Journal of
Water & Climate Change

Check for updates

IWA PUBLISHING

© 2025 The Authors

Journal of Water and Climate Change Vol 00 No 0, 1 doi: 10.2166/wcc.2025.216

Predicting future groundwater levels in Bangladesh under climate change using boosting algorithms

Sakib Hosan, Showmitra Kumar Sarkar* and Mafrid Haydar

Department of Urban and Regional Planning, Khulna University of Engineering & Technology (KUET), Khulna 9203, Bangladesh
*Corresponding author. E-mail: mail4dhrubo@gmail.com

SKS, 0000-0001-5928-7577

ABSTRACT

Initially, 26 parameters were assessed, and key factors were determined using collinearity assessments and the variance inflation factor (VIF). Precipitation and land use and land cover were identified as the primary determinants of groundwater levels in Bangladesh. The dataset consists of 2,145 data points gathered from 2023 to 2024. Five advanced boosting algorithms, such as adaptive boosting, categorical boosting, gradient boosted decision trees, light gradient boosting machine (LightGBM), and extreme gradient boosting, were evaluated, whereas LightGBM exhibited the highest performance (AUC: 0.92, Accuracy: 0.85), attributed to its efficacy in managing extensive datasets and class imbalances. The model predicts that low groundwater potential zones will represent 1.33–11.70% of the total land area in Bangladesh, whereas high-potential zones will comprise 29.43–37.65%. The Shared Socioeconomic Pathways (SSPs) framework was utilized to forecast future groundwater availability under scenarios SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5, with precipitation data projected for the years 2040, 2060, 2080, and 2100. The estimates indicate significant spatiotemporal variations in groundwater potential, with high-potential regions spanning 13.84–80.59%, while low-potential regions vary from 0.00 to 22.35%. Some places see increased precipitation and enhanced groundwater recharge, while others are vulnerable to groundwater scarcity due to consistently low precipitation.

Key words: climate change, ensemble machine learning, GIS and remote sensing, groundwater level, sustainable groundwater management

HIGHLIGHTS

- Applied five boosting ML models for groundwater potential mapping in Bangladesh.
- LightGBM showed best results (AUC = 0.92, Accuracy = 0.85).
- Precipitation and LULC identified as key groundwater drivers.
- Predicted groundwater for 2040–2100 under SSP1–2.6 to SSP5–8.5 scenarios.
- High zones: Barisal 80.27%, Khulna 71.83%; low: Chittagong 27.43%.

1. INTRODUCTION

Water is essential to the hydrological process and one of the most significant natural resources for human survival and progress (Arabameri *et al.* 2019; Ferozur *et al.* 2019; Fatema *et al.* 2023; Dey *et al.* 2023). Worldwide, water is needed by 785 million people, or one in nine (Babu *et al.* 2020; Howlader *et al.* 2024). Groundwater is the most important part of water resources (Kumar *et al.* 2016). The majority of households, businesses, and farms use groundwater, with 36% for domestic, 42% for agricultural, and 27% for industrial use (Taylor *et al.* 2013; Nampak *et al.* 2014). More than 2.5 billion people depend on groundwater, and demand has increased by 600% in the past century (Wang-Erlandsson *et al.* 2022; WBG 2022). But 70% of people on Earth live in areas with limited water resources, which might result in a water deficit for 1.8 billion people in the years to come (Mekonnen & Hoekstra 2016). Actually, groundwater levels are falling every day due to urbanization and development, largely harming agricultural output (Cha *et al.* 2014; Adiguzel *et al.* 2022; Tekin *et al.* 2022; Varol *et al.* 2022). Climate change will also provide major issues for future water supply and groundwater levels (Mallick *et al.* 2015). Bangladesh also encounters distinct issues related to the worldwide relevance of this resource. Bangladesh has a severe dry season water scarcity despite its tropical climate and riverine topography (Saha *et al.* 2019; Quino Lima *et al.* 2021;

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Rana *et al.* 2022; Sarkar *et al.* 2022a). Nine major droughts have threatened horticulture, aquaculture, special crop agriculture, and commercial vegetable output in the last 45 years (Ferozur *et al.* 2019; Sresto *et al.* 2021). It is also unclear how much groundwater recharge occurs in Bangladesh, which ranges from 21 to 65 billion m^3 per year (Rahman *et al.* 2022). 60% of the population works in agriculture, and 70% of irrigation projects in Bangladesh rely on groundwater (Hasan *et al.* 2007; Shahid & Hazarika 2010; Khan & Haque 2023). As Bangladesh is an agricultural country, overexploitation is lowering groundwater levels due to unregulated extraction (Salvadore *et al.* 2015). As a result, groundwater is often exploited, making long-term use dangerous (Pradhan 2009; Ayazi *et al.* 2010). So, providing enough water for everyone is difficult, especially in developing nations like Bangladesh, if we do not manage the groundwater properly.

Remote sensing (RS) and geographic information systems (GIS) have been used to map groundwater resources management, which helps to manage the groundwater properly (Corsini *et al.* 2009). This technique necessitates the identification and integration of multiple variables that affect groundwater potential (GWP). Previous research reveals that many variables affect GWP. Different categories, such as elevation, slope, curvature, aspect, and topographical indices like the topographical roughness and topographical position, affect GWP according to Maskooni *et al.* (2020), Prasad *et al.* (2020), Park & Kim (2021), Sachdeva & Kumar (2021), and Halder *et al.* (2024). Topography has a great influence on groundwater recharge and groundwater flow systems, as observed by Sidle & Onda (2004). GWP is significantly governed by hydrological parameters like the topographic wetness index (TWI), stream power index, flow accumulation (FA), stream distance, and drainage density (DD) affect surface water flow and infiltration, as observed by Maskooni *et al.* (2020), Prasad *et al.* (2020), Park & Kim (2021), and Halder *et al.* (2024). As discussed by Sachdeva & Kumar (2021), Rao *et al.* (2022), and Rasool *et al.* (2022), meteorological conditions play a crucial role in controlling the groundwater availability, while temperature and precipitation both directly affect recharge rates and the water balance as a whole. According to Al-Fugara *et al.* (2020), Park & Kim (2021), Rasool *et al.* (2022), and Halder *et al.* (2024), the geologic parameters are geomorphology, distance to faults (DTF), geology, lineament density (LD), and soil texture; all these factors contribute to groundwater flow and storage. In GWP zone assessment, studies by Maskooni *et al.* (2020), Prasad *et al.* (2020), Park & Kim (2021), Sachdeva & Kumar (2021), and Halder *et al.* (2024) have highlighted the significance of land use and land cover, the normalized difference vegetation index (NDVI), the normalized difference water index (NDWI), and road distance. Evapotranspiration (EV) was used in several studies, e.g., Wang *et al.* (2023) and Nguyen *et al.* (2024), for the forecasting of groundwater future potential, while Zzaman *et al.* (2022) and Sarkar *et al.* (2024a) have used the sediment transport index (STI), and Singha *et al.* (2024) have used soil moisture. Humidity was shown to play a role in groundwater potentiality (Sarkar *et al.* 2024b). A number of researchers used various parameters; however, it should be kept in mind that not all parameters are universally applicable to groundwater level (GWL) prediction. Jou *et al.* (2014) and Mazumder & Saroar (2025) used the variance inflation factor (VIF) for the identification of the most effective variables. Having chosen the significant variables, we must now forecast future groundwater. For future groundwater data prediction, according to Al-Fugara *et al.* (2020), machine learning (ML) algorithms are the best method. Various ML models have been integrated with RS and GIS methods, according to Oh *et al.* (2011), Al-Abadi *et al.* (2017), and Al-Shabeeb *et al.* (2018). Based on studies, adaptive boosting (AdaBoost) (Chen *et al.* 2020; Yen *et al.* 2021; Zhao *et al.* 2024), extreme gradient boosting (XGBoost) (Naghibi *et al.* 2020; Park & Kim 2021; Rao *et al.* 2022; Rasool *et al.* 2022; Dey *et al.* 2023; Wang *et al.* 2023; Xiong *et al.* 2023, 2024; Halder *et al.* 2024), random forest (Al-Fugara *et al.* 2020; Prasad *et al.* 2020; Mosavi *et al.* 2021; Nugroho *et al.* 2024; Roy *et al.* 2024), light gradient boosting machine (LightGBM) (Guo *et al.* 2023), gradient boosted decision trees (GBDT) (Sachdeva & Kumar 2021; Wei *et al.* 2022a), and deep boosting (Maskooni *et al.* 2020; Chen *et al.* 2022) are the most suitable models for GWP prediction. To get the best results, several researchers have utilized categorical boosting (CatBoost) (Xiong *et al.* 2023, 2024), GamBoost (Mosavi *et al.* 2021), and SVM (Al-Fugara *et al.* 2020; Prasad *et al.* 2020; Mallick *et al.* 2022; Rasool *et al.* 2022; Dey *et al.* 2023; Roy *et al.* 2024). Other models have also been used, including Artificial Neural Network (ANN) (Mallick *et al.* 2022; Rasool *et al.* 2022; Sarkar *et al.* 2024b) and k-nearest neighbor (KNN) (Rasool *et al.* 2022; Dey *et al.* 2023). Past studies have used neural networks, tree-based approaches, bagging, and boosting methods, whereas boosting has most often been the best for groundwater projection. The majority of recent studies also rely on well-established global or regional hydrologic models such as MODFLOW, SWAT, and MIKE SHE, which are scientifically robust but reflect fundamental limitations at the local scale in data-limited environments (Sarker & Leta 2025). For instance, SWAT+ can be appropriately used in long-term land use and climate change impacts estimation, but is limited for small-scale watershed calibration with limited geospatial and hydrometeorological data and weak surface water–groundwater interaction representation unless it is coupled with MODFLOW (Liang *et al.* 2023; Sánchez-Gómez *et al.* 2024). Similarly, MIKE

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

SHE offers an entirely coupled simulation of surface and subsurface processes with high spatial resolution, albeit with extremely high computation requirements, high learning curve, and high licensing costs (Gallego Ortega Ventzi Bojkov 2023; Sarker & Leta 2025). Hydrological Simulation Program - FORTRAN (HSPF) and Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS), while convenient for some applications (long-term water quality simulation and event-based runoff modeling, respectively), are equally inconvenient to utilize for long-term climate change and land use analysis (Chathuranika *et al.* 2022; Vogeti *et al.* 2023). There are several significant issues with traditional process-based hydrologic models. To function properly, they require extensive calibration, precise high-resolution data, and trustworthy flow records. Since rivers and watersheds are frequently ungauged or inadequately monitored, this is an issue in many developing nations. These models may consequently become ambiguous and less trustworthy. For instance, frequently miss the small-scale activities in tiny basins, even when detailed spatial data is available. Furthermore, without discharge data, it is challenging to predict site-specific parameters, which undermines the validity of model conclusions. However, ML models that rely on boosting are far more useful. Since they do not require complete physical system parameterization, they can continue to function even in the case of missing or ambiguous inputs. ML algorithms have been increasingly applied in the last few years to bridge this gap, and boosting algorithms have been shown to outperform the rest (Naghibi *et al.* 2016; Ibrahem Ahmed Osman *et al.* 2021; Halder *et al.* 2024). Boosting method benefits include low variance, bias reduction, iterative error correction, and adaptation to complex hydrological and geographical data (Mosavi *et al.* 2021; Xiong *et al.* 2023; Razavi-Termeh *et al.* 2024). Boosting is suitable for noisy and data-poor environments such as Bangladesh, whose available hydrological data are often sparse, imprecise, or irregularly disseminated. The more notable these benefits are, the more inadequately applied are the boosting algorithms in GWP mapping and prediction in Bangladesh. Most of the earlier studies have already tried predicting GWP in Bangladesh. But some of them had some serious problems. First, they took into account too many predictor variables that were not really important for Bangladesh's own hydrogeological and environmental conditions. Because of this, their models were not that useful and interpretable. Second, no one, or a few one, used boosting algorithms in predicting groundwater in Bangladesh for future climate change. This is a significant gap because climate change is already affecting groundwater levels in an unpredictable way. Without using future climate projections, studies are not able to show the actual risks and challenges for Bangladesh's groundwater.

The current study builds upon previous work in this area by bridging the void in the existing body of knowledge using the three particular objectives, aimed at maximizing the accuracy, interpretability, and usefulness of groundwater prediction for Bangladesh under a climate change scenario. First, methodologically applies feature selection by keeping only the most explanatory features, from topography, hydrology, climate, geology, and land use, with emphasis on Bangladesh. This eliminates redundancy and enhances model interpretability. Second, generate boosting-based prediction algorithms that reflect Bangladesh's specific hydrogeological zones, enhancing mapping accuracy by considering local variation. Third, it combines projected future climate to forecast groundwater availability at spatial and temporal resolutions, combining long-term climate impacts. By combining high-quality variable selection, context-specific modeling, and climate-adjusted prediction, the research formulates a high-resolution analytical framework for exploring groundwater dynamics in Bangladesh. The results of this research are likely to provide policy-relevant insights on groundwater depletion, recharge potential, and contamination risks. Besides enhancing scientific understanding, the findings give policymakers, resource managers, and city planners a decision-support tool in the form of a pragmatic contribution, enabling evidence-based policy for optimizing water allocation, reducing unsustainable use, and safeguarding groundwater resources in the face of ongoing climate change.

2. METHODS AND MATERIALS

2.1. Comprehensive characterization of the study area

The research area encompasses Bangladesh, located between 20°34′ and 26°38′N latitude and 88°01′ and 92°41′E longitude in South Asia, bordered by the Bay of Bengal to the south, India to the north and west, and Myanmar to the east (Kamruzzaman *et al.* 2019; Uddin *et al.* 2020). Figure 1 illustrates the geographical position of Bangladesh, highlighting its climate zones and topographical variances. Bangladesh comprises seven distinct climatic zones: The southeastern zone has substantial precipitation exceeding 2,540 mm, whereas the northeastern zone is characterized by dense cloud cover, elevated humidity, recurrent winter fog, and considerable winter precipitation; the northern section of the northern region experiences sweltering summers exceeding 32 °C and frigid winters below 10 °C; the northwestern region has comparable temperature extremes but receives less precipitation; the western zone is the driest, particularly during summer; the transitional south-central zone,

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

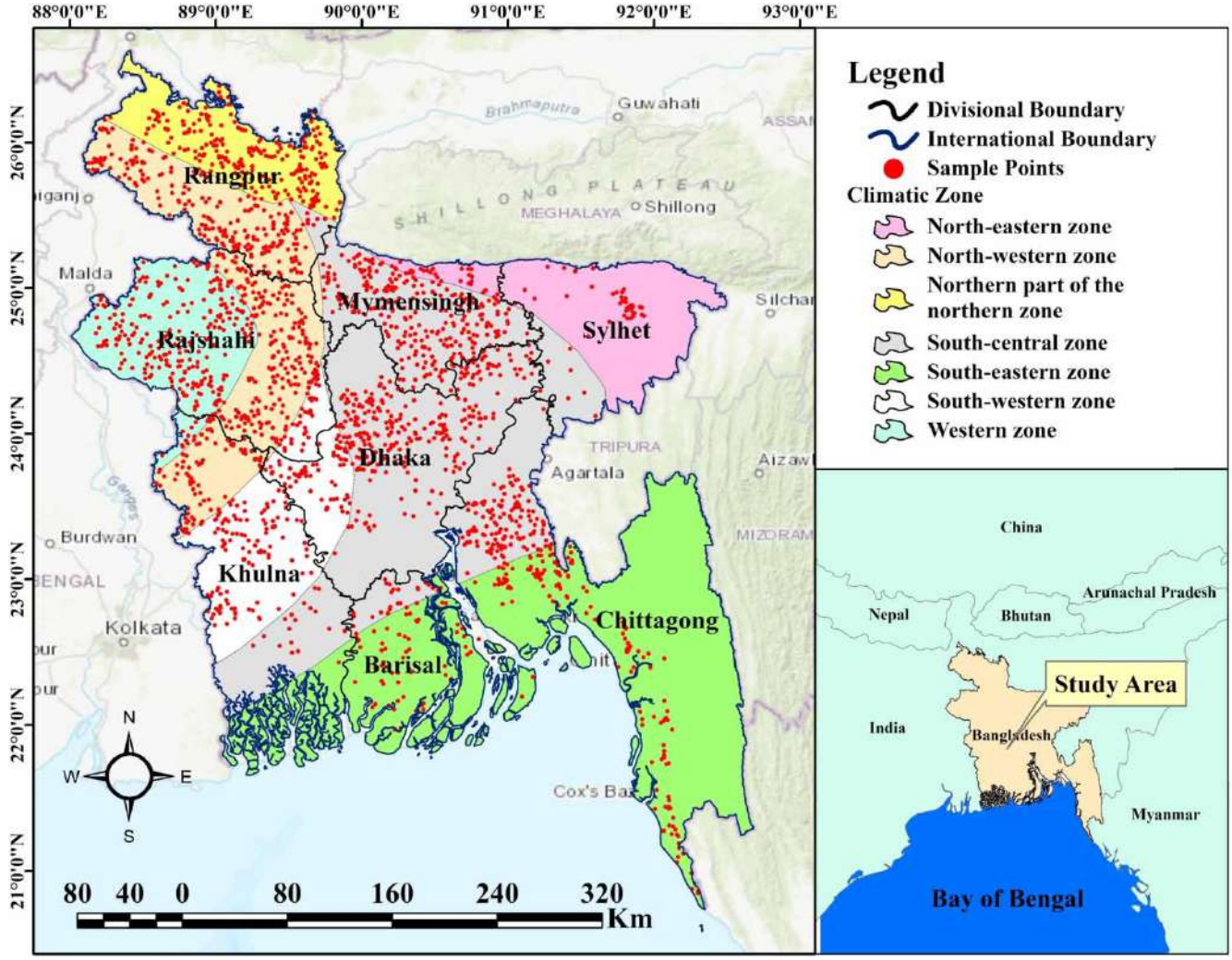

Figure 1 | Geographical location of the research area (Bangladesh).

over 1,900 mm of precipitation occurs, frequently accompanied by severe storms; the southwestern region experiences significant dew and moderated seasonal variations (Banglapedia 2014). Bangladesh has a tropical humid climate with seasonal precipitation changes, moderate temperatures, and high humidity, and it has nearly 170 million people, making it the eighth most populous and densest nation (Fattah *et al.* 2023). Bangladesh encompasses an area of 147,570 km^2 and is administratively segmented into eight divisions: Barisal (13,225.20 km^2), Chattogram (33,908.55 km^2), Dhaka (20,593.74 km^2), Khulna (22,284.22 km^2), Mymensingh (10,584.06 km^2), Rajshahi (18,153.08 km^2), Rangpur (16,184.99 km^2), and Sylhet (12,635.22 km^2) (Bangladesh National Portal 2020). Due to its elevated population density, Bangladesh encounters considerable issues with water scarcity, with specific areas having more acute shortages. Groundwater is essential to Bangladesh's water supply, fulfilling 90% of irrigation requirements, with other inland freshwater sources providing the remaining 10% (Anwar & Ahmed 2006). GWL depths were measured at 2,145 monitoring points by the Bangladesh Water Development Board (BWDB) to assess groundwater availability nationwide. According to Marjuanto *et al.* (2019), areas where the groundwater depth is greater than 6.85 m are considerably less vulnerable to groundwater depletion.

