The Egyptian Journal of Remote Sensing and Space Sciences 27 (2024) 178–191

Contents lists available at ScienceDirect

The Egyptian Journal of Remote Sensing and Space Sciences

journal homepage: www.sciencedirect.com

Research Paper

# Hybrid deep learning and remote sensing for the delineation of artificial groundwater recharge zones

Rami Al-Ruzouq [a,b], Abdallah Shanableh [a,b], Ratiranjan Jena [a,*], Sunanda Mukherjee [a], Mohamad Ali Khalil [a], Mohamed Barakat A. Gibril [a], Biswajeet Pradhan [c], Nezar Atalla Hammouri [a]

[a] GIS & Remote Sensing Center, Research Institute of Sciences and Engineering, University of Sharjah, Sharjah 27272, United Arab Emirates
[b] Civil and Environmental Engineering Department, University of Sharjah, Sharjah 27272, United Arab Emirates
[c] Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), School of Civil and Environmental Engineering, Faculty of Engineering and Information Technology, University of Technology Sydney, Australia

ARTICLE INFO

*Keywords:*
Artificial groundwater recharge
Remote sensing (RS)
GIS
Multicriteria analysis
CNN-XGB model
United Arab Emirates

ABSTRACT

The increase in water demand and the scarcity of fresh water in arid regions have contributed to the depletion of groundwater. Artificial Groundwater Recharge (AGR) is an advanced strategy that contributes to combating water shortage issues. Limited efforts have been exerted to evaluate and demarcate AGR potential zones in the United Arab Emirates (UAE). The current study aims to delineate AGR potential zone mapping using the traditional analytical hierarchy process (AHP) and a hybrid deep learning model namely, Convolutional Neural Network-Xtreme Gradient Boosting (CNN-XGB) was used for the optimal prediction-based suitability assessment. A total of nine hydrogeological factors were considered for AGR mapping. First, the influence of each parameter was determined based on expert opinion and literature reviews for the AHP approach (0.007 consistency ratio). Second, a hybrid CNN-XGB model (90.8 % accuracy) predicted the AGR and non-AGR classes as part of binary classification and generated an AGR potential zone map. Moreover, the contributing factors were analyzed deeply for the AGR site selection to understand the intercorrelation, importance, and prediction interaction. Using both approaches, a comparative assessment was conducted in the eastern, central, and western parts of Sharjah. The AGR zone based on the CNN-XGB model achieved a precision of (0.8168), recall (0.7873), and F1-score (0.8018). The critical contributing factors for AGR mapping were found to be geology (20%), geomorphology (15%), rainfall (10%), and groundwater level (10%). The AGR map is expected to help explore new sites with potentially higher favourability to retain water, deal with water scarcity, and improve water management in the UAE.

## 1. Introduction

Globally, groundwater is a prime reserve of available freshwater that accounts for 30.1 % of the available groundwater (Richey et al., 2015). Due to rapid urbanization and excessive consumption patterns for socioeconomic needs, groundwater is becoming scarce in many parts of the world, specifically in arid and semiarid climatic countries (Richey et al., 2015). The scarcity of fresh groundwater resources has been a significant issue affecting all continents. Therefore, sustainable water management, reclamation, and awareness campaigns are crucial to address these issues. The abstraction of groundwater, mainly for irrigation, has led to ecologically problematic declines in groundwater tables. This process contributes to the problem of water scarcity around the world.

One of the sustainable management practices for groundwater adopted by many Arabic countries is artificial groundwater recharge (AGR). AGR can be beneficial for storing fresh water for potable use, especially in semiarid and arid nations. AGR can also be utilized to control seawater intrusion in coastal aquifers, regulate land subsidence caused by declining groundwater levels (GWLs), and maintain base flow in some streams. However, the evaluation and analysis of the recharging process become crucial because contaminant spreading is possible while

* Corresponding author.
*E-mail addresses:* ralruzouq@sharjah.ac.ae (R. Al-Ruzouq), shanableh@sharjah.ac.ae (A. Shanableh), rjena@sharjah.ac.ae (R. Jena), mkhalil@sharjah.ac.ae (M. Ali Khalil), mbgibril@sharjah.ac.ae (M.B.A. Gibril), Biswajeet.Pradhan@uts.edu.au (B. Pradhan), nhammouri@sharjah.ac.ae (N. Atalla Hammouri).

https://doi.org/10.1016/j.ejrs.2024.02.006
Received 5 June 2023; Received in revised form 19 October 2023; Accepted 14 February 2024
Available online 24 February 2024

recharging (Murad et al., 2007).

For decades, remotely sensed data and derivatives have been incorporated to determine the potential AGR sites for further field investigations (Mahmoud and Alazba, 2014; Sherif et al., 2021; da Costa et al., 2019; Shelar et al., 2022). For instance, high-resolution satellite imagery enables the recognition of surface features (streams and rivers) associated with groundwater storage and recharge (da Costa et al., 2019). Several thematic layers are incorporated to delineate AGR potential zones, including topographic slope, natural vegetation, geologic landforms, drainage pattern, density, and lineaments (Anbazhagan et al., 2005; Kazakis, 2018). Other parameters include precipitation, runoff curve number, geomorphology, total dissolved solids (TDS), elevation, and fractures (Anbazhagan et al., 2005; Selvarani et al., 2017). In recent decades, ML models have created powerful algorithms to automate tasks or improve processes, predict outcomes, and make better decisions based on past experiences (Huang et al., 2019; Pourghasemi et al., 2020).

Several studies have been conducted to map groundwater recharge prospect zones using various machines and deep learning algorithms. For instance, Huang et al. (2019) developed linear regression, multilayer perception (MLP), and Long Short-Term Memory (LSTM) models for groundwater recharge zone prediction. They compared the performance and estimated as 0.19, 0.11, and 0.12. Pourghasemi et al. (2020) estimated the groundwater recharge zones by introducing a new factor named the multiresolution index using three machine learning algorithms Support Vector Machine (SVM AUC = 0.963), Multivariate Adaptive Regression Spline (MARS = 0.962), and Random Forest (RF AUC = 0.987). Another work by Martinsen et al. (2022) developed a high-resolution groundwater recharge map using machine learning techniques in the pan-European region and achieved RMSE varying from 0.12 to 0.14. Table 1 summarizes the research carried out on AGR zonation in the recent past five to seven years globally using traditional techniques. Table 2 summarizes several statistical, dynamic, futuristic, machine learning (ML), and deep learning (DL) models used for AGR zonation and groundwater studies.

The United Arab Emirates (UAE) is located in a semiarid climatic condition, which was chosen to investigate potential sites of AGR as only a handful of research has been carried out on the UAE soil on AGR (Machiwal and Jha, 2015; Sherif et al., 2021). The UAE ranks amongst the top 10 countries globally with water scarcity problems (Al-Taani et al., 2021; Sherif et al., 2018). The water consumption per capita in the UAE is considered amongst the highest in the world. Given the limited annual rainfall (about 110–150 mm/year) (Sherif et al., 2014), the UAE depends on desalinating sea water and groundwater. Over the years, excessive groundwater abstraction (decreased from 238 $km^3$ in 1969 to around 10 $km^3$ in 2015) occurred in the UAE, leading to a substantial depletion in its levels (Sherif et al., 2014).

The literature review suggests that the major problem is the scarcity of fresh groundwater resources due to the continuous groundwater abstraction in the UAE. Thus, this research aims to develop a GIS-based analysis (AHP) and hybrid deep learning (CNN-XGB) to identify, update, and compare the potential AGR sites in Sharjah. Therefore, this study intends to address the research questions as follows: (i) what are the advantages of the developed hybrid model over AHP? (ii) What is the threshold of factors to attain accurate results? and (iii) how the factors that interact in the current AI model helped to obtain excellent output. To achieve this, the main objectives are as follows: i) evaluate the key factors affecting AGR in semiarid regions, ii) develop and validate an AGR zonation map for Sharjah City, UAE, and iii) Compare the decision-making method (AHP) with hybrid deep learning (CNN-XGB) model. The main contribution deals with the uniqueness of the study area (Sharjah, UAE) characterized by low groundwater storage and the first-time implementation of the CNN-XGB hybrid model in AGR mapping. The novelty of this paper deals with a recently developed hybrid CNN-XGB model for more clarity on feature learning and optimal prediction-based suitability mapping. The CNN-XGB model implements auto-feature learning and predicts classes more accurately than the two individual models. However, the contributing factors were analyzed deeply for the AGR site selection to understand the intercorrelation, importance, and prediction interaction. In the end, the threshold of factors was derived for the AGR and non-AGR mapping in Sharjah. Therefore, the lessons learned from exploring AGR sites for Sharjah might benefit similar studies for other areas.

**Table 1**
Review of literature for recent AGR studies using the AHP technique.

| References | Location of study area | Factors used | Techniques used |
|---|---|---|---|
| Zaidi et al., (2015) | North-western part of Kingdom of Saudi Arabia | Slope, Soil Type, Vadose Zone Thickness, TDS, Water Bearing Formation, LU/LC | Boolean logic |
| Agarwal and Garg, (2016) | Loni and Morahi watersheds, Unnao and Rae Bareli district, Uttar Pradesh, India | Rainfall, Geology, Soil Type, Drainage Density, LU/LC, Geomorphology, Slope | AHP. |
| Machiwal and Singh, (2015) | Udaipur, India | Geomorphology, Soil Type, LU/LC, Elevation, Slope, Drainage Density, Transmissivity, Groundwater Fluctuation | AHP and Boolean logic |
| Chandramohan et al., (2017) | Palani Taluk, Dindigul District, Tamil Nadu, India | Geomorphology, Lineament, Geology, LU/LC, Groundwater Depth Level Data Collected From Existing Bore Wells, Rainfall, Slope, Soil type | WLC. |
| Farhadian et al., (2017) | Kerman Province, Iran | Precipitation, Vegetation, Distance from Roads, Soil Type, Distance From Rivers, Geology, Terrain Slope, LU/LC | Nash Conflict Resolution Method and AHP. |
| Sherif et al., (2021) | United Arab Emirates (UAE.) | Infiltration Rate, Soil Type, Slope, Geology, Hydrogeology, Rainfall, Harvesting Infrastructure, Water Table | recharge potential method & water table fluctuation method |
| Costa et al., (2019) | Jequitiba River basin, Minas Gerais, Brazil | Rainfall, Hydrometric, Relief, LU/LC, Soil Type, Lithology | WO & Hydrograph method |
| Mokarram et al., (2021) | Shiraz watershed, Iran | Slope, Lithology, Drainage density, Distance to fault, Precipitation, Elevation, LU/LC | AHP, Mathematical Dempster-Shafer theory |
| Abijith et al., (2020) | Ponnaniyaru, Tamil Nadu, India | Geomorphology, Geology, LU/LC, Slope, Soil, Drainage density, Rainfall | A.H.P., Multi-influence factor (MIF), WO. |
| Xu et al., (2021) | Xiong, China | Distance to canal, Water quality, Slope, Vadose infiltration, Drainage density, Groundwater level, Aquifer hydraulic conductivity, Aquifer thickness, Soil, Distance to sensitive areas | AHP. |

