The bootstrap samples are then used to anticipate the missing OOB samples, and the mean square error (MSEOOB) is calculated by summing the OOB projections from the entire model trees (Wiesmeier et al., 2011; Prasad et al., 2020). **Eq. 8**, which integrates the mean square error of every model decision tree with its associated OOB samples for appraisal, was used to get the learning error Eoob. These features represent an estimation of ambiguity surrounding the assessment of prospective groundwater at a grid cell.

 (8)

The total number of OOB samples is denoted by "n," the obtained outcome by y*i*, and the model result from RF produced during training by (Prasad et al., 2020). The mean across the trees is the model's outcome (Rahmati et al., 2016).

3.3.3. Support vector machine (SVM)

SVM is a controlled ML algorithm that is implemented using the structural risk minimization (SRM) concept and theoretical framework of statistical learning (e.g., Tehrany et al., 2015; Kavzoglu et al., 2019). SVM converts the initially provided input field into a feature space with greater dimensions in order to locate the best segmenting hyperplane. Marjanović et al. (2011) confirmed that a hyperplane that separates is formed in the initially created space of n dimensions between vertices from two independent classifications. If the location lies above the hyperplane, it is designated as +1; otherwise, it is labelled as -1. It is often used to resolve classification and regression difficulties, reducing algorithmic excessive fitting (Gayen et al., 2019).

When addressing linearly separable data, **Eq. 9** is easily used to calculate a separating hyperplane (Hong et al., 2017):

 (9)

The coefficient vector w indicates the hyperplane's inclination in the characteristic space, whereas b is the hyperplane's distance from its starting point, and ζi determines the positive slack components. **Eqs. 10a** and **b**: The optimization problem is potentially addressed by establishing an ideal hyperplane (Samui, 2008).

 (10a)

 (10b)

In this case, C stands for the penalty and ai for the Lagrange multiplier. The choice of kernel format affects the SVM model's accuracy of efficient categorization (Yao et al., 2008). Many scholars (e.g., Tehrany et al., 2015; Gayen et al., 2019) have claimed that the radial basis function (RBF) has a stronger interpolation effectiveness, which is why it was used in this study. Eq. 11 typically defines how RBF can be determined:

 (11)

where the RBF kernel function is denoted by y and the kernel function by K(xi, xj). Using the technique was intended to lower the model's complexity and inaccuracy (Naghibi et al., 2018). The composite model emerged by merging two or more distinct prediction models.

3.3.4. Adaptive boosting (AdaBoost)

Adaptive boosting is an efficient composite learning approach that updates the weight of training scenarios to fully use a restricted number of instances for learning. After every round of weak learners, the AdaBoost method maintains a set of values and adjusts them to create poor learners within the training dataset (e.g., Freund and Schapire 1995, 1997). All weights get configured uniformly at the start. In each cycle, the weight of the appropriately categorized samples decreases, whereas the weight of the samples that were misidentified increases. However, this boosting technique does not require previous experience of the efficacy of the weak algorithm in operation (Jennifer, 2022).

The starting weights for every sample are assigned at the start of each step using the formula , where N is the number of evaluations. After training the weak discriminant ht with weights on the training collections, the weighted error (errt) and discriminant weight (at) are calculated (**Eqs**. **12a** and **b**).

 (12a)

 (12b)

In this instance, t is a number between 1 and T, while T is the total quantity of repetitions. y*i* represents the real classification label of input *i*, ht (xi) represents the assumption made by the *t*h weak classifier for the identical sample, and is the weight given to the sample at repetition t. In the case of an incorrect classification, where the real label is different from the label anticipated by the weak classifier, the indicating functionsequate to 1.

At every round t, the algorithm (AdaBoost) determines the weighted error and total weight of the weak classifier before updating the sample weighting. In order to focus the next iterations on cases that the existing ensembles of inadequate classifiers were incapable of sufficiently training, the modification aims to give the wrongly categorized samples bigger weights. The sample weight adjustment is expressed using **Eq. 13**:

 (13)

Integrating the weak classifiers yields the ultimate predictions for the binary category issue (**Eq. 14**): (14)

The projection made by the *t*h poor classifier given a sample is represented by h*t* (x) in this expression. The ultimate projection is guaranteed to be between +1 and -1, signifying the +ve or -ve categories, respectively, due to the sign () operator.

3.3.5. eXtreme gradient boosting (XGB) algorithm

The XGB technique is a sophisticated gradient boosting solution developed for machine learning assignments that require great rapidity and effectiveness. It implements visual illustrations to assess preference answers and expands upon decision tree approaches. By arbitrarily selecting attributes to create a forest of trees, XGB's bagging combination approach improves predictability by combining forecasts from several decision trees by dominant voting. The basic premise is simply to construct classification or regression trees one at a time, then train the next model using the remaining data from the prior tree. To improve the training operation, this approach can use values from already trained trees (e.g., Niazkar et al., 2024). By eliminating decision points that make minimal contributions to goal values, the pruning process minimizes the size of a decision tree and helps prevent overfitting. Even with values not present in the datasets, XGB performs more efficiently.

The XGB and AdaBoost approaches constitute part of the collaborative learning framework, which involves training weak learners sequentially to generate an effective model. In comparison to AdaBoost's concentration on deciding on a stump for weak learners, XGB takes a more adaptable strategy, regularly employing decision trees and incorporating normalization algorithms to thwart excessive fitting (**Eq. 15**).

 (15)

In this instance, every weak learner's aggregate input results in the ultimate projection F(X), which is represented as f*m*(X), where *M* is the entire set of weak learners. By deliberately training a single tree to minimize the diminution of its predecessors, a method of adaptation is created that raises the accuracy of the entire model. The incorporation of regularity elements into the objective function is a crucial element in XGB's effectiveness. Regularity in XGB is accomplished by the use of penalty features in the objective function, which is composed of two fundamental parts: the term that regulates and prohibits overwhelmingly complicated models and the loss term that evaluates how well the model matches the training data. **Eq. 16** provides the following expression for the objective function:

 (16)

represents the regularization term given to the *m*th tree, and  represents the loss function, which quantifies the difference between the true label () and the anticipated label ().

3.4. Model performance evaluation

The verification of every model is an essential phase in empirical research (Naghibi et al., 2016; Ozegin et al., 2024a, b, c). Evidently, there are numerous methods for determining the performance of MLAs, notably the ROC curve, considered an accurate depiction of how effectively the algorithms perform, particularly in binary categorization tasks (Sachdeva and Kumar, 2021; Singha et al., 2024; Rana et al., 2025). The graph plots a classifier's accuracy leveraging the true positive rate (TPR) vs. the false positive rate (FPR) at various threshold settings. The horizontal represents TPR, often known as sensitivity or recollection (**Eq. 17a**). It refers to the ratio of the number of reliably anticipated observed positives to the entire number of occurrences in the real class. Nevertheless, FPR, calculated as (1—specificity) and plotted on the horizontal axis (**Eq. 17b**), estimates the fraction of incorrectly expected positive findings compared with the total number of real negative evaluations (Prasad et al., 2020; Saha et al., 2022). TN depicts true negative, FP means false positive, TP means true positive, and FN means false negative (Bai et al., 2022; Braham et al., 2022).

 (17a)

 (18b)

The receiver operating characteristic (ROC) curve was used to analyze the GWP map's performance across various techniques. The area under the receiver-operating characteristic (AU-ROC) curve is often employed for assessing model prediction capacity, with a greater AUC indicating a superior model (Rahmati and Melesse, 2016; Echogdali et al., 2022). The area under the receiver-operating characteristic (AU-ROC) curve, often known as the area under the curve (AUC), was calculated to assess the reliability of predictions.

According to Senapati and Das (2021) and Ozegin et al. (2024b), the AUC values varied from 0 to 1 (**Table 9**). While a number of 0 denotes no difference between the likelihood of groundwater (weak relationship) and the available data, a value of 1 denotes the highest degree of accuracy (e.g., Masroor et al., 2023; Ali et al., 2023; Ozegin et al., 2024b; Sharma et al., 2024). A high ROC value suggests that the model is very effective.

**Table 9.**

AUC values

|  |  |  |
| --- | --- | --- |
| S/No. | Range | Description |
| 1. | 0.50–0.60 | poor |
| 2. | 0.61–0.70 | average |
| 3. | 0.71–0.80 | good |
| 4. | 0.81–0.90 | very good |
| 5. | 0.91–1.00 | excellent |

**4. Results and discussion**

The study of geospatial evaluation of geological features is critical to determining where and what dimensions of things are, as well as how they connect to one another. This study is critical for the incorporation of field data and the generation of various conceptual layers regarding the environment using GIS platforms and remote sensing. These data sets are merged and leveraged to adequately identify the GWP zone and aid in long-term water asset development, planning, and administration. The choice of decisive variables is determined by the anthropological, hydrogeological, hydrological, and topographical characteristics, as well as the accessibility of appropriate data for the study area. As a result, in this study, the GWP zone map is generated from nine (9) thematic maps: normalized difference vegetation index (NDVI), geology (GY), lineament density (LD), rainfall distribution (RD), proximity to surface water bodies (PSW), aspect (AP), drainage density (DD), slope (SP), and topographic wetness index (TWI). These maps and their particular attributes are extremely important since they indicate the geographical effect on GW distribution. The study aims to estimate GWP zones in the Edo North district using AHP and machine learning (ML) methods such as RF, SVM, AdaBoost, and XGB. The subsequent subsections provide insights on each concept map related to freshwater potential for the present work.

4.1.Variables influencing groundwater multicollinearity evaluation

Seven (7) of the sixteen (16) groundwater-predicting indicators considered in the study exhibit multicollinearity (**Table 10a**). The statistics identified seven influencing factors with collinearity issues: land use/land cover (LL), soil type (ST), geomorphology (GM), elevation (EL), ruggedness index (TRI), soil permeability (SB), and stream power index (SPI). The VIF and tolerance metrics for the aforementioned variables are larger than 10 and less than 0.1, respectively, and hence were rejected for the development of the various models utilized in the study. And the following parameters (as demonstrated by T and VIF values > 0.1 and < 10, respectively) were subsequently used for the modelling (**Table 10b),** viz., normalized difference vegetation index (NDVI), geology (GY), lineament density (LD), rainfall distribution (RD), proximity to surface water bodies (PSW), aspect (AP), drainage density (DD), slope (SP), and topographic wetness index (TWI).

The statistical results show that RD possesses the least T value (0.1496) and the largest VIF value (6.6860). Conversely, LD has the greatest T value of 0.6608 and the smallest VIF score of 1.5132. In the other groundwater-influencing parameters, the values of T and VIF were observed to lie between RD and LD, as shown in **Table 10b**. Consequently, the outcomes demonstrated that there was indeed no evidence of MC among the chosen indicators, and no ambiguity was infused into the model's outcomes due to the issue of multicollinearity. Therefore, the assessments show that each of the parameters chosen had an effect on the GWP zones; consequently, all of these criteria (see **Table 10b**) were incorporated in the modelling process.

**Table 10a.**

MC statistics (T and VIF) of the groundwater potential-influencing determinants

|  |  |  |  |
| --- | --- | --- | --- |
| S/NO | Criteria | VIF | T |
| 1. | ***LL*** | ***16.4885*** | ***0.0606*** |
| 2. | NDVI | 2.1910 | 0.4564 |
| 3. | GY | 2.2275 | 0.4489 |
| 4. | ***ST*** | ***13.214*** | ***0.0712*** |
| 5. | LD | 1.5132 | 0.6608 |
| 6. | RD | 6.6860 | 0.1496 |
| 7. | PSW | 2.2299 | 0.4485 |
| 8. | SP | 1.6540 | 0.6046 |
| 9. | ***GM*** | ***11.375*** | ***0.0787*** |
| 10. | DD | 2.6276 | 0.3806 |
| 11. | ***EL*** | ***infinity*** | ***infinity*** |
| 12. | TWI | 3.1316 | 0.3193 |
| 13. | ***TRI*** | ***14.636*** | ***0.0699*** |
| 14. | ***SB*** | ***infinity*** | ***infinity*** |
| 15. | ***SPI*** | ***infinity*** | ***infinity*** |
| 16. | AP | 1.9128 | 0.5228 |

**Table 10b**

MC statistics (T and VIF) used for the modelling

|  |  |  |  |
| --- | --- | --- | --- |
| S/NO | Criteria | VIF (< 10) | T (> 0.1) |
| 1. | NDVI | 2.1910 | 0.4564 |
| 2. | GY | 2.2275 | 0.4489 |
| 3. | LD | 1.5132 | 0.6608 |
| 4. | RD | 6.6860 | 0.1496 |
| 5. | PSW | 2.2299 | 0.4485 |
| 6. | SP | 1.6540 | 0.6046 |
| 7. | DD | 2.6276 | 0.3806 |
| 8. | TWI | 3.1316 | 0.3193 |
| 9. | AP | 1.9128 | 0.5228 |

***land use/land cover (LL)****,* normalized difference vegetation index (NDVI), geology (GY), ***soil type (ST)****,* lineament density (LD), rainfall distribution (RD), proximity to surface water bodies (PSW), slope (SP), ***geomorphology (GM)****,* drainage density (DD), ***elevation (EL)***, topographic wetness index (TWI) Topography ***ruggedness index (TRI)***, ***soil permeability (SB)***, ***stream power index (SPI)***, and aspect (AP).

4.2. Sensitivity analysis

In an MCDA, sensitivity analysis is required to ensure that the result is dependable notwithstanding the variable nature of expert assessments. **Table 11** summarizes the overall implication of each theme layer to the GWP zone modelling map. Regardless of the variations in mean percentage change (MPC), removing a thematic layer has a substantial effect on the final map. Thus, each determining component included in AHP analysis has a unique function in defining the GWP zones. In this analysis, the most significant sensitivity index is obtained by deleting the geological layer with the highest comparative score of 21.45%. Furthermore, groundwater is slightly susceptible to LD, LD, DD, TWI, PSW, NDVI, AP, and SP, with MPC values of 21.045, 12.083, 9.5126, 5.6305, 4.6139, 3.9813, 2.6305, and 1.4579%, respectively (**Fig. 4**).