2.2. Analytical overview of collected data

The research utilized 26 parameters systematically classified into five major domains: topographical, hydrological, meteorological, geological, and land use and human activity. The geological domain included six parameters: DTF, geological composition, geomorphological features, LD, soil moisture content, and soil texture. The hydrological domain contained

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

five parameters: distance to stream (DTS), DD, FA, stream power index, and TWI. The land use and human activity domain contained four parameters: distance to roads (DTR), land use and land cover (LULC), NDVI, and NDWI. The meteorological domain contained four parameters: EV, humidity, precipitation, and temperature. The topographical parameter consisted of seven parameters: aspect, curvature, elevation, STI, slope, topographic position index (TPI), and topographic roughness index (TRI). Faults affect GWP through alterations in permeability, enhancing recharge with increased permeability or arresting flow with impermeable barriers, hence affecting groundwater storage and recharge (Al-Fugara *et al.* 2020; Cook *et al.* 2022). The faults' shapefiles were downloaded and then projected into the coordinate system of the study area. The Euclidean distance tool in ArcGIS was used to create the raster. Groundwater storage and flow are predominantly affected by geological attributes, such as weathered materials, source rocks, sedimentary conditions, hydraulic conductivity, and porosity, with permeable formations augmenting storage capacity (Xie *et al.* 2013; Abdullateef *et al.* 2021; Jhariya *et al.* 2021). Lithological layers were categorized by rock type and clipped to the study region. Geomorphology profoundly affects GWP by determining infiltration, storage, and movement; valleys and alluvial plains facilitate recharging, whereas steep and rocky terrains restrict infiltration due to heightened runoff (Huggett 2016; Halder *et al.* 2024). Geomorphic units represented by polygons were downloaded as shapefiles and categorized into typical geomorphic forms. Lineaments, such as faults, fractures, and joints, facilitate groundwater circulation and storage by augmenting infiltration and subsurface connectivity, with greater LD enhancing permeability and groundwater flow (Falah *et al.* 2017; Cook *et al.* 2022). Soil moisture is essential for groundwater dynamics since it governs infiltration, retention, and recharge; elevated moisture levels facilitate infiltration, while diminished levels signify arid, less permeable regions (Wei *et al.* 2022b; Singha *et al.* 2024). The data were extracted from soil moisture active passive (SMAP). Soil texture affects water potential by regulating groundwater infiltration into the aquifer system, altering water retention capacity, and influencing aquifer production (Oh *et al.* 2011; Melese & Belay 2022). Types of textures were used to categorize soil polygons and convert them to a raster. The distance to streams markedly affects GWP, as areas in closer proximity experience enhanced infiltration and lateral seepage, rendering stream data crucial for evaluating recharge and designating potential zones (Ahmed *et al.* 2021; Zeydalinejad *et al.* 2024; Porun *et al.* 2025). The data were collected from the Local Government Engineering Department (LGED). Groundwater supply and pollution are significantly influenced by DD, enabling the measurement of characteristics such as permeability, runoff, infiltration, and topography (Rana *et al.* 2022). Waikar & Nilawar (2014) demonstrated that hilly areas have high drainage densities, while dense vegetation in flat areas has low drainage densities. Low DD areas can have greater GWP than those of high DD (Biswas *et al.* 2020). After the drainage lines have been derived, density is obtained by the total stream length per unit area. FA is significant in water availability assessment as it indicates the runoff accumulation to streams and rivers, with the high-accumulation areas being open for infiltration and groundwater recharge (Schäuble *et al.* 2008; Newman *et al.* 2025). Extraction was conducted from the digital elevation model (DEM) using ArcGIS. The stream power index (SPI) quantifies the erosive power of a stream near flow turbidity, and infiltration so that GWP areas can be determined with negative SPI values indicating good aquifer locations for water harvesting (Golkarian & Rahmati 2018; Ahmad *et al.* 2020; Porun *et al.* 2025). The TWI is an important measure of water availability as it indicates water concentration due to topography, with higher values of TWI depicting better GWP (Mokarram *et al.* 2015; Qadir *et al.* 2020). DTR decreases GWP through enhanced surface runoff and reduced infiltration, especially in urban environments (Uliasz-Misiak *et al.* 2022; Singha *et al.* 2024). The LGED prepared road networks and created a Euclidean distance raster. LULC also significantly influence the groundwater distribution and availability through influencing infiltration, soil moisture, and surface water dependence, as natural ecosystems like forests and wetlands permit recharge, whereas urbanization detains infiltration (Rana *et al.* 2022; Hall 2023; Porun *et al.* 2025). The ESRI Sentinel-2 Global Land Cover 2024 data, which provides land cover at a 10-m spatial resolution, was utilized in this study. With robust techniques involving extensive training data and validation with high-resolution satellite images, this dataset is globally renowned (Karra *et al.* 2021). The NDVI is utilized extensively to determine the vegetation versus water relation, where high NDVI suggests higher GWP with increased plant cover (Hasegawa *et al.* 2017; Mallick *et al.* 2019; Porun *et al.* 2025). NDVI calculated from Landsat data based on NIR and Red bands. The Normalized Difference Vegetation Index (NDWI) is also used extensively for the identification of water bodies and quantification of vegetation and soil moisture content, and in determining water potential using spectral reflectance analysis. It identifies characteristics such as dams, rivers, and streams, with the NDWI value being more in cases of surface water bodies, permitting recharging of groundwater (Chatterjee & Dutta 2022; Porun *et al.* 2025). Computed NDWI from Landsat imagery based on Green and NIR bands. Groundwater is significantly affected by EV, which rises in warmer regions due to enhanced evaporative demand; however, enhanced EV causes more water loss, thus decreasing groundwater recharge (Condon *et al.*

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

2020). Humidity affects groundwater levels by influencing EV, precipitation, and infiltration; increased humidity diminishes evaporation losses and facilitates greater water percolation into the soil (Farhat 2018; Sarkar *et al.* 2024b). The data were taken from the Global Land Data Assimilation System, or GLDAS. Data on relative humidity were reprojected and resampled to the resolution of the research area. Precipitation serves as the principal supply of water for replenishing rivers, lakes, and reservoirs, and is essential for groundwater recharge; thus, higher precipitation augments recharging, whereas less precipitation restricts infiltration (Sarkar *et al.* 2024b; Sobaga *et al.* 2024). CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) is the source of the data. Elevated temperatures expedite evaporation and EV, diminishing surface water infiltration and groundwater recharge, while temperature fluctuations can influence precipitation patterns, so altering the timing and volume of groundwater replenishment (Xiong *et al.* 2023). Both temperatures and EV were taken from satellite products of the Moderate Resolution Imaging Spectroradiometer (MODIS). Aspect, a crucial determinant, affects vegetation, precipitation, wind orientation, atmospheric humidity, and solar irradiance, hence influencing water availability; north-facing slopes retain greater moisture and facilitate recharge, whereas south-facing slopes undergo increased evaporation (Zabihi *et al.* 2016). Curvature denotes local topography, with positive values signifying convexity, negative values suggesting concavity, and zero representing a smooth surface (Bertamini & Wagemans 2013; Minár *et al.* 2020). Water accumulates in concave profiles and decelerates in convex ones, hence affecting water infiltration and flow dynamics (Arulbalaji *et al.* 2019; Xiong *et al.* 2023; Porun *et al.* 2025). Elevations forecast groundwater flow and delineate water table maps by affecting vegetation, soil composition, and climate, with low-lying regions adjacent to rivers and valleys promoting recharging, whereas elevated places facilitate runoff (Al-Abadi & Shahid 2015; Aloui *et al.* 2024; Porun *et al.* 2025), utilizing digital elevation data from the Shuttle Radar Topography Mission (SRTM). The DEM was clipped to the study area edge and reprojected to the study coordinate system after being downloaded. Slopes are vital in determining high water absorption areas; gentle slopes provide for percolation and diminish runoff, whereas steep slopes enhance the flow velocity, producing faster drainage and decreased infiltration (Magesh *et al.* 2012; Porun *et al.* 2025). The STI quantifies sediment transport capacity by overland flow using slope steepness and length. Higher STI values refer to erosion-prone areas that decrease infiltration and recharge of groundwater (Wischmeier & Smith 1978). The TPI is critical in determining groundwater resources because it gives information on elevation and landform. Bare conditions are associated with reduced water production capacity, where low values of TPI (valleys) correspond to areas of high recharge and positive values (ridges) to limited infiltration (Knitter *et al.* 2019; Mosavi *et al.* 2021; Porun *et al.* 2025). The TRI affects the identification of GWP zones by indicating topographical irregularities, in which a high TRI value signifies low water potential due to a higher terrain variability (Panahi *et al.* 2020; Porun *et al.* 2025). In our study, SRTM DEM with 30-m spatial resolution was used as the primary data source for extracting most of the hydrological and topographic parameters. These included DD, FA, SPI, TWI, aspect, curvature, STI, slope, TPI, and TRI. Different indices were derived through DEM-based raster operations and geospatial analysis procedures indicated in Table 1, and hydrological variables such as flow direction and FA were derived through the ArcGIS Spatial Analyst hydrology model tools. Dataset description and the corresponding sources used to evaluate the groundwater potentiality are explained in Table 2. To ensure the obtained results are reliable, accurate, and consistent, the output was thoroughly tested with cross-verification from publicly available datasets, existing topographic maps, and *in*

Table 1 | Indices derived from DEM for hydrological and topographical analyses (Chowdhury 2023)

Indices	Equation
TWI	$\ln \frac{\text{Radian of slope}}{Tan \text{ of slope}}$
TRI	$\frac{\text{MEAN}_{\text{DEM}} - \text{MIN}_{\text{DEM}}}{\text{MAX}_{\text{DEM}} - \text{MIN}_{\text{DEM}}}$
SPI	$\ln(\text{Flow accumulation} + 0.001) \times \left(\frac{\text{Slope}}{100} + 0.001\right)$
STI	$\left(\frac{\text{Flow accumulation}}{22.12}\right)^{0.6} \times Sin\left(\frac{\text{Slope}}{0.0896}\right)^{1.3}$
TPI	$\text{DEM} - \text{MEAN}_{\text{DEM}}$

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Table 2 | Detailed explanation of the sources and datasets used to find groundwater potentiality

Indicators types	Indicators	Abbr.	Data sources	Data type	Unit	Scale	Time
Geological	Distance to faults	DTF	USGS	Shape file	Degree (°)	–	2024
	Geology	GL	USGS Earth Explorer	Vector	Lithological classes	–	2024
	Geomorphology	GM	USGS Earth Explorer	Vector	Geomorphic units	–	2024
	Lineament density	LD	USGS Earth Explorer	Vector	Km/km^2	–	2024
	Soil moisture	SM	SMAP	Raster	Percentage (%)	30 m	2024
	Soil texture	ST	FAO Global Soil Map	Vector	Soil classes	–	2024
Hydrological	Distance to stream	DTS	LGED Datasets	Raster	Degree (°)	30 m	2024
	Drainage density	DD	Derived from DEM	Raster	Km/km^2	30 m	2024
	Flow accumulation	FA	Derived from DEM	Raster	–	30 m	2024
	Stream power index	SPI	Derived from DEM	Raster	–	30 m	2024
	Topographic wetness index	TWI	Derived from DEM	Raster	–	30 m	2024
Land use and human activity	Distance to roads	DTR	LGED Datasets	Vector	Degree (°)	–	2024
	Land use/land cover	LULC	ESRI	Raster	Land cover classes	30 m	2024
	Normalized difference vegetation index	NDVI	Landsat-9	Raster	–	30 m	2024
	Normalized difference water index	NDWI	Landsat-9	Raster	–	30 m	2024
Meteorological	Evapotranspiration	EV	MODIS	Raster	mm/day	500 m	2024
	Humidity	HUM	GLDAS	Raster	%	0.25°	2024
	Precipitation	PR	CHIRPS	Raster	mm/day	0.05°	2024
	Temperature	TEM	MODIS	Raster	(°C)	1,000 m	2024
Topographical	Aspect	AS	Derived from DEM	Raster	–	30 m	2024
	Curvature	CUV	Derived from DEM	Raster	–	30 m	2024
	Elevation	EL	SRTM	Raster	Meter	30 m	2024
	Sediment transport index	STI	Derived from DEM	Raster	–	30 m	2024
	Slope	SL	Derived from DEM	Raster	Degree (°)	30 m	2024
	Topographic position index	TPI	Derived from DEM	Raster	–	30 m	2024
	Topographic roughness index	TRI	Derived from DEM	Raster	–	30 m	2024
GWP inventory	–	–	Bangladesh Water Development Board	Vector	Meter	–	2023 – 2024
Future climate scenario (precipitation)	–	–	UCAR	Vector	mm/year	–	2040, 2060, 2080, 2100

situ measurements within the study region. Following the derivation, all raster layers were normalized through reclassification using the reclassify tool of ArcGIS. All parameters were reclassified into suitability classes based on scientifically derived criteria and value ranges for application in GWP evaluation. The reclassified raster was then combined into a multi-criteria evaluation framework using the weighted overlay analysis tool of ArcGIS. In this process, weights for each parameter were found based on its relative relevance and effect on the occurrence and distribution of groundwater, as established by literature review and expert opinion. Finally, all the weighted and reclassified layers were merged spatially to produce a composite GWP map. The composite map provided a methodologically robust and geographically coherent representation of groundwater availability, which would serve as a sound basis for sustainable management of the resource as well as future hydrological analysis.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

2.3. Selection and justification of variables

Multicollinearity occurs when the multiple linear regression analysis is done with a number of variables that have a high correlation with the dependent variable, and also among themselves, making some of the significant variables under study statistically insignificant (Maity & Mandal 2019; Shrestha 2020). The two detection methods of multicollinearity are explained in the present study. The two techniques are the VIF and the Pearson correlation coefficient. Multicollinearity can be identified through a Pearson correlation matrix by the exhibition of correlation coefficients between independent variables. If two variables have a high correlation of near −1 or 1, then it indicates potential multicollinearity (Wagavkar 2023). But for a sounder judgment, VIFs are used to measure the impact of multicollinearity (Kalantar *et al.* 2019; Shrestha 2020). VIF values above 10 mean that there is serious multicollinearity, meaning that the variable may be redundant because it correlates strongly with other predictors (Jou *et al.* 2014; Vörösmarty & Dobos 2020; Mazumder & Saroar 2025), and another p-value less than 0.05 was taken to be a statistically significant relationship between the independent variable and groundwater levels (Mazumder & Saroar 2025). To increase a model's stability and dependability, it is advisable to remove one of the variables showing high correlation with any other. Hence, the variables with p-values >0.05 are not included in the study model as, at a significance level of 5%, they are not statistically significant in the context that their contribution toward explaining the variation in the dependent variable. And then the variables for which VIF is more than 10 are eliminated from the research study as they indicate the presence of serious multicollinearity.

2.4. Development of boosting-based predictive models

Initial studies on groundwater prediction have primarily shown that boosting models perform better compared with other ML algorithms in terms of accuracy and reliability. In this research, boosting models alone are employed for increasing GWP mapping. Boosting approaches such as AdaBoost, CatBoost, GBDT, LightGBM, and XGBoost are employed for their better performance in structured tabular data and imbalanced datasets (Figure 3).

2.4.1. Adaptive boosting

AdaBoost, or adaptive boosting, is a popular ensemble learning algorithm that boosts the performance of decision trees through iterative methods of weight adjustment to training samples (Dinakaran & Ranjit Jeba Thangaiah 2017). One of its most outstanding features is its power to detect outliers in data, though this comes at the cost of increased sensitivity to noise (Kégl & Busa-Fekete 2009; Mosavi *et al.* 2021). The training process in AdaBoost begins with uniform weighting of all the data points. A base classifier, typically a decision tree, is then trained. Following this initial iteration, the algorithm detects misclassified samples and increases their weights to ensure future classifiers pay greater heed to the challenging cases. The process is repeated, each weak learner being trained to fix mistakes made by previous learners. The resultant model is a weighted sum of all learners, where improved-performing classifiers have increased influence (Saini 2024). Through altering their weights, AdaBoost (Figure 2(a)) builds a series of iterative weak classifiers in a manner that each subsequent one puts greater emphasis on the samples misclassified in the previous step. Weighted voting is used to aggregate all weak classifiers to a final result. In this current research, an AdaBoost classifier was implemented with random_state = 42 for reproducibility purposes and n_estimators = 100 to specify the number of weak classifiers. Accuracy, F1-score, feature importance, and area under the curve (AUC) are all controlled by hyperparameters and also control model complexity and learning speed. In this study, the AdaBoost algorithm was used with 70% data allocated for training and 30% for testing.

2.4.2. Categorical boosting

CatBoost is a new boosting gradient algorithm that is efficient in handling complex datasets with numerous explanatory variables, including those with categorical variables, noisy data, and mixed distributions (Hancock & Khoshgoftaar 2020; Ibrahim *et al.* 2020). CatBoost introduces an efficient mechanism known as target statistics (or target encoding) that uses the relation between the categorical variable and the target variable in order to develop meaningful numeric representations (Ptr *et al.* 2024). CatBoost (Figure 2(b)) uses symmetric tree splitting and ordered boosting to deal effectively with category variables. On high-cardinality categorical attribute datasets, it shows high performance by preventing prediction shift and overfitting. A CatBoost classifier was constructed in this study using random_state = 42 for reproducibility. These hyperparameters shape the process of training, learning evolution, and general predictive accuracy of the model. This study employs a target-based statistics approach that entails randomly permuting 70% of the train set, followed by calculating the average label value by category and inserting it before the provided label in the permutation.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

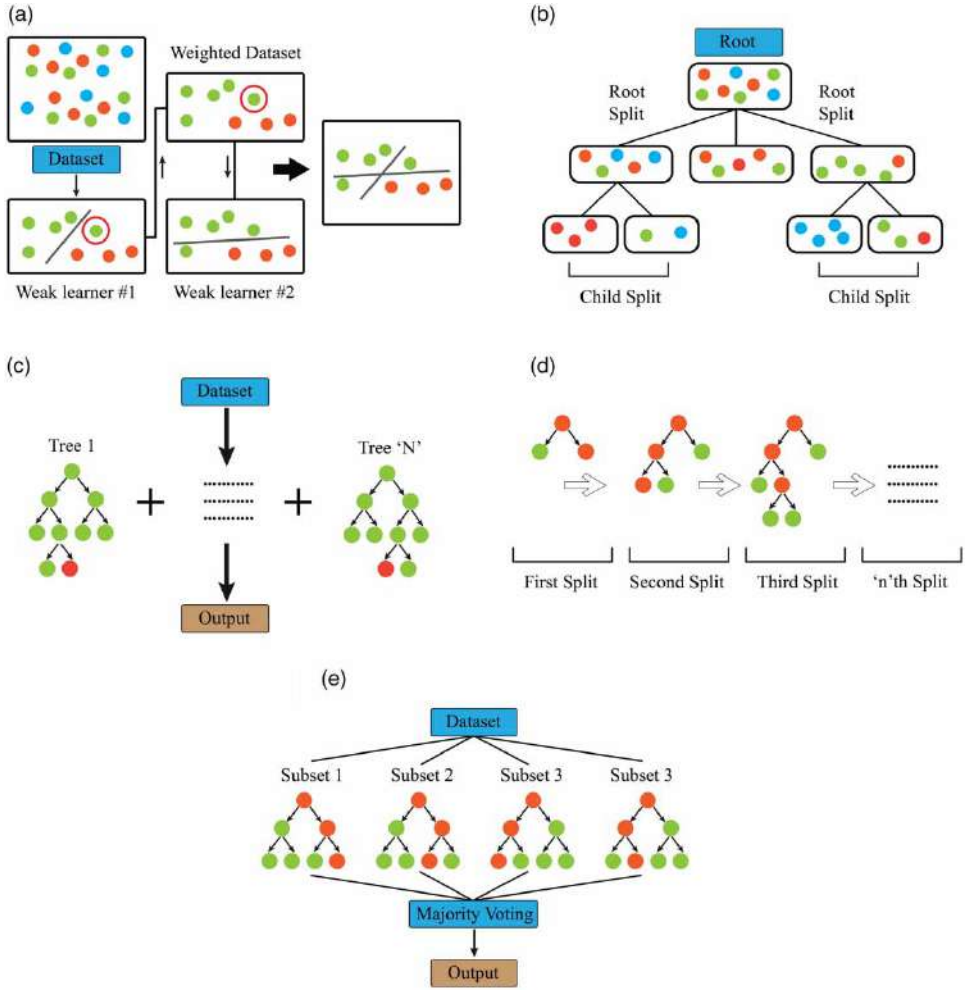

Figure 2 | Schematic diagrams of (a) AdaBoost, (b) CatBoost, (c) GBDT, (d) LightGBM, and (e) XGBoost.