**Table 2**
Several statistical, dynamic, futuristic, machine learning (ML), and deep learning (DL) models are used for groundwater vulnerability, potential, and recharge zone mapping.

| Model/Technique for AGR Zone Mapping | Chronology | Advantages | Disadvantages |
|---|---|---|---|
| Statistic Models | | | |
| DRASTIC | Aller and Thornhill, 1987; Fannakh and Farsang, 2022 | Relatively simple and widely used. | Limited spatial and temporal dynamics. |
| SINTACS | 2007 | Improved accuracy over DRASTIC, better data utilization | Determining parameters can be complex and data-intensive |
| GOD Method | Expósito et al. (2010) | Incorporates multiple factors, considers spatial variability | Calibration can be challenging, data requirements |
| Dynamic Models | | | |
| HELP Model | 1994 | Simulates the hydrological cycle, dynamic nature | Data and parameter requirements, complexity |
| MODFLOW | Harbaugh et al. (2000) | Comprehensive, physically-based model. | Requires extensive data and expertise. |
| SWAT | Neitsch et al. (2011) | Captures the hydrological cycle. | Data-intensive, complex to set up. |
| Machine Learning (ML) | | | |
| Random Forest Linear regression | Breiman (2001) Huang et al. (2019) | Handles complex, non-linear relationships. | Prone to overfitting with small datasets. |
| Support Vector Machine (SVM) | Lee et al. (2018) | Effective for classification tasks. | Sensitivity to hyperparameters. |
| Deep Learning (DL) | | | |
| Convolutional Neural Networks (CNNs) | Madanan et al. (2021) | Can capture spatial features from images. | Large datasets and computational power. |
| Muli-Layer Perceptron, Recurrent Neural Networks (RNNs), LSTM Networks | Raihan et al. (2022)Huang et al. (2019) Zaresefat et al. (2023) | Useful for time series data. | May struggle with long sequences, data requirements. |
| GANs for Data Augmentation | Ongoing research | Enhances data quality, better training data | Complex to implement, computationally intensive |
| Futuristic Models | | | |
| Scenario-Based Modeling (Various) | Current research | Adaptable to future climate scenarios. | Uncertainties in long-term predictions. |
| Climate Change Projections (Various) | Ongoing research | Incorporates future climate projections. | Uncertainty in climate models. |
| Artificial Intelligence & IoT-based Models | Ongoing research | Real-time monitoring, predictive capabilities | Evolving technology, data security, cost |

## 2. Materials and methods

### 2.1. Study area

The Emirate of Sharjah, UAE, is located between approximately 25.2550° North latitude and 55.3188° East longitude to 25.4372° North latitude and 55.6112° East longitude, encompassing the entire emirate. The geographical extent of the Emirate of Sharjah encompasses an area of approximately 2,590 sq.km (about 1000 sq. miles) (Al-Busaid, 1995). The study area is located in Sharjah, which is one of the UAE emirates with the highest population of 1,628,932 and density (540.71/sq.km) (Sherif et al., 2014). Sharjah is located between Dubai and Ajman cities, as shown in Fig. 1. Given the high water demand in Sharjah, the Government has taken the issue of water sustainability seriously and initiated a pilot AGR study in the Nazwa area (Fig. 1) (Sherif et al., 2014). Sharjah's hydro-geological character is influenced by a complex aquifer system, with both shallow and deep aquifers, including unconfined and confined aquifers (Bruno and Vesnaver, 2021). The predominant soil type in Sharjah is sandy soil, often mixed with gravel and silt. Sharjah's topography is relatively flat, with gentle slopes throughout the emirate. The predominantly flat landscape impacts surface water runoff and natural groundwater recharge rates (Pande et al., 2022). All these factors influence both natural recharge processes and the success of artificial recharge initiatives.

Nazwa is a small village in the southwestern part of UAE characterized by limestone outcrop and surrounded by desert (Sherif et al., 2014). The study region has an arid climate with extremely high temperatures during summer (close to 50 °C) and pleasant winter (from as low as 12 °C to 25 °C) (Seasons of the UAE 2020) (Alsharhan and Rizk 2020a). The precipitation pattern of the UAE is highly irregular and varies from 60 mm to approximately less than 200 mm in the mountains in the northeastern part of the country. More than 70 % of the land composition is covered with sand (Alsharhan and Rizk 2020a). Vegetation is scarce in the country, and the surface freshwater is minimal; thus, the urge sustains to depend on groundwater.

### 2.2. Methodology framework

Fig. 2 presents the methodological flowchart utilized to demarcate AGR zones in Sharjah. A total of nine thematic layers were considered based on an extensive literature review, expert opinion, and data availability. These thematic layers are precipitation, elevation, drainage density, geomorphology, geology, groundwater levels (GWL), total dissolved solids (TDS), lineament density, and distance to residential areas. The current study did not use these factors (hydraulic conductivity, aquifer thickness) due to data unavailability. The layers were prepared based on the natural break technique scaled at 30 m resolution (Agarwal and Garg, 2016). Accordingly, we utilized several data sources to develop these layers in a GIS environment.

In this study, geospatial multicriteria analysis and a hybrid machine learning model were used to delineate the potential AGR zones. First, we utilized a geospatial multicriteria analysis to derive weights for each parameter and applied overlay analysis to map the potential zones of AGR in Sharjah. The modeling approach relied on each thematic layer's influence on the possibility of locating a suitable AGR zone. Second, a machine learning model that integrates a convolutional neural network (CNN) and extreme gradient boosting (XGBoost) (CNN–XGBoost) was also used to delineate AGR prospect sites. The hybrid deep learning model was trained using the previously published groundwater potential map based on arbitrary point selection. The AGR maps were categorized into five classes: very high, high, moderate, low, and very low.

### 2.3. Description of the thematic layers

Precipitation is the primary factor for any water resource study. UAE receives a mean rainfall of 80–100 mm/yr (Sherif et al., 2021). This

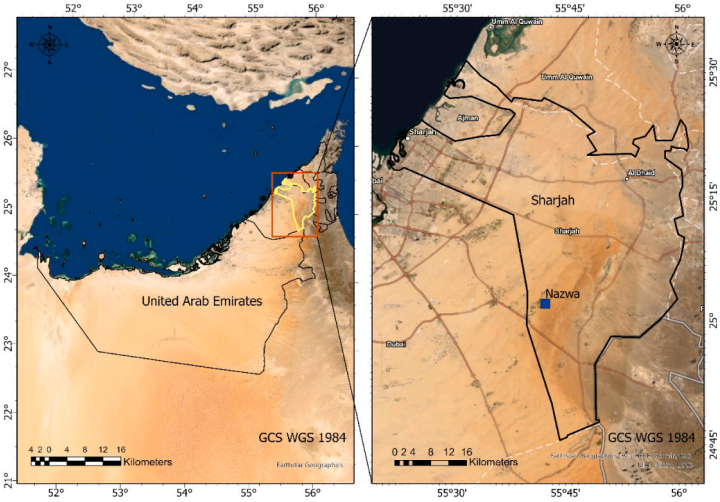


**Fig. 1.** Location map of the United Arab Emirates and the Emirate of Sharjah.

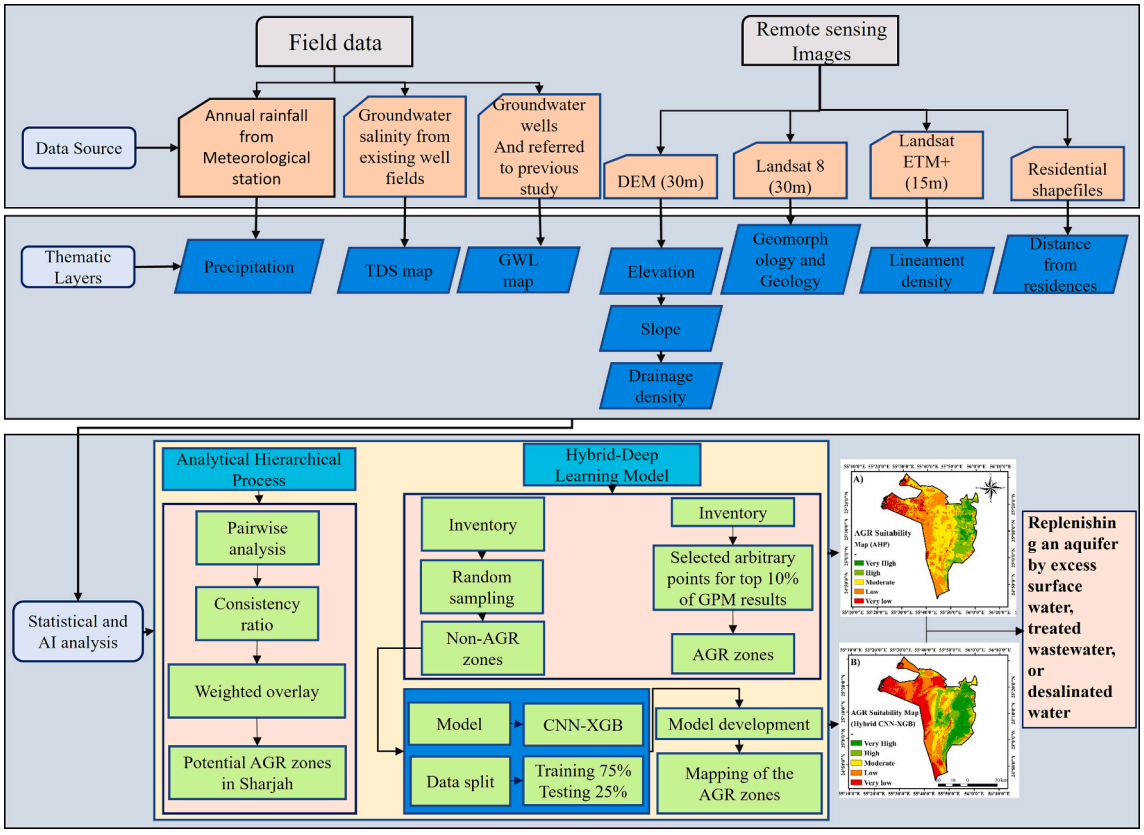


**Fig. 2.** Methodology framework for AGR suitability mapping in the Emirate of Sharjah.

thematic layer was prepared by using historical rainfall records (2003–2017) from 34 stations within the UAE, adopted from the National Centre of Meteorology, UAE as shown in Fig. 3A. Higher rainfall is more suitable for AGR potential zones due to the higher availability of water for recharging.