**Table 11**

Statistical study of map removal SA.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| S/NO. | Removed Factor | Mean Abs. Change | Max. Abs. change | Min. Abs. change | Std. Abs. change | Mean %\_change |
| 1. | NDVI | 0.0172 | 0.0816 | 0.0000 | 0.0147 | 5.6305 |
| 2. | GY | 0.0358 | 0.1225 | 0.0000 | 0.0293 | 12.0830 |
| 3. | LD | 0.0137 | 0.0513 | 0.0000 | 0.0101 | 4.6139 |
| 4. | RD | 0.0647 | 0.1467 | 0.0021 | 0.0286 | 21.0451 |
| 5. | PSW | 0.0639 | 0.1845 | 0.0000 | 0.0420 | 21.4592 |
| 6. | SP | 0.0254 | 0.0825 | 0.0003 | 0.0187 | 9.5126 |
| 7. | DD | 0.0110 | 0.0332 | 0.0001 | 0.0086 | 3.9813 |
| 8. | TWI | 0.0069 | 0.0205 | 0.0000 | 0.0050 | 2.6305 |
| 9. | AP | 0.0045 | 0.0163 | 0.0001 | 0.0031 | 1.4579 |



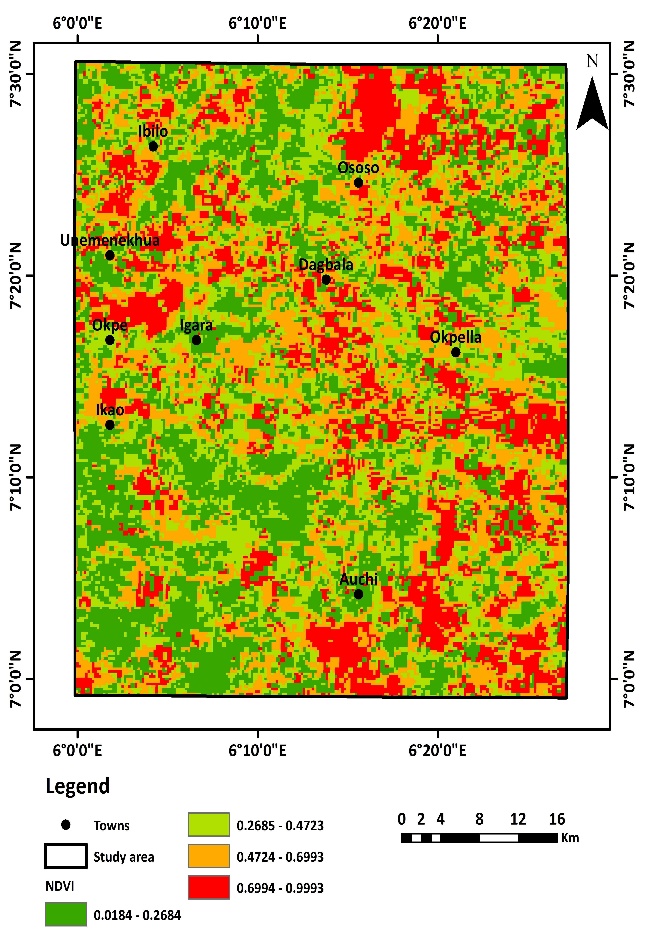
Fig. 4: A bar graph illustrating the mean percentage of the map removal analysis's AHP score

4.3. Analysis of thematic layers

In general, GW production for a given aquifer is determined by a number of factors, including anthropology, hydrogeology, hydrology, and topography, all of which are dependent on data that is readily available and of good quality. RS data are often large in size and contain detailed geographical information. The following characteristics provide a quick overview of the theme layers employed in the study.

4.3.1. *Anthropological (NDVI) characteristics and groundwater potential*

**NDVI:** The NDVI is a useful metric for measuring vegetative component activity in ecosystem patterns, while the productiveness of vegetative coverings pertains to water loss and precipitation (e.g., Hussain et al., 2022; Ozegin and Ilugbo, 2025). It is a broadly utilized measurement that monitors changing vegetation on both a global and regional basis. Consequently, it has profound effects with regard to groundwater accumulation and accessibility. Tucker (1979) developed this metric, which spans from -1 to 1. Numbers below 0 indicate no vegetative cover, while values above 0 indicate accessible cover. In this study, areas' NDVI was classified into four groups using themed mapping: 0.0184-0.02684, occupying 32%; 0.02685-0.4723, occupying 25%; 0.4724-0.6993, occupying 25%; and 0.6994-0.9993, occupying 18% (**Fig. 5** and **Table 5**). An abundance of flourishing greenery in a specific location emerges to be strongly associated with optimal groundwater recharge—vegetation designations range from 0.2 to +1 (Goward et al., 1991; Ozegin and Ilugbo, 2025). Consequently, the higher the score, the more significance is assigned, while minimal vegetation and grasses receive only a small percentage of weight. A greater NDVI implies extensive vegetation cover, given vegetation decreases runoff (e.g., rainwater) and helps to recharge groundwater reservoirs (Hasan­uzzaman et al. 2022). The volume of precipitation, temperature, soil utilization, vegetation moisture content, soil moisture, and evaporation are all controlled by groundwater depth.



**Fig. 5.** Anthropological factor—normalized difference vegetation index

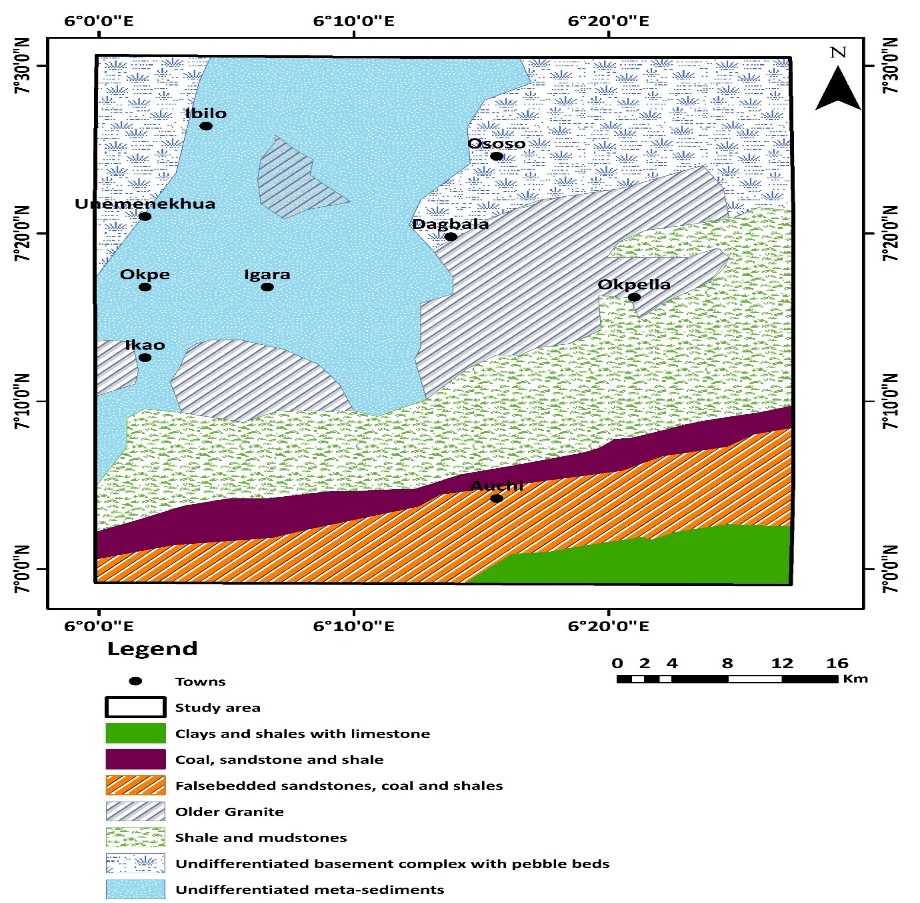
4.3.2. *Hydrogeological (LD and GY) characteristics and groundwater potential*

**LD:** Any linear geological characteristic, including faults and fractures, that develops from the pressing and shearing forces connecting geologic formations is called a lineament. It is a geological framework that can be used as an effective guide when studying the hydrogeology of a basement. As a result, lineaments are capable of showing surface structure, including fractures, fissures, and defects, among others. The arrangement of structural elements that produce secondary porosity in the rocks determines hydrogeological processes and the presence of groundwater in the rocky landscape (e.g., Al-Djazouli et al., 2020). Lineaments affect subsurface passage of water and spread to optimize flowing paths. Groundwater activity can be better understood by methodically examining the concentration, direction, and intersection patterns within lineaments (e.g., Ilugbo et al., 2023a, b).

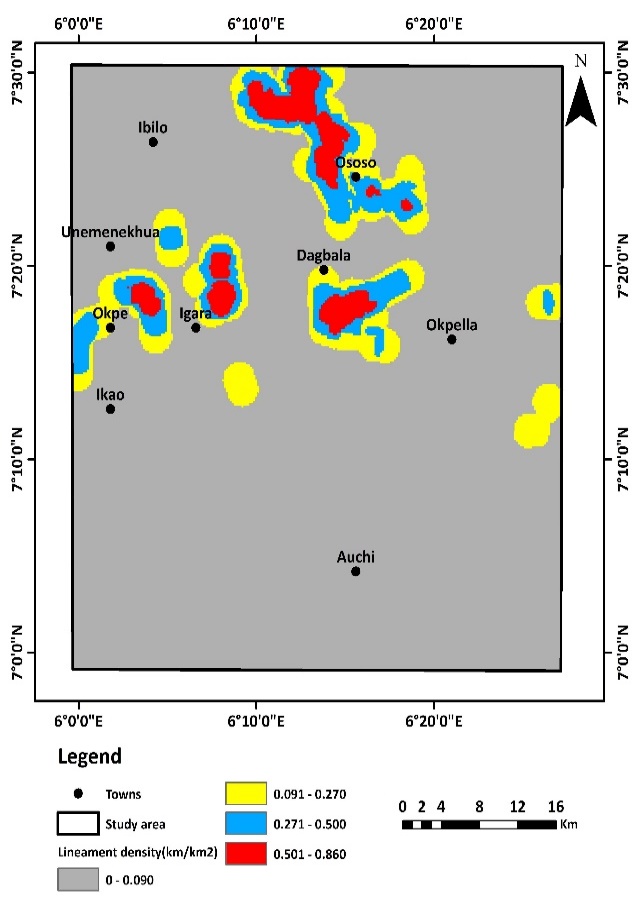
The lineament map for the investigation region is shown in **Fig. 6a.** The northern section of this research region is notable for fairly substantial lineament density. Higher lineament densities are indicative of optimized porosity, facilitating groundwater supplies. Areas with a substantial lineament density are considered to have good prospects for groundwater and receive a greater ranking based on good pore size and permeation, while areas with low to moderate lineament density are considered to have poor groundwater possibilities and are given lower rankings.

**GY:** Geology constitutes one of the most important hydrogeological criteria since it indicates the structural characteristics of the geographic rocks that shape the accumulation of groundwater. The properties of the major lithological development in the geographic region worthwhile have a significant impact on the pattern of distribution, happenstance, and condition of groundwater. GY completely controls groundwater migration and absorption since the number of pores and permeability in aquifer rocks are inherent features (Ilugbo et al., 2023b). The hydraulic conductivity within a lithologic structure or its weathered consequence increases the pace of water penetration and flow through it, as well as its capacity for storing underground water (e.g., Rajendran et al., 2020). The explored location's groundwater availability is largely determined by the hydraulic characteristics of the crystalline basement rocks; therefore, older granites are capable of simply behaving as groundwater storage facilities when they are weakened and/or fragmented.

In this study, geological structures have changed from the undifferentiated basement complex and undifferentiated meta-sediments in the northern section of this research location, whose components are interspersed with granite gneiss, shale, and mudstones (**Fig. 6b**), to the distinguishing sedimentary composition formation in the southern part, which is composed primarily of clays and shales alongside limestone, coal, sandstone, and shale; false-bedded sandstones; coal; and shales. The mild slope decreases in the direction of the southern regions of the research zone, and the deposition of the basement rocks found in the northern mountainous terrain could potentially be the cause of the phenomenon. The research area's lithologic components were classified as 1 (false-bedded sandstone, coal, and shale), 2 (coal, sandstone, and shale), 3 (clays, shales with limestone, Older granite, and undifferentiated basement with complex pebbles), and 4 (shale and mudstones and undifferentiated meta-sediments), which indicate very low, low, moderate, and high groundwater potentialities, respectively (**Table 5**). These covered 505 km² (21%), 529 km² (22%), 739 km² (21%), and 265 km² (11%), respectively. The lithology rated 3 and 4 has substantial consequences on groundwater spread and accessibility (Rahmati et al., 2016; Ozegin et al., 2023; 2024b). Every lithologic category was assigned a weight based on its relevance for groundwater availability and hydrodynamic qualities. Groundwater phenomena are determined by the interwoven interplay of lithological properties, usage of land, terrain, and watershed environment.