2.4.3. Gradient boosting decision trees

Gradient boosting machine (GBM), or GBDT, is a robust ensemble learning strategy that builds prediction models sequentially through the combination of decision trees (Natekin & Knoll 2013; Park & Kim 2021). In contrast to bagging techniques, where multiple learners are trained independently and predictions are combined, boosting is interested in incrementally creating trees one after another, with every new tree focused on correcting the errors of the previous ones. In GBM, the process is guided by the gradient descent optimization rule, where the algorithm reduces the general loss function through the incorporation of each new tree into the model residuals constructed so far (Zanotti *et al.* 2019). The method develops in a step-by-step manner an ensemble of weak predictors, hallow regression trees in most applications that, since they are combined, generate a strong and accurate prediction model (Haidong *et al.* 2020). Each tree is trained in sequence, and instead of reweighting the misclassified samples as in AdaBoost, GBM emphasizes the gradient of the loss function, thereby addressing directly the errors that had remained unresolved by the past trees (Zhang *et al.* 2019). Such successive refinement allows GBM to handle both regression and classification tasks with unprecedented effectiveness. A collection of trees is created with GBDT (Figure 2(c)), where every new tree is trained in order to reduce the residual errors of the previous. As per a selected loss function, the predictions are merged in an effort to maximize the prediction. In the spirit of reproducibility, for this research, the gradient boosting classifier was utilized with random_state = 42. These parameters govern the general predictive capability, complexity, and iterative learning of the model. The original dataset is divided into two portions: 30% of the samples are reserved for validation purposes and 70% for training.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

2.4.4. Light gradient boosting machine

LightGBM is an advanced gradient boosting library that was created to improve the efficiency and scalability of typical GBDT models (Ke *et al.* 2017). Unlike normal boosting algorithms, LightGBM employs a histogram-based learning mechanism that bins continuous feature values, thereby registering extremely low memory usage and accelerating the training process (Sivanandam *et al.* 2024). Another important feature of LightGBM is that it applies a leaf-wise tree growth strategy in which the tree grows by splitting the most varying leaf, as compared with the traditional level-wise strategy. This allows the model to recognize more complex patterns from fewer iterations, which usually results in higher accuracy (Machado *et al.* 2019). LightGBM is an efficient gradient boosting algorithm that constructs trees in a leaf-wise manner, uses histogram-based binning for optimization, exclusive feature bundling for dimensionality reduction, and gradient-based one-sided sampling for improving training accuracy (ESRI 2023). LightGBM (Figure 2(d)), utilizing leaf-wise tree growth and histogram-based techniques, is more efficient than the legacy GBDT. It speeds up training, minimizes processing cost, and enables the algorithm to process colossal datasets effectively. In the interest of reproducibility, the LightGBM classifier was constructed in this research with random_state = 42. Such hyperparameters play a part in performance metrics like accuracy, F1-score, feature importance, and AUC, and influence the tree-based learning process, complexity, and global predictability of the model. Two sets of data are created, with 70% training and 30% testing.

2.4.5. Extreme gradient boosting

XGBoost is a supervised learning algorithm belonging to the family of GBDT (Ptr *et al.* 2024; Ok & Emmanuel 2025). XGBoost improves upon the conventional gradient boosting framework by adding a series of innovations that make it faster, bigger scale, and more accurate. Unlike other boosting methods that handle weak learners sequentially in a cumulative way, XGBoost leverages parallelization, advanced regularization strategies, and strategic handling of sparse data to achieve better accuracy in large-scale prediction problems (Xiong *et al.* 2023). It is a very scalable and effective tree-based ensemble learning algorithm with gradient boosting to achieve a variety of supervised learning tasks, including regression, classification, ranking, and other user-defined prediction problems. It is able to handle high-dimensional features and large datasets efficiently with high predictability (Arif Ali *et al.* 2023). XGBoost is a boosting algorithm that utilizes numerous Central Processing Unit (CPU) cores for parallel computations, enabling fast training on large datasets and making it suitable for big data applications (Chen & Guestrin 2016). XGBoost (Figure 2(e)) improves performance using regularization and second-order gradient information in building several boosted trees simultaneously in parallel data subsets. Efficiency and accuracy are guaranteed by using aggregation (regression) or majority vote (classification) in a bid to uncover the ultimate outcome. To ensure consistency, this research constructed the XGBoost classifier with random_state = 42. These hyperparameters determine the model's prediction capability overall, the complexity of the model, and the process of boosting learning. The XGBoost algorithm was used in this study, where 70% of the data was used for training purposes and 30% for testing.

2.5. Performance evaluation of developed models

2.5.1. ROC-AUC

Model performance was evaluated on the test set with metrics including the receiver operating characteristic curve (ROC), AUC. The AUC may be computed by integrating the features of the ROC curve. The ROC curve is drawn by computing the true positive rate (TPR) and the false positive rate (FPR) at all possible thresholds and plotting TPR against FPR (Sonego *et al.* 2008). One of the thoroughly tested techniques to enable groundwater zone mapping is the ROC. Tests (1978), Sonego *et al.* (2008), Haydar *et al.* (2024), and Porun *et al.* (2025) stated that the AUC 0.5 indicates random guessing, but anything above 0.5 indicates better predictive accuracy of the model. For specificity, the area under the ROC (AUC) has the following values: poor (0.5–0.6), average (0.6–0.7), good (0.7–0.8), very good (0.8–0.9), and excellent (0.9–1) (Porun *et al.* 2025).

2.5.2. Confusion matrix

A confusion matrix evaluates the performance of a model with classification by indicating true positives (the number of appropriately recognized positive nodes), true negatives (the number of appropriately recognized negative nodes), false positives (the number of inappropriately recognized positive nodes), and false negatives (the number of inappropriately recognized negative nodes), thus providing an overall comparison of the performance of the model for more than one class (Riehl *et al.* 2023; Mazumder & Saroar 2025). This matrix allows for model accuracy to be measured, misclassifications to be detected, and the computation of crucial parameters such as the F1-score for enhancing model performance. Accuracy

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

is defined as correct predictions divided by total predictions, where overall accuracy is high when accuracy is high; precision is the ratio of true positives to predicted positives, where high precision is equivalent to low false positives; recall is the ratio of true positives to all actual positives, where recall is high when false negatives are low; the F1-score is an equal trade-off between precision and recall as their average and it is useful in evaluating models, particularly in imbalanced data (Riehl *et al.* 2023). An F1-score of almost 1 in case of high values usually indicates a balanced performance, wherein the model can achieve high precision along with high recall simultaneously. An F1-score of almost 0 usually indicates a trade-off between precision and recall, i.e., the model cannot achieve the balance (Elmahdy *et al.* 2021; Madani & Niyazi 2023; Gómez-Escalonilla & Martínez-Santos 2024).

2.6. Projection models for future climate scenarios

This research will contrast precipitation factor projections to gauge the potential for future water resources based on groundwater levels under the condition of an altering climate because precipitation is the dominant climatic factor influencing water potential (Sarkar *et al.* 2024b). Projections from global climate models are produced through the utilization of large climate models such as CMIP6. These frameworks provide a solid foundation for assessing climate variability and its impact on water resources (Sarkar *et al.* 2024b). These pathways bring varied projections regarding future climatic situations and their consequences (Ghazi *et al.* 2021). Shared Socioeconomic Pathways (SSPs), which analyze the interdependence of socioeconomic and climatic change, are part of the CMIP6 framework (Roshani & Hamidi 2022). This research considers four SSP scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 based on the four SSPs as SSP1 (Sustainability), SSP2 (Middle of the Road), SSP3 (Regional Rivalry), and SSP5 (Fossil-Fueled Development). The research examines precipitation among selected SSP scenarios based on data projections for the years 2040, 2060, 2080, and 2100. Data used in the research are sourced from the NCAR GIS Initiative Climate Change Scenario website (https://gisclimatechange.ucar.edu), providing publicly available climate projections. It is an inclusive method that enables the accurate examination of how future climate change is set to influence water supply.

2.7. Projection models for groundwater availability under future climate scenarios

The global climate model was employed to produce 16 GWP maps for the years 2040, 2060, 2080, and 2100 with four SSP scenarios (Figure 3). The maps were produced with a range of climate change scenarios applying the SSP paradigm. Several SSP scenarios were each linked to a given temporal interval, enabling an overall assessment of GWP under diverse future climatic conditions. A raster calculator was utilized to introduce the impact of climate change in the study. The climatic scenario-associated variables were combined with results from the best model boosting using the software. The study was able to estimate the impact of predicted climate variations on groundwater availability by combining these factors. GWP maps obtained from the study were categorized into levels, from very low to high potential. This classification yielded a clear graphic interpretation of forecasted groundwater distribution patterns, enabling future planning and management of water resources.

3. RESULTS

3.1. Systematic overview of key indicators

3.1.1. Geological

Geological features are depicted in Figure 4, whereas the DTF (Figure 4(a)) illustrates the differing proximities to fault lines around the nation, as the southwest and northwest regions exhibit greater distances, whereas the southeast and northeast regions display lesser distances. The geology (Figure 4(b)) of Bangladesh mostly consists of alluvial sediments from the Ganges-Brahmaputra-Meghna River system. The nation comprises four geological units. Alluvium predominates in Bangladesh. The Chittagong Hill Tracts have lower groundwater reserves compared with alluvial lowlands, which are characterized by Pleistocene, Miocene, and Pliocene deposits. Water traverses both surface and subsurface environments influenced by geomorphology (Figure 4(c)), whereas the Chittagong and some portions of the Sylhet have hilly regions. The soil moisture in Bangladesh (Figure 4(d)) exhibits periodic fluctuations, particularly in the Khulna and Sylhet regions. In Bangladesh, water penetration and retention are contingent upon soil texture (Figure 4(e)), while the majority of the floodplains consist of silty soils, and most of the regions consist of clay soils. Groundwater flow is influenced by LD (Figure 4(f)), facilitating the migration of groundwater from areas with lower LD to regions of higher LD, such as Khulna and Barisal.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

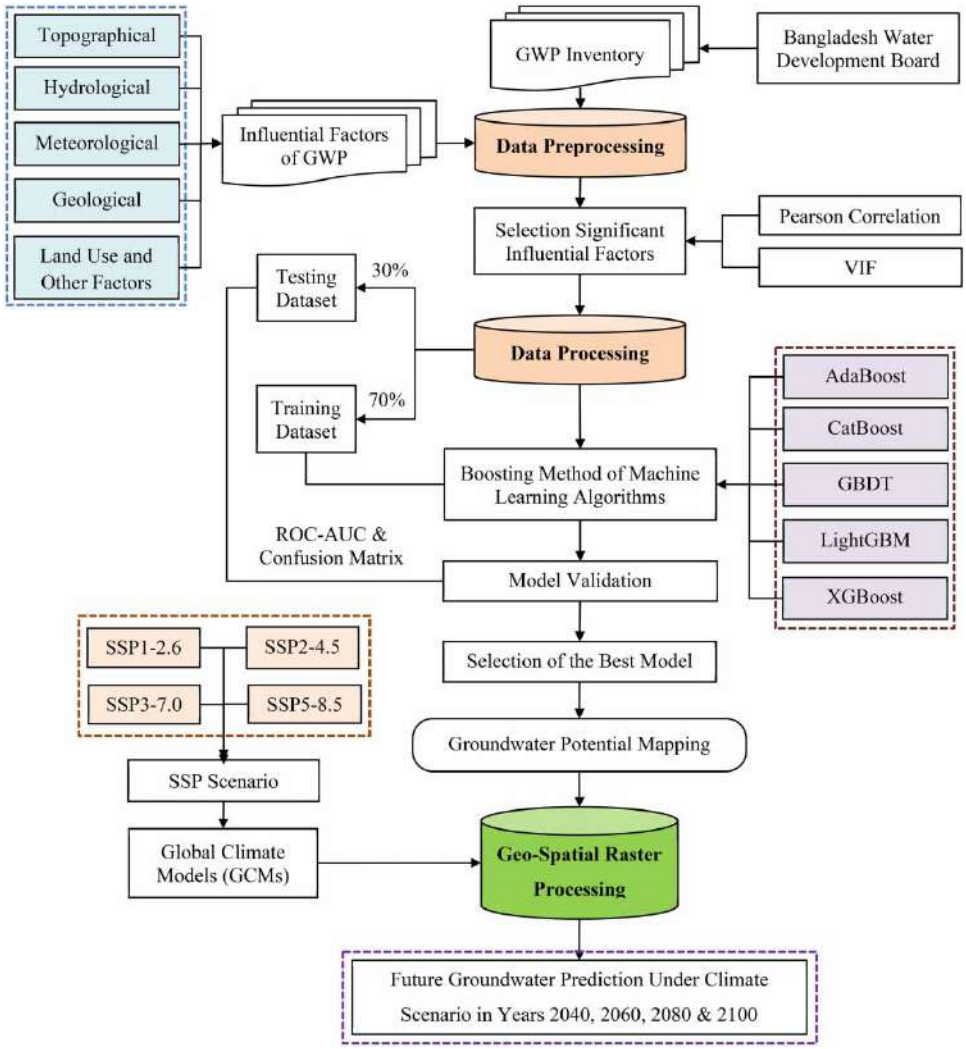

Figure 3 | Methodological framework for the design and execution of research approaches.

3.1.2. Hydrological

Figure 5 illustrates hydrological variables. The availability of surface water and groundwater recharge is contingent upon DD (Figure 5(a)), and the floodplains and low-lying deltaic areas of Bangladesh exhibit considerable DD, whereas certain areas of the Chittagong Hill Tracts exhibit poor DD. Most places in Bangladesh are situated near water due to the extensive stream network (Figure 5(b)). Regions near the coastline experience a decrease in the distances from perennial streams. Bangladesh possesses an intricate river system characterized by substantial FA (Figure 5(c)) in its principal river basins. Most of Bangladesh exhibits elevated SPI (Figure 5(d)) values, while mountainous regions with steep gradients have both high and low SPI values. The TWI (Figure 5(e)) quantifies areas susceptible to water accumulation. The floodplains and low-lying deltaic areas of Bangladesh exhibit elevated TWI values, while coastal regions such as the Sundarbans and southwestern districts also demonstrate high TWI values. Elevated regions in the southeast exhibit reduced TWI values.

3.1.3. Land use and other

Figure 6 depicts land use and environmental factors that significantly influence Bangladesh's groundwater recharge and water resource management. In Bangladesh, road distance (Figure 6(a)) influences land utilization and groundwater recharge, particularly in urbanized regions adjacent to major roads such as Dhaka-Chittagong and Dhaka-Khulna. In LULC (Figure 6(b)), most of Bangladesh's land is allocated for agriculture and barren areas. Water tables diminish in agricultural regions. The

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

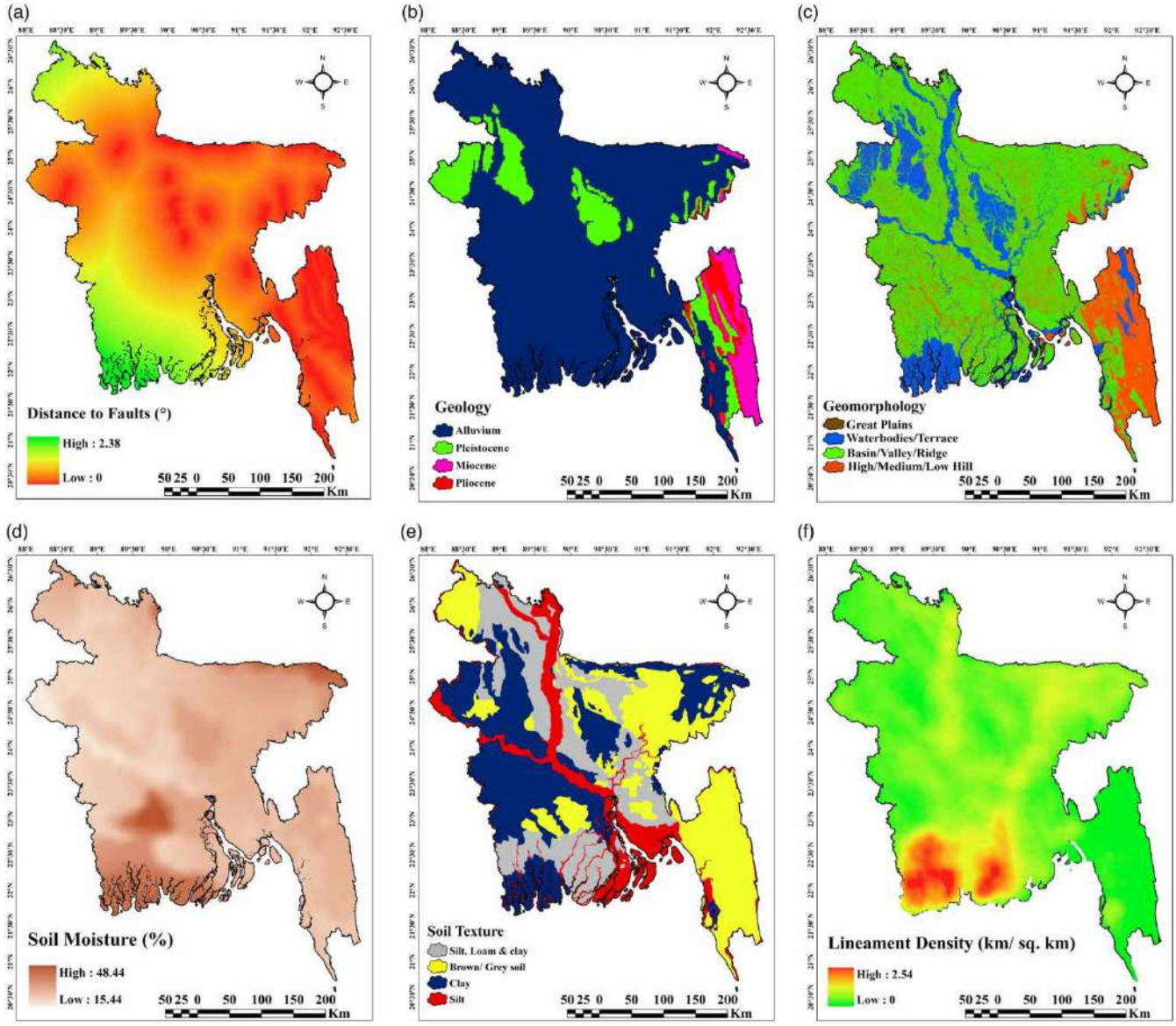

Figure 4 | Geological indicators (a) distance to faults, (b) geology, (c) geomorphology, (d) soil moisture, (e) soil texture, and (f) LD.