Elevation is expressed in meters. The highest elevation in Sharjah is approximately 413 m (Fig. 3B). The lower elevation is more suitable for AGR potential sites to simplify the construction and accessibility of the supplied water (Machiwal and Singh, 2015).

Drainage density has a negative correlation with groundwater recharge (Fig. 3-C). Drainage density has a direct relation with slope and, therefore, is inversely proportional to permeability (Selvarani et al., 2017; Al-Ruzouq et al., 2018). Generally, low permeability occurs due to steep slopes where the drainage density is high, thus generating more runoff and less infiltration (Selvarani et al., 2017).

Geomorphological features are classified as structural hills, residual hills, piedmont alluvium, high dunes, low dunes, flood plains, braided bars, vegetation, and urban areas, as mentioned in previous studies as shown in Fig. 3-D (Bhunia, 2020). The presented map was developed by a joint team between the Sharjah Electricity, Water and Gas Authority (SEWA) and Boston University (Al-Ruzouq et al., 2019a). High dunes are suitable for AGR implementation.

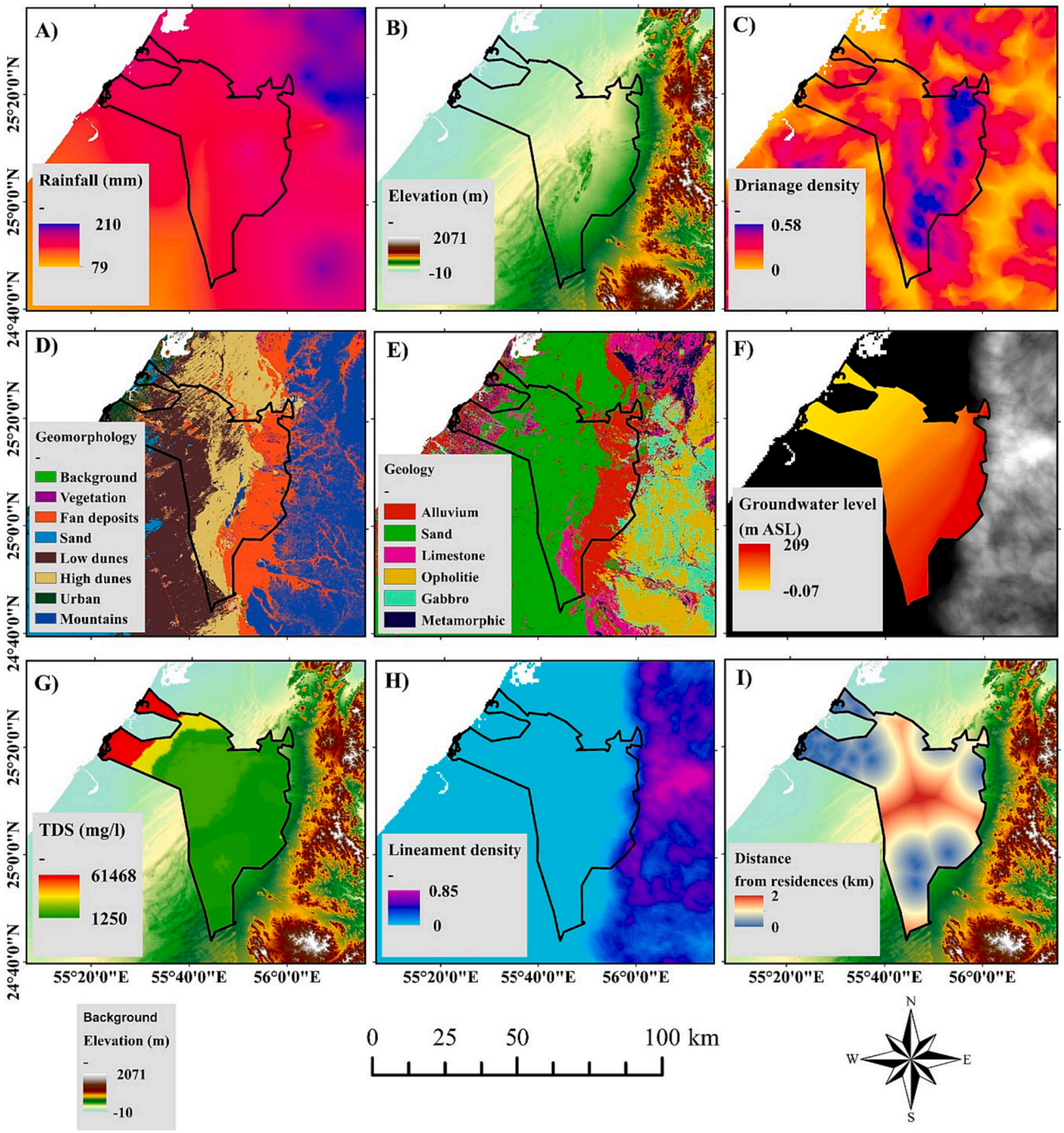


**Fig. 3.** Primary thematic layers were used for AGR suitability mapping for both AHP and CNN-XGB models.

Geology formations are categorized into consolidated and unconsolidated formations. The classes define the water holding capacity and the feasibility of injecting artificial groundwater. Alluvium, specifically the newer one, is considered to have a high water-holding capacity (Alsharhan and Rizk 2020b; Sherif et al., 2021).

Components of rocks, such as carbonates, sulfates, chlorides, and iron, are highly soluble in water and can lead to aquifer dissolution if they are ignored during recharging (Ghayoumian et al., 2005). More than 60 % of the area is covered with desert sand for this thematic layer (Fig. 3E).

Groundwater level is the upper layer of the saturated zone under the ground (named the water table). The AGR is infeasible in areas with a depth of water table less than 3–4 m (Riad et al., 2011). This study utilized conventional in-situ data on available groundwater levels (Fig. 3F). Also, a potential groundwater map was referred to from previous studies of Sharjah (Al-Ruzouq et al., 2019a). The groundwater level increases with an increase in precipitation as Sharjah is dry land and runoff is low (Fig. 4A).

Total dissolved solids help to define water quality. As shown in Fig. 3G, Sharjah appears in the Arabian Gulf, which poses a significant challenge given sweater intrusion (Murad et al., 2007). In Sharjah, the TDS level decreases with rainfall, and the groundwater level increases because the proportion of TDS with the water level goes down (Fig. 4B). The TDS map by the UAE Ministry of Water and Environment (2015) was referred to prepare the TDS layer shown for Sharjah using digitization and georeferencing by selecting 300 observation points.

Lineaments are naturally occurring linear features formed from faults or large fractures. They are potential sites for the accumulation of

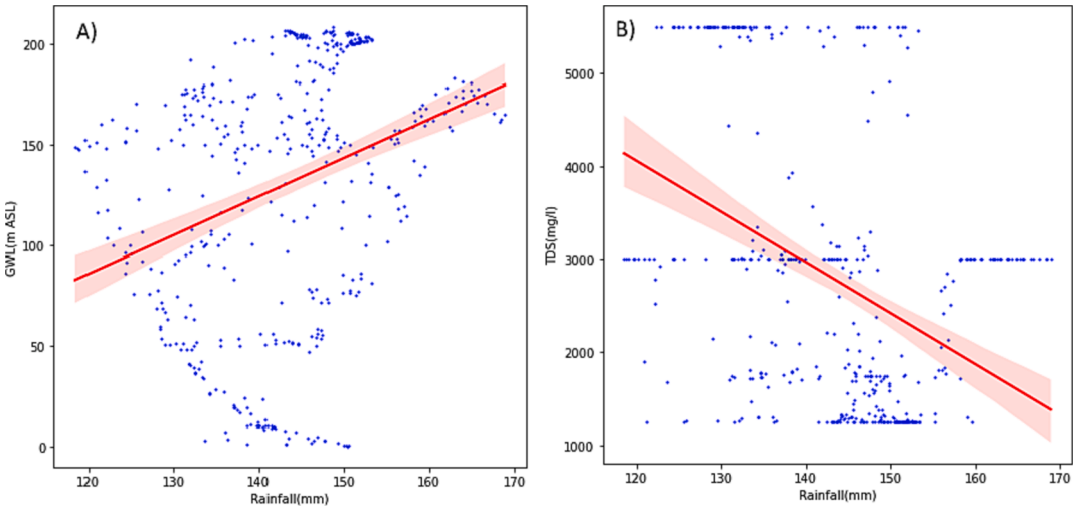


**Fig. 4.** Regression analysis of rainfall with respect to: (A) groundwater level, and (B) TDS (mg/l).

groundwater to help increase the water table levels downstream (Selvarani et al., 2017). The lineament map of Sharjah was extracted from Landsat 8 ETM+, and then Euclidean distances were calculated (Fig. 3H).

Distance from the residence explains how far each location is from urban areas, as shown in Fig. 3I. This layer is essential because AGR cannot be constructed at the heart of any residence cluster.

### 2.4. Geospatial multi-criteria analysis

Weighted overlay (WO) helps to organize all thematic layers with the highest percentage influence for most priority layers. Each layer provides specific information for groundwater holding and recharging capacity. The conclusive output map for AGR categorically shows areas with the highest to least probable sites. The technique is expressed as follows in Eq. (1):

$$S = Æ©\,(w_i * x_i) \tag{1}$$

where $w_i$ is the weight of the $i$th factor, $x_i$ is the criteria score of the class of factor $i$, and S is the suitability index for each pixel.

The AHP technique is used to assign weights to the parameters based on their importance to the groundwater recharge process (Kazakis et al. 2018; Moharir et al., 2023).

#### 2.4.1. Assigning weights to the parameters

The weights of the parameters were considered concerning various literature on AGR demarcation in semiarid and arid areas. Referring to these papers the experts' opinions were considered for the pairwise analysis and were compared with the literature to ensure the estimated weights are lying within the range. Satty's intensity scale was used to assign values based on the importance of factors according to experts' opinions (Chaudhari et al., 2018). A total of 18 experts were chosen from multidisciplinary expertise from industry and local UAE universities. The weights of all the nine parameters are placed in a square matrix considering all the diagonal elements as 1(Chaudhari et al., 2018; Al-Ruzouq et al., 2019b; Kazakis et al. 2018). The relative importance is examined by the principal Eigenvalue and the normalized eigenvector in the matrix (Chowdhury et al., 2010; Skilodimou et al., 2019; Jena et al., 2020).

#### 2.4.2. Pairwise comparison matrix

The consistency ratio (CR) has been calculated to verify the square matrix's consistency. The consistency ratio (CR) was calculated to verify the consistency of the square matrix. The CR can be derived by using the following Eqs. (2) to (4)(Saaty n.d. (1990)):

$$CI = \frac{\lambda max - n}{n-1} \tag{2}$$

$$\text{Random Consistency Index } (RI) = \frac{1.98 \times (n-1)}{n} \tag{3}$$

$$CR = \frac{\text{CI}}{\text{RI}} \tag{4}$$

where $CI$ is the consistency index, $RI$ is a randomized index, $CR$ is the consistency ratio, and $n$ is the order of the matrix. The RI is predetermined and depends on the number of criteria or alternatives being compared. Table 3 shows the RI values available for different matrix sizes. Table 4 represents the pairwise comparison matrix of the research. The CR was calculated as 0.007<0.01, which validates the criteria to be a good model.