(6a)



(6b)

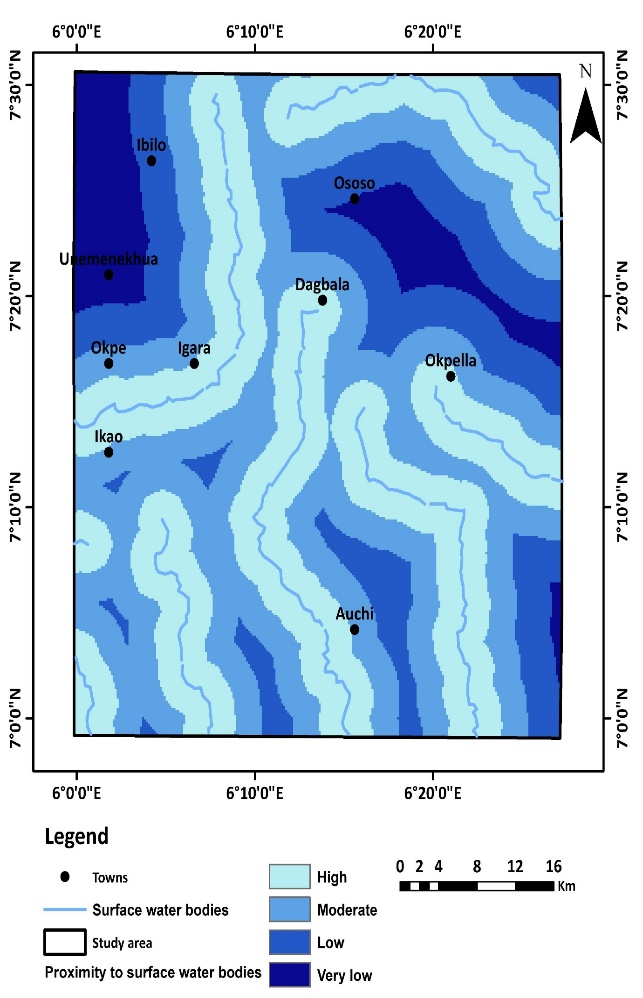
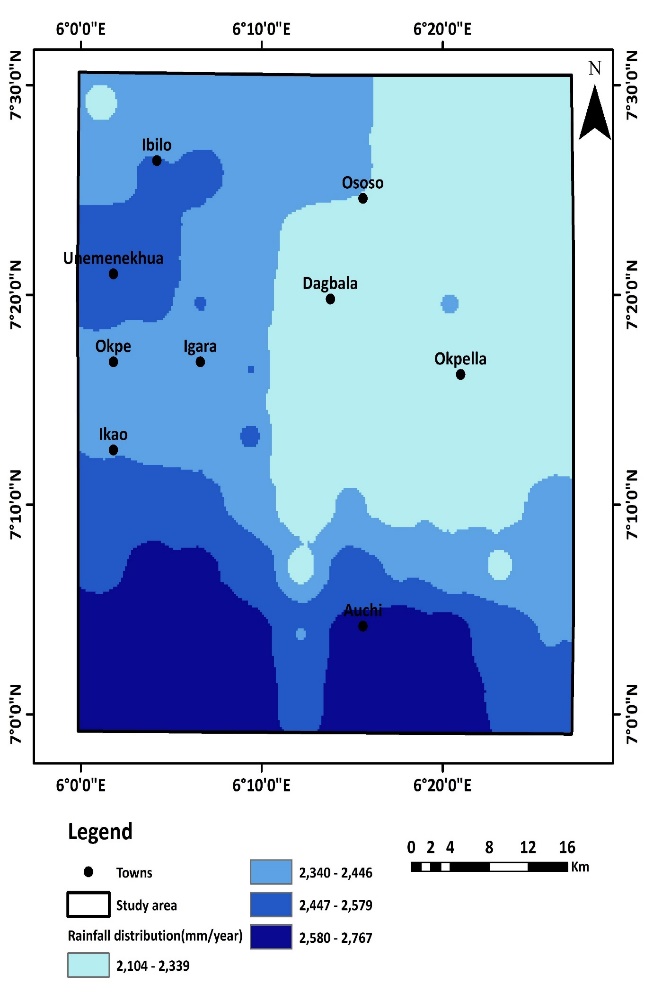
**Fig. 6.** Hydrogeological factors—a. lineament density and b. geology

4.3.3. *Hydrological (RD and PSW) characteristics and groundwater potential*

**RD:** The hydrology and hydrogeological functions of an area are greatly influenced by both the spatial and time course of rainfall (Ozegin et al., 2024a, b). It is the main and potentially important mechanism of groundwater restoration. With yearly projections varying from 2,104 mm to 2,767 mm, rainfall is the primary contributor to groundwater replenishment in the research location (**Fig. 7a**). The research region is separated into four rainy categories based on the amount of rainfall obtained: 34% of the total area is classified as very low, 32% as low, 17% as moderate, and 17% as high (**Table 5**). The amount, length, and extent of rainwater all affect the pace of penetration and subsequent discharge. Low absorption and high surface discharge are impacted by prolonged low precipitation. The annual mean rainfall has an impact on groundwater recharge as well. While less precipitation indicates a low groundwater possibility, more rainfall indicates higher groundwater prospects. There were considerable amounts of rain in the southwest area. When defining groundwater zones, heavier precipitation zones are given more weight, and vice versa.

**PSW:** Groundwater replenishment processes are directly impacted by how far it is to surface bodies of water, which indicates the nearness of every location in the research region to the closest body of water (e.g., Arabameri et al. 2021; Ozegin et al., 2024a). It is an essential variable in the mechanisms of surface-groundwater and aquifer recharge prospects. Those for influent flow facilitate groundwater runoff, while those for effluent flow maximize groundwater recharging. When compared with places more remote, groundwater potential is typically higher near water sources (rivers or streams) because they serve as a source of a refill for the nearby aquifers, especially during times of excessive flow (e.g., Maity and Mandal, 2019; Naghibi et al., 2020; Mathewos et al., 2024). As a result, groundwater prospect delineation is correlated with the proximity to water bodies.

The buffered areas nearest to water bodies received the maximum score of 4, while those situated most remotely received a lesser score of 1, signifying the greatest and smallest prospects for replenishment of groundwater, respectively (**Table 5**). The flow condition in the research area can be regarded as effluent. The buffer regions are categorized into four categories: "high (888 km), moderate (872 km), low (330 km), and very low (300 km)" (**Fig. 7b)**. Inevitably, when proximity to water bodies (such as streams) decreases, the weight value rises, indicating enhanced groundwater possibilities (e.g., Arabameri et al., 2021; Ozegin et al., 2024a).



(7b)

(7a)

**Fig. 7.** Hydrological factors—a. rainfall distribution and b. proximity to surface water bodies

4.3.4. *Topographical factors (SP, GM, DD, TWI, and AP) characteristics and groundwater potential*

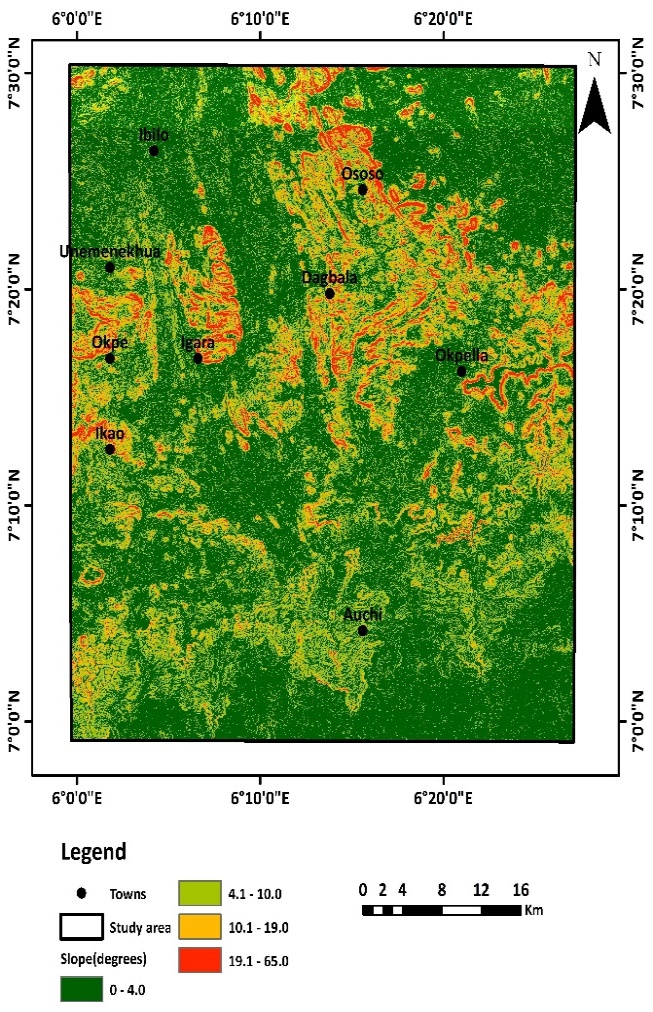
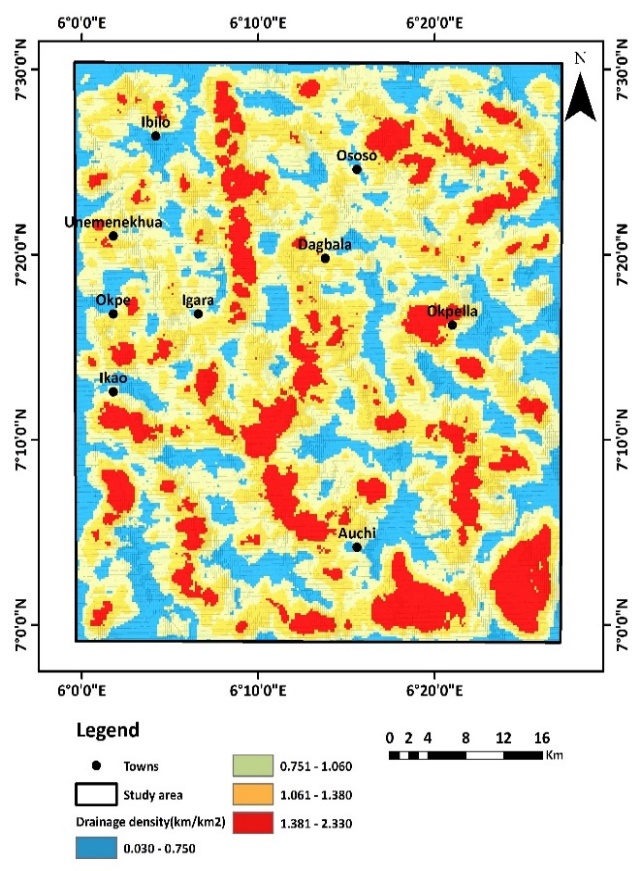
**SP:** Typically, SP specifies the vertical extent of every geologic framework across the earth and consequently plays a decisive role in influencing the gravity-driven passage of water (Chen et al., 2019; Shahinuzzaman et al., 2021; Ilugbo et al., 2023). It provides vital information about regionally focused geologic and geodynamic phenomena. The SP of an area, usually determined by elevation, entails the sharp descent from a lateral surface. Slope influences the passage of water based on gravitational pull, revealing the speed of subterranean longitudinal transmittance. As the SP rises, the water makeup of the soil becomes exceedingly hard to maintain, as water runs off faster and gets less soaked by the earth's pores.

During a rain, rainwater rapidly sweeps along steep slopes, resulting in reduced recharge at greater slope inclinations. The research area map's slope is displayed in **Fig. 8a**. Four categories—very low (19.1–65.0°), low (10.1–19.0°), moderate (4.1–10.0°), and high (0–4.0°)—were created from the reclassification of the slope values. A more significant weight is assigned to mild and even slopes, while less weight is assigned to steep and exceedingly steep slopes (**Table 5**). 25% (minimum area) of the study region consists principally of 0-4.0° values rated high (4). Consequently, the minimal area has a level slope that retains groundwater. The aquifer recharging decreases as the slope gets steeper. It occurs because rainfall creates an intense rush of water downward on the steep gradient, resulting in minimal absorption. Table 5 shows that slope inclination is inversely correlated to surface runoff permeation (Wang et al., 2018; Morbidelli et al., 2018; Ozegin and Ilugbo, 2024; Zheng et al., 2024).

**DD:**Drainage density, which represents the number of networks moving surface water, is regarded as one of the primary signals of potential groundwater reserves in a given location. DD increases groundwater outflow while lowering recharge potentials. As a result, land with a compact watercourse exhibits a rapid recharge rate, but the contrary is similarly applicable (Roy et al., 2019). Four drainage densities were identified for the groundwater potential in the research region based on the drainage map (**Fig. 8b**): high (0.0300-0.750 km/km²), moderate (0.751-1.060 km/km²), low (1.061-1.380 km/km²), and extremely low (1.381-2.330 km/km²). 1,261 km² (53%) of the research region had high and moderately substantial groundwater prospects (**Table 5**). The DD of the hydrographic arrangement indicates various physical conditions, including the proportion of surface and subsurface movements. This threshold, which favours slope drainage, allows us to appreciate the significance of surface drainage. This signifies the water-striking system's mean length per kilometre. Thus, increased DD reduces the likelihood of absorption and recharging of groundwater. Determinants like geology, penetration, runoff, waters, plant diversity, and climate have an immediate effect on DD (e.g., Al-Djazouli et al., 2020).