NDVI (Figure 6(c)), the lowest NDVI values were observed in the Sundarbans mangrove forest. The urban growth of Dhaka, Chittagong, and Khulna has escalated. The principal rivers, wetlands, and coastal areas of Bangladesh exhibit the highest NDWI (Figure 6(d)) values.

3.1.4. Meteorological

Meteorological indicators in Figure 7 influence groundwater recharge. EV (Figure 7(a)) refers to the cumulative loss of water from terrestrial surfaces and plant transpiration, with hilly locations exhibiting the highest EV values, and Sylhet and Khulna demonstrate significantly low rates. Bangladesh's southwest and mid-south region experiences a high amount of humidity (Figure 7(b)) and also, and coastal regions such as Khulna have elevated humidity levels, while the drier western parts of Rajshahi and Dinajpur display reduced humidity. Bangladesh depends on precipitation (Figure 7(c)) to replenish groundwater, with average annual precipitation varying from 3.02 mm per day in the west to over 15.84 mm per day in northeastern Sylhet. The western and central portions of Bangladesh experience elevated temperatures (Figure 7(d)).

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

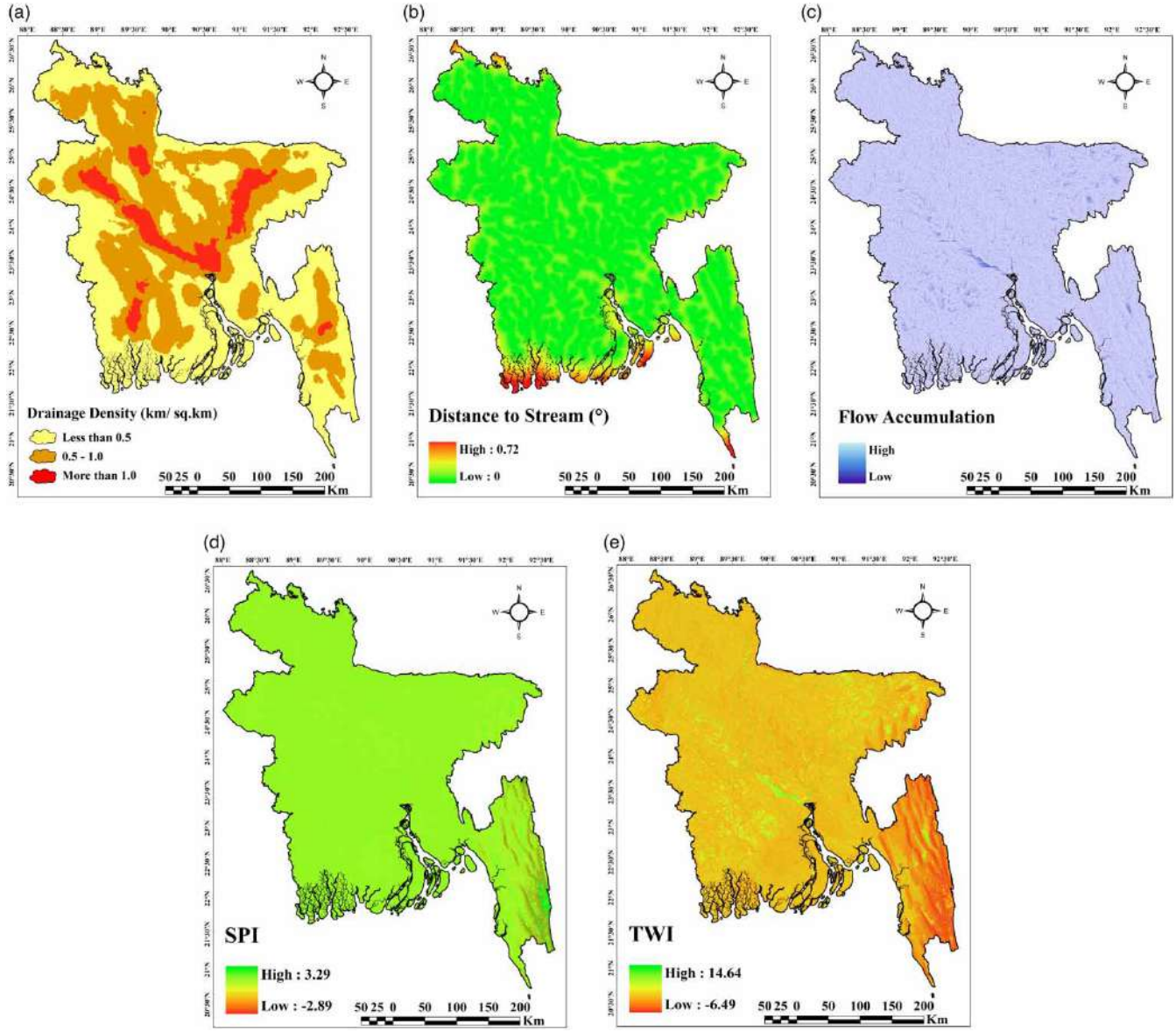

Figure 5 | Hydrological indicators (a) drainage density, (b) distance to stream, (c) flow accumulation, (d) stream power index (SPI), and (e) topographic wetness index (TWI).

3.1.5. Topographical

Figure 8 depicts topographical factors, whereas the aspect (Figure 8(a)) affects sunshine exposure, evaporation, and water retention. The hilly regions of Bangladesh, such as Chittagong and Sylhet, have a larger variability in aspect. Curvature (Figure 8(b)) refers to the concave or convex terrain that affects water accumulation and flow direction in Bangladesh. Most areas of Bangladesh feature concave topographies. The southeast Chittagong Hill Tracts have steep convex terrains. Elevation (Figure 8(c)) affects Bangladesh's topography and hydrology. The northern portions of the north region are slightly raised, whereas in the southeast Chittagong Hill Tracts, the altitudes exceed 1,000 m. The flow of water in Bangladesh across the landscape depends on the slope (Figure 8(d)). The predominant topography of Bangladesh features a gradual slope. Nonetheless, the steep gradients in the southeast hilly regions. In the hilly and upland areas of Bangladesh, STI (Figure 8(e)) values are noted, attributed to steep gradients in hilly regions. Most areas of Bangladesh have diminished STI values. TPI (Figure 8(f)) distinguishes among mountains, valleys, and plains. The topography of Bangladesh has low TPI values. Elevated TPI values

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

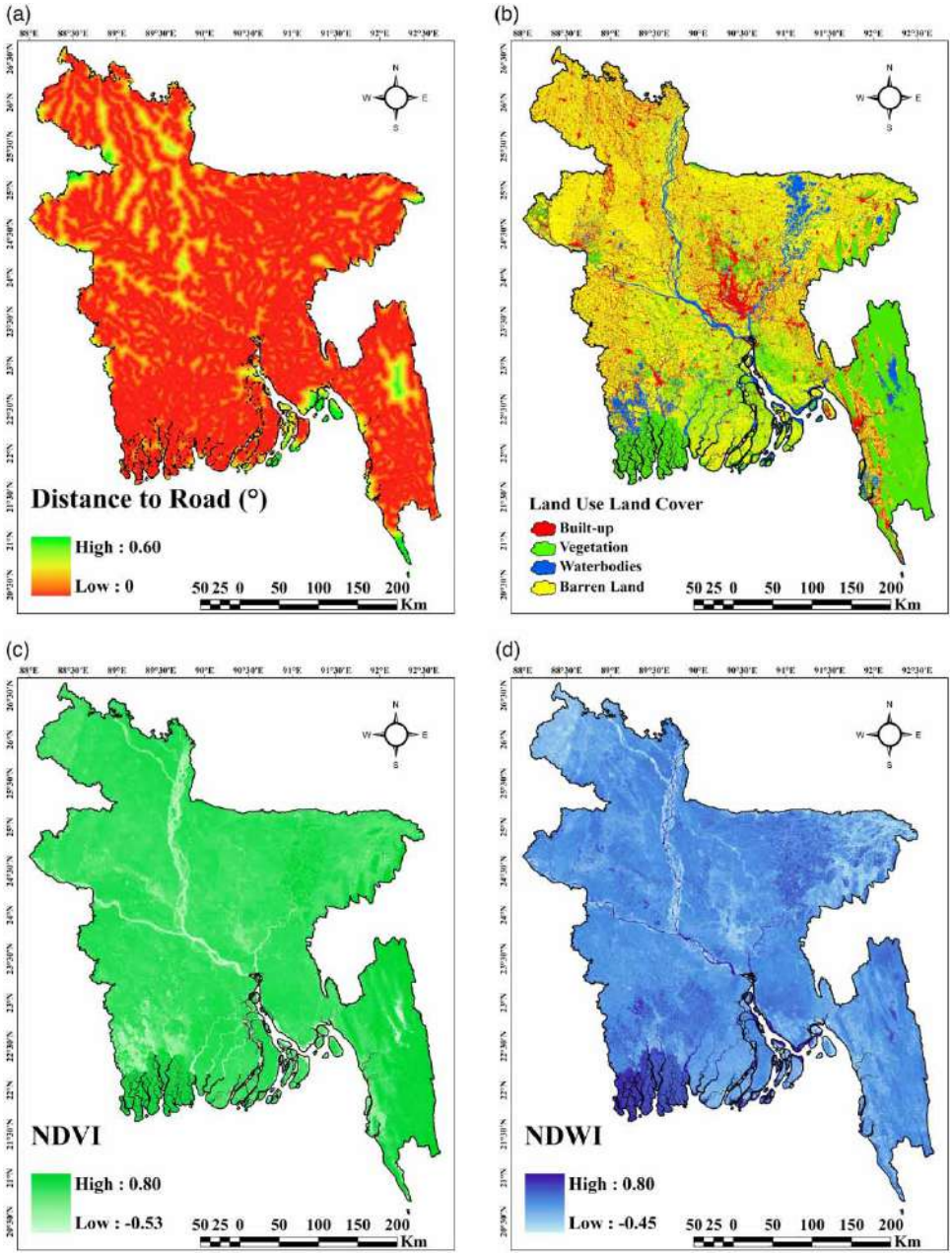

Figure 6 | Land use and other indicators (a) distance to road, (b) land use/land cover (LULC), (c) normalized difference vegetation index (NDVI), and (d) normalized difference water index (NDWI).

signify prominent ridges in the Chittagong Hill Tracts. TRI (Figure 8(g)) measures local elevation variations in Bangladesh, whereas only the rough Chittagong Hill Tracts have elevated TRI values, and the rest of the territory is largely flat.

3.2. Collinearity-informed determination of analytical variables

Table 3 examines 27 explanatory factors, with GWL as the dependent variable and the other 26 as independent variables. The VIF, tolerance, Pearson correlation, and *p*-values are employed to evaluate these variables, determining significance levels and probable collinearity issues. In this database, the majority of variables such as AS (3.84), Curvature (CUV) (1.75), DD (5.08), DTF (4.82), DTR (2.33), DTS (3.07), FA (5.82), GL (1.50), PR (2.63), SL (4.89), SPI (3.62), TPI (2.35), and TWI (2.77) are within acceptable range, so mild multicollinearity is not probable to considerably skew regression estimates.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

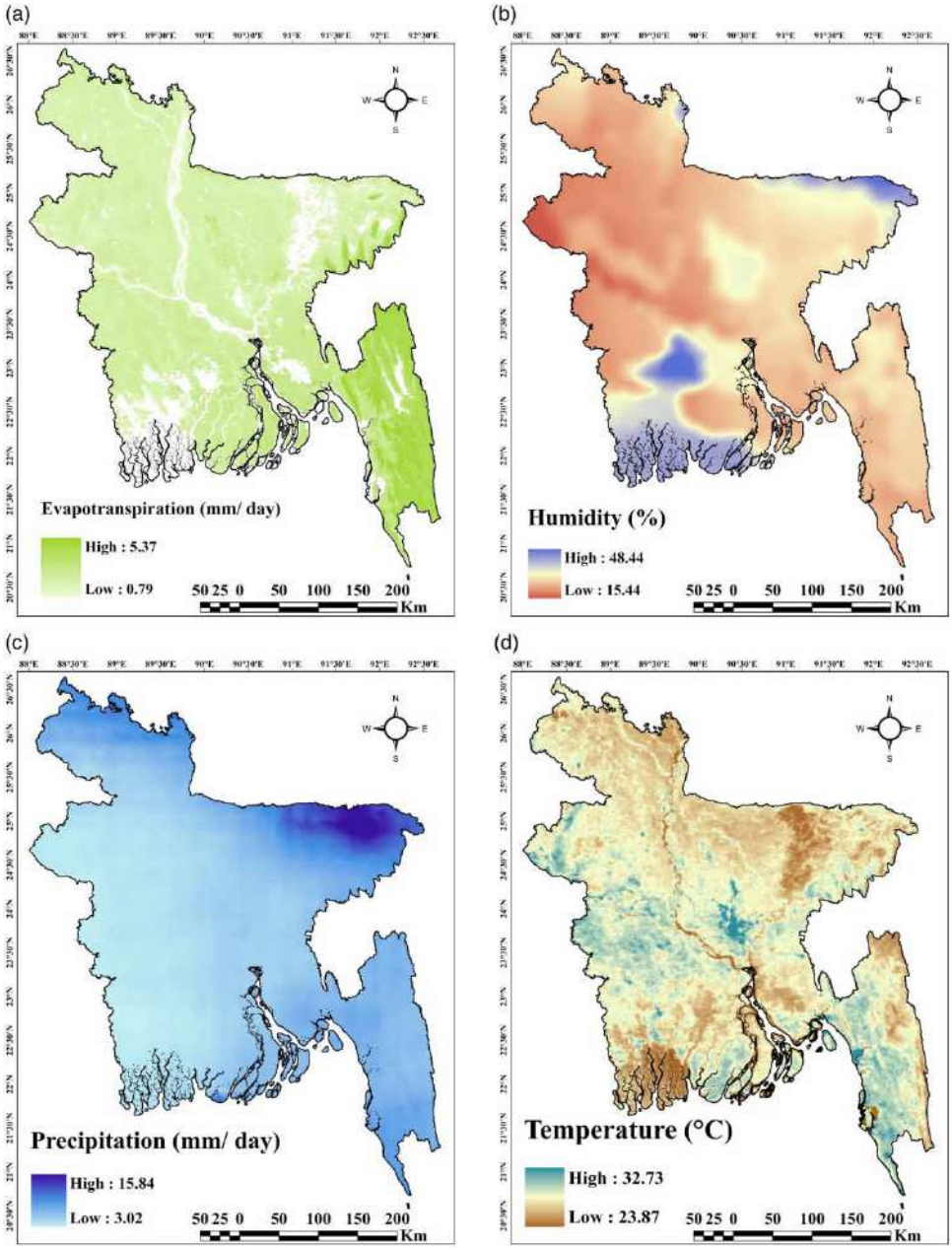

Figure 7 | Meteorological indicators (a) evapotranspiration, (b) humidity, (c) precipitation, and (d) temperature.

Alternatively, variables such as EL (9.61), EV (7.73), ST (8.03), GM (8.96), LD (7.99), LULC (7.69), NDVI (6.95), SM (9.99), and STI (6.52) are near the upper limit, for which there may be a need to exercise caution when interpreting their individual coefficients. It should be noted that Humidity (HUM) (733.42), NDWI (21.65), TEM (546.20), and TRI (36.86) possess very high values of VIF, which indicate serious multicollinearity, which can compromise their independent predictive power. Pearson correlation coefficients inform us regarding the direction and magnitude of linear relationships with GWL. All variables have fairly weak correlations with GWL. Stronger negative correlations are observed for DTF, EV, and HUM, which reflect those higher values in these variables are linked to gradual groundwater declines. Positive relations are also observed in GL and ST, reflecting a bias toward increasing groundwater levels with growing values of the parameters. The *p*-values also stratify statistical significance since most variables have values below the conventional 0.05 threshold, confirming their utility in

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

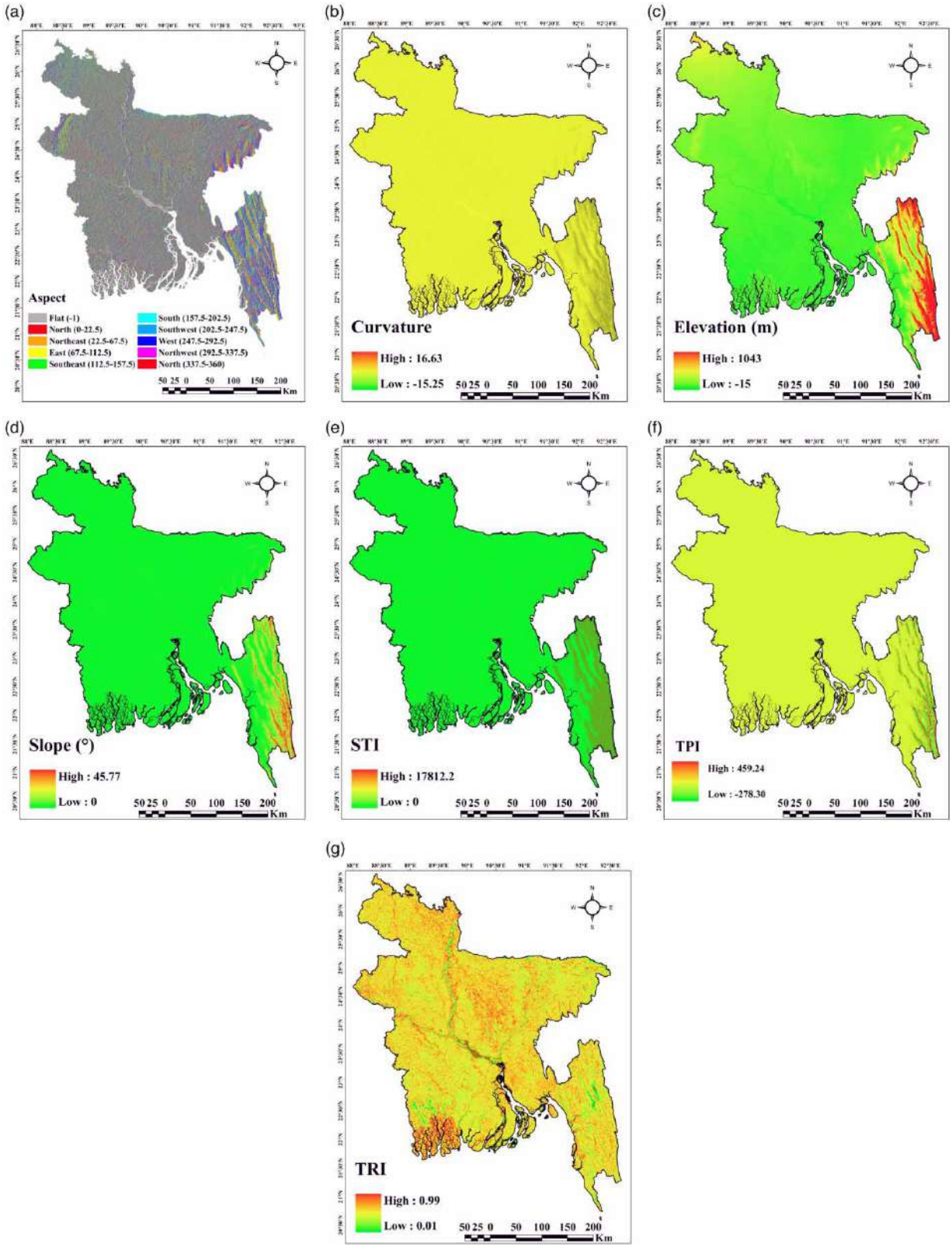

Figure 8 | Topographical indicators (a) aspect, (b) curvature, (c) elevation, (d) slope, (e) sediment transport index, (f) topographic position index (TPI), and (g) topographic roughness index (TRI).