The considered weights were then assigned utilizing weighted overlay (WO) in ArcGIS. Each class was classified under Natural Breaks (Jenks), and individual ranks were assigned, as shown in Table 5.

### 2.5. Hybrid CNN-XGB model

The architecture of ConvXGB has several layers as described below (Thongsuwan et al., 2021): The model goes deeply into the features to automate tasks or improve processes, predict outcomes, and make decisions based on past experiences.

Input and output layer: Input and output layers are responsible for model input and output. In the input layer, the assumption of a training data set, X, characterized by a set of tuples, $(x_j, y_j)$, where j is the dataset index, $x_j$ is a $\sqrt{N} \times \sqrt{N}$ feature matrix and $y_j$ is the class label.

Data pre-processing layer: A common square matrix was used to make a flexible system and to handle multi-source data for the tensors in the convolution layers (Madanan et al., 2021). In this layer, padding can be applied if the input data don't follow a standard square dimension, $\sqrt{N} \times \sqrt{N}$.

Input data scaling: To normalize the pixel values (e.g., [0, 1] or [-1, 1]) scaling was performed. TensorFlow and Keras libraries were utilized to perform scaling (Madanan et al., 2021).

Pooling Layers: After convolutional layers, pooling layers were added to reduce spatial dimensions. Common pooling methods include Max Pooling and Average Pooling.

Flattening: Flattening generally flattens the feature maps into a 1D vector and feeds it into the XGBoost model.Convolutional filters were attached to the CNN model, where h denotes length. The $X_{[i:i+h]}$ signifies

**Table 3**
Random consistency index used for different matrix sizes in the AHP method.

| n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| R.I. | 0 | 0 | 0.52 | 0.89 | 1.12 | 1.26 | 1.36 | 1.41 | 1.46 | 1.49 |

**Table 4**
Pairwise-comparison matrix for ranking the factors.

| | Precipitation | Elevation | DSD * | Geomorphology | Geology | G.W.L. ** | TDS | LMD + | RED ++ |
|---|---|---|---|---|---|---|---|---|---|
| Precipitation | 1 | 2 | 0.9 | 0.7 | 0.3 | 0.9 | 0.9 | 2 | 0.9 |
| Elevation | 0.5 | 1 | 0.5 | 0.3 | 0.3 | 0.5 | 0.5 | 0.9 | 0.5 |
| DSD* | 1.1 | 2 | 1 | 0. 7 | 0.3 | 0.9 | 0.9 | 2 | 0.9 |
| Geomorphology | 1.5 | 3 | 1.5 | 1 | 0.8 | 1.5 | 1.5 | 3 | 1.5 |
| Geology | 4 | 4 | 4 | 1.3 | 1 | 2 | 2 | 4 | 2 |
| Groundwater Level | 1.1 | 2 | 1.1 | 0. 7 | 0.5 | 1 | 0.9 | 2 | 0.9 |
| TDS | 1.1 | 2 | 1.1 | 0. 7 | 0.5 | 1 | 1 | 2 | 0.9 |
| Lineament Density | 0.5 | 1 | 0.5 | 0.3 | 0.3 | 0.5 | 0.5 | 1 | 0.5 |
| Residential ED | 1.1 | 2 | 1.1 | 0. 7 | 0.5 | 1.1 | 1.1 | 2 | 1 |

* DSD: Drainage Stream Density.
** GWL: Groundwater Level.
\+ LMD: Lineament Density.
\++ RED: Residential Euclidean Distance.

the merging of all data $X_{[i]}$ to $X_{[i+h]}$. Therefore, the feature $C_i$ can be utilized with a filter F by employing the following Eq. (5).

$$C_i = \sum_{k,j} (X_{[i+h]})_{k,j} . F_{k,j} \quad (5)$$

XGBoost layer: An XGBoost model was implemented as a hidden layer at the end of the four convolutional layers. This model is an ensemble model characterized by a gradient boost algorithm, which helps improve the model performance and generate a strong hybrid model with CNN (Chen et al., 2015). Generally, the ensemble models offer excellent outputs compared to a single model (Fig. 5).

The model's predicted values $\mathring{A}\cdot_i^{(t)}$ can be expressed as $\mathring{A}\cdot_i^{(t)} = \mathring{A}\cdot_i^{(t-1)} + f_2(x_i)$, where $\mathring{A}\cdot_i^{(t)}$ represents the predicted value of t trees on sample $x_i$. This procedure establishes the t *th* tree as shown in Eq. (6).

$$p_i = \frac{1}{1 + e^{-\mathring{A}\cdot_i^{(t)}}} \quad (6)$$

The formula that is applicable in determining the classification output is to achieve the probability by transforming the ultimate predicted value $\mathring{A}\cdot_i^{(t)}$ of the sample. When $p_i \geq 0.5$ the probability is 1, otherwise, it is 0.

Softmax: A softmax regression layer s linked to the obtained outputs of the preceding layer $x \in R^m$, which yields the highest probability values as $\mathring{A}\cdot \in [1, K]$ and can be represented as shown in Eqs. (7) and (8):

$$\mathring{A}\cdot = \text{argmax}_j P(y = j|x, w, a) \quad (7)$$

$$\text{argmax}_j \frac{e^{x^t_{wj+aj}}}{\sum_{k=1}^{k} e^{x^t_{wj+aj}}} \quad (8)$$

where wj denotes the weight vector of class j, from which the dot product can be computed using the input, whereas aj is the bias.

### 2.5.1. Feature learning

Feature extraction is a crucial stage when training a machine learning model, which directly influences the accuracy and predictability of the developed model (Madanan et al., 2021). The major advantage of deep learning algorithms lies in the automatic extraction of additional abstract and effective features during the training process, which provides an automatic and end-to-end classification approach (Huang et al., 2019). This led to the development of a hybrid CNN–XGBoost model, which swapped the fully connected layer with an XGBoost classifier. In this study, a convolutional neural network (CNN) and extreme gradient boosting (XGBoost) (CNN–XGBoost) were integrated to delineate AGR potential sites, as shown in Fig. 5.

A CNN model was developed and the extracted features by the CNN were fed to the XGBoost algorithm followed by a fully connected layer. The output units in the last layer are the estimated probabilities classified as 0 (AGR Non-suitability) and 1(AGR suitability). The output probability is considered by the SoftMax activation function. The output of the hidden layers only makes sense to the CNN model; however, these values are considered features for XGBoost classifiers. The training and validation prediction values are input feature vectors to fit the XGBoost model. In this study, the XGBoost model was trained for 100 iterations, and early stoppings were set to avoid overfitting. The predicted values for the testing dataset are used for classification reports and final mapping.

### 2.5.2. Data preparation and parameter selection

In this study, various data sources including remote sensing (i.e., DEM, Landsat 8, LandsatETM+, etc.) and observed data (i.e., meteorological data from Nazwa) obtained to delineate AGR potential sites (Machiwal and Singh, 2015). The top 10 % of the very high and high areas of previously published groundwater potential maps were used for training data preparation given the scarcity of the AGR data in Sharjah (Al-Ruzouq et al., 2019a). Datasets utilized for correlation were converted into multi-values points (Fig. 6).

The CNN-XGB model was developed and randomly trained with 75 % (training) and 25 % (validation) of the total data. During training, 2,542,194 samples were prepared, out of which 1,141,749 points were for Non-AGR and 1,400,445 for AGR using the dataset mentioned above. However, 75 % (1,906,646) and 25 % (6,35,548) points of the total data points were selected for training and validation purposes, respectively. Finally, the test was conducted with 2.5 million points to derive the AGR suitability zone map. The assigned model hyperparameters include the sigmoid optimizer, batch size (100), and verbose (1). Then, predicted values were converted into pixels after post-processing, and the spatial AGR suitability map was generated. In the end, a correlation among the top 6 factors was derived as presented below, and the detailed correlation values are shown in Table 6 and Fig. 6. The top four factors are Geology, Geomorphology, Rainfall, and Groundwater level which are comparatively different from the AHP correlation values. The top two factors according to AHP are geology and geomorphology whereas,

**Table 5**
Ranks and weights estimated for thematic layers and their subclasses.

| Thematic layer | Weight | Classes | Ranks |
|---|---|---|---|
| Precipitation (mm) | 10 % | 74.84–79.14 | 1 |
| | | 79.15–82.46 | 3 |
| | | 82.47–85.28 | 5 |
| | | 85.29–87.82 | 7 |
| | | 87.83–92.90 | 9 |
| Elevation (m) | 5 % | 0–56 | 9 |
| | | 57–109 | 7 |
| | | 110–156 | 5 |
| | | 157–201 | 3 |
| | | 202–413 | 1 |
| Drainage stream density (km/km$^2$) | 10 % | 0.01–0.14 | 9 |
| | | 0.15–0.23 | 7 |
| | | 0.24–0.32 | 5 |
| | | 0.33–0.42 | 3 |
| | | 0.43–0.58 | 1 |
| Geomorphology | 15 % | Fan deposits | 9 |
| | | High dunes | 8 |
| | | Sand | 7 |
| | | Low dunes | 4 |
| | | Vegetation | 3 |
| | | Mountains | 2 |
| | | Urban | 1 |
| Geology | 20 % | Alluvium | 9 |
| | | Sand | 8 |
| | | Limestone | 5 |
| | | Ophiolite | 3 |
| | | Gabbro | 2 |
| | | Metamorphic | 1 |
| Potential Groundwater | 10 % | Very High | 9 |
| | | High | 8 |
| | | Moderate | 7 |
| | | Low | 6 |
| | | Very Low | 3 |
| TDS (mg/l) | 10 % | 658.68–973.01 | 9 |
| | | 973.02–1301.95 | 7 |
| | | 1301.96–1594.35 | 5 |
| | | 1594.36–2003.71 | 3 |
| | | 2003.72–2530.05 | 1 |
| Lineament density (km/km$^2$) | 5 % | 0–0.03 | 1 |
| | | 0.04–0.11 | 3 |
| | | 0.12–0.21 | 5 |
| | | 0.22–0.33 | 7 |
| | | 0.34–0.56 | 9 |
| Residential Euclidean distance (m) | 10 % | 0–647.61 | 1 |
| | | 647.62–1786.45 | 3 |
| | | 1786.46–2972.57 | 5 |
| | | 2972.58–4340.78 | 7 |
| | | 4340.79–8406.96 | 9 |

rainfall and groundwater level show equal importance.