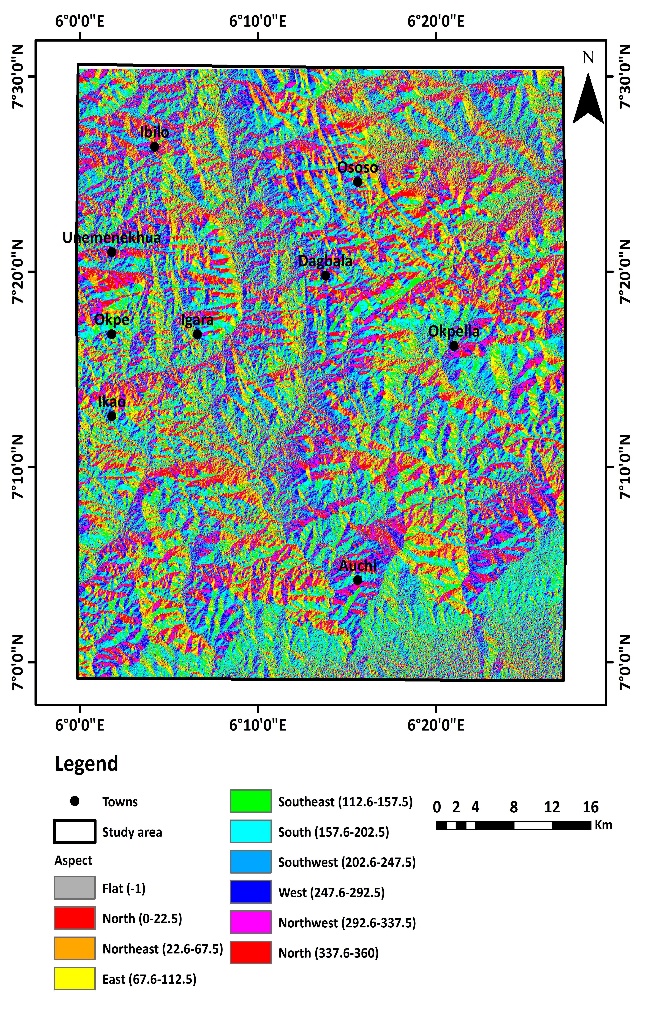
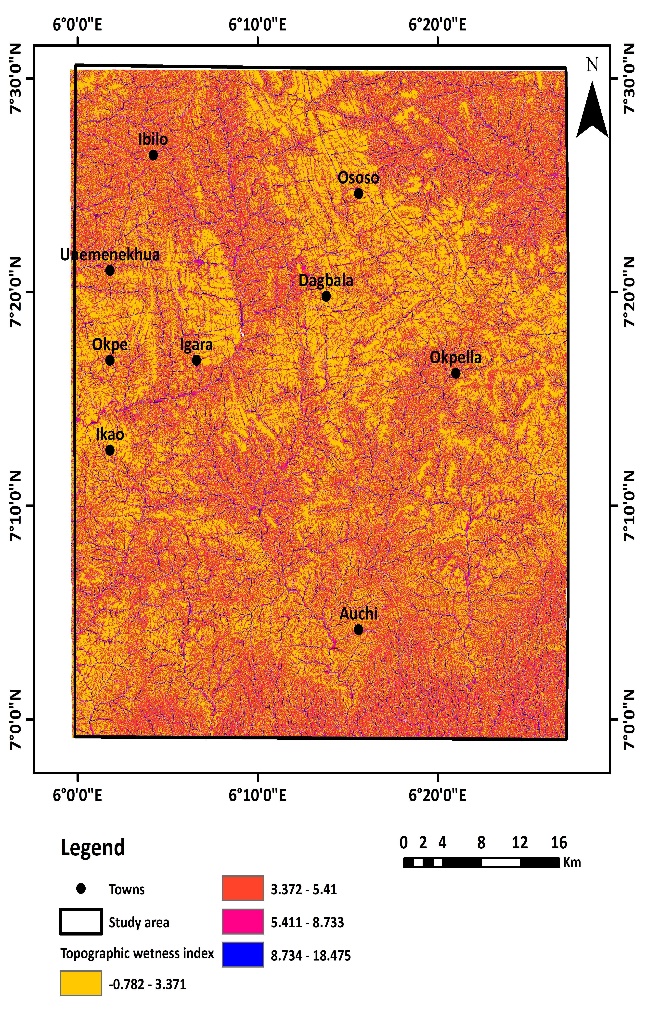
**TWI:** The TWI is a frequently used metric to assess the influence of geographical features on watershed dynamics (Sørensen et al., 2006; Ozegin and Ilugbo, 2025). It is an indicator that describes the amount of hydration or saturated states in a given region. In the course of an intense downpour occurrence, areas exhibiting high TWI values retain greater amounts of water than other areas, resulting in a saturated state (Ozegin and Ilugbo, 2025). **Fig. 8c** depicts the TWI map for the study region, which gives an illustration of the geographical spread of TWI values and the consequences for GWP across the entire study area. Recognizing that hillslope impact, the TWI aids in finding locations with characteristics that are more probable to create floods than to encourage replenishment of groundwater. Low (-0.782 and 3.371), moderate (3.372 and 4.410), high (4.411 and 8.733), and very high (8.733 and 18.475) constitute the four categories into which the TWI metrics in this research are reclassified. Lower weighting is assigned to locations with lower TWI ratings in the distribution of GWP, indicating a limited capability for groundwater recharging. Locations that have larger TWI values, which indicate suitable replenishment conditions, are given greater significance (e.g., Zhao et al., 2024). This metric takes into account simultaneously the inclination and the upstream influencing location per unit width perpendicular to the course orientation. It provides useful details about the hydrology dynamics in a given locale. There exists a positive correlation since the highest TWI value indicates a high subsurface water prospect and vice versa.

**AP:** Aspect is the slope's foremost orientation and defines the perspective of the watercourse (Razavi-Termeh et al., 2019). This regulates slope development, including lineament, precipitation, wind consequences, and direct sunlight (Solomon and Quiel, 2006; Zabihi et al., 2016). It is commonly employed in mountainous and rugged places, considering the period of sunlight or obscurity has an important influence in determining soil wetness (Sinha et al., 2012). The aspect additionally influences the formation of runoff through the development of vegetation and GW enhancement (Ahmed and Sajjad, 2018). This component is categorized into ten groups, as shown in **Fig. 8d**. Evidently, the area of a shady slope with significant soil humidity has an abundance of flow. For flat-facing sides to the north, northeast, northwest, and north, aspect-influencing feature classifications are predominant, suggesting a high probability of groundwater prospects. Conversely, the other aspect classes have the lowest scores (**Table 5**), which suggests that there is minimal likelihood of groundwater prospects.

(8b)

(8a)



(8d)

(8c)

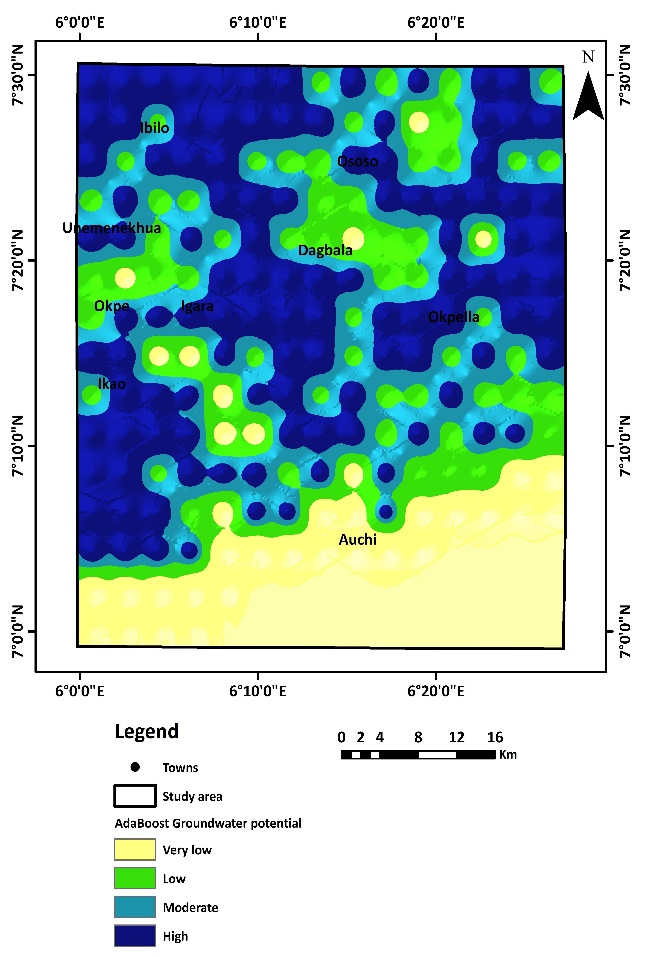
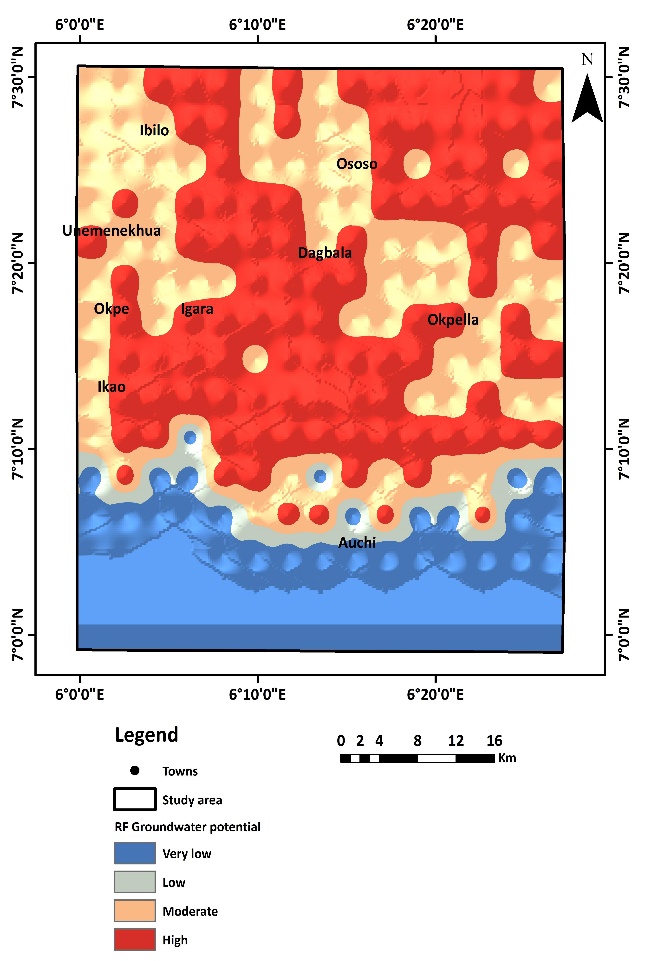
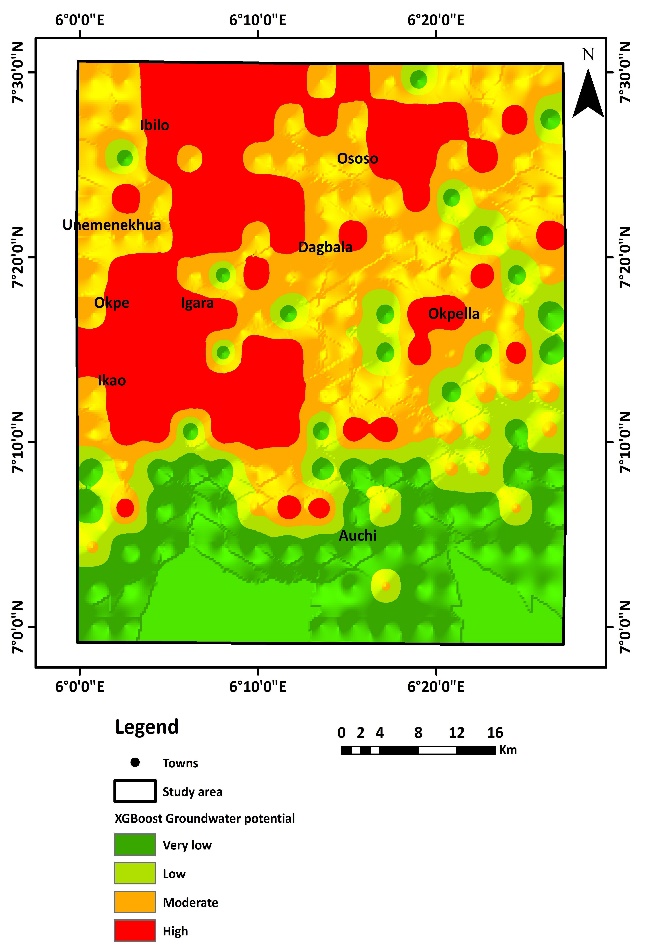
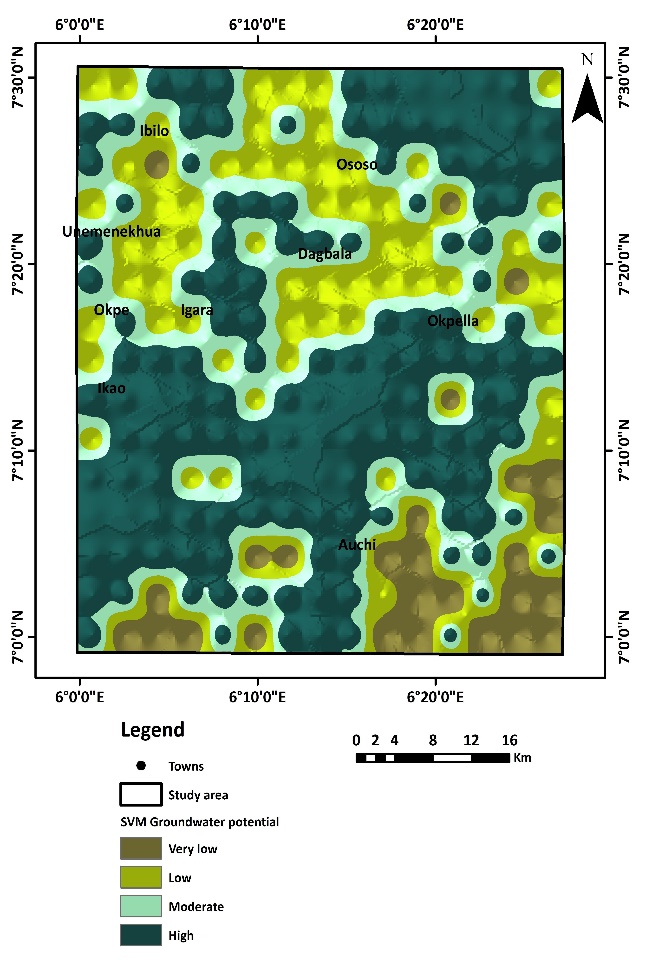
**Fig. 8.** Topographical factors—a. slope; b. drainage density; d. aspect; and e. topographic wetness index

4.4. GW potentiality modelling

Long-term development requires proper GWP zone characterization. Given that groundwater is an inherent asset, its amount has been decreasing due to a variety of anthropological and natural phenomena. Groundwater management plays a critical role in regional planning and long-term sustainability. As an effect, the GWP zones in the study area have been established based on the research data. This work prioritizes projecting GWP zones using AHP-based MCDA, and ML approaches include AdaBoost, RF, SVM, and XGB. These methods were used to examine groundwater geoenvironmental factors, which included topographical, hydrological, geological, and anthropogenic features critical for determining groundwater distribution. **Fig. 9** depicts the geographical spread of GWP zones calculated using AHP-based MCDA and four contemporary ML algorithms: AdaBoost, RF, SVM, and XGB. ‘’Jenks' natural breaks classifier’’ in a GIS context was leveraged to classify each of the GWP maps into four categories (high, moderate, low, and extremely low). A greater GWPZ rating implies a significant quantity of groundwater potential, whereas a lower GWPZ index figure suggests a smaller quantity of water beneath the ground surface.