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Table 3 | The predictors yielded VIF and *p*-value

Feature	VIF	Tolerance	Pearson correlation	*p*-value
GWL			1	0.000
AS	3.84	0.26	0.01	0.781**
CUV	1.75	0.56	0.00	0.825**
DD	5.08	0.19	0.07	0.002
DTF	4.82	0.20	− 0.26	0.001
DTR	2.33	0.42	0.01	0.628**
DTS	3.07	0.32	− 0.07	0.002
EL	9.61	0.10	− 0.15	0.001
EV	7.73	0.12	− 0.19	0.001
FA	5.82	0.17	0.02	0.005
GL	1.50	0.66	0.29	0.001
GM	8.96	0.11	− 0.11	0.001
HUM	733.42***	0.00	− 0.23	0.001
LD	7.99	0.12	− 0.17	0.001
LULC	7.69	0.13	0.01	0.008
NDVI	6.95	0.14	− 0.13	0.001
NDWI	21.65***	0.04	0.03	0.215**
PR	2.63	0.37	− 0.17	0.001
SL	4.89	0.20	0.04	0.001
SM	9.99	0.10	− 0.15	0.001
SPI	3.62	0.27	− 0.01	0.589**
ST	8.03	0.12	0.26	0.001
STI	6.52	0.15	0.00	0.823**
TEM	546.20***	0.00	0.09	0.001
TPI	2.35	0.42	0.04	0.006
TRI	36.86***	0.02	0.03	0.259**
TWI	2.77	0.36	− 0.03	0.001

Note: Variables with *p*-values beyond 0.05 are denoted with (**), and those with a VIF surpassing 10 are indicated with (***), and they are omitted from the model as they are statistically insignificant.

explaining differences in GWL. Notably, variables such as DD (0.002), DTF (0.001), DTS (0.002), EL (0.001), EV (0.001), FA (0.005), GL (0.001), GM (0.001), Hum (0.001), LD (0.001), LULC (0.008), NDVI (0.001), PR (0.001), SL (0.001), SM (0.001), ST (0.001), TEM (0.001), TPI (0.006), and TWI (0.001) are identified as having satisfactory statistical significance. Conversely, in direct contrast to the above, variables with *p*-values >0.05, i.e., AS (0.781), CUV (0.825), DTR (0.628), NDWI (0.215), SPI (0.589), STI (0.823), and TRI (0.259), do not reveal any significant linear correlations with GWL, implying that their impact may be minimal or non-linear.

Figure 9 shows the correlation matrix offers significant insights into the correlations between variables, depicting their degree of association via a color-coded heatmap. Robust positive correlations, denoted by red hues, imply a substantial association between variables, whereas robust negative correlations, illustrated in blue, signify an inverse relationship as the STI and FA exhibit a strong correlation of 0.89.

3.3. Assessment of key factors derived through modeling

Figure 10 illustrates the relevance scores from five boosting-based ML models, indicating the most significant parameters in GWP mapping. Precipitation (PR) is the most significant element, attaining the highest scores across all models, from 11

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

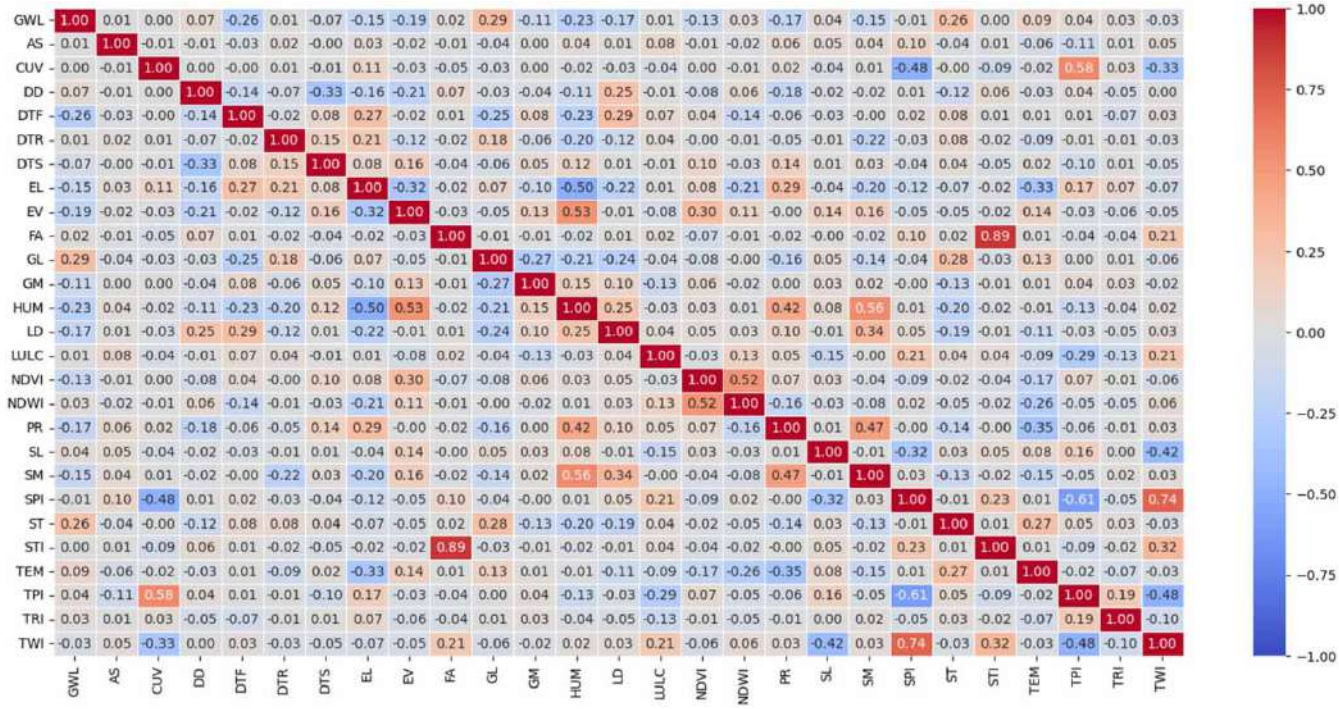

Figure 9 | Correlation among the qualities and the goal variable, as well as among the features.

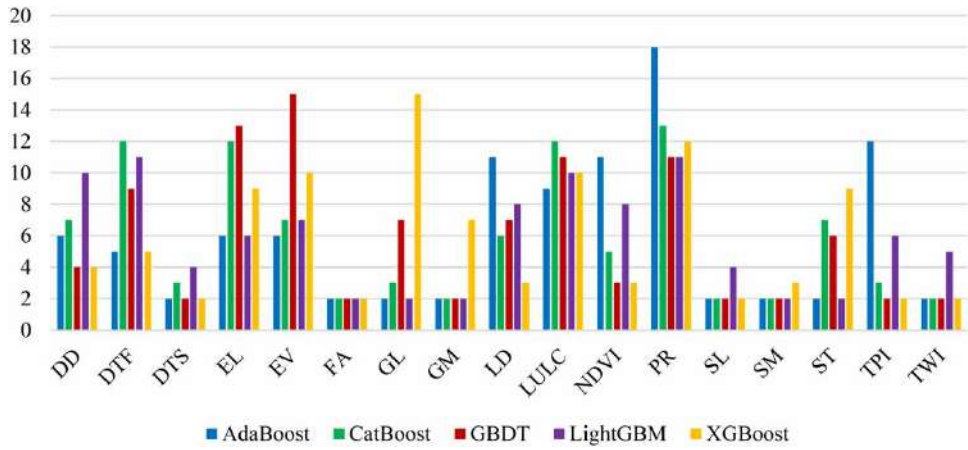

Figure 10 | Importance score of the influential factors of GWP.

(GBDT, LightGBM) to 18 (AdaBoost). Precipitation proved to be the paramount element affecting GWP, underscoring the strong correlation between precipitation and aquifer recharge. Likewise, LULC serves as a crucial factor, with scores ranging from 9 (AdaBoost) to 12 (CatBoost). Additional significant parameters encompass elevation (EL) and EV, which exhibit differing levels of relevance among models. Elevation ratings vary from 6 (AdaBoost, LightGBM) to 13 (GBDT), while EV scores range from 6 (AdaBoost, CatBoost) to 15 (GBDT), underscoring their critical functions in water dynamics and depletion. DD and LD contribute moderately, achieving scores of 11 (AdaBoost) for LD and 10 (LightGBM) for DD. Conversely, indicators such as the NDVI and soil texture (ST) exhibit diminished significance, with NDVI registering scores ranging from 3 (GBDT) to 11 (AdaBoost) and ST sustaining values between 2 and 9. Likewise, the DTF, the TPI, the

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

TWI, and the DTS demonstrate negligible influence, regularly receiving scores between 2 and 12. Various additional characteristics, such as FA, slope (SL), geology (GL), geomorphology (GM), and soil moisture (SM), typically exhibit minimal significance.

3.4. Spatial distribution of GWP based on ML models

Table 4 and Figure 11 depict the examination of GWP zones, classified into low, moderate, and high potential, employing various ML techniques. CatBoost assessed that 1.33% (1,965.13 km^2) of the study area is classified as low potential, while the majority, 69.24% (102,171.28 km^2), is classified as moderate potential, and 29.43% (43,433.59 km^2) as high potential. AdaBoost identified a larger percentage of 9.20% (13,575.38 km^2) in the low potential category, with 53.15% (78,427.30 km^2) classified as intermediate and 37.65% (55,567.31 km^2) as high. XGBoost results classified 11.70% (17,266.81 km^2) as low, 52.88% (78,038.28 km^2) as moderate, and 35.42% (52,264.91 km^2) as high. LightGBM produced estimates similar to CatBoost, with 6.28% (9,268.09 km^2) in low-potential zones, 60.09% (88,677.14 km^2) in moderate potential zones, and 33.63% (49,624.78 km^2) in high-potential zones. GBDT determined that the high-potential group included 33.07% (48,799.53 km^2), whereas 10.53% (15,544.34 km^2) was categorized as low and 56.40% (83,226.13 km^2) as moderate.

3.5. Accuracy assessment and validation of models

3.5.1. ROC-AUC

Figures 12 and 13 illustrate the predicted accuracy of various boosting-based ML models in GWP mapping, with an analysis of the AUC values. Of the five models evaluated, XGBoost, CatBoost, and LightGBM had superior predictive performance, each attaining an AUC value of above 0.90. GBDT exhibited commendable performance, achieving an AUC of 0.89. AdaBoost, although somewhat inferior in performance, attained an AUC of 0.85; however, it is less effective than the other ensemble models. AdaBoost demonstrated reduced accuracy owing to its susceptibility to noisy data. Of all evaluated models, LightGBM and CatBoost exhibited the highest performance, presumably owing to their capacity to manage unbalanced data.

3.5.2. Confusion matrix

Figure 14 illustrates that the confusion matrix analysis offers a comprehensive assessment of the classification efficacy of several boosting-based ML models in GWP mapping. AdaBoost demonstrated an accuracy of 0.78 and an F1-score of 0.74. CatBoost and XGBoost exhibited superior performance, attaining an accuracy of 0.84 and F1-scores of 0.81 and 0.82. GBDT, although marginally less effective, exhibited consistent predictive capability with an accuracy of 0.80 and an F1-score of 0.77. LightGBM proved to be the most effective model, achieving a maximum accuracy of 0.85 and an F1-score of 0.82.

3.6. Spatial distribution of GWP based on the optimal ML model

Figure 15 and Table 5 depict the evaluation of GWP distribution throughout various divisions in Bangladesh, employing the ideal model, LightGBM, which categorized areas into low, moderate, and high-potential zones. Barisal contained no areas classified as low potential; yet, 19.73% (2,609.65 km^2) of its land was classified as moderate potential, while 80.27% (10,615.55 km^2) was labeled as high potential. In Chittagong, 27.43% (9,301.02 km^2) was classified as low potential, 63.95% (21,683.43 km^2) as moderate potential, and 8.62% (2,924.10 km^2) as high potential. Dhaka comprised 0.10% (21.45 km^2) of its area categorized as low potential, while 74.03% (15,245.02 km^2) was classed as intermediate, and

Table 4 | Distribution of GWP according to different ML models in Bangladesh

	Low		Moderate		High	
Models	**Percentage**	**Area (km^2)**	**Percentage**	**Area (km^2)**	**Percentage**	**Area (km^2)**
CatBoost	1.33	1,965.13	69.24	102,171.28	29.43	43,433.59
AdaBoost	9.20	13,575.38	53.15	78,427.30	37.65	55,567.31
XGBoost	11.70	17,266.81	52.88	78,038.28	35.42	52,264.91
LightGBM	6.28	9,268.09	60.09	88,677.14	33.63	49,624.78
GBDT	10.53	15,544.34	56.40	83,226.13	33.07	48,799.53

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

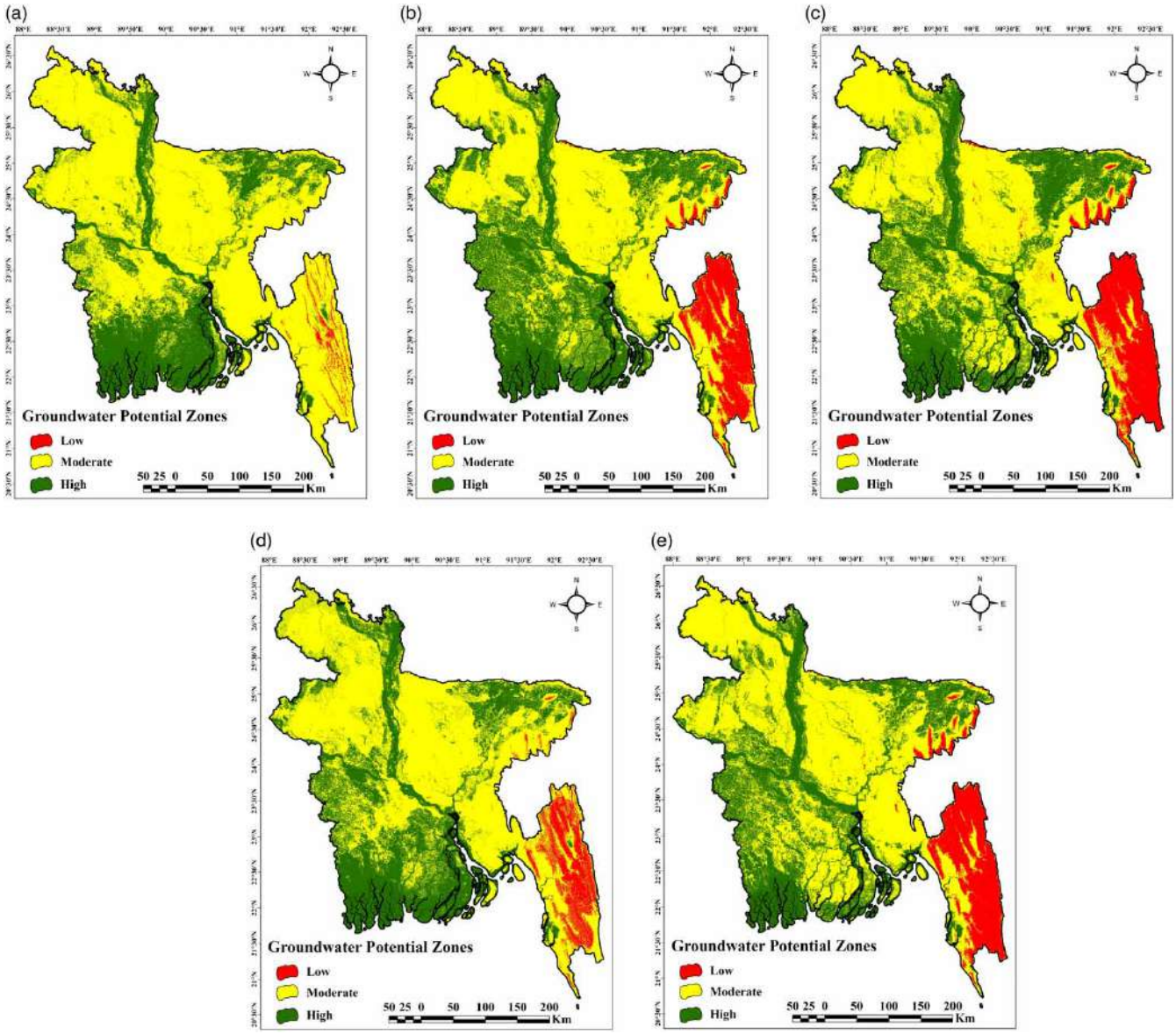

Figure 11 | GWP maps generated based on ML models (a) AdaBoost (0.80), (b) CatBoost (0.88), (c) GBDT (0.85), (d) LightGBM (0.93), and (e) XGBoost (0.88).