#### 2.5.3. Classification analysis

Performance metrics are used to compute the accuracy of the classification. To appraise the predictive dimensions, the metrics used are as follows: recall (*TPR*), $F_1$-score ($F_1$), precision (*PPV*), support, and accuracy (*ACC*) (Aafaq et al., 2019). The mathematical representation of the performance mentioned above metrics is presented in Eqs. (9) to (12).

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN} = 1 - FNR \tag{9}$$

$$F_1 = 2 \times \frac{PPV \times TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN} \tag{10}$$

where TPR represents the hit rate can also be called a true positive rate while $F_1$-score is the harmonic mean of precision and sensitivity that can be expressed as:

$$PPV = \frac{TP}{TP + FP} \tag{11}$$

$$ACC = \frac{TP + TN}{P + N} \tag{12}$$

The Precision, Recall, F1 score, and ACC for all target cells are estimated. As shown in the CNN-XGBoost model, the hybrid approach demonstrates outstanding performance in groundwater recharge zone mapping with an accuracy of 90 %.

## 3. Results

### 3.1. Results based on the AHP approach

The AGR map for Sharjah was produced by applying the GIS overlay analysis based on the weights observed from the AHP. The contribution of all utilized factors remains highly debated in AGR mapping. The results, shown in Table 4, stipulate that geology, geomorphology, precipitation (mm), groundwater level, and TDS (mg/l) are the most valuable factors estimated as the mean effective weight of 20 %, 15 %, 10 %, and 10 %, respectively. On the other hand, elevation (m) and lineament density (km/km$^2$) are the least important factors, with a mean effective weight of 5 %. Fig. 7(A) shows the AGR zonation map that is classified into five classes: very high, high, moderate, low, and very low. Fig. 7(A) shows the extent of the area percentage for each of these classes. The lowest class is the very high class with only 3 % of the study area, whereas the highest area is moderate, which constitutes approximately 44 % of the study area. The urban center of the study area (the one close to the coast) has the lowest possibility of having an AGR site.

The central region of the study area is mixed with moderate and high classes. The very high class is located at the eastern boundaries of the study area. The driving force behind this condition is the high precipitation rates (relative to the study area). For further analysis, the resultant AGR zones as shown in Fig. 7(A) was validated through the existing AGR site in Nazwa in Sharjah. To understand the accuracy of the AHP results, an advanced artificial intelligence decision-making algorithm (CNN-XGB) was adapted to attain better accuracy of results.

### 3.2. Results based on CNN-XGB model

The AGR map was also developed based on the CNN-XGB hybrid deep learning-based binary classification technique. Fig. 7(B) shows the AGR map classified into five classes, and the extent of area percentage for all the classes is derived. The lowest class is the very high class with only 22 % of the study area, whereas the highest area is very high, where it constitutes approximately 40 % of the study area. This condition is mainly due to its dense residential areas and a high TDS value, given its proximity to the coast. The central region of the study area is mixed with low to high classes. The very high class can be observed in the eastern parts. The driving force behind this condition is the high precipitation rates and elevation. The accuracy and loss of the hybrid model are plotted in Fig. 8.

### 3.3. Performance results of CNN-XGB model

Proper training is a major component of deep learning models. The

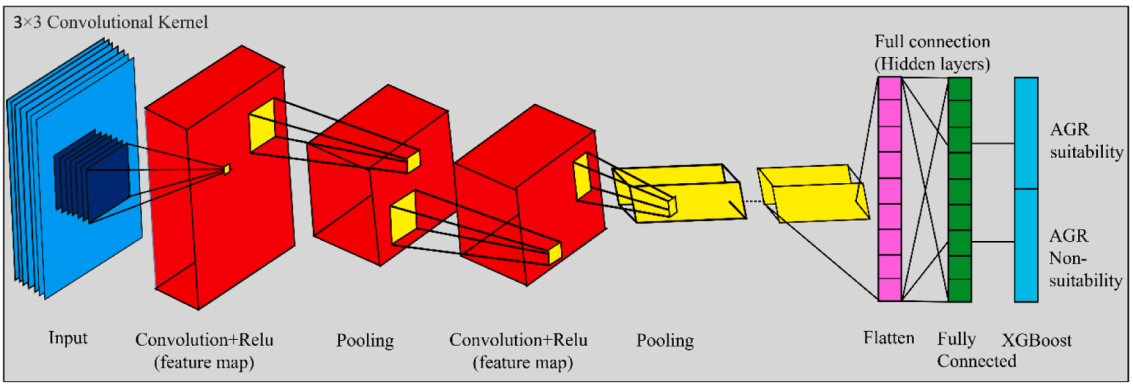


**Fig. 5.** Developed hybrid CNN-XGB model for AGR suitability zone mapping.

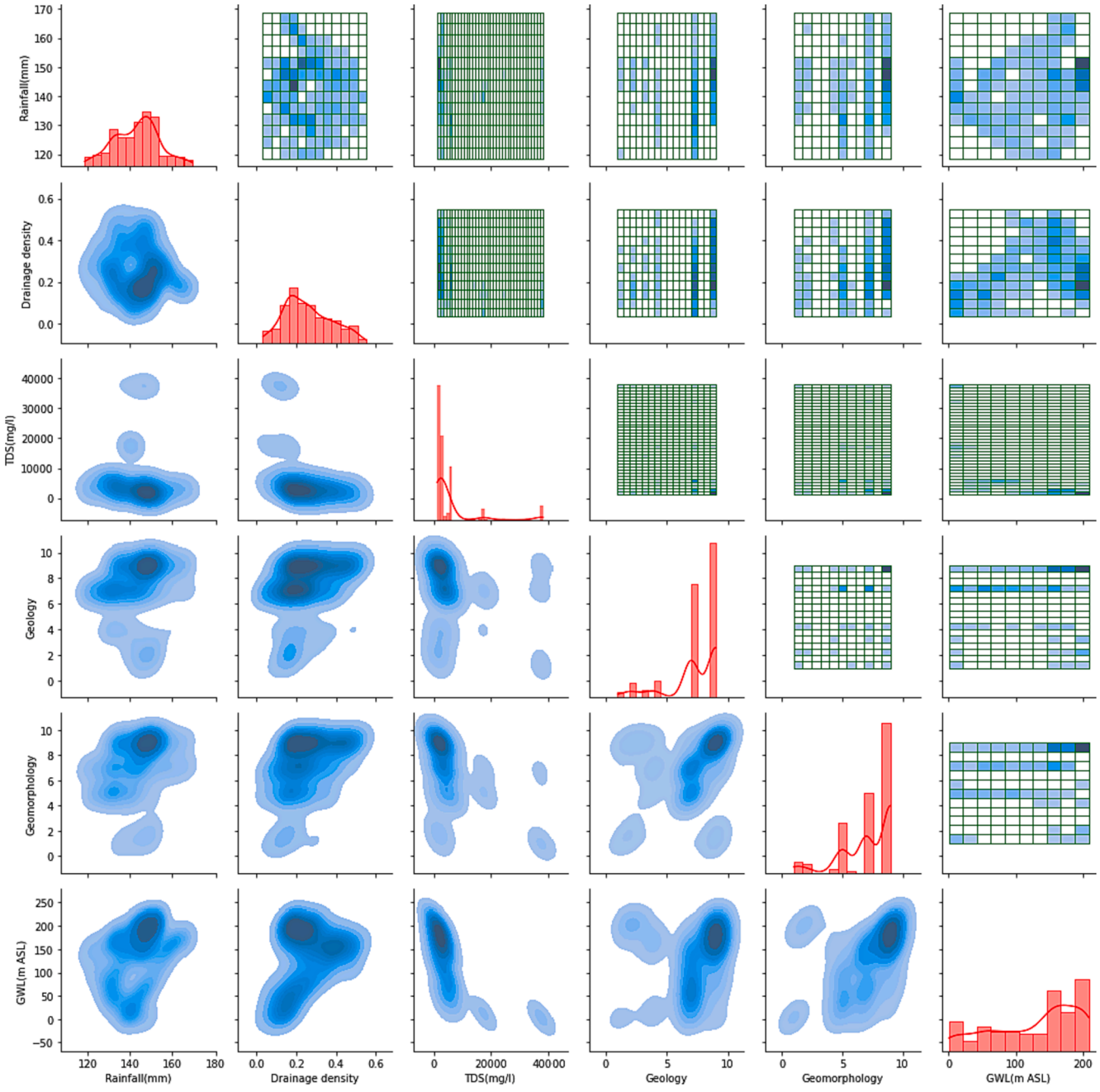


**Fig. 6.** Pearson correlation among top six factors based on hybrid CNN-XGB model.

**Table 6**
Correlation values for the top six factors according to CNN-XGB model.

| | Rainfall(mm) | Drainage density | TDS (mg/l) | Geology | Geomorphology | GWL (m ASL.) |
|---|---|---|---|---|---|---|
| Rainfall(mm) | 1.000000 | -0.104120 | -0.062427 | 0.156790 | 0.258264 | 0.313012 |
| Drainage density | -0.104120 | 1.000000 | -0.376844 | 0.291707 | 0.378057 | 0.361504 |
| TDS (mg/l) | -0.062427 | -0.376844 | 1.000000 | -0.232462 | -0.430490 | -0.644455 |
| Geology | 0.156790 | 0.291707 | -0.232462 | 1.000000 | 0.361709 | 0.169391 |
| Geomorphology | 0.258264 | 0.378057 | -0.430490 | 0.361709 | 1.000000 | 0.438604 |
| GWL (m ASL.) | 0.313012 | 0.361504 | -0.644455 | 0.169391 | 0.438604 | 1.000000 |