The AHP model demonstrates exclusively that 14.03% of the entire watershed area has "high" groundwater potential, with a further 32.30% rated to exhibit "moderate" potential. The GWP zone ratings, attributed to the Random Forest results, indicate areas covered of 23.50, 4.00, 30.83, and 41.59% for very low, low, moderate, and high potential, respectively. According to the SVM model, there is "high" groundwater potential in 45.93% of the basin's total area and "moderate" potential in a further 21.54%. Based on the AdaBoost model, 36.42% of the study's entirety has "high" groundwater potential, and a further 23.29% has "moderate" potential. In contrast, 59.23% of the area coverage in the XGB model is made up of the high and moderate potential categories (**Fig. 10**). For GWP zones with descending high prospect values, the SVM, RF, AdaBoost, XGB, and AHP modes generally show a spatial extent of 45.93, 41.59, 36.42, 25.30, and 14.03%, respectively. Moreover, it is evident that there is considerable potential for underground water in parts of the study location with high rainfall, lineament density values, and maximum geology rating. On the other hand, high slope and PSB values are indicative of regions with very low to low groundwater potential. According to the models (e.g., SVM, RF, and AdaBoost), the northern portion of the research zone has a larger potential for GW and should be explored for GW development. Conversely, the study demonstrates that the southern section of the research area has low GW potential, which is attributed to longer distances to wetlands and lower lineament density. The SVM model estimates that the high GWP zone occupies approximately 1098 km², or 45.93% of the overall research area (**Fig. 10** and **Table 12**). The existence of high lineament density indicates that there is substantial hydro-potential development; however, the entire lineament density in the evaluated terrain is low (Ozegin et al., 2023). Research indicates that diverse morphologies on earth's surface led to distinct types of freshwater accumulation (Rajaveni et al., 2017; Prasad et al., 2020). The differences in the availability of GWP zones resulting from the modelling approaches are highlighted by the geographical contrast shown in **Figs. 9a-e**. SVM and RF performed better than other models in establishing GWP zones, which can be utilized to evaluate regulating land use and water supply.

The combination of powerful ML algorithms and AHP with an extensive repertoire of contextual variables resulted in nuanced and remarkably accurate approaches to estimating GW potential zones in the research area. The aforementioned technique can considerably improve long-lasting GW planning and oversight. The various influencing features gave an in-depth understanding of the groundwater context, including both natural and human-induced aspects. The study highlights the effectiveness of ML approaches in environmental and governance of resources, laying the groundwork for subsequent studies in related areas. Prospective studies might extend the framework by including contemporaneous data and variable land-use patterns to improve the models' predictability and application for effective watershed management.

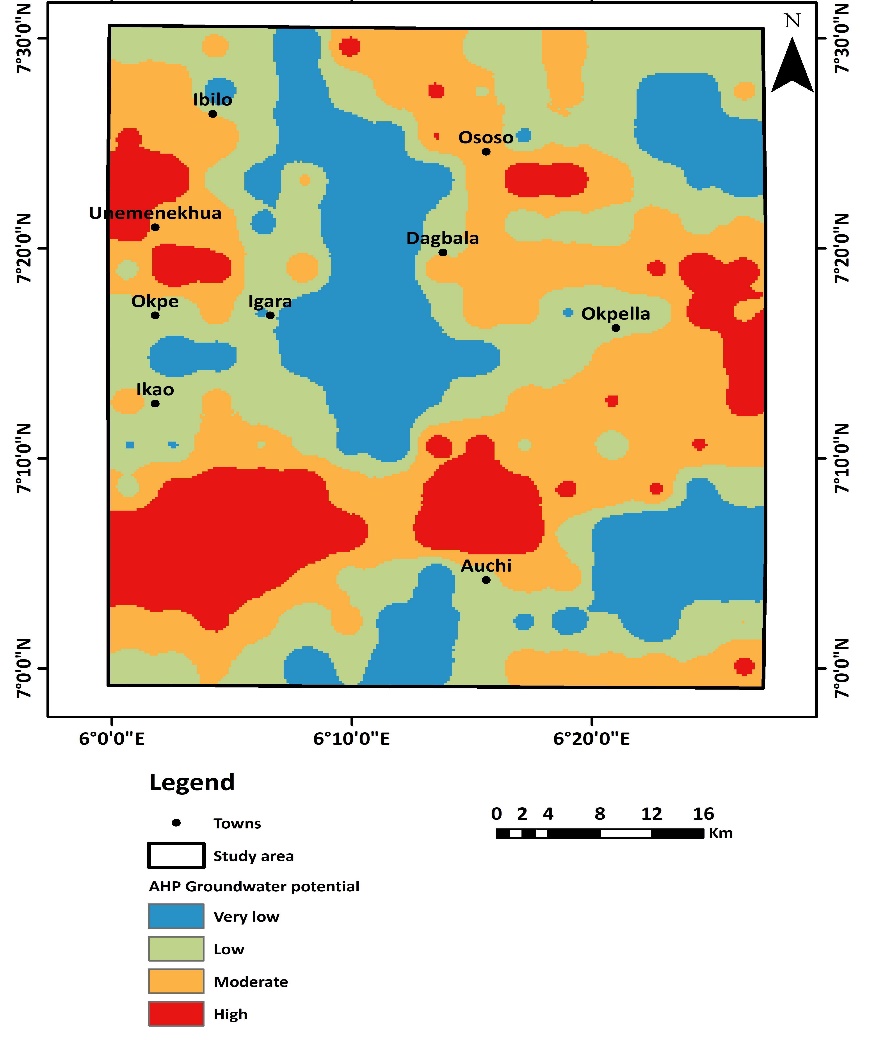


(**9c**)

(**9d**)

(**9a**)

(**9b**)



(**9e**)

**Fig. 9.** GWP zone: a. SVM b. XGB c. RF d. AdaBoost, and e. AHP

**Fig. 10.** Area-wise component of GWP classification: AHP, RF, SVM, AdaBoost, and XGB

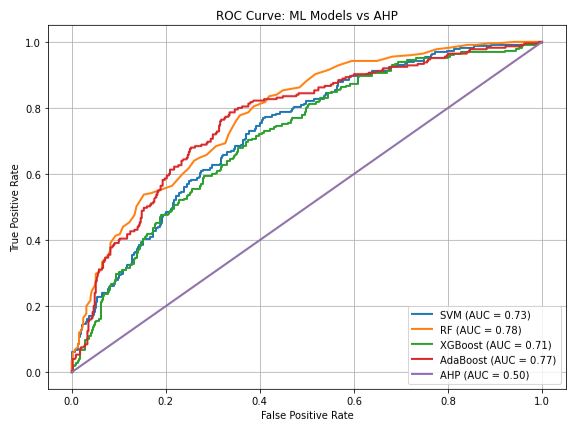
**Table 12.**

Area under GWP zones of distinct models in square kilometres and percentages

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **GWP class** | AHP (km2) | AHP (%) | RF  (km2) | RF  (%) | SVM (km2) | SVM (%) | AdaBoost (km2) | AdaBoost (%) | XGB (km2) | XGB  (%) |
| **Very low** | 519.00 | 21.70 | 562.00 | 23.50 | 204.00 | 8.54 | 542.00 | 22.69 | 654.00 | 27.36 |
| **Low** | 764.00 | 31.96 | 96.00 | 4.00 | 573.00 | 23.99 | 420.00 | 17.59 | 320.00 | 13.40 |
| **Moderate** | 772.00 | 32.30 | 737.00 | 30.83 | 515.00 | 21.54 | 557.00 | 23.29 | 811.00 | 33.93 |
| **High** | 335.00 | 14.03 | 994.00 | 41.59 | 1098.00 | 45.93 | 871.00 | 36.42 | 605.00 | 25.30 |

4.5. Evaluation of models and contrasting the suitable frameworks

The ROC curve is an effective approach for evaluating the categorical outcome of AHP and ML models. It gives information on a model's ability to distinguish between instances that are positive and negative. The AU-ROC indicator measures a model's general success, with a greater AUC value indicating better predictive capacity. The AUC scores for the ML models and AHP are shown subsequently (**Fig. 11**): RF had an optimal AUC of 0.78, preceded by AdaBoost (0.77), SVM (0.73), XGB (0.71), and AHP (0.50). These AUC values indicate the models' capacity to distinguish between prospects for GW and non-prospect areas. The study of the ROC curve yielded an AUC value of 0.78 to 71 for ML, suggesting a good outcome because it fits within the spectrum of 0.71-0.80, as shown in **Fig. 11.** As a result, the use of ML algorithms in this work produced adequate spatial predictions of groundwater potential. Conversely, AHP ranges from 0.50 to 0.60, indicating minimal efficiency. Prior studies (e.g., Masroor et al., 2023; Ali et al., 2023; Ozegin et al., 2024b; Sharma et al., 2024) have also used the AUC values from the ROC curve to assess the exactitude of the created GW prospective zone.

****

**Fig. 11.** ROC curve-based prospect appraisal for a GW map.

4.6. Limitations

Certain limitations must be taken into consideration, even if the study uses a scientific method that is budget-friendly for analyzing different geoenvironmental criteria. The following defines the primary constraints:

The multifaceted geological nature of the subject areas is an important bottleneck to this research. The variability caused by different rock categories, geologic frameworks, and features, including folding and fractures, makes it difficult to determine GWP zones. This research's satellite and field information are heavily reliant on satellites provided by a variety of categories (e.g., agencies). Even so, the map's greatest strength is its incredibly extensive regional coverage. A significant variation in all of these factors can be found at small scales. As a result of this issue, the resulting GWP mapping could prove inaccurate on the local level. There are some limitations associated with predicting groundwater modelling. These are often related to inadequate quantity and quality of data, as well as mistakes in the model's intrinsic configuration and designations. Considering this study, having 36 ground-truth indications, an implication of sample size was incapable of being eliminated under this threshold. A detailed choice of attributes and adjustments can assist in improving prediction power by reducing potential constraints.

One limitation is that fishnets impose an artificial grid structure that may not align perfectly with the actual geographic patterns. Also, the cell size can drastically impact the analysis results. Choosing an appropriate cell size is a key factor in effective analysis.

4.7. Interrogating the potential impacts of the SDGs

Lack of data and monitoring limitations continually inhibit precise evaluation of the remaining SDG 6 targets, which include managing water resources, water quality, aquatic ecosystems, and a supportive environment. To attain the SDGs, it is critical to highlight the significance of this study within the overall framework of a sustainability strategy, notably within the framework of managing water resources. The SDGs represent an extensive ensemble of global initiatives aimed at ensuring a sustainable future for human civilization. This entails a wide range of the financial, ecological, and interpersonal phenomena. In the setting of this current study, using the ML and AHP models to identify the prospect for GW zones would provide reliable data on the state of groundwater that eventually contributes to the achievement of SDG Goal 6 (clean water and sanitation). Besides, the accurate and reliable outcomes of the ML-based GIS model in locating appropriate agriculturally championed GW resources will strengthen its long-term viability to aid in the achievement of the Sustainable Development Goals: zero hunger (Goal 2) and good health and well-being (Goal 3) through sustainable management of terrestrial ecosystems and their services (SDG 15) and mitigating the effects of climate change (SDG 13). Thereby strengthening sustainable water, food, and energy nexuses and fostering perspectives for a reliable water future.

**5. Conclusion**

Defining appropriate areas for GWP is critical to ensuring the effective and long-lasting utilization of existing water assets in the research area. In locations where data is limited, RS-based data sources can give insightful knowledge. AHP, SVM, XGB, RF, and AdaBoost were employed in this work. Nine thematic maps (normalized difference vegetation index, geology, lineament density, rainfall distribution, proximity to surface water bodies, slope, aspect, drainage density, and topographic wetness index) were developed with an impact on regional groundwater. The study area is characterized by a wetland environment, a minimal number of lineaments, a progressive slope, and homogeneous alluvial deposit geology; every one of these factors influences GW. Based on the greater density of lineaments, the precipitous slope, and the availability of geological features, the northeastern portion of the study region presents certain challenges. Geology, rainfall, and lineament density had the largest impact on GWP zone delineation throughout the entire model. Four plausible zones are highlighted on the GWP zone map created by ArcGIS spatial analysis tools: "very low, low, moderate, and high." Wetlands and agriculture-related areas have a high and moderate GWP zone. Every model, consisting of AHP, SVM, XGB, RF, and AdaBoost, identifies high GWP zones within the range of 14.03-45.93% of the overall area, while very low groundwater potential zones cover 8.54-27.36%. The effectiveness of the model was verified using ROC to corroborate the GWP zones. The findings show that RF and AdaBoost surpass GWP zone estimation. The outcomes of verification considerably improve the dependability of the methods used. These frameworks will be beneficial for effectively assessing groundwater replenishment and directing the best place for artificially constructed replenishing mechanisms and other watershed planning projects. The approaches used in this study, which is based on conceptualization needs and is systematic in nature, can be easily employed everywhere data is available, independent of the alterations needed to tackle concerns such as water scarcity and changing climates. The GW recharging prospect map serves as an archive for resource knowledge, which may be updated on a regular basis by combining fresh data and other thematic map kinds. The study demonstrates dependable modelling of GW prospects gets established by combining remotely observed features, groundwater bore data, and themed data. Furthermore, the study improves our comprehension of major discoveries identified during expeditions with local populations. However, the widespread harvesting of groundwater for use in factories using deep-water motors, especially during the warmer seasons of November to February, offers a serious challenge to users of shallow, privately tubed wells. It is difficult for these people to get enough water to satisfy their regular requirements. Additional practical consequences of this study include addressing a shortage of water, enhancing water usage and preservation oversight, assisting in the development of a strategic plan to address the enduring issue of water resource preservation, and giving priority to strategies for guaranteeing sustainable utilization of GW in regions with a significant amount of farming activity and comparable geographical and climatic features. The future goal for the GWP study is to look into how changes in climate, land use planning, and farming operations impact recharging zones. When developing recharging zone maps, use AI—machine learning and deep learning technologies—to improve modelling, estimation, and decision assistance. Prospective GWP studies must take into account the specific chemical makeup of various kinds of rocks, the trend of groundwater flow, the complex setting, and, certainly, the high installation expenses. The uncovered GWP zone maps of the data-scarce and disadvantaged area are going to offer the best solution for the public and private sectors to properly oversee and strategically organize the resource. Considering their superior and speedy effectiveness, the methodologies used show the efficacy of MLAs, remote sensing, and GIS in spatial multi-decision-making processes, particularly in addressing groundwater issues.