25.87% (5,327.27 km^2) was identified as high potential. Khulna exhibited a similar pattern, with no areas classified as low potential; 28.17% (6,276.81 km^2) was classified as moderate potential, and 71.83% (16,007.41 km^2) as high potential. In Mymensingh, only 0.05% (5.07 km^2) of the land was classified as low potential, while 79.45% (8,408.62 km^2) was deemed moderate, and 20.51% (2,170.38 km^2) was identified as high potential. Rajshahi labeled a mere 0.05% (8.83 km^2) as low potential, whereas 74.19% (13,467.94 km^2) was classified as moderate potential, and 25.76% (4,676.32 km^2) as high potential. Rangpur displayed 0.01% (1.72 km^2) of its area in the low category, 72.91% (11,800.12 km^2) in the midrange category, and 27.08% (4,383.15 km^2) in the high-potential category. Sylhet exhibited 2.20% (278.42 km^2) in the low-potential category, 58.78% (7,427.37 km^2) in the moderate potential category, and 39.01% (4,929.43 km^2) in the high-potential category.

3.7. Climate change projections for future scenarios

The annual average precipitation in Bangladesh from 2040 to 2100, under various SSPs, has considerable regional and temporal differences, as seen in Figure 16. The distribution of precipitation throughout the nation fluctuates according to the

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

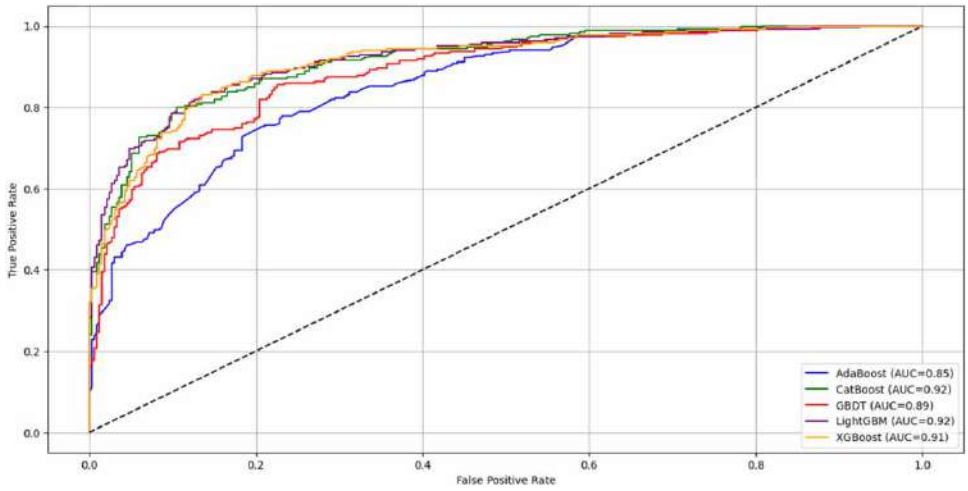

Figure 12 | ROC-AUC for different boosting models.

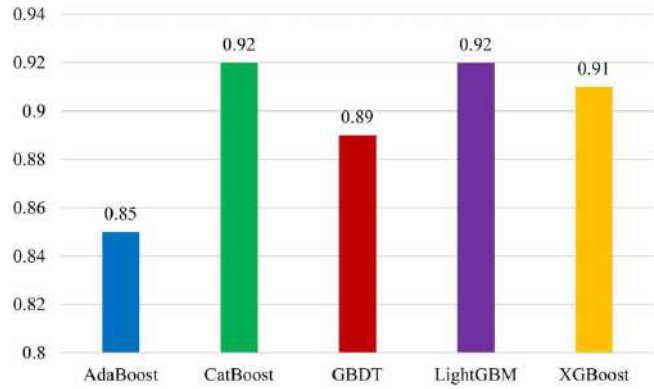

Figure 13 | AUC scores in different models.

emission scenario, with elevated precipitation levels being seen in the northeastern and southeastern areas, while the central and western regions frequently encounter reduced levels. Under SSP1-2.6, a drying trend is apparent, with maximum precipitation decreasing from 2,165.35 mm in 2040 to 1,954.04 mm in 2100. SSP2-4.5 exhibits modest variability, with a slight increase in precipitation. SSP3-7.0 experiences a significant increase, reaching a maximum of 2,571.69 mm in 2060. The most extreme scenario, SSP5-8.5, indicates a significant escalation, with maximum precipitation rising from 1,983.7 mm in 2040 to 2,982.67 mm in 2100.

3.8. Spatial analysis of GWP under multiple future climate change projections

GWP distribution across several SSPs from 2040 to 2100 indicates significant variations throughout Bangladesh, as illustrated in Figure 17 and Table 6. The allocation of land classified as low, medium, and high GWP fluctuates throughout time, shaped by climatic conditions in each case. Under SSP1-2.6, the high GWP category initially constitutes 25.27% (37,288.42 km^2) of the total land area in 2040, rising to 34.89% (51,494.21 km^2) by 2060. Nevertheless, it decreases to 13.84% (20,420.67 km^2) by the year 2100. The moderate potential continues to prevail at approximately 63%, whereas the low-potential category shows a substantial rise from 11.45% (16,893.47 km^2) in 2040 to 22.35% (32,984.01 km^2) in 2100. In SSP2-4.5, the high GWP region varies over time, accounting for 34.65% (51,136.33 km^2) in 2040 and rising to 40.21% (59,336.68 km^2) by 2100. The low-potential group is projected to be small, at 19.61% (28,941.75 km^2) in 2080, before vanishing completely by 2100. The moderate potential continues to be the prevailing category, including more than 50% of the land area in the majority of scenarios. In SSP3-7.0, the high GWP category exhibits relative stability, with a slight increase from 37.55% (55,417.20 km^2) in 2040 to

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

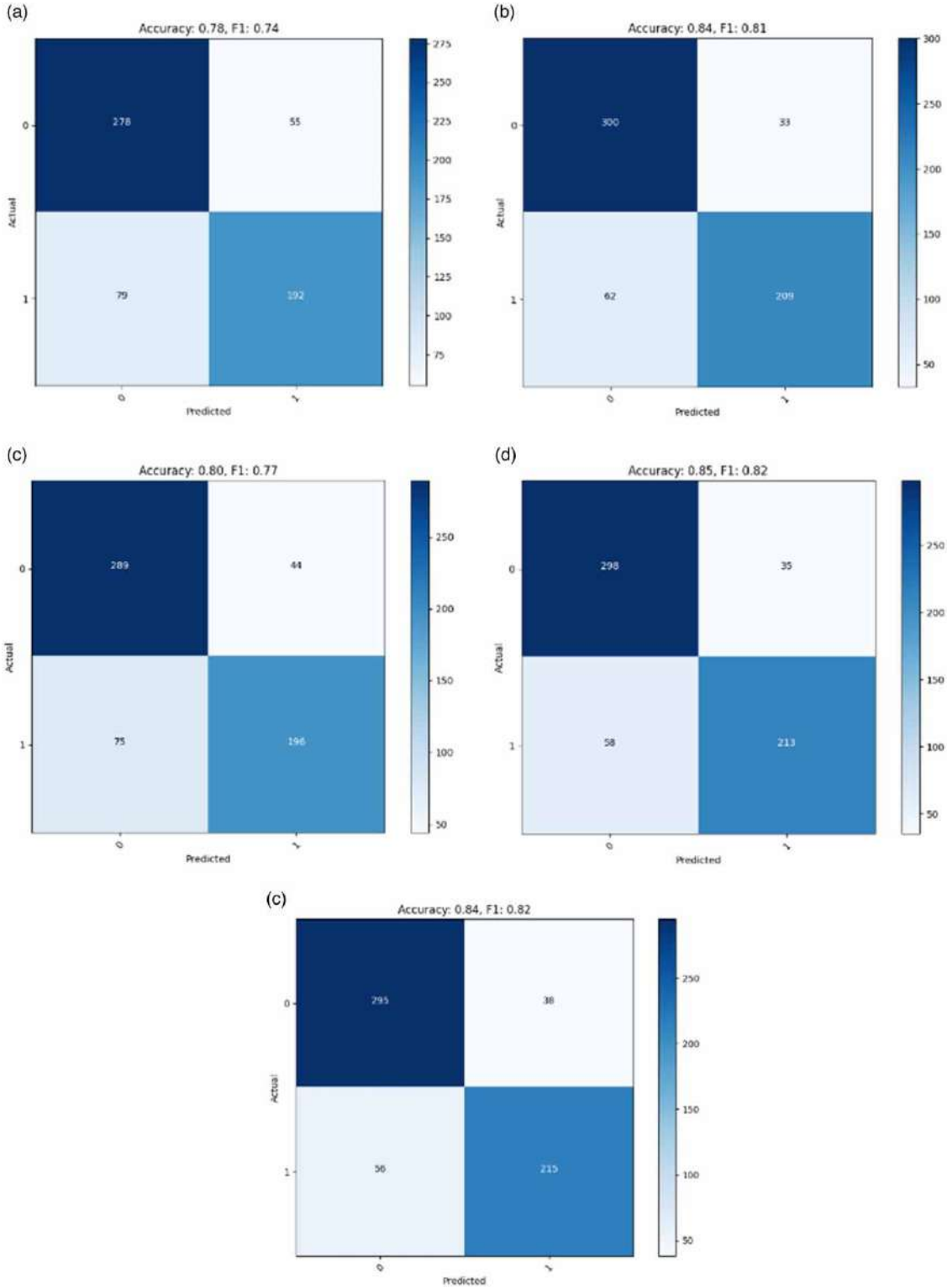

Figure 14 | Confusion matrix for boosting-based ML models (a) AdaBoost (0.78), (b) CatBoost (0.84), (c) GBDT (0.80), (d) LightGBM (0.85), and (e) XGBoost (0.84).

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

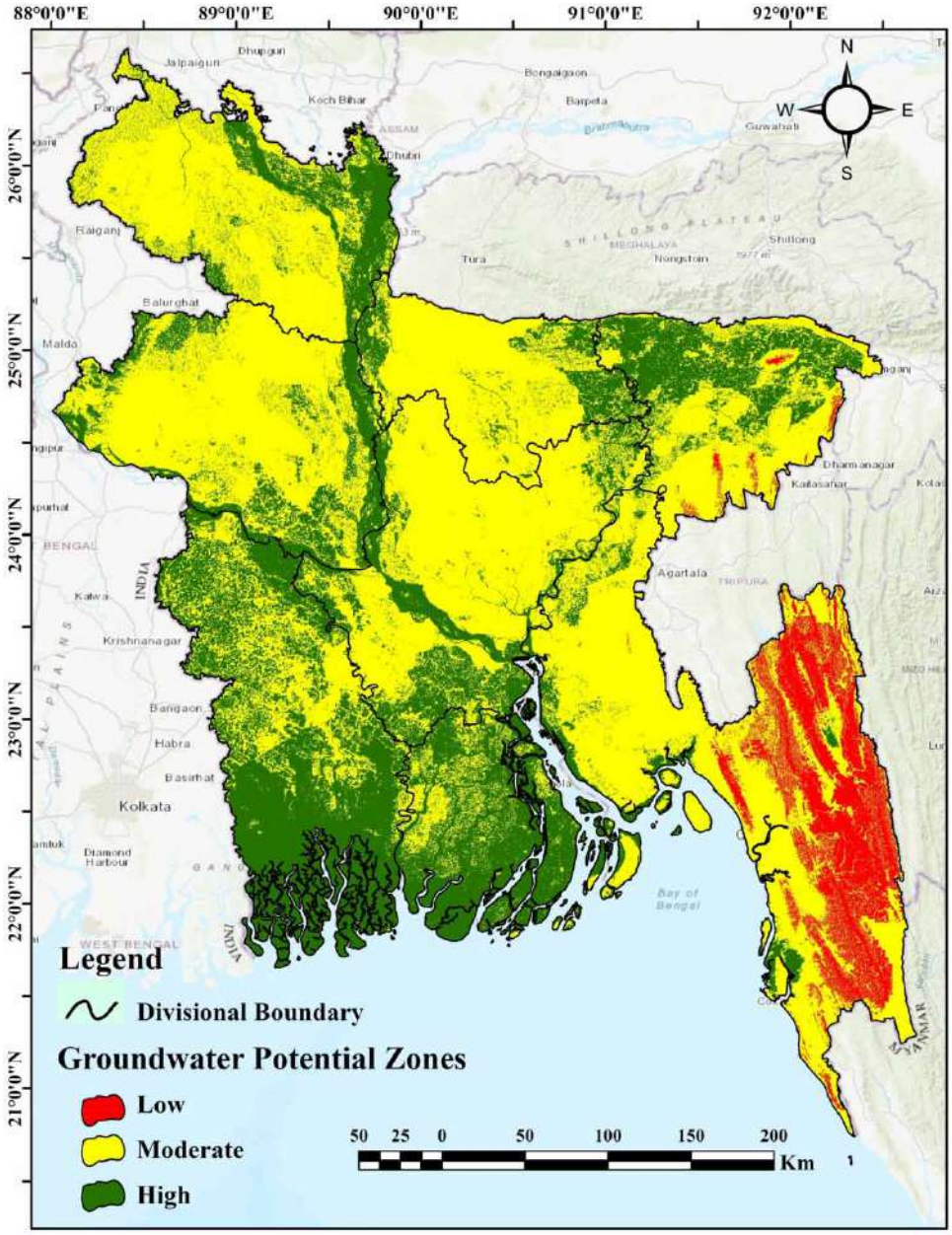

Figure 15 | Division-wise GWP by the best model (LightGBM).

47.32% (69,824.72 km^2) in 2100. The intermediate group exhibits slight variations, decreasing from 62.45% (92,152.80 km^2) in 2040 to 52.68% (77,745.28 km^2) in 2100. Significantly, there is no terrain classified as having low GWP for the whole studied period. In SSP5-8.5, the high GWP category exhibits a substantial rise, increasing from 35.64% (52,587.68 km^2) in 2040 to 65.12% (96,097.57 km^2) in 2100. The moderate category predominates for the majority of the period, reaching a zenith of 64.36% (94,982.32 km^2) in 2040, followed by a slow decline thereafter. The low-potential group emerges briefly in 2080, encompassing 5.07% (7,483.97 km^2) before vanishing in 2100.

4. DISCUSSION

The study presents a novel approach for effectively forecasting the spatiotemporal dynamics of GWP under projected climate change scenarios by combining important variables with sophisticated boosting-based ML. Geographical heterogeneity in

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Table 5 | Distribution of GWP among divisions in Bangladesh by the best model (LightGBM)

	Low		Moderate		High	
Model	**Percentage**	**Area (km²)**	**Percentage**	**Area (km²)**	**Percentage**	**Area (km²)**
Barisal	0.00	0.00	19.73	2,609.65	80.27	10,615.55
Chittagong	27.43	9,301.02	63.95	21,683.43	8.62	2,924.10
Dhaka	0.10	21.45	74.03	15,245.02	25.87	5,327.27
Khulna	0.00	0.00	28.17	6,276.81	71.83	16,007.41
Mymensingh	0.05	5.07	79.45	8,408.62	20.51	2,170.38
Rajshahi	0.05	8.83	74.19	13,467.94	25.76	4,676.32
Rangpur	0.01	1.72	72.91	11,800.12	27.08	4,383.15
Sylhet	2.20	278.42	58.78	7,427.37	39.01	4,929.43

GWP aligns with prior research, including Nowreen *et al.* (2020), Hasan *et al.* (2021), Elbeltagi *et al.* (2022), Sarkar *et al.* (2022b), and Fatema *et al.* (2023), conducted in certain locations of Bangladesh. Our study builds upon this knowledge at the national level, employing certain boosting algorithms and future climatic conditions to evaluate GWP in Bangladesh. This study emphasizes the intricate interdependencies among several factors influencing groundwater supply at the national level in Bangladesh. A collinearity check was also conducted to further enhance predictive modeling, resulting in the selection of 17 suitable parameters out of an initial 26 for approximating GWP, excluding variables with *p*-values of over 0.05 or VIF values of over 10 due to being statistically insignificant. Among the 17 suitable parameters, LULC and precipitation are the primary drivers affecting GWP, with a mean of 13% contributed by precipitation and a mean of 10% contributed by LULC in terms of overall contribution. Topological and hydrological parameters, which increase the precision of water resources assessment, have been shown by Singhal *et al.* (2024). LULC and precipitation were also presented as the primary drivers affecting GWP (Díaz-Alcaide & Martínez-Santos 2019). The contribution of LULC is significant since urbanization has intensified at the cost of reduced infiltration rates and surface runoff, thus enhancing groundwater depletion (Siddik *et al.* 2022). Conversely, precipitation is the most dominant driver in recharging groundwater, with regions having higher precipitation exhibiting improved groundwater retention capacity as principles of climatology (Sarker 2022; Sadeak *et al.* 2023). Internationally, LULC modifications and precipitation regimes are known to have a substantial influence on GWP. Agricultural intensification and urbanization have been found to decrease infiltration and increase surface runoff, which ultimately decreases groundwater recharge in Spain (Ávila-Marín *et al.* 2025). The widespread impact of LULC on aquifer sustainability is also supported by studies conducted in Southeast Asia and South Asia, as attested by the fact that excessive groundwater depletion has been induced by urbanization at a fast rate through increasing impervious surface area, changing natural land to built-up environment, and also furthermore (Haque *et al.* 2013; Siddik *et al.* 2022; Salim *et al.* 2024). Our findings of the main parameters impacting GWP estimation concur with recent research, and hence, validate the validity and reliability of our results and methodology. Such consistency with existing literature validates the methodology and offers proof of its suitability for use in the estimation of groundwater in other hydrogeological settings. All the boosting algorithms had a very good prediction capacity with AUC values of greater than 0.80, as AdaBoost (0.85), CatBoost (0.92), GBDT (0.89), LightGBM (0.92), and XGBoost (0.91). The high performance of these models in groundwater assessment conforms to the findings of earlier studies (e.g., Maskooni *et al.* 2020; Naghibi *et al.* 2020; Park & Kim 2021; Sachdeva & Kumar 2021; Rao *et al.* 2022; Rasool *et al.* 2022; Dey *et al.* 2023; Halder *et al.* 2024), indicating the strength of boosting-based models with respect to hydrogeological modeling stability. Among five ML boosting algorithms, LightGBM was found to perform the best, which aligns with current studies (Mohtaram *et al.* 2024), who highlighted the importance of LightGBM over other techniques in long hydrological research due to its ability to handle high-dimensional environmental and hydroclimatic data well. LightGBM is an efficient and accurate ML technique superior to groundwater studies by efficiently managing large data, describing intricate non-linear processes, and improving computational performance because of its advanced gradient boosting architecture (Zegaar *et al.* 2024). Additionally, another recent research has further highlighted LightGBM's superiority in GWP mapping (Guo *et al.* 2023; Xiong *et al.* 2023), which improves the validity and consistency of our outcomes. There are significant regional disparities in GWP throughout Bangladesh, according to the analysis of spatial distribution. Khulna (71.83%) and