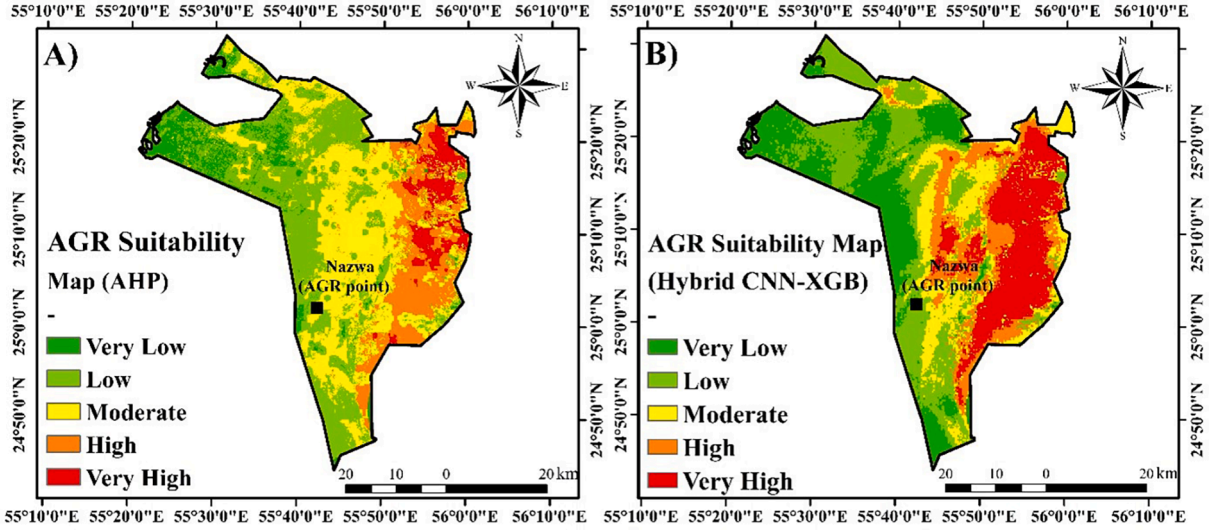


**Fig. 7.** (A) Developed AGR suitability maps based on AHP and (B) hybrid CNN-XGB model.

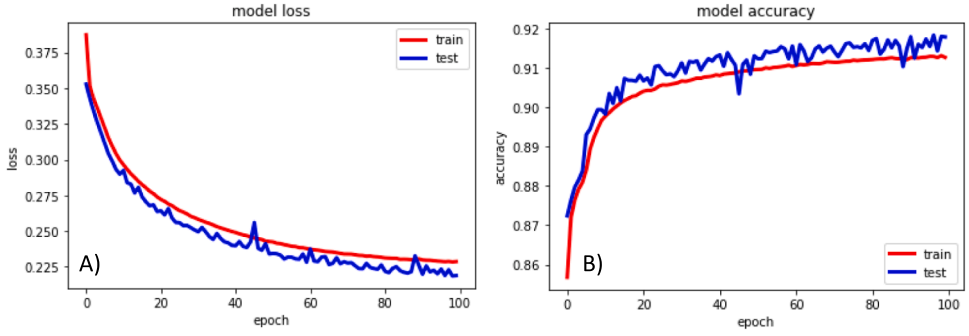


**Fig. 8.** (A) Developed AGR suitability maps based on AHP and (B) hybrid CNN-XGB model.

AGR assessment was performed using the CNN-XGB model based on the classifications of AGR (1) and non-AGR (0) points. In total, 31,500 points were derived using various factors, of which 15,700 are AGR points, while 15,800 are non-AGR points. The hybrid model recorded 90 % accuracy for the AGR zone mapping. The precision was estimated as 0.81, the recall was 0.78, F1-score was 0.80, and support was 88575. The estimated model loss was compared with respect to the epochs (Fig. 8A). Loss decreases with an increase in the number of epochs. The

obtained baseline error was 9.19 %, and the final testing accuracy was 90.81 % (Fig. 8B).

The classification report (Table 7) and the confusion matrix (Table 8) targeting two classes are presented, respectively. The false negatives in the confusion matrix did not follow any patterns. The result shows that the precision, recall, F1-score, and support for the non-AGR zone prediction are 0.93, 0.94, 0.94, and 286425, respectively. The precision and recall for the micro average are 0.87 and 0.86. However, the weighted average was summarized as 0.90 and 0.90, respectively. Table 3 and Table 5 present the details of the accuracy, error, and confusion matrices. The predicted zones were classified as very low to very high, indicating that the results are correct and similar to the AHP results.

## 4. Discussion

The study derived the suitability for artificial groundwater recharge (AGR) zones in the UAE using AHP and a hybrid CNN-XGB model. The study estimated that the northeastern parts are the most suitable regions for AGR. The density of AGR prediction is plotted where the model predicted more AGR points than non-AGR points (Fig. 9A and B). The values ranging between 0 and 0.25 will be considered very low AGR probability, while the values range between 0.75 and 1 will be regarded as having very high AGR probability (Fig. 9C) and moderate range (0.25 to 0.75) can be seen in (Fig. 9D). The prediction density indicates that the model well interprets the classes.

Eastern, central, and western parts of Sharjah were compared for both the AHP (Fig. 10A – C) and CNN-XGB (Fig. 10D – F) approaches. Eastern parts of Sharjah fall under very high suitability for both approaches. Central parts show low to high suitability, however, based on AHP, moderate to high, whereas more moderate zones can be observed in the CNN-XGB results. Results of the eastern region show similar results for both approaches, where the Nazwa region falls under moderate according to AHP and low suitability zone as per CNN-XGB model, respectively. Therefore, the Nazwa area might not be fully suitable for AGR.

The areas with optimal suitability are majorly concentrated in the eastern and northeastern parts of the Sharjah Emirate. A small portion of the central part of Sharjah is located under a highly suitable zone for AGR zonation. The elevation parameters within this zone have diversified values ranging from low (80 m) to high (190 m). TDS of 1000–1500 mg/l and groundwater level of 120–140 mASL are obtained. Other parts are observed as medium to low suitability zones. The southwestern part of Sharjah falls within low to moderately suitable zones. The regions mostly comprise low dunes, with moderate elevation ranging from 130 to 160 m, groundwater level of approximately 100–140 mASL, and drainage density of 0.13–0.29 km/km$^2$. The resulting AGR map in this study is based on local parameters and a very high area in this AGR map can be different from another very high class in other regions.

The AGR spatial probability was plotted against all nine significant factors presented in Fig. 11 for the CNN-XGB model. The thresholds for all nine factors such as Precipitation (mm), Elevation (m), Drainage Stream Density (km/km2), Geomorphology, Geology, Potential Groundwater, TDS (mg/l), Lineament Density (km/km2), Residential Euclidean Distance (m) (Fig. 11A – I). A high AGR possibility zone can be observed at high elevations because the groundwater movement occurs from hills towards the lowland therefore, the coastal areas might be the discharge zones. The most important factors are geomorphology, geology, rainfall density, and groundwater level. Fan deposits and sand in the mountain areas are suitable for AGR zone selection. Gabbro, limestone, sand, and alluvium are suitable geological conditions for AGR site selection. The rainfall threshold for the best AGR site is 142 to 155 mm. These plots show a comparative assessment of a possible threshold of factors that is responsible for 90 % (CNN-XGB accuracy). However, the interactions among the factors are responsible for the observed accuracy and the final output map.

**Table 7**
Classification report for hybrid CNN-XGB model.

| Classification report | Precision | Recall | F1-score | Support |
|---|---|---|---|---|
| Non AGR zone | 0.9349 | 0.9454 | 0.9401 | 286,425 |
| AGR zone | 0.8168 | 0.7873 | 0.8018 | 88,575 |
| Accuracy | | | 0.9081 | 375,000 |
| Macro avg | 0.8759 | 0.8663 | 0.8710 | 375,000 |
| Weighted avg | 0.9070 | 0.9081 | 0.9075 | 375,000 |
| Classification accuracy | 0.908059 | | | |

**Table 8**
Confusion matrix for the hybrid CNN-XGB model.

| Confusion matrix | Positive (1) | Negative (0) |
|---|---|---|
| Positive (1) | 270,788 | 15,637 |
| Negative (0) | 18,841 | 69,734 |
| 11719/11719–14 s 1 ms/step | | |
| 46875/46875–108 s 2 ms/step – loss: 0.2363 – accuracy: 0.9081 | | |
| [0.23634308576583862, 0.9080586433410645] | | |
| Baseline Error: 9.19 % | | |

To validate the result of the AGR map, we compared our output with the existing AGR pilot study implemented by SEWA, named the Nazwa AGR project (Fig. 10). The Nazwa AGR project was selected based on previous groundwater and hydrological studies in the study area. Nazwa area in Sharjah, UAE is the first cost-effective operational AGR system, which has the potential to recharge and store groundwater generated from the desalination plant (Sherif et al., 2014; Şener, 2021). As shown in Fig. 10, the Nazwa region falls within moderate-high suitable zones. The TDS of the zone is around 1200 mg/l. SEWA aims to expand Nazwa AGR to reach a 500 million imperial gallon capacity. Zaresfat et al. (2023) conducted a study utilizing artificial intelligence for the identification of optimal sites for AGR in the Iranshahr Basin and achieved an overall accuracy of 97 %. Huang et al. (2019) forecasted AGR through linear regression, multi-layer perceptron network (MLP), and long short-term memory (LSTM) techniques. The linear regression model was achieved a root mean square error of 0.19 whereas the RMSE for the MLP model was lower than 0.20. The deep learning model LSTM was found to have the best performance with RMSE less than 0.12. Comparatively the applied CNN-XGB model achieved great accuracy with excellent AGR suitability map.

## 5. Conclusions

The current study addressed the demarcation of potential AGR zones in Sharjah, UAE. The prime objective of the study was to develop and compare the results obtained from AHP and CNN-XGB approaches. Geology and geomorphology are the most decisive factors with the highest weights as per the AHP approach. However, based on the CNN-XGB model the most important factors are geology, geomorphology, rainfall, and groundwater level. The CNN-XGBoost model avoids overfitting, provides clarity in feature learning, and performs well in comparison to individual models. This model requires huge data, knowledge of topology, parameters, and training methods.

According to the results, the northeastern part of Sharjah is constituted as a high AGR suitability zone. The central and portions of the eastern parts of the study area show moderate-high and moderate suitability for AGR, respectively. A previous pilot study by SEWA conducted in Nazwa, Sharjah was considered to highlight the importance of the AGR suitability map. Nazwa lies within the moderate- as per the AHP and low zone as per the CNN-XGB-based AGR map, respectively. This research has several practical implications, such as helping deal with water scarcity, improving water management, storage management, and reserving stormwater. The outcomes of this research could help in preparing a management action plan to tackle the challenge of water