**Abbreviation**

AdaBoost Adaptive Boosting

AHP Analytical Hierarchy Process

AUC Area Under Curve

AUROC Area Under the Receiver Operating Characteristic

BH Borehole

CI Consistency Index

CR Consistency Ratio

CRU Climatic Research Unit

DD Drainage Density

DEM Digital Elevation Model

FN Fishnet

FPR False Positive Rate

GIS Geographic Information System

GW Groundwater

GWP Groundwater Potential

GWPZ Groundwater Potential Zones

GY Geology

LD Lineament Density

LL Land use/Land cover

MC Multicollinearity

MCDA Multi-Criteria Decision Analysis

ML Machine Learning

MLAs ML Algorithms

MPC mean percentage change

NDVI Normalized Difference Vegetation Index

OOB Out-of-Bag

PSW Proximity to Surface Water Bodies

RBF Radial Basis Function

RD Rainfall Distribution

RF Random Forest

RI Random Index

ROC Receiver Operating Characteristics

RS Remote sensing

SA Sensitivity Analysis

SDGs Sustainable Development Goals

SI Sensitivity Index

SP Slope

SRTM Shuttle Radar Topography Mission

SVM Support Vector Machines

T Tolerance

TPR True Frue Positive Rate

TWI Topographic Wetness Index

VIF Variance Inflation Factor

XGB eXtreme gradient boost

**References**

Abrams, W., Ghoneim, E., Shew, R., LaMaskin, T., Al-Bloushi, K., Hussein, S., AbuBakr, M., Al-Mulla, E., Al-Awar, M., El-Baz, F., 2018. Delineation of groundwater potential (GWP) in the northern United Arab Emirates and Oman using geospatial technologies in conjunction with Simple Additive Weight (SAW), Analytical Hierarchy Process (AHP), and Probabilistic Frequency Ratio (PFR) techniques. J. Arid. Environ. 157, 77–96. <https://doi.org/10.1016/j.jaridenv.2018.05.005>.

Abrar, H., Kura, A.L., Dube, E.E. Beyene, D.L., 2021. AHP based analysis of groundwater potential in the western escarpment of the Ethiopian rift valley, Geol. Ecol. Landsc. 7 (3), 175-188. [https://10.1080/24749508.202 1.1952761](https://10.1080/24749508.202%201.1952761).

Adewumi, J.R., Ajibade, T.F., Ajibade, F.O., 2020. Appraisal of on-site sanitation facilities and solid waste management in public places within Akure municipality, Nigeria. J Civil Eng, Sci Technol. 11 (1), 1–20. <https://doi.org/10.33736/jcest.1872.2020>.

Agarwal, E., Agarwal, R., Garg, R.D., Garg, P.K., 2013. Delineation of groundwater potential zone: An AHP/ANP approach. J Earth Syst Sci 122, 887–898. <https://doi.org/10.1007/s12040-013-0309-8>.

Ahmed, R., Sajjad, H., 2018. Analyzing factors of groundwater potential and its relation with population in the Lower Barpani Watershed, Assam, India. Nat. Resour. Res. 27, 503–515.  <https://doi.org/10.1007/s11053-017-9367-y>.

Ajay Kumar, V., Mondal, N.C., Ahmed, S., 2020. Identification of groundwater potential zones using RS, GIS and AHP techniques: a case study in a part of Deccan Volcanic Province (DVP), Maharashtra, India. J Indian Soc Remote Sens 48, 497–511. <https://doi.org/10.1007/s12524-019-01086-3>.

Akıncı, H., Özalp, A.Y., Bülent, T., 2013. Agricultural Land Use Suitability using GIS and AHP Technique. Comput. Electron. Agric. 97, 71–82. [htt.ps://doi.org/10.1016/j.compag.2013.07.006](https://doi.org/10.1016/j.compag.2013.07.006).

Al-Djazouli, M.O., Elmorabiti, K., Rahimi, A., Amellah, O., Fadil, O.A.M., 2020. Delineating of groundwater potential zones based on remote sensing, GIS and analytical hierarchical process: a case of Waddai, eastern Chad. GeoJournal. <https://doi.org/10.1007/s10708-020-10160-0>.

Ali, R., Sajjad, H., Saha, T.K., Roshani, Masroor, M., Rahaman, M.H., 2023. Effectiveness of machine learning ensemble models in assessing groundwater potential in Lidder watershed, India. Acta Geophys. 72, 2843–2856. <https://doi.org/10.1007/s11600-023-01237-8>.

Alonso, J.A., Lamata, M.T., 2006. Consistency in the analytic hierarchy process: a new Approach. Int. J. Uncertain. Fuzziness Knowledge-Based Syst. 14 (4), 445–459. <https://doi.org/10.1142/S0218488506004114>.

Al-Shabeeb, AA-R., Al-Adamat, R., Al-Fugara, A., Al-Amoush, H., AlAyyash, S., 2018. Delineating groundwater potential zones within the Azraq Basin of Central Jordan using multi-criteria GIS analysis. Groundw Sustain Dev 7, 82–90. <https://doi.org/10.1016/j.gsd.2018.03.011>.

Arabameri, A., Pal, S.C., Rezaie, F., Nalivan, O.M., Chowdhuri, I., Saha, A., Lee, S., Moayedi, H., 2021. Modeling groundwater potential using novel GIS-based machine-learning ensemble techniques. J Hydrol Reg Stud 36, 100848. <https://doi.org/10.1016/j.ejrh.2021.100848>.

Arefin, R., 2020. Groundwater potential zone identification using an analytic hierarchy process in Dhaka City, Bangladesh. Environ. Earth Sci. 79 (11), 268. <https://doi.org/10.1007/s12665-020-09024-0>.

Arshad, A., Zhang, Z., Zhang, W., Dilawar, A., 2020. Mapping favorable groundwater potential recharge zones using a GIS-based analytical hierarchical process and probability frequency ratio model: A case study from an agro-urban region of Pakistan. Geosci. Front. 11, 1805–1819. <https://doi.org/10.1016/j.gsf.2019.12.013>.

Avand, M., Janizadeh, S., Tien Bui, D., Pham, V.H., Ngo, P.T.T., Nhu, V.H., 2020. A tree-based intelligence ensemble approach for spatial prediction of potential groundwater. Int. J. Digit. Earth 13 (12), 1408–1429. <https://doi.org/10.1080/17538947.2020.1718785>.

Baghel, S., Tripathi, M.P., Khalkho, D., Al-Ansari, N., Kumar, A., Elbeltagi, A., 2023. Delineation of suitable sites for groundwater recharge based on groundwater potential with RS, GIS, and AHP approach for Mand catchment of Mahanadi Basin. Sci. Rep. 13 (1), 9860. <https://doi.org/10.1038/s41598-023-36897-5>.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.

Brunelli, M., 2015. Introduction to the Analytic Hierarchy Process. SpringerBriefs in Operations Research. Springer Cham. <https://doi.org/10.1007/978-3-319-12502-2>.

Castillo, J.L., Martínez Cruz, D.A., Ramos Leal, J.A., Tuxpan Vargas, J., Rodríguez Tapia, S. A., Marín Celestino, A.E., 2022. Delineation of groundwater potential zones (GWPZs) in a semi-arid basin through remote sensing, GIS, and AHP approaches, Water 14 2138. <https://doi.org/10.3390/w14132138>

Chatterjee, D., Singh, P.K., Singh, D., Das, D.B., 2024. Implementation of Satellite-based Water Accounting Plus (WA+) Framework for Estimating Groundwater Balance and Utilization in a Semi-Arid River Basin of India. 28, 101391. <https://doi.org/10.1016/j.gsd.2024.101391>

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. J Hydrol 572, 435–448. <https://doi.org/10.1016/j.jhydrol.2019.03.013>.

Chen, W., Li, H., Hou, E., Wang, S., Wang, G., Panahi, M., Li, T., Peng, T., Guo, C., Niu, C., 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. <https://doi.org/10.1016/j.scitotenv.2018.04.055>.

Danumah, J.H., Odai, S.N., Saley, B.M., Szarzynski, J., Thiel, M., Kwaku, A., Kouame, F.K., Akpa, L.Y., 2016. Flood risk assessment and mapping in Abidjan district using multi-criteria analysis (AHP) model and geoinformation techniques, (Cote d’Ivoire). Geoenviron. Disasters 3, 10. <https://doi.org/10.1186/s40677-016-0044-y>.

Dar, T., Rai, N., Bhat, A., 2021. Delineation of potential groundwater recharge zones using analytical hierarchy process (AHP). Geol Ecol Landsc 5(4), 292–307. <https://doi.org/10.1080/24749508.2020.1726562>.

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., 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, 110886. <https://doi.org/10.1016/j.ecolind.2023.110886>.

Echogdali, F.Z., Boutaleb, S., Bendarma, A., Saidi, M.E., Aadraoui, M., Abioui, M., Ouchchen, M., Abdelrahman, K., Fnais, M.S., Sajinkumar, K.S., 2022. Application of analytical hierarchy process and geophysical method for groundwater potential mapping in the Tata basin, Morocco. Water 14(15), 2393. <https://doi.org/10.3390/w14152393>.

Fajemilo, P.T., Ozegin, K.O., 2024. Evaluation of the groundwater potential of Ogbomoso, Southwestern Nigeria, using an adaptive neuro-fuzzy inference system optimized by three metaheuristic algorithms. Iran J Geophys. 8 (6), 47-78 <https://doi.org/10.30499/ijg.2024.431927.1559>.

Freund, Y., Schapire, R.E., 1995 A decision-theoretic generalization of on-line learning and an application to boosting. In: Vitányi, P. (eds) Computational Learning Theory. EuroCOLT 1995. Lecture Notes in Computer Science, 904. Springer, Berlin, Heidelberg. <https://doi.org/10.1007/3-540-59119-2_166>.

Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences 55, 119–139. <https://doi.org/10.1145/2818346.2823306>.

Gayen, A., Pourghasemi, H.R., Saha, S., Keesstra, S., Bai, S., 2019. Gully Erosion Susceptibility Assessment an­ Management of Hazard­-prone Areas in Ind­ia Using Different Machine Learning Algorithms. Science of the Total Environment. 668, 124–138. <https://doi.org/10.1016/j.scitotenv.2019.02.436>.

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. Environ. Monit. Assess. 190 (3), 149. <https://doi.org/10.1007/s10661-018-6507-8>.

Goward, S.N., Markham, B., Dye, D.G., Dulaney, W., Yang, J., 1991. Normalized difference vegetation index measurements from the advanced very high-resolution radiometer. Remote Sensing of Environment 35 (2-3), 257-277. <https://doi.org/10.1016/0034-4257(91)90017-Z>.

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. <https://doi.org/10.1016/j.jksus.2016.08.003>.

Hasan­uzzaman, M., Mandal, M.H., Hasnine, M., Shit, P.K., 2022. Groundwater potential mapping ­sing multi-criteria decision, bivariate statistic and machine learning algorithms: evidence from Chota Nagp­r Platea­, India. Appl Water Sci. 12(4), 1–16. <https://doi.org/10.1007/s13201-022-01584-9>.

Hong, H., Pra­dhan, B., Bui, D.T., Xu, C., Youssef, A.M., Chen, W., 2017. Comparison of Four Kernel Functions Used in Support Vector Machines for Landslide Susceptibility Mapping: A Case Study at Suichuan Area (China). Geomatics, Natural Hazards and Risk. 8(2), 544–569. <https://doi.org/10.1080/19475705.2016.1250112>.

Horn, R.A., Johnson, C.R., 1985. Matrix Analysis. Cambridge University Press. <https://assets.cambridge.org/97805218/39402/frontmatter/9780521839402_frontmatter.pdf>.

Hussain, S., Mubeen, M., Ahmad, A., Majeed, H., Qaisrani, S.A., Hammad, H.M., Amjad, M., Ahmed, I., Fahhad, S., Ahmed, N., Nasim, W., 2022. Assessment of land use/land cover changes and its effect on land surface temperature using remote sensing techniques in Southern Punjab, Pakistan. Environ. Sci. Pollution Res., 30, 99202–99218. <https://doi.org/10.1007/s11356-022-21650-8>.

Hutchinson, M.L., Antono, E., Gibbons, B.M., Paradiso, S., Ling, J., Meredig, B., 2017. Overcoming data scarcity with transfer learning. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. <https://arxiv.org/abs/1711.05099v1>.

Ikrri, M, Boutaleb, I, Abiou, M, Zahra, F.E, Abdelrahman, K, Mouna, I, Tamer, A, Hasna, E.A, Sara, E, Farid, F. 2023. Delineation of groundwater potential area using an AHP, remote sensing, and GIS Techniques in the Ifni Basin, Western Anti-Atlas, and Morocco. Water 15(7), 1436. <https://doi.org/10.3390/w15071436>.

Ilugbo, S.O., Aigbedion, I., Ozegin, K.O., Bawallah, M.A., 2023a. Assessment of groundwater occurrence in a typical schist belt region in Osun state, southwestern Nigeria using VES, aeromagnetic dataset, remotely sensed data and MCDA approaches. Sustain. Water Resourc. Manag. 9, 29. <https://doi.org/10.1007/s40899-022-00810-1>.

Ilugbo, S.O., Aigbedion, I., Ozegin, K.O., 2023b. Structural mapping for groundwater occurrence using remote sensing and geophysical data in Ilesha Schist Belt, Southwestern Nigeria. Geology, Ecology, and Landscapes. <https://doi.org/10.1080/24749508.2023.2182063>.

Ilugbo, S.O.,Ozegin, K.O., 2018. Significances of Deep-Seated Lineament in Groundwater Studies around Ilesha, Southwestern Nigeria. Asian Journal of geological Research 1 (1), 1-16.<https://doi.org/10.9734/AJOGER/2018/42078>.

Jennifer, J.J., 2022. Feature elimination and comparison of machine learning algorithms in landslide susceptibility mapping. Environ Earth Sci. 81, 489. <https://doi.org/10.1007/s12665-022-10620-5>.

Kalura, P., Pandey, A., Chowdary, V.M., Raju, P.V., 2022. Land Use Land Cover Change Detection of the Tons River Basin Using Remote Sensing and GIS. In: Pandey, A., Chowdary, V.M., Behera, M.D., Singh, V.P. (eds) Geospatial Technologies for Land and Water Resources Management. Water Science and Technology Library, 103, 53–65. Springer, Cham. <https://doi.org/10.1007/978-3-030-90479-1_4>.

Karimi-Rizvandi, S., Goodarzi, H.V., Afkoueieh, J.H., Chung, I-M., Kisi, O., Kim, S., 2021. Groundwater-potential mapping using a self-learning bayesian network model: a comparison among metaheuristic algorithms. Water 13, 658. <https://doi.org/10.3390/w13050658>.

Kavzoglu, T., Colkesen, I., Sahin, E.K., 2019. Machine Learning Techniques in Landslide Susceptibility Mapping: a Survey and a Case Study. Springer, Cham, 50, 283–301. <https://doi.org/10.1007/978-3-319-77377-3_13>.