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

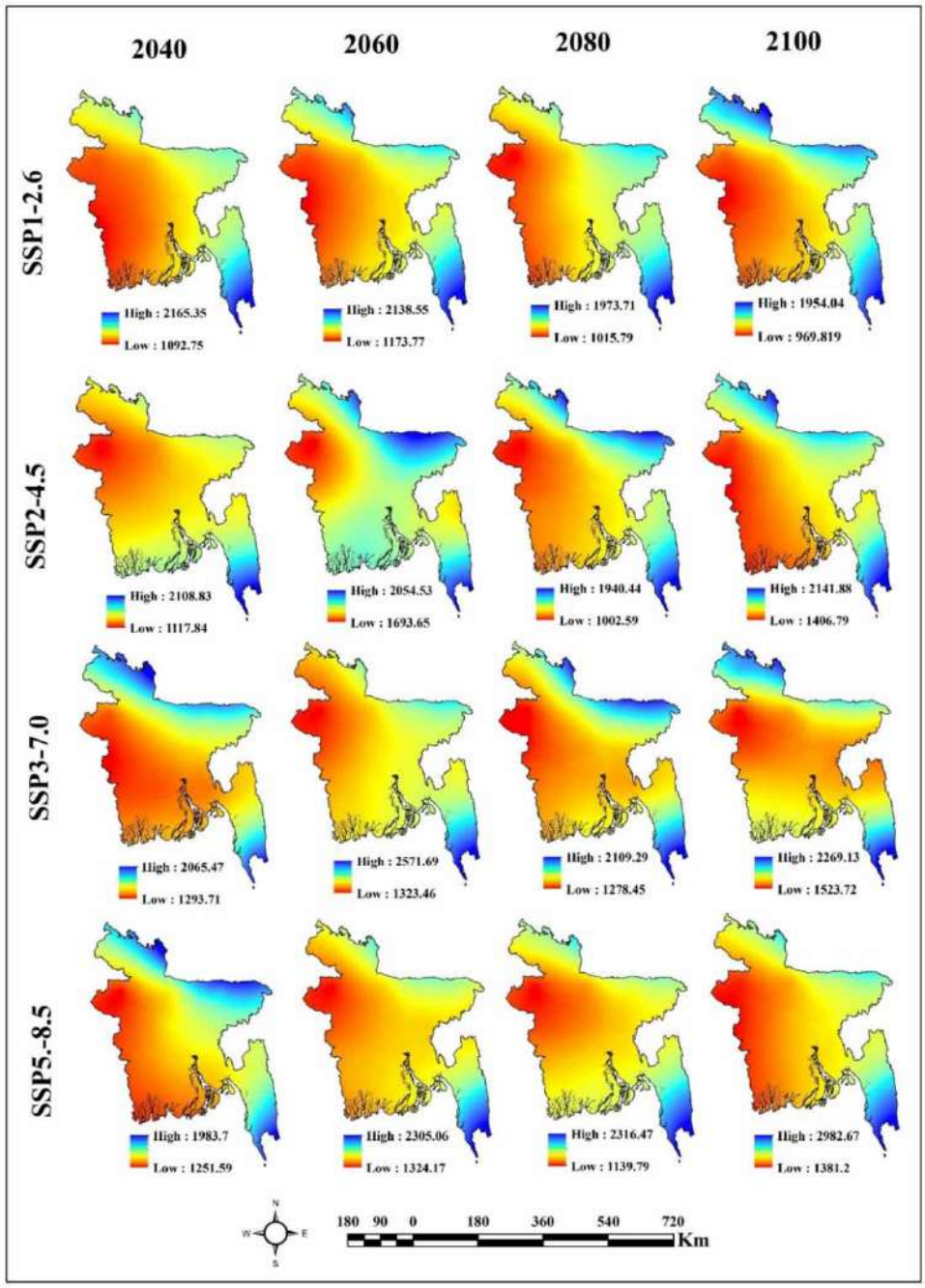

Figure 16 | Annual average precipitation for different SSP scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5) from 2040 to 2100.

Barishal (80.27%) have the largest proportions of areas with high GWP. Surprisingly, both divisions had an extremely small percentage of low GWP areas, almost zero, due to the favorable alluvial geology, gentle topography, and soil texture for high infiltration, topography of flat, highly permeable aquifers, and high monsoonal and fluvial recharge from rivers like the Ganges-Brahmaputra system. In contrast, with the highest percentage of low GWP zones (27.43%) and the lowest percentage of high-potential zones (8.62%), Chittagong ranked worst due to its hilly, sloping relief, high surface runoff, and few areas of natural recharge, exacerbated by urbanization and the dominance of impervious surfaces. Confirming the overall spatial variation of GWP in Khulna, Barisal, and Chittagong Divisions, these results are similar to those obtained using Sarkar *et al.*

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

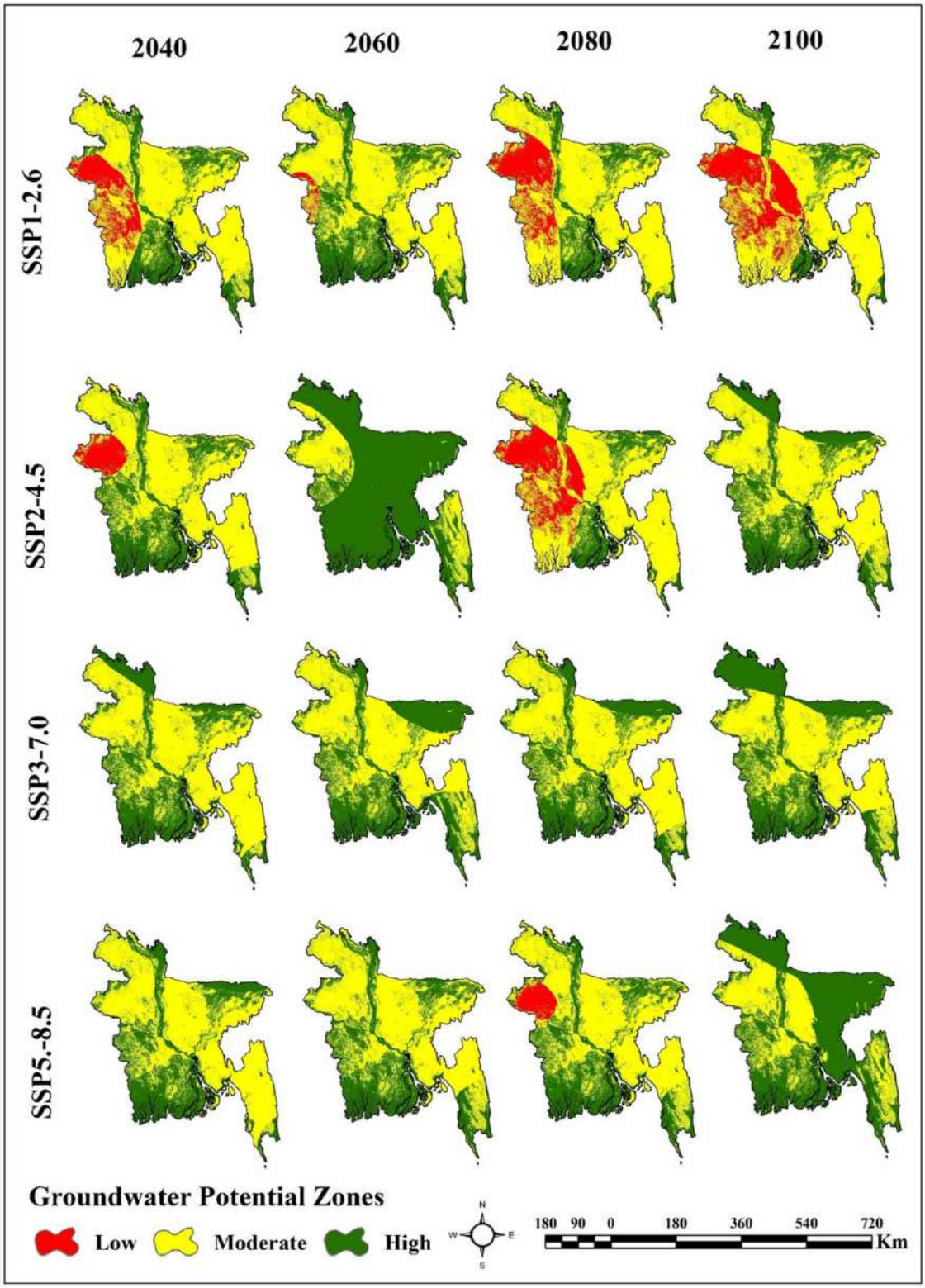

Figure 17 | GWP for different SSP scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5) from 2040 to 2100 with the best model (LightGBM).

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Table 6 | Distribution of GWP according to different SSP scenarios

		Low		Moderate		High	
Model	**Year**	**Percentage**	**Area (km²)**	**Percentage**	**Area (km²)**	**Percentage**	**Area (km²)**
SSP1-2.6	2040	11.45	16,893.47	63.28	93,388.11	25.27	37,288.42
	2060	1.32	1,948.74	63.78	94,127.05	34.89	51,494.21
	2080	14.61	21,557.91	64.31	94,897.81	21.08	31,114.28
	2100	22.35	32,984.01	63.81	94,165.32	13.84	20,420.67
SSP2-4.5	2040	6.29	9,283.71	59.06	87,149.96	34.65	51,136.33
	2060	0.00	0.00	19.41	28,642.22	80.59	118,927.78
	2080	19.61	28,941.75	62.53	92,278.17	17.86	26,350.08
	2100	0.00	0.00	59.79	88,233.32	40.21	59,336.68
SSP3-7.0	2040	0.00	0.00	62.45	92,152.80	37.55	55,417.20
	2060	0.00	0.00	55.25	81,535.78	44.75	66,034.22
	2080	0.00	0.00	61.70	91,044.35	38.30	56,525.65
	2100	0.00	0.00	52.68	77,745.28	47.32	69,824.72
SSP5-8.5	2040	0.00	0.00	64.36	94,982.32	35.64	52,587.68
	2060	0.00	0.00	62.13	91,686.45	37.87	55,883.55
	2080	5.07	7,483.97	58.37	86,136.47	36.56	53,949.56
	2100	0.00	0.00	34.88	51,472.43	65.12	96,097.57

(2024a). Precipitation conditions were used as the prevailing climate change proxies in our research, as it has been used in past and ongoing research, such as Lee *et al.* (2019) and Wu *et al.* (2025), who have results that precipitation patterns are one of the factors most susceptible to predicting future climate changes on groundwater levels. As LightGBM provided the best output, a GWP map under climate change scenarios was prepared using this algorithm with four varied SSP scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5) to analyze the response of precipitation and zones of GWP to various climate change scenarios across various years (2040, 2060, 2080, and 2100) and to identify high temporal and spatial variations in groundwater occurrence owing to climate-induced changes in the precipitation rates. SSPs offer consistent projections of future socioeconomic and emission pathways, allowing integrated climate impact evaluation and efficient decision-making for sustainable resource management (O'Neill *et al.* 2014; Frame *et al.* 2018). SSP scenarios indicate that precipitation plays an important role in influencing GWP in Bangladesh. Heavy precipitation enhances groundwater recharge in the northern and southeastern regions of the country, thereby expanding high-potential areas, particularly under SSP5-8.5, where high GWP is likely to be 65.12% by 2100. Low precipitation under SSP1-2.6 expands the zone of low potential to 22.35% by 2100, indicating the likelihood of sustainability problems, which noted substantial geographical and temporal fluctuations in groundwater availability resulting from climate-induced alterations. These are also supported by hydrological climate impact studies, which indicate that sustainable management strategies might mitigate groundwater depletion risks (Kumar *et al.* 2024; Rahman *et al.* 2024).

This research has significant methodological and practical significance, thereby contributing valuably to groundwater studies and climate change adaptation techniques in Bangladesh. Practically, it makes a contribution to hydro-environmental studies by demonstrating the increased predictive capability of boosting algorithms. The research identifies the most important determinants of groundwater levels, thereby maximizing interpretability and minimizing redundancy. By linking climate change scenarios to boosting models, it develops high-resolution projections of groundwater availability in Bangladesh until 2100. Practically, it provides an evidence-based decision-support system for resource managers, planners, and policymakers. The outcomes enable prevention of groundwater depletion, water allocation optimization, designing recharge interventions, and management of spatial variations in vulnerability, especially in drought- and flood-prone areas. Importantly, it furnishes the foundation for climate-resilient groundwater policies for long-term sustainability. Despite these contributions, there are some limitations. The study did not leverage Sentinel-2 high-resolution imagery that would have increased land-use classification and GWP mapping resolution, particularly in urban fringes and agricultural areas. Future research must include Sentinel-2 data for more precise recharge zone delineation. Another limitation is the exclusion of socioeconomic variables such as population growth, industrial water demand, and urbanization. These exert significant influences on groundwater abstraction and recharge, and their exclusion can result in an underestimation of depletion risks in the future. Research in

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

the future must integrate hydrological, demographic, and socioeconomic drivers so as to better capture real-time pressures and ensure projections are more policy-relevant. In the research, precipitation was only used as the climate variable in the SSP scenarios. Despite its importance, groundwater also reacts to temperature, humidity, and extreme events such as floods, droughts, and cyclones. Data availability prevented their use, which may have reduced accuracy. Incorporating these datasets will produce more accurate and richer predictions of GWP for different conditions of climate.

5. CONCLUSION

In Bangladesh, groundwater is an essential resource that supports industrial growth, domestic consumption, and agricultural output. However, unplanned urbanization, climate change, and rising water demand are posing a growing threat to it. To address these issues, this study methodically analyzes important variables and evaluates sophisticated boosting-based predictive models, and projects spatiotemporal fluctuations in groundwater levels under future climate scenarios. After evaluating five sophisticated boosting models, LightGBM performed the best (AUC: 0.92, Accuracy: 0.85). The most influential parameters governing GWP in Bangladesh were found to be precipitation (average contribution: 13%) and LULC (10%). The maximum percentage of high GWP zones was reported by Barisal (80.27%) and Khulna (71.83%), whereas Chittagong recorded the greatest percentage of low GWP zones (27.43%) and the smallest percentage of high-potential zones (8.62%). Rajshahi division is likely to face severe groundwater scarcity based on future climatic conditions estimated, owing to the continuous low precipitation. This study offers a comprehensive and organized body of material that contributes to our understanding of hydrogeology and provides the groundwork for future research. It contributes scientifically by improving predictive modeling of groundwater distribution and expanding the methodological toolkit for hydrogeological studies by defining regional GWP and identifying the most significant factors. Besides its contributions to science, the study has extensive policy and practical implications. The GWP maps generated provide spatially explicit data on areas of high and low groundwater availability, enabling policymakers and water managers to target interventions and resource allocation. They can aid the design of focused conservation activities, optimize the utilization of groundwater in agriculture and industry, and guide urban planning to prevent overexploitation in risky zones. Furthermore, the predictive model created here helps the development of climate-resilient water management policies, improves the ability to monitor groundwater in real-time, and fortifies the basis for early-warning systems for drought or water scarcity. This research contributes to various of the Sustainable Development Goals (SDGs), including SDG 6 (Clean Water and Sanitation) specifically through promoting sustainable groundwater management; support SDG 11 (Sustainable Cities and Communities) by informing adaptive and resilient urban planning with proper water allocation; and also promotes SDG 13 (Climate Action) by ensuring climate projections are integrated into GWP evaluations and adaptive management strategies.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Abdullateef, L., Tijani, M. N., Nuru, N. A., John, S. & Mustapha, A. (2021) Assessment of groundwater recharge potential in a typical geological transition zone in Bauchi, NE-Nigeria using remote sensing/GIS and MCDA approaches, *Heliyon*, **7** (4), e06762. https://doi.org/10.1016/j.heliyon.2021.e06762.

Adiguzel, F., Cetin, M., Dogan, M., Gungor, S., Kose, M., Bozdogan Sert, E. & Kaya, E. (2022) The assessment of the thermal behavior of an urban park surface in a dense urban area for planning decisions, *Environmental Monitoring and Assessment*, **194** (7). https://doi.org/10.1007/s10661-022-10172-y.

Ahmad, I., Dar, M. A., Andualem, T. G. & Teka, A. H. (2020) Groundwater development using geographic information system, *Applied Geomatics*, **12** (1), 73–82. https://doi.org/10.1007/s12518-019-00283-6.

Ahmed, N., Hoque, M. A. A., Pradhan, B. & Arabameri, A. (2021) Spatio-temporal assessment of groundwater potential zone in the drought-prone area of Bangladesh using GIS-based bivariate models, *Natural Resources Research*, **30** (5), 3315–3337. https://doi.org/10.1007/s11053-021-09870-0.

Al-Abadi, A. M. & Shahid, S. (2015) A comparison between index of entropy and catastrophe theory methods for mapping groundwater potential in an arid region, *Environmental Monitoring and Assessment*, **187** (9). https://doi.org/10.1007/s10661-015-4801-2.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Al-Abadi, A. M., Pourghasemi, H. R., Shahid, S. & Ghalib, H. B. (2017) Spatial mapping of groundwater potential using entropy weighted linear aggregate novel approach and GIS, *Arabian Journal for Science and Engineering*, **42** (3), 1185–1199. https://doi.org/10.1007/s13369-016-2374-1.

Al-Fugara, A., Pourghasemi, H. R., Al-Shabeeb, A. R., Habib, M., Al-Adamat, R., Al-Amoush, H. & Collins, A. L. (2020) A comparison of machine learning models for the mapping of groundwater spring potential, *Environmental Earth Sciences*, **79** (10), 1–19. https://doi.org/10.1007/s12665-020-08944-1.

Aloui, S., Zghibi, A., Mazzoni, A., Elomri, A. & Al-Ansari, T. (2024) Identifying suitable zones for integrated aquifer recharge and flood control in arid Qatar using GIS-based multi-criteria decision-making, *Groundwater for Sustainable Development*, **25** (March), 101137. https://doi.org/10.1016/j.gsd.2024.101137.

Al-Shabeeb, A. A. 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, *Groundwater for Sustainable Development*, **7**, 82–90. https://doi.org/10.1016/j.gsd.2018.03.011.

Anwar, Z. & Ahmed, U. (2006) *Groundwater Resources Development in Bangladesh: Contribution to Irrigation for Food Security and Constraints to Sustainability*. International Water Management Institute (IWMI), Colombo, Sri Lanka. Available at: https://publications.iwmi.org/pdf/H039306.pdf.

Arabameri, A., Rezaei, K., Cerda, A., Lombardo, L. & Rodrigo-Comino, J. (2019) GIS-based groundwater potential mapping in Shahroud plain, Iran. A comparison among statistical (bivariate and multivariate), data mining and MCDM approaches, *Science of the Total Environment*, **658**, 160–177. https://doi.org/10.1016/j.scitotenv.2018.12.115.

Arif Ali, Z., Abduljabbar, Z. H., Tahir, H., Bibo Sallow, A. & Almufti, S. M. (2023) Extreme gradient boosting algorithm with machine learning: a review, *Academic Journal of Nawroz University*, **12** (2), 320–334. https://doi.org/10.25007/AJNU.V12N2A1612.

Arulbalaji, P., Padmalal, D. & Sreelash, K. (2019) GIS and AHP techniques based delineation of groundwater potential zones: a case study from Southern Western Ghats, India, *Scientific Reports*, **9** (1), 1–17. https://doi.org/10.1038/s41598-019-38567-x.

Ávila-Marín, J., Gil-Márquez, J. M. & Andreo, B. (2025) Evaluating the feasibility of managed aquifer recharge techniques as a drought mitigation strategy for the Seville water supply system (Southern Spain), *Science of the Total Environment*, **983** (November 2024), 179636. https://doi.org/10.1016/j.scitotenv.2025.179636.