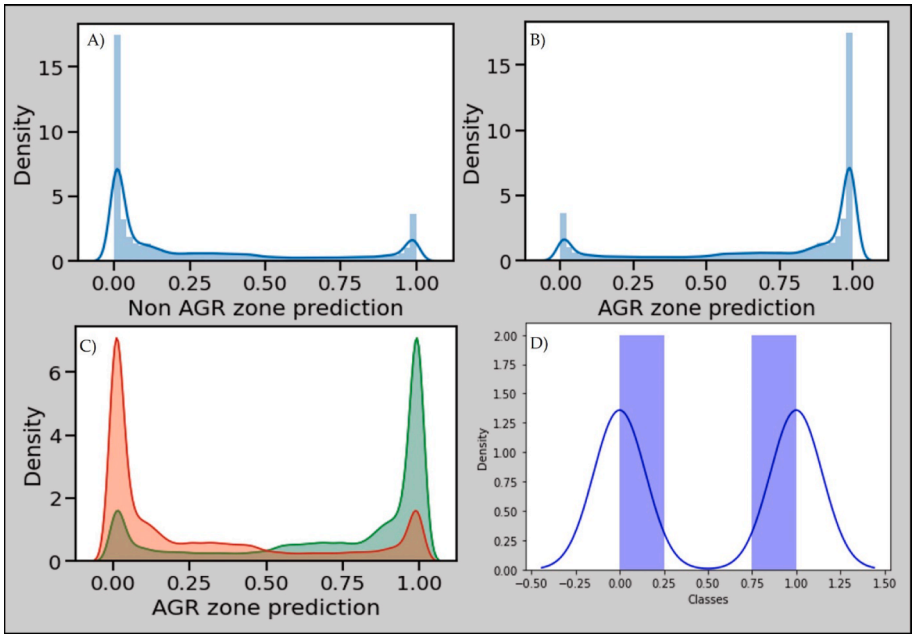


**Fig. 9.** Predicted class density for non-AGR zone (A), (B) AGR zone, and (C) comparison.

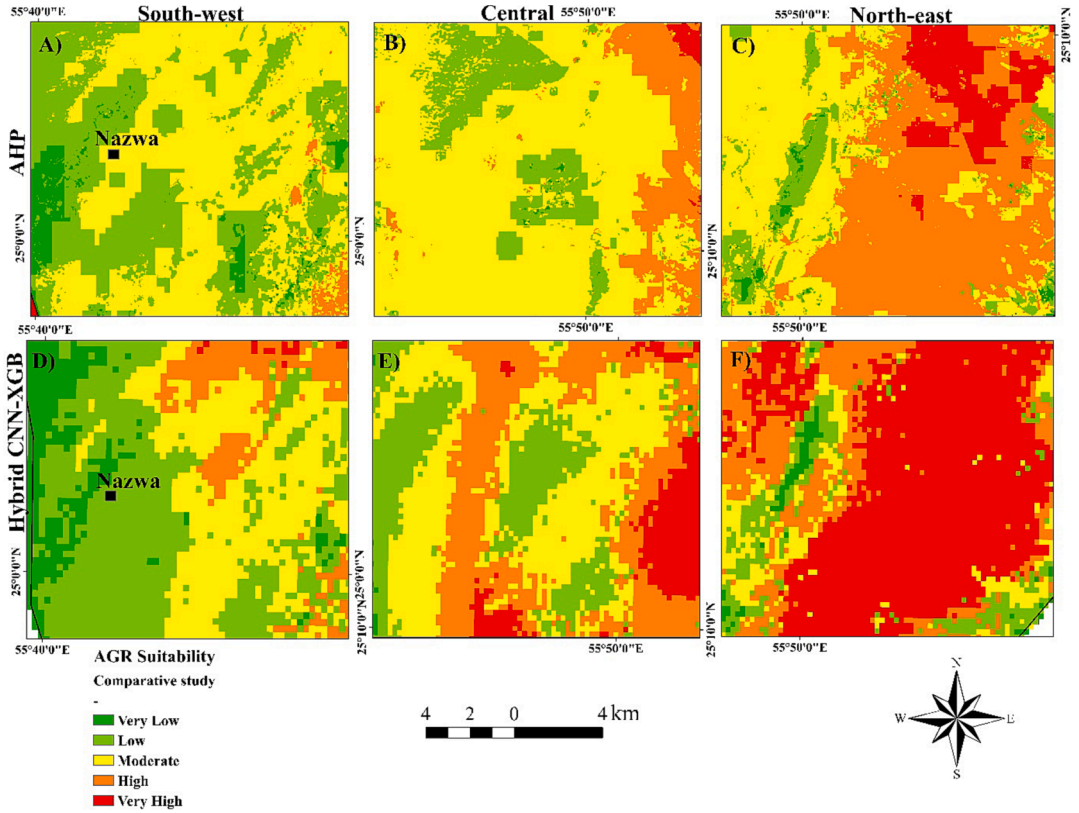


**Fig. 10.** Comparative study for AGR suitability maps based on AHP (A-Western, B-Central, C-Eastern) and hybrid CNN-XGB model (D-Western, E-Central, F-Eastern) in Sharjah.

resource storage for the long term. This method demands significant computational resources, and less suitable for limited resources. Moreover, it might struggle when dealing with small datasets, ultimately result in overfitting. To enhance its effective performance, regularization, data augmentation techniques can be adopted. Thorough selection of features and fine-tuning can help enhance the predictive power by minimising possible limitations. The future recommendation for AGR research leads to investigate how climate change, land use planning, and agricultural practices affect recharge zones. Leverage the capabilities of AI and ML methods to enhance modelling, prediction, and decision support when creating recharge zone maps. The chemical composition of rock types, underground water movement direction, complicated

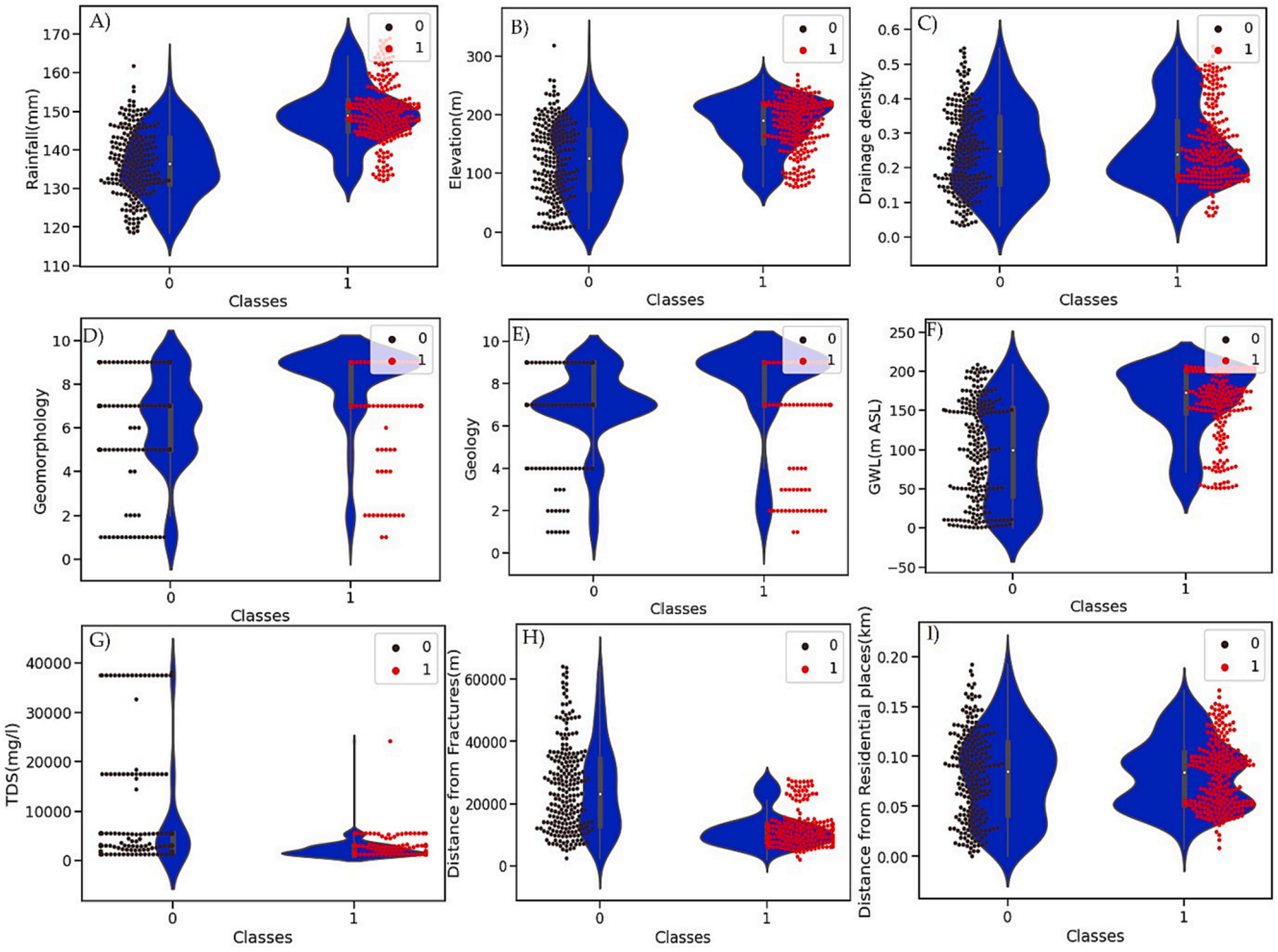


**Fig. 11.** Threshold of factors estimation for all the factors using hybrid CNN-XGB model.

environment, and of course the high installation cost need to be considered in future AGR research.

## CRediT authorship contribution statement

Conceptualization, R.A.R., R.J., M.A.K. and A.S.; methodology, R.J., R.A.R., M.B.G., and A.S.; software, R.J.; validation, R.A.R., R.J., and A. S.; formal analysis, S.M. and M.A.K.; investigation, R.A.R.; resources, A. S.; data curation, M.A.K., A.S. and R.A.R.; writing—original draft preparation, R.A.R., R.J., S.M. and M.A.K.; writing—review and editing, B.P., S.M., M.B.G., N.A.H.; visualization, R.J., R.A.R, and B.P.; supervision, R.A.R. and A.S.; project administration, R.A.R., and A.S.; funding acquisition, R.A.R.. All authors have read and agreed to the published version of the manuscript.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

The authors would like to acknowledge the funds received by the University of Sharjah (Grant No. 1902041134-P) that helped to facilitate this research.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejrs.2024.02.006.

## References

Aafaq, N., Mian, A., Liu, W., Gilani, S.Z., Shah, M., 2019. Video description: A survey of methods, datasets, and evaluation metrics. ACM Comput. Surv. 52, 1–37.

Abijith, D., Saravanan, S., Singh, L., Jennifer, J.J., Saranya, T., Parthasarathy, K.S.S., 2020. GIS-based multi-criteria analysis for identification of potential groundwater recharge zones-a case study from Ponnaniyaru watershed, Tamil Nadu, India. HydroResearch 3, 1–14.

Agarwal, R., Garg, P.K., 2016. Remote sensing and GIS based groundwater potential & recharge zones mapping using multi-criteria decision making technique. Water Resour. Manag. 30, 243–260.

Al-Busaid, K.M., 1995. *The* Gulf Cooperation Council security and the Arab order: Relations in question. United States International University.

Aller, L., Thornhill, J., 1987. *DRASTIC:* a standardized system for evaluating ground water pollution potential using hydrogeologic settings. Robert S. Kerr Environmental Research Laboratory, Office of Research and Development, US Environmental Protection Agency.

Al-Ruzouq, R., Shanableh, A., Omar, M., Al-Khayyat, G., 2018. Macro and micro geo-spatial environment consideration for landfill site selection in Sharjah, United Arab Emirates. Environ. Monit. Assess. 190, 1–15.