Kaya, Y.; Sanli, F.B.; Abdikan, S., 2023. Determination of long-term volume change in lakes by integration of UAV and satellite data: The case of Lake Burdur in Türkiye. Environ. Sci. Pollut. Res. Int. 30, 117729–117747. <https://doi.org/10.1007/s11356-023-30369-z>.

Khan, Q., Liaqat, M.U., Mohamed, M.M., 2021. A comparative assessment of modeling groundwater vulnerability using DRASTIC method from GIS and a novel classification method using machine learning classifiers. Geocarto International. 37(20), 5832–5850. <https://doi.org/10.1080/10106049.2021.1923833>.

Khan, Z.A., Jhamnani, B., 2023. Identification of groundwater potential zones of Idukki district using remote sensing and GIS-based machine-learning approach. Water Supply 23 (6), 2426. <https://doi.org/10.2166/ws.2023.134>.

Kolawole, A.O., 2015. Nigeria’s Water Resources and Spatial Variation in Water Quality. Department of Environmental Quality Sciences, The Hebrew University of Jerusalem. <https://doi.org/10.13140/RG.2.1.2843.9209>. Technical Report, March.

Kroll, C.N., Song, P., 2013. Impact of multicollinearity on small sample hydrologic regression models. Water Resour. Res. 49, 3756–3769. <https://doi.org/10.1002/wrcr.20315>.

Lee, S., Hong S-M., Jung, H.S., 2018. GIS-based groundwater potential mapping using artificial neural network and support vector machine models: the case of boryeong city in Korea. Geocarto Int 33(8), 847–861. <https://doi.org/10.1080/10106049.2017.1303091>.

Lee, S., Song, K-Y., Kim, Y., Park, I., 2012. Regional groundwater productivity potential mapping using a geographic information system (GIS) based artificial neural network model. Hydrogeol J 20, 1511–1527. <https://doi.org/10.1007/s10040-012-0894-7>.

Lee, S., Hyun, Y., Lee, S., Lee, M.J., 2020. Groundwater potential mapping using remote sensing and GIS-based machine learning techniques. *Remote Sens. 12*(7), 1200. <https://doi.org/10.3390/rs12071200>.

Liaw, A., Wiener., M., 2024. Breiman and Cutler’s Random Forests for Classification and Regression. Available online: <https://cran.r-project.org/web/packages/randomForest/randomForest.pdf> (accessed on 15 April 2025).

Maity, D.K., Mandal, S., 2019. Identification of Groundwater potential zones of the Kumari River Basin, India: an RS & GIS based Semi-quantitative Approach. Environ Dev Sustain 21(2), 1013– 1034. <https://doi.org/10.1007/s10668-017-0072-0>.

Malczewski, J., 2010. Multiple Criteria Decision Analysis and Geographic Information Systems. In: Ehrgott, M., Figueira, J., Greco, S. (eds) Trends in Multiple Criteria Decision Analysis. International Series in Operations Research & Management Science 142, 369–395. Springer, Boston, MA. <https://doi.org/10.1007/978-1-4419-5904-1_13>.

Mandal, P., Saha, J., Bhattacharya, S., Paul, S., 2021. Delineation of groundwater potential zones using the integration of geospatial and MIF techniques: A case study on Rarh region of West Bengal, India. Environ. Chall. 5, 100396. <https://doi.org/10.1016/j.envc.2021.100396>.

Mansour, S., Al-Awhadi, T., Al-Hatrushi, S., 2019. Geospatial based Multi-criteria Analysis for Ecotourism Land Suitability using GIS and AHP: A Case Study in Masirah Island, Oman. J. Ecotourism. 19(2), 148-167. <https://doi.org/10.1080/14724049.2019.1663202>.

Marjanović, M., Kovačević, M., Bajat, B., Vozenilek, V., 2011. Landslide Susceptibility Assessment Using SVM Machine Learning Algorithm. Engineering Geology. 123(3), 225–234. <https://doi.org/10.1016/j.enggeo.2011.09.006>.

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 Sens 12, 2742. <https://doi.org/10.3390/rs12172742>.

Masroor, M., Sajjad, H., Kumar, P., Saha, T.K., Rahaman, M.H., Choudhari, P., Kulimushi, L.C., Pal, S., Saito, O., 2023. Novel ensemble machine learning modeling approach for groundwater potential mapping in Parbhani District of Maharashtra, India. Water 15 (3), 419. <https://doi.org/10.3390/w15030419>.

Mathewos, Y., Abate, B., Dadi, M., Mathewos, M., 2024. Evaluation of the groundwater prospective zone by coupling hydro-meteorological and geospatial evidence in Wabe River Catchment Omo Gibe River Basin, Ethiopia. Water Cycle 5, 37–58. <https://doi.org/10.1016/j.watcyc.2024.01.002>.

McKay, G., Harris, J.R., 2016. Comparison of the data-driven random forests model and a knowledge-driven method for mineral prospectively mapping: a case study for gold deposits around the Huritz Group and Nueltin Suite, Nunavut, Canada. Nat Resour Res.  25, 125–143. <https://doi.org/10.1007/s11053-015-9274-z>.

Moghaddam, D.D., Rahmati, O., Haghizadeh, A., Kalantari, Z.A., 2020. A modeling comparison of groundwater potential mapping in a mountain bedrock aquifer: QUEST, GARP, and RF models. Water 12(3), 679. <https://doi.org/10.3390/w12030679>.

Mohammadi-Behzad, H.R., Charchi, A., Kalantari, N., Nejad, A.M., Vardanjani, H.K., 2019. Delineation of groundwater potential zones using remote sensing (RS), geographical information system (GIS) and analytic hierarchy process (AHP) techniques: A case study in the Leylia–Keynow watershed, southwest of Iran. Carbonates Evaporites. 34, 1307–1319. <https://doi.org/10.1007/s13146-018-0420-7>.

Morales, F., Jr., de Vries, W.T., 2021. Establishment of Land Use Suitability Mapping Criteria Using Analytic Hierarchy Process (AHP) with Practitioners and Beneficiaries. Land 10, 235. <https://doi.org/10.3390/land10030235>.

Morbidelli, R., Saltalippi, C., Flammini, A., Govindaraju R.S., 2018. Role of slope on infiltration: a review. J. Hydrol. 557, 878-886. <https://doi.org/10.1016/j.jhydrol.2018.01.019>.

Mukherjee, I., Singh, U., 2020. Delineation of groundwater potential zones in a drought prone semi-arid region of east India using GIS and analytical hierarchical process techniques. Catena 194, 104681. <https://doi.org/10.1016/j.catena.2020.104681>.

Muthu, K., Sudalaimuthu, K., 2021. Integration of Remote sensing, GIS, and AHP in demarcating groundwater potential zones in Pattukottai Taluk, Tamilnadu, India, Arabian J. Geosci. 14, 1748. <https://doi.org/10.1007/s12517-021-08110-2>.

Naghibi, S.A., Pourghasemi, H.R., 2015. A comparative assessment between three machine learning models and their performance comparison by bivariate and multivariate statistical methods in groundwater potential mapping. Water Resour Manag. 29, 5217–5236. <https://doi.org/10.1007/s11269-015-1114-8>.

Naghibi, S.A., Pourghasemi, H.R., Dixon, B., 2016. GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. Environ. Monit. Assess. 188, 44. <https://doi.org/10.1007/s10661-015-5049-6>.

Naghibi, S.A., Kourosh Ahmadi, K., Daneshi, A., 2017. Application of Support Vector Machine, Random Forest, and Genetic Algorithm Optimized Random Forest Models. Water Resour Manage 31, 2761–2775. <https://doi.org/10.1007/s11269-017-1660-3>.

Naghibi, S.A., Pourghasemi, H.R., Abbaspour, K., 2018. A Comparison between Ten A­vance­ an­ Soft Computing Models for Groundwater Qanat Potential Assessment in Iran Using R an­ GIS. Theoretical and Applied Climatology. 131, 967–984. ­ <https://doi.org/10.1007/s00704-016-2022-4>.

Naghibi, S.A., Hashemi, H., Berndtsson, R., Lee, S., 2020. Application of extreme gradient boosting and parallel random forest algorithms for assessing groundwater spring potential using DEM-derived factors, J. Hydrol. 589, 125197. <https://doi.org/10.1016/j.jhydrol.2020.125197>.

Niazkar, M., Menapace, A., Brentan, B., Piraei, R., Jimenez, D., Dhawan, P., Righetti, M., 2024. Applications of XGBoost in water resources engineering: a systematic literature review (Dec 2018–May 2023). Environ. Model. Softw. 174, 105971. <https://doi.org/10.1016/j.envsoft.2024.105971>.

Obaje, N.G. 2009. The Basement Complex. In: Geology and Mineral Resources of Nigeria. Lecture Notes in Earth Sciences. 120. Springer, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-540-92685-6_2>.

Odeyemi, I.B., 1981. A review of orogenic events in the Precambrian Basement of Nigeria, West Africa. Geol Rundsch. 70(3):897–909. <https://doi.org/10.1007/BF01820170>.

Oh, H. J., Kim, Y.S., Choi, J.K., Park, E., Lee, S., 2011. GIS Mapping of Regional Probabilistic Groundwater Potential in the Area of Pohang City, Korea. Journal of Hydrology 399 (3–4), 158–172. <https://doi.org/10.1016/j.jhydrol.2010.12.027>.

Olden, J.D., Lawler, J.J., Poff, N.L., 2008. Machine learning methods without tears: a primer for ecologists. Quarterly Review of Biology 83, 171–193. <https://doi.org/10.1086/587826>.

Opoku, P.A., Shu, L., Amoako-Nimako, G.K., 2024. Assessment of Groundwater Potential Zones by Integrating Hydrogeological Data, Geographic Information Systems, Remote Sensing, and Analytical Hierarchical Process Techniques in the Jinan Karst Spring Basin of China. Water 16, 566. <https://doi.org/10.3390/w16040566>.

Ozdemir, A., 2011.GIS-based groundwater spring potential mapping in the Sultan Mountains (Konya, Turkey) using frequency ratio, weights of evidence and logistic regression methods and their comparison. J. Hydrol. 411 (3-4), 290–308. <https://doi.org/10.1016/j.jhydrol.2011.10.010>.

Ozegin, K.O., Ilugbo, S.O., Ogunseye, T.T., 2023. Groundwater exploration in a landscape with heterogeneous geology: an application of geospatial and analytical hierarchical process (AHP) techniques in the Edo north region, in Nigeria. Groundwater for Sustainable Development 20, 100871. <https://doi.org/10.1016/j.gsd.2022.100871>.

Ozegin, K.O., Ilugbo, S.O., 2024. A triangulation approach for groundwater potential evaluation using geospatial technology and multi-criteria decision analysis (MCDA) in Edo State, Nigeria. J. Afr. Earth Sci. 209, 105101. <https://doi.org/10.1016/j.jafrearsci.2023.105101>.

Ozegin, K.O., Ilugbo, S.O., Akande, O.N., 2024a. Leveraging geospatial technology and AHP for groundwater potential zonation in parts of south and north-Central Nigeria. Sustainable Water Resources Management. 10, 146. <https://doi.org/10.1007/s40899-024-01124-0>.

Ozegin, K.O., Ilugbo, S.O., Alile, O.M., Iluore, K., 2024b. Integrating in-situ data and spatial decision support systems (SDSS) to identify groundwater potential sites in the Esan plateau, Nigeria. Groundwater for Sustainable Development 26, 101276. <https://doi.org/10.1016/j.gsd.2024.101276>.

Ozegin, K.O., Ilugbo, S.O., Adebo, B., 2024c. Spatial evaluation of groundwater vulnerability using the DRASTIC-L model with the analytic hierarchy process (AHP) and GIS approaches in Edo State, Nigeria. Phys. Chem. Earth 134, 103562. <https://doi.org/10.1016/j.pce.2024.103562>.

Ozegin, K.O., Ilugbo,S.O.,2025. Evaluation of potentially susceptible flooding areas leveraging geospatial technology with multicriteria decision analysis in Edo State, Nigeria. Natural Hazards Research. 5 (1), 109-133. <https://doi.org/10.1016/j.nhres.2024.07.002>

Pande, C.B., Moharir, K.N., Singh, S.K., Varade, A.M., 2020. An integrated approach to delineate the groundwater potential zones in Devdari watershed area of Akola district, Maharashtra, Central India. Environ. Dev. Sustain. 22 (5), 4867–4887. <https://doi.org/10.1007/s10668-019-00409-1>.

Patidar, R., Pingale, S.M. Khare, D., 2021. An integration of geospatial and machine learning techniques for mapping groundwater potential: a case study of the Shipra River basin, India. Arab J Geosci 14, 1645. <https://doi.org/10.1007/s12517-021-07871-0>.

Patra, S., Mishra, P., Mahapatra, S.C., 2018. Delineation of groundwater potential zone for sustainable development: A case study from Ganga Alluvial Plain covering Hooghly district of India using remote sensing, geographic information system and analytic hierarchy process. Journal of Cleaner Production,172, 20 2485-2502. <https://doi.org/10.1016/j.jclepro.2017.11.161>.

Pham, B.T., Jaafari, A., Phong, T.V., Mafi-Gholami, D., Amiri, M., Van Tao, N., Duong, V.-H., Prakash, I., 2021. Naïve Bayes ensemble models for groundwater potential mapping. Ecol. Inform. 64, 101389. <https://doi.org/10.1016/j.ecoinf.2021.101389>.

Pou­rtaghi, Z.S., Pou­rghasemi, H.R., 2014. GIS-based groundwater spring potential assessment and mapping in the Birjand Township, southern Khorasan Province, Iran. Hydrogeol J. 22(3), 643–662. <https://doi.org/10.1007/s10040-013-1089-6>.

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. Gisci. Remote Sens. 57 (6), 735–752. <https://doi.org/10.1080/15481603.2020.1794104>.

Rahmati, O., Pourghasemi, H.R., Melesse, A.M., 2016. Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: a case study at Mehran region, Iran. Catena 137, 360–372. <https://doi.org/10.1016/j.catena.2015.10.010>.

Rahmati, O., Melesse, A.M., 2016. Application of Dempster-Shafer theory, spatial analysis and remote sensing for groundwater potentiality and nitrate pollution analysis in the semi-arid region of Khuzestan, Iran. Sci. Total Environ. 568, 1110–1123.

<https://doi.org/10.1016/j.scitotenv.2016.06.176>.

Rahmati, O., Pourghasemi, H.R., Melesse, A.M., 2016. Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: A case study at Mehran Region, Iran. Catena, 137, 360–372. <https://doi.org/10.1016/j.catena.2015.10.010>.

Rajaveni, S.P., Brind­ha, K., Elango, L., 2017. Geological and Geomorphological Controls on Groundwater Occurrence in a Hard­ Rock Region. Applied Water Science. 7(3), 1377–1389. <https://doi.org/10.1007/s13201-015-0327-6>.

Rajendran, G., Mohammed, M., Shivakumar, S., Merera, W., Taddese, K., 2020. Geospatial techniques amalgamated with two-dimensional electrical resistivity imaging for delineation of groundwater potential zones in West Guji Zone, Ethiopia, Groundwater Sustain. Dev. 11 (2020) 100407, <https://doi.org/10.1016/j.gsd.2020.100407>.

Rana, M.M.S.P., Rahman, M.T., Hassan, M.F., 2025. Mapping groundwater potential zone by robust machine learning algorithms & remote sensing techniques in agriculture dominated area, Bangladesh. Cleaner Water 3, 100064. <https://doi.org/10.1016/j.clwat.2025.100064>.

Rao, N.S., Gugulothu, S., Das, R., 2022. Deciphering artificial groundwater recharge suitability zones in the agricultural area of a river basin in Andhra Pradesh, India using geospatial techniques and analytical hierarchical process method. Catena 212, 106085. <https://doi.org/10.1016/j.catena.2022.106085>

Razavi-Termeh, S.V., Sadeghi-Niaraki, A., Choi, S.M., 2019. Groundwater potential mapping using an integrated ensemble of three bivariate statistical models with random forest and logistic model tree models. Water 11(8), 1596. <https://doi.org/10.3390/w11081596>.

Reza, A., Seiyed, S., Hosseini, M., 2023. Assessment and delineation of potential groundwater recharge zones in areas prone to saltwater intrusion hazard: A case from Central Iran. Environ. Monit. Assess. 195, 203. <https://doi.org/10.1007/s10661-022-10778-2>.

Roy, A., Keesari, T., Sinha, U.K., Sabarathinam, C., 2019. Delineating groundwater prospect zones in a region with extreme climatic conditions using GIS and remote sensing techniques: A case study from central India. Journal of Earth System Science, 128, 1–19. <https://doi.org/10.1007/s12040-019-1205-7>.

Saaty, T.L., 1990. How to make a decision: the Analytic hierarchy process. Eur. J. Oper. Res. 48, 9–26. <https://doi.org/10.1016/0377-2217(90)90057-I>.

Sachdeva, S., Kumar, B., 2021. Comparison of gradient boosted decision trees and random forest for groundwater potential mapping in Dholpur (Rajasthan), India. Stochastic Environmental Research and Risk Assessment 35 (2), 287–306. [https://doi.org/10.1007/S00477- 020-01891-0/TABLES/6](https://doi.org/10.1007/S00477-%20020-01891-0/TABLES/6).

Saha, S., 2017. Groundwater potential mapping using analytical hierarchical process: a study on Md. Bazar Block of Birbhum District, West Bengal. Spatial Information Research 25 (4), 615–626. <https://doi.org/10.1007/s41324-017-0127-1>.

Saha, R., Baranval, N.K., Das, I., Kumaranchat, V.K., Reddy, K.S., 2022. Application of machine learning and geospatial techniques for groundwater potential mapping. J Ind Soc Remote Sensing 50(10), 1995–2010. <https://doi.org/10.1007/s12524-022-01582-z>.

Saha, B., Acharya, T., Sur, T., 2024. Delineation of groundwater potential zones in coastal alluvial region using remote sensing, geographic information system and analytical hierarchy process techniques. 26, 101238. <https://doi.org/10.1016/j.gsd.2024.101238>.

Sahoo, S., Russo, T.A., Elliott, J., Foster, I., 2017. Machine learning algorithms for modeling groundwater level changes in agricultural regions of the U.S. Water Resour. Res. 53 (5), 3878–3895. <https://doi.org/10.1002/2016WR019933>.

Sameen, M.I., Pradhan, B., Lee, S., 2019. Self-learning random forests model for mapping groundwater yield in data-scarce areas. Nat. Resour. Res. 28, 757–775. <https://doi.org/10.1007/s11053-018-9416-1>.

Samui, P., 2008. Slope Stability Analysis: A Support Vector Machine Approach. Environ Geol. 56(2), 255–267. <https://doi.org/10.1007/s00254-007-1161-4>.

Scanlon, B.R., Fakhreddine, S., Rateb, A., de Graaf, I., Famiglietti, J., Gleeson, T., Grafton, R.Q., Jobbagy, E.C., Kebede, S. Kolusu, S.R., 2023. Global water resources and the role of groundwater in a resilient water future, Nature Reviews Earth & Environment, 4 (2), 87–101. <https://doi.org/10.1038/s43017-022-00378-6>.

Senanayake, I.P., Dissanayake, D.M.D.O.K., Mayadunna, B.B., Weerasekera, W.L., 2016. An approach to delineate groundwater recharge potential sites in Ambalantota, Sri Lanka using GIS techniques. Geosci. Front. 7, 115–124. <https://doi.org/10.1016/j.gsf.2015.03.002>.

Senapati, U., Das, T.K., 2021. Assessment of basin-scale groundwater potentiality mapping in drought-prone upper Dwarakeshwar River basin, West Bengal, India, using GIS-based AHP techniques. Arabian J. Geosci. 14 (11), 1–22. <https://doi.org/10.1007/s12517-021-07316-8>.

Shahinuzzaman, M., Haque, M.N., Shahid, S., 2021. Delineation of groundwater potential zones using a parsimonious concept based on catastrophe theory and analytical hierarchy process. Hydrgeol. J. 29, 1091–1116. <https://doi.org/10.1007/s10040-021-02322-2>.

Shandu, I.D., Atif, I., 2023. An Integration of Geospatial Modelling and Machine Learning Techniques for Mapping Groundwater Potential Zones in Nelson Mandela Bay, South Africa. Water 15, 3447. <https://doi.org/10.3390/w15193447>.

Sharma, R., Pandey, R.K., Singh, A.K., 2024. A Multi-criterian Approach to Identify Groundwater Potential Zones in the Subarnarekha River Basin Using Integrated Analytical Hierarchy Process and Geospatial Technology, Journal of Geography, Environment and Earth Science International, 28 (10), 78–100. <https://doi.org/10.9734/jgeesi/2024/v28i10827>.

Sharma, Y., Ahmed, R., Saha, T.K., Bhuyan, N., Kumari, G., Roshani, Pal S., Sajja, 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. Groundw. Sustain. Dev. 25, 101094 <https://doi.org/10.1016/j.gsd.2024.101094>.

Singh, L.K., Jha, M.K., Chowdary, V.M., 2016. Multi-criteria analysis and GIS modeling for identifying prospective water harvesting and artificial recharge sites for sustainable water supply. J. Clean. Prod. 142, 1436–1456. <https://doi.org/10.1016/j.jclepro.2016.11.163>.

Singha, C., Swain, K.C., Pradhan, B., Rusia D.K., Moghimi, A., Ranjgar, B., 2024. Mapping groundwater potential zone in the subarnarekha basin, India, using a novel hybrid multi-criteria approach in Google earth Engine. Heliyon 10, e24308. <https://doi.org/10.1016/j.heliyon.2024.e24308>.

Sinha, D.D., Mohapatra, S.N., Pani, P., 2012. Mapping and assessment of groundwater potential in Bilrai watershed (Shivpuri District, M.P.) a geomatics approach. J. Indian Soc. Remote Sens. 40, 649–668. <https://doi.org/10.1007/s12524-011-0175-2>.

Solomon, S., Quiel, F., 2006. Groundwater study using remote sensing and geographic information systems (GIS) in the central highlands of Eritrea. Hydrogeol. J. 14 (6), 1029–1041. <https://doi.org/10.1007/s10040-006-0096-2>.

Sørensen, R., Zinko, U., Seibert, J., 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrol. Earth Syst. Sci. 10, 101–112. <https://doi.org/10.5194/hess-10-101-2006>.

Tehrany, M.S., Prad­han, B., Mansor, S., Ahma­d, N., 2015. Flood­ Susceptibility Assessment Using GIS-based Support Vector Machine Mo­el with Different Kernel Types. Catena. 125, 91–101. <https://doi.org/10.1016/j.catena.2014.10.017>.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ 8(2), 127–150. <https://doi.org/10.1016/0034-4257(79)90013-0>.

Thanh, N.N., Chotpantarat, S., Trung, N.H., Ngu, N.H., Muoi, L.N., 2022. Mapping groundwater potential zones in Kanchanaburi Province, Thailand by integrating of analytic hierarchy process, frequency ratio, and random forest. Ecol. Indicat. 145, 109591. <https://doi.org/10.1016/j.ecolind.2022.109591>.

UN DESA, 2022. United Nations department of economic and social affairs, population division. In: World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/NO.3. <https://www.un.org/development/desa/pd/content/World-Population-Prospects-2022>. (accessed on 5 April 2025).

UN DESA. SDGs Report 2023. The Sustainable Development Goals Report 2023: Special Edition. 2023, p. 80. Available online: <https://unstats.un.org/sdgs/report/2023/> (accessed on 9 February 2024).

United Nations, 2015. The UN 2030 agenda for sustainable development. Sustainable Development Goals Series. <https://doi.org/10.1007/978-3-031-07461-5_1>.

Unted Nations, 2023. The sustainable development goals report 2023: special Edition. <https://unstats.un.org/sdgs/report/2023/>.

UNWWDR (UN World Water Development Report), 2022. Groundwater: Making the Invisible Visible. <https://www.unesco.org/reports/wwdr/2022/en/download>. (Accessed 17 January 2024).

UNWWDR (The United Nations World Water Development Report) 2025. Mountains and glaciers: water towers. UNESCO, Paris. <https://doi.org/10.54679/LHPJ5153>.

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. J. Hydrol. 624, 129952. <https://doi.org/10.1016/j.jhydrol.2023.129952>.

van der Gun, J., 2012. Groundwater and Global Change: Trends, Opportunities and Challenges. United Nations Educational, Scientific, and Cultural Organization, Paris, France, 44. <https://unesdoc.unesco.org/ark:/48223/pf0000215496>.

Venkatesh, S., Parimalarenganayaki, S., 2023. Multi-collinearity-based parameter optimization and comparison among multi-criteria decision analysis to map groundwater potential zones. J. Geol. Soc. India 99 (8), 1158–1164. <https://doi.org/10.1007/s12594-023-2441-7>.

Venkatramanan, S., Prasanna, M.V., Chung, S.Y., 2019. GIS and geostatistical techniques for groundwater science. In: introduction and history of GIS and Geostatistical Techniques for Groundwater Science. <https://doi.org/10.1016/C2017-0-02667-8>.

Wang, J., Chen, L., Yu, Z., 2018. Modeling rainfall infiltration on hillslopes using flux-concentration relation and time compression approximation. J. Hydrol. 557, 243-253. <https://doi.org/10.1016/j.jhydrol.2017.12.031>.

Wiesmeier, M., Barthold, F., Blank, B., Kögel-Knabner, I., 2011. Digital Mapping of Soil Organic Matter Stocks Using Random Forest Modeling in a Semi-Arid Steppe Ecosystem. Plant Soil 340, 7–24. <https://doi.org/10.1007/s11104-010-0425-z>.

Wind, Y., Saaty, T.L., 1980. Marketing applications of the analytic hierarchy process. Manag Sci 26(7), 641–658 <https://doi.org/10.1287/mnsc.26.7.641>.

World Bank Group, 2017. Report: a wake-up call: Nigeria water supply, sanitation, and hygiene poverty diagnostic. In: WASH Poverty Diagnostic Series. 2017 International Bank for Reconstruction and Development/The World Bank. <http://documents.worldbank.org/curated/en/56638.15023.79018.228/pdf/118241-WP-PUBLICWASH-series.pdf>. (Accessed 7 July 2022)

WWQA (World Water Quality Alliance). 2021. Assessing Groundwater Quality: A Global Perspective: Importance, Methods and Potential Data Sources. A Report by the Friends of Groundwater in the World Water Quality Alliance. Information document annex for display at the 5th session of the United Nations Environment Assembly, Nairobi. 2021. <https://un-igrac.org/data/resources/assessing-groundwater-quality-a-global-perspective/>.

Yao, X., Tham, L.G., Dai., F.C., 2008. Landslide Susceptibility Mapping Based­ on Support Vector Machine: A Case Stu­dy on Natural Slopes of Hong Kong, China. Geomorphology. 101(4), 572–582. <https://doi.org/10.1016/j.geomorph.2008.02.011>.

Zabihi, M., Pourghasemi, H.R., Pourtaghi, Z.S., Behzadfar, M., 2016. GIS-based multivariate adaptive regression spline and random forest models for groundwater potential mapping in Iran. Environ. Earth Sci. 75 (8), 1–19. <https://doi.org/10.1007/s12665-016-5424-9>.

Zewdie, M.M., Kasie, L.A., Bogale, S., 2024. Groundwater potential zones delineation using GIS and AHP techniques in upper parts of Chemoga watershed, Ethiopia. Appl. Water Sci. 14 (4), 85. <https://doi.org/10.1007/s13201-024-02119-0>.

Zhang, Q., Zhang, S., Zhang, Y., Li, M.,Wei, Y., Chen, M., Zhang, Z., Dai, Z., 2021. GIS-based groundwater potential assessment in varied topographic areas of Mianyang city, Southwestern China, using AHP, Rem. Sens. 13, 4684. <https://doi.org/10.3390/rs13224684>.

Zhao, R., Fan, C. Arabameri, A., Santosh, M., Mohammad, L., Mondal, I., 2024. Groundwater spring potential mapping: Assessment the contribution of hydrogeological factors. Advances in Space Research 74, 48–64. <https://doi.org/10.1016/j.asr.2024.03.038>.

Zheng, Y., Hu, Z., Wang, R., Wen, X., Ren, X., Pan, R., 2024. Slope Stability Analysis under Heavy Rainfall Conditions Based on a Modified Green-Ampt Model. Water Resour Manage 38, 2627–2646. <https://doi.org/10.1007/s11269-024-03790-3>.