Al-Ruzouq, R., Shanableh, A., Merabtene, T., Siddique, M., Khalil, M.A., Idris, A., Almulla, E., 2019a. Potential groundwater zone mapping based on geo-hydrological considerations and multi-criteria spatial analysis: North UAE. Catena 173, 511–524.

Al-Ruzouq, R., Shanableh, A., Yilmaz, A.G., Idris, A., Mukherjee, S., Khalil, M.A., Gibril, M.B.A., 2019b. Dam site suitability mapping and analysis using an integrated GIS and machine learning approach. Water 11, 1880.

Alsharhan, A.S., Rizk, Z.E., 2020a. Climate Conditions and Their Impact on Water Resources. In: Water Resources and Integrated Management of the United Arab Emirates. Springer, pp. 115–175.

Alsharhan, A.S., Rizk, Z.E., 2020b. Water resources and integrated management of the United Arab Emirates. Springer Nature.
Al-Taani A.A., Nazzal Y., Howari F.M., 2021. Groundwater scarcity in the Middle East, in: Global Groundwater. Elsevier. 163–175.
Anbazhagan, S., Ramasamy, S.M., Das Gupta, S., 2005. Remote sensing and GIS for artificial recharge study, runoff estimation and planning in Ayyar basin, Tamil Nadu. India. Environ. Geol. 48, 158–170.
Bhunia, G.S., 2020. An approach to demarcate groundwater recharge potential zone using geospatial technology. Appl. Water Sci. 10, 1–12.
Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32.
Bruno, P.P.G., Vesnaver, A., 2021. Groundwater characterization in arid regions using seismic and gravity attributes: Al Jaww Plain. UAE. Front. Earth Sci. 8, 575019.
Chaudhari, R.V., Lal, D., Dutta, S., Umrikar, B., Halder, S., 2018. Weighted overlay analysis for delineation of groundwater potential zone: A case study of pirangut river basin. Int. J. Remote Sens. Geosci 7, 2319–3484.
Chen T., He T., Benesty M., Khotilovich V., Tang Y., Cho H., Chen K., Mitchell R., Cano I., Zhou T., 2015. Xgboost: extreme gradient boosting. R package version 0.4-2. 1. 1-4.
Chowdhury, A., Jha, M.K., Chowdary, V.M., 2010. Delineation of groundwater recharge zones and identification of artificial recharge sites in West Medinipur district, West Bengal, using RS, GIS and MCDM techniques. Environ. Earth Sci. 59, 1209–1222.
da Costa, A.M., de Salis, H.H.C., Viana, J.H.M., Pacheco, F.A.L., 2019. Groundwater recharge potential for sustainable water use in urban areas of the Jequitiba River Basin, Brazil. Sustainability 11, 1–20.
Expósito, J.L., Esteller, M.V., Paredes, J., Rico, C., Franco, R., 2010. Groundwater protection using vulnerability maps and wellhead protection area (WHPA): a case study in Mexico. Water Resour. Manag. 24, 4219–4236.
Fannakh, A., Farsang, A., 2022. DRASTIC, GOD, and SI approaches for assessing groundwater vulnerability to pollution: a review. Environ. Sci. Eur. 34, 1–16.
Farhadian, H., Nikvar Hassani, A., Katibeh, H., 2017. Groundwater inflow assessment to Karaj Water Conveyance tunnel, northern Iran. KSCE J. Civ. Eng. 21, 2429–2438.
Ghayoumian, J., Ghermezcheshme, B., Feiznia, S., Noroozi, A.A., 2005. Integrating GIS and DSS for identification of suitable areas for artificial recharge, case study Meimeh Basin, Isfahan. Iran. Environ. Geol. 47, 493–500.
Harbaugh, A.W., Banta, E.R., Hill, M.C., McDonald, M.G., 2000. Modflow-2000, the U. S. Geological survey modular ground-water model-user guide to modularization concepts and the ground-water flow process. Open File Report 00-92.
Huang, X., Gao, L., Crosbie, R.S., Zhang, N., Fu, G., Doble, R., 2019. Groundwater recharge prediction using linear regression, multi-layer perception network, and deep learning. Water 11, 1879.
Jena, R., Pradhan, B., Beydoun, G., Alamri, A.M., Sofyan, H., 2020. Earthquake hazard and risk assessment using machine learning approaches at Palu. Indonesia. Sci. Total Environ. 749, 141582.
Kazakis, N., 2018. Delineation of suitable zones for the application of managed aquifer recharge (MAR) in coastal aquifers using quantitative parameters and the analytical hierarchy process. Water 10, 804.
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, 847–861.
Machiwal, D., Jha, M.K., 2015. Identifying sources of groundwater contamination in a hard-rock aquifer system using multivariate statistical analyses and GIS-based geostatistical modeling techniques. J. Hydrol. Reg. Stud. 4, 80–110.
Machiwal, D., Singh, P.K., 2015. Comparing GIS-based multi-criteria decision-making and Boolean logic modelling approaches for delineating groundwater recharge zones. Arab. J. Geosci. 8, 10675–10691.
Madanan, M., Sayed, B., Akhmal, N., Velayudhan, N., 2021. An Artificial Intelligence Approach Based on Hybrid CNN-XGB Model to Achieve High Prediction Accuracy through Feature Extraction, Classification and Regression for Enhancing Drug Discovery in Biomedicine. Int. J. Biol. Biomed. Eng 15, 190–201.
Mahmoud, S.H., Alazba, A.A., 2014. Identification of potential sites for groundwater recharge using a GIS-based decision support system in Jazan region-Saudi Arabia. Water Resour. Manag. 28, 3319–3340.
Martinsen, G., Bessiere, H., Caballero, Y., Koch, J., Collados-Lara, A.J., Mansour, M., Sallasmaa, O., Pulido-Velazquez, D., Williams, N.H., Zaadnoordijk, W.J., 2022. Developing a pan-European high-resolution groundwater recharge map–Combining satellite data and national survey data using machine learning. Sci. Total Environ. 822, 153464.
Moharir, K.N., Pande, C.B., Gautam, V.K., Singh, S.K., Rane, N.L., 2023. Integration of hydrogeological data, GIS and AHP techniques applied to delineate groundwater potential zones in sandstone, limestone and shales rocks of the Damoh district,(MP) central India. Environ. Res. 228, 115832.
Mokarram, M., Negahban, S., Abdolali, A., Ghasemi, M.M., 2021. Using GIS-based order weight average (OWA) methods to predict suitable locations for the artificial recharge of groundwater. Environ. Earth Sci. 80, 1–17.
Murad, A.A., Al Nuaimi, H., Al Hammadi, M., 2007. Comprehensive assessment of water resources in the United Arab Emirates (UAE). Water Resour. Manag. 21, 1449–1463.
Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and water assessment tool theoretical documentation version 2009. Texas Water Resources Institute.
Pande, C.B., Moharir, K.N., Singh, S.K., Elbeltagi, A., Pham, Q.B., Panneerselvam, B., Varade, A.M., Kouadri, S., 2022. Groundwater flow modeling in the basaltic hard rock area of Maharashtra. India. Appl. Water Sci. 12, 1–14.
Pourghasemi, H.R., Sadhasivam, N., Yousefi, S., Tavangar, S., Nazarlou, H.G., Santosh, M., 2020. Using machine learning algorithms to map the groundwater recharge potential zones. J. Environ. Manage. 265, 110525.
Raihan, A.T., Bauer, S., Mukhopadhaya, S., 2022. An AHP based approach to forecast groundwater level at potential recharge zones of Uckermark District, Brandenburg. Germany. Sci. Rep. 12, 6365.
Riad, P.H.S., Billib, M., Hassan, A.A., Salam, M.A., El Din, M.N., 2011. Application of the overlay weighted model and Boolean logic to determine the best locations for artificial recharge of groundwater. J. Urban Environ. Eng. 5, 57–66.
Richey, A.S., Thomas, B.F., Lo, M., Famiglietti, J.S., Swenson, S., Rodell, M., 2015. Uncertainty in global groundwater storage estimates in a T otal G roundwater S tress framework. Water Resour. Res. 51, 5198–5216.
Saaty T.L., n.d. How to make a decision: the analytic hierarchy process (1990). Eur. J. Oper. Res. (48). 9–26.
Selvarani, A.G., Maheswaran, G., Elangovan, K., 2017. Identification of artificial recharge sites for Noyyal River Basin using GIS and remote sensing. J. Indian Soc. Remote Sens. 45, 67–77.
Şener, E., 2021. Intrinsic groundwater vulnerability assessment, comparison of different methodologies and validation. Arab. J. Geosci. 14, 1–17.
Shelar, R.S., Shinde, S.P., Pande, C.B., Moharir, K.N., Orimoloye, I.R., Mishra, A.P., Varade, A.M., 2022. Sub-watershed prioritization of Koyna river basin in India using multi criteria analytical hierarchical process, remote sensing and GIS techniques. Phys. Chem. Earth, Parts a/b/c 128, 103219.
Sherif, M.M., Ebraheem, A.M., Al Mulla, M.M., Shetty, A.V., 2018. New system for the assessment of annual groundwater recharge from rainfall in the United Arab Emirates. Environ. Earth Sci. 77, 1–17.
Sherif, M., Chowdhury, R., Shetty, A., 2014. Rainfall and intensity-duration-frequency (IDF) curves in the United Arab Emirates, in: Proceedings of the World Environmental and Water Resources Congress, Portland, OR, USA. pp. 1–5.
Sherif, M., Sefelnasr, A., Ebraheem, A.A., Al Mulla, M., Alzaabi, M., Alghafli, K., 2021. Spatial and Temporal Changes of Groundwater Storage in the Quaternary Aquifer. UAE. Water 13, 864.
Skilodimou, H.D., Bathrellos, G.D., Chousianitis, K., Youssef, A.M., Pradhan, B., 2019. Multi-hazard assessment modeling via multi-criteria analysis and GIS: a case study. Environ. Earth Sci. 78, 1–21.
Thongsuwan, S., Jaiyen, S., Padcharoen, A., Agarwal, P., 2021. ConvXGB: A new deep learning model for classification problems based on CNN and XGBoost. Nucl. Eng. Technol. 53, 522–531.
Xu, B., Liu, S., Zhou, J.L., Zheng, C., Weifeng, J., Chen, B., Zhang, T., Qiu, W., 2021. PFAS and their substitutes in groundwater: Occurrence, transformation and remediation. J. Hazard. Mater. 412, 125159.
Zaidi, F.K., Nazzal, Y., Ahmed, I., Naeem, M., Jafri, M.K., 2015. Identification of potential artificial groundwater recharge zones in Northwestern Saudi Arabia using GIS and Boolean logic. J. African Earth Sci. 111, 156–169.
Zaresefat, M., Derakhshani, R., Nikpeyman, V., GhasemiNejad, A., Raoof, A., 2023. Using artificial intelligence to identify suitable artificial groundwater recharge areas for the Iranshahr basin. Water 15, 1182.