Ayazi, M. H., Pirasteh, S., Arvin, A. K. P., Pradhan, B., Nikouravan, B. & Mansor, S. (2010) Disasters and risk reduction in groundwater: Zagros Mountain Southwest Iran using geoinformatics techniques, *Disaster Adv*, **3** (1), 51–57.

Babu, M. A. H., Islam, M. R., Farzana, F., Uddin, M. J. & Islam, M. S. (2020) Application of GIS and remote sensing for identification of groundwater potential zone in the Hilly Terrain of Bangladesh, *Grassroots Journal of Natural Resources*, **3** (3), 16–27. https://doi.org/10.33002/nr2581.6853.03032.

Bangladesh National Portal (2020) *Divisions of Bangladesh*. Bangladesh National Portal. Available at: http://www.bangladesh.gov.bd/site/view/division-list/বিভাগসমূহ.

Banglapedia (2014) *Climatic Zone – Banglapedia*. Asiatic Society of Bangladesh, Dhaka, Bangladesh. Available at: http://en.banglapedia.org/index.php?title=Climatic_Zone.

Bertamini, M. & Wagemans, J. (2013) Processing convexity and concavity along a 2-D contour: figure-ground, structural shape, and attention, *Psychonomic Bulletin and Review*, **20** (2), 191–207. https://doi.org/10.3758/s13423-012-0347-2.

Biswas, S., Mukhopadhyay, B. P. & Bera, A. (2020) Delineating groundwater potential zones of agriculture dominated landscapes using GIS based AHP techniques: a case study from Uttar Dinajpur district, West Bengal, *Environmental Earth Sciences*, **79** (12), 1–25. https://doi.org/10.1007/s12665-020-09053-9.

Cha, Y., Park, S. S., Kim, K., Byeon, M. & Stow, C. A. (2014) Probabilistic prediction of cyanobacteria abundance in a Korean reservoir using a Bayesian Poisson model, *Water Resources Research*, **50** (3), 2518–2532. https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2013WR014372.

Chathuranika, I. M., Gunathilake, M. B., Baddewela, P. K., Sachinthanie, E., Babel, M. S., Shrestha, S., Jha, M. K. & Rathnayake, U. S. (2022) Comparison of two hydrological models, HEC-HMS and SWAT in runoff estimation: application to Huai Bang Sai Tropical Watershed, Thailand, *Fluids*, **7** (8), 267. https://doi.org/10.3390/fluids7080267.

Chatterjee, S. & Dutta, S. (2022) Assessment of groundwater potential zone for sustainable water resource management in south-western part of Birbhum District, West Bengal, *Applied Water Science*, **12** (3), 1–16. https://doi.org/10.1007/s13201-021-01549-4.

Chen, T. & Guestrin, C. (2016) 'XGBoost: a scalable tree boosting system', *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13 – 17 August 2016, pp. 785–794.

Chen, W., Zhao, X., Tsangaratos, P., Shahabi, H., Ilia, I., Xue, W., Wang, X. & Ahmad, B. B. (2020) Evaluating the usage of tree-based ensemble methods in groundwater spring potential mapping, *Journal of Hydrology*, **583** (December 2019), 124602. https://doi.org/10.1016/j.jhydrol.2020.124602.

Chen, Y., Chen, W., Chandra Pal, S., Saha, A., Chowdhuri, I., Adeli, B., Janizadeh, S., Dineva, A. A., Wang, X. & Mosavi, A. (2022) Evaluation efficiency of hybrid deep learning algorithms with neural network decision tree and boosting methods for predicting groundwater potential, *Geocarto International*, **37** (19), 5564–5584. https://doi.org/10.1080/10106049.2021.1920635.

Chowdhury, M. S. (2023) Modelling hydrological factors from DEM using GIS, *MethodsX*, **10** (December 2022), 102062. https://doi.org/10.1016/j.mex.2023.102062.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Condon, L. E., Atchley, A. L. & Maxwell, R. M. (2020) Evapotranspiration depletes groundwater under warming over the contiguous United States, *Nature Communications*, **11** (1). https://doi.org/10.1038/s41467-020-14688-0.

Cook, P. G., Banks, E. W., Marshall, S. K., Harrington, G. A., Batlle-Aguilar, J., Dogramaci, S. & Turnadge, C. (2022) Inferring fault hydrology using groundwater age tracers, *Journal of Hydrology*, **610**, 127905. https://doi.org/10.1016/J.JHYDROL.2022.127905.

Corsini, A., Cervi, F. & Ronchetti, F. (2009) Weight of evidence and artificial neural networks for potential groundwater spring mapping: an application to the Mt. Modino area (Northern Apennines, Italy), *Geomorphology*, **111** (1–2), 79–87. https://doi.org/10.1016/j.geomorph.2008.03.015.

Dey, B., Abir, K. A. M., Ahmed, R., Salam, M. A., Redowan, M., Miah, M. D. & Iqbal, M. A. (2023) Monitoring groundwater potential dynamics of north-eastern Bengal Basin in Bangladesh using AHP-machine learning approaches, *Ecological Indicators*, **154** (August), 110886. https://doi.org/10.1016/j.ecolind.2023.110886.

Díaz-Alcaide, S. & Martínez-Santos, P. (2019) Review: advances in groundwater potential mapping, *Hydrogeology Journal*, **27** (7), 2307–2324. https://doi.org/10.1007/s10040-019-02001-3.

Dinakaran, S. & Ranjit Jeba Thangaiah, P. (2017) Ensemble method of effective AdaBoost algorithm for decision tree classifiers, *International Journal on Artificial Intelligence Tools*, **26** (3), 1750007. https://doi.org/10.1142/S0218213017500075.

Elbeltagi, A., Salam, R., Pal, S. C., Zerouali, B., Shahid, S., Mallick, J., Islam, M. S. & Islam, A. R. M. T. (2022) Groundwater level estimation in northern region of Bangladesh using hybrid locally weighted linear regression and Gaussian process regression modeling, *Theoretical and Applied Climatology*, **149** (1–2), 131–151. https://doi.org/10.1007/s00704-022-04037-0.

Elmahdy, S., Ali, T. & Mohamed, M. (2021) Regional mapping of groundwater potential in Ar Rub Al Khali, Arabian Peninsula using the classification and regression trees model, *Remote Sensing*, **13** (12), 2300. https://doi.org/10.3390/rs13122300.

ESRI (2023) *How LightGBM Algorithm Works – ArcGIS Pro | Documentation*. Esri, Redlands, California, USA. Available at: https://pro.arcgis.com/en/pro-app/latest/tool-reference/geoai/how-lightgbm-works.htm.

Falah, F., Ghorbani Nejad, S., Rahmati, O., Daneshfar, M. & Zeinivand, H. (2017) Applicability of generalized additive model in groundwater potential modelling and comparison its performance by bivariate statistical methods, *Geocarto International*, **32** (10), 1069–1089. https://doi.org/10.1080/10106049.2016.1188166.

Farhat, N. (2018) Effect of relative humidity on evaporation rates in Nabatieh Region, *Lebanese Science Journal*, **19** (1), 59–66. https://doi.org/10.22453/lsj-019.1.059-066.

Fatema, K., Joy, M. A. R., Amin, F. M. R. & Sarkar, S. K. (2023) Groundwater potential mapping in Jashore, Bangladesh, *Heliyon*, **9** (3), e13966. https://doi.org/10.1016/j.heliyon.2023.e13966.

Fattah, M. A., Gupta, S. D., Faroque, M. Z., Ghosh, B., Morshed, S. R., Chakraborty, T., Kafy, A. A. & Rahman, M. T. (2023) Spatiotemporal characterization of relative humidity trends and influence of climatic factors in Bangladesh, *Heliyon*, **9** (9), e19991. https://doi.org/10.1016/j.heliyon.2023.e19991.

Ferozur, R. M., Jahan, C. S., Arefin, R. & Mazumder, Q. H. (2019) Groundwater potentiality study in drought prone barind tract, NW Bangladesh using remote sensing and GIS, *Groundwater for Sustainable Development*, **8**, 205–215. https://doi.org/10.1016/j.gsd.2018.11.006.

Frame, B., Lawrence, J., Ausseil, A. G., Reisinger, A. & Daigneault, A. (2018) Adapting global shared socio-economic pathways for national and local scenarios, *Climate Risk Management*, **21**, 39–51. https://doi.org/10.1016/J.CRM.2018.05.001.

Gallego Ortega Ventzi Bojkov, R. (2023) *Hydrologic Modeling With MIKE SHE*. June. Available at: https://upcommons.upc.edu/server/api/core/bitstreams/3832c2f1-2aa4-404d-a6e7-b28643c70975/content.

Ghazi, B., Jeihouni, E., Kouzehgar, K. & Haghighi, A. T. (2021) Assessment of probable groundwater changes under representative concentration pathway (RCP) scenarios through the wavelet–GEP model, *Environmental Earth Sciences*, **80** (12), 1–15. https://doi.org/10.1007/s12665-021-09746-9.

Golkarian, A. & Rahmati, O. (2018) Use of a maximum entropy model to identify the key factors that influence groundwater availability on the Gonabad Plain, Iran, *Environmental Earth Sciences*, **77** (10), 1–20. https://doi.org/10.1007/s12665-018-7551-y.

Gómez-Escalonilla, V. & Martínez-Santos, P. (2024) A machine learning approach to map the vulnerability of groundwater resources to agricultural contamination, *Hydrology*, **11** (9), 153. https://doi.org/10.3390/hydrology11090153.

Guo, X., Gui, X., Xiong, H., Hu, X., Li, Y., Cui, H., Qiu, Y. & Ma, C. (2023) Critical role of climate factors for groundwater potential mapping in arid regions: insights from random forest, XGBoost, and LightGBM algorithms, *Journal of Hydrology*, **621**, 129599. https://doi.org/10.1016/J.JHYDROL.2023.129599.

Haidong, C., Shangqi, D., Xingke, G., Shuangde, H., Tao, W., Debin, X. & Baoyu, X. (2020) 'Multi-temporal remote sensing fire detection based on GBDT in Yunnan area', *Proceedings – 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence, MLBDBI 2020*, pp. 469–473.

Halder, K., Srivastava, A. K., Ghosh, A., Nabik, R., Pan, S., Chatterjee, U., Bisai, D., Pal, S. C., Zeng, W., Ewert, F., Gaiser, T., Pande, C. B., Islam, A. R. M. T., Alam, E. & Islam, M. K. (2024) Application of bagging and boosting ensemble machine learning techniques for groundwater potential mapping in a drought-prone agriculture region of eastern India, *Environmental Sciences Europe*, **36** (1). https://doi.org/10.1186/s12302-024-00981-y.

Hall, J. (2023) *The Impacts of Urbanization on Aquifer Recharge*. KrakenSense. Available at: https://krakensense.com/blog/the-impacts-of-urbanization-on-aquifer-recharge.

Hancock, J. T. & Khoshgoftaar, T. M. (2020) CatBoost for big data: an interdisciplinary review, *Journal of Big Data*, **7** (1). https://doi.org/10.1186/S40537-020-00369-8.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Haque, S. J., Onodera, S. i. & Shimizu, Y. (2013) An overview of the effects of urbanization on the quantity and quality of groundwater in South Asian megacities, *Limnology*, **14** (2), 135–145. https://doi.org/10.1007/S10201-012-0392-6.

Hasan, M. A., Ahmed, K. M., Sracek, O., Bhattacharya, P., Brömssen, M., Broms, S., Fogelström, J., Mazumder, M. L. & Jacks, G. (2007) Arsenic in shallow groundwater of Bangladesh: investigations from three different physiographic settings, *Hydrogeology Journal*, **15** (8), 1507–1522. https://doi.org/10.1007/s10040-007-0203-z.

Hasan, K., Paul, S., Chy, T. J. & Antipova, A. (2021) Analysis of groundwater table variability and trend using ordinary kriging: the case study of Sylhet, Bangladesh, *Applied Water Science*, **11** (7), 1–12. https://doi.org/10.1007/s13201-021-01454-w.

Hasegawa, T., Fujimori, S., Ito, A., Takahashi, K. & Masui, T. (2017) Global land-use allocation model linked to an integrated assessment model, *Science of the Total Environment*, **580**, 787–796. https://doi.org/10.1016/j.scitotenv.2016.12.025.

Haydar, M., Hosan, S. & Rafi, A. H. (2024) Assessment of urban expansion susceptibility in major urban units of Bangladesh leveraging machine learning and geostatistical approach, *Journal of Urban Management*, **14**, 451–467. https://doi.org/10.1016/j.jum.2024.11.011.

Howlader, R., Chowdhury, M. M. A., Jahan, C. S., Hossain, M. A., Rahaman, M. F., Ghose, B. K. & Islam, M. (2024) Delineation of fresh groundwater potentiality zones in saline coastal aquifers, Southwest Bangladesh using remote sensing and GIS approaches, *Environmental Geochemistry and Health*, **46** (11), 1–19. https://doi.org/10.1007/s10653-024-02237-3.

Huggett, R. J. (2016) *Fundamentals of geomorphology (4th ed.)*. Routledge. https://doi.org/10.4324/9781315674179.

Ibrahem Ahmed Osman, A., Najah Ahmed, A., Chow, M. F., Feng Huang, Y. & El-Shafie, A. (2021) Extreme gradient boosting (XGBoost) model to predict the groundwater levels in Selangor Malaysia, *Ain Shams Engineering Journal*, **12** (2), 1545–1556. https://doi.org/10.1016/j.asej.2020.11.011.

Ibrahim, A. A., Ridwan, R. L., Muhammed, M. M., Abdulaziz, R. O. & Saheed, G. A. (2020) Comparison of the CatBoost classifier with other machine learning methods, *International Journal of Advanced Computer Science and Applications*, **11** (11), 738–748. https://doi.org/10.14569/IJACSA.2020.0111190.

Jhariya, D. C., Mondal, K. C., Kumar, T., Indhulekha, K., Khan, R. & Singh, V. K. (2021) Assessment of groundwater potential zone using GIS-based multi-influencing factor (MIF), multi-criteria decision analysis (MCDA) and electrical resistivity survey techniques in Raipur city, Chhattisgarh, India, *Aqua Water Infrastructure, Ecosystems and Society*, **70** (3), 375–400. https://doi.org/10.2166/aqua.2021.129.

Jou, Y. J., Huang, C. C. L. & Cho, H. J. (2014) A VIF-based optimization model to alleviate collinearity problems in multiple linear regression, *Computational Statistics*, **29** (6), 1515–1541. https://doi.org/10.1007/s00180-014-0504-3.

Kalantar, B., Al-Najjar, H. A. H., Pradhan, B., Saeidi, V., Halin, A. A., Ueda, N. & Naghibi, S. A. (2019) Optimized conditioning factors using machine learning techniques for groundwater potential mapping, *Water (Switzerland)*, **11** (9), 1909. https://doi.org/10.3390/w11091909.

Kamruzzaman, M., Hwang, S., Cho, J., Jang, M. W. & Jeong, H. (2019) Evaluating the spatiotemporal characteristics of agricultural drought in Bangladesh using effective drought index, *Water (Switzerland)*, **11** (12), 2437. https://doi.org/10.3390/W11122437.

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M. & Brumby, S. P. (2021) 'Global land use/land cover with Sentinel 2 and deep learning', *International Geoscience and Remote Sensing Symposium (IGARSS)*, July 2021, pp. 4704–4707.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. & Liu, T. Y. (2017) LightGBM: a highly efficient gradient boosting decision tree, *Advances in Neural Information Processing Systems*, **2017** (February 2024), 3147–3155.

Kégl, B. & Busa-Fekete, R. (2009) Boosting products of base classifiers, *ACM International Conference Proceeding Series*, **382** (May), 497–504. https://doi.org/10.1145/1553374.1553439.

Khan, M. N. U. & Haque, M. N. (2023) Evaluation of aquifer properties to categorize groundwater potential in a part of southwest Bangladesh using catastrophe theory, *European Journal of Environment and Earth Sciences*, **4** (6), 21–31. https://doi.org/10.24018/ejgeo.2023.4.6.435.

Knitter, D., Brozio, J. P., Hamer, W., Duttmann, R., Müller, J. & Nakoinz, O. (2019) Transformations and site locations from a landscape archaeological perspective: The case of Neolithic Wagrien, Schleswig-Holstein, Germany, *Land*, **8** (4), 68. https://doi.org/10.3390/land8040068.

Kumar, P., Herath, S., Avtar, R. & Takeuchi, K. (2016) Mapping of groundwater potential zones in Killinochi area, Sri Lanka, using GIS and remote sensing techniques, *Sustainable Water Resources Management*, **2** (4), 419–430. https://doi.org/10.1007/s40899-016-0072-5.

Kumar, R., Bhattacharjee, R., Gaur, S. & Ohri, A. (2024) *Groundwater Sustainability in the Varuna River Basin: Impacts of Climate Change and Population Growth*, pp. 1–21.

Lee, J., Jung, C., Kim, S. & Kim, S. (2019) Assessment of climate change impact on future groundwater-level behavior using SWAT groundwater-consumption function in Geum River Basin of South Korea, *Water (Switzerland)*, **11** (5), 949. https://doi.org/10.3390/w11050949.

Liang, F., Li, S., Jie, F., Ge, Y., Liu, N. & Jia, G. (2023) The development of a coupled Soil Water Assessment Tool-MODFLOW model for studying the impact of irrigation on a regional water cycle, *Water (Switzerland)*, **15** (20), 3542. https://doi.org/10.3390/w15203542.

Machado, M. R., Karray, S. & De Sousa, I. T. (2019) 'LightGBM: an effective decision tree gradient boosting method to predict customer loyalty in the finance industry', *14th International Conference on Computer Science and Education, ICCSE 2019*, pp. 1111–1116.

Madani, A. & Niyazi, B. (2023) Groundwater potential mapping using remote sensing and random forest machine learning model: a case study from lower part of Wadi Yalamlam, Western Saudi Arabia, *Sustainability (Switzerland)*, **15** (3), 2772. https://doi.org/10.3390/su15032772.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Magesh, N. S., Chandrasekar, N. & Soundranayagam, J. P. (2012) Delineation of groundwater potential zones in Theni district, Tamil Nadu, using remote sensing, GIS and MIF techniques, *Geoscience Frontiers*, **3** (2), 189–196. https://doi.org/10.1016/j.gsf.2011.10.007.

Maity, D. K. & Mandal, S. (2019) Identification of groundwater potential zones of the Kumari River basin, India: an RS & GIS based semi-quantitative approach, *Environment, Development and Sustainability*, **21** (2), 1013–1034. https://doi.org/10.1007/s10668-017-0072-0.

Mallick, J., Singh, C. K., Al-Wadi, H., Ahmed, M., Rahman, A., Shashtri, S. & Mukherjee, S. (2015) Geospatial and geostatistical approach for groundwater potential zone delineation, *Hydrological Processes*, **29** (3), 395–418. https://doi.org/10.1002/hyp.10153.

Mallick, J., Khan, R. A., Ahmed, M., Alqadhi, S. D., Alsubih, M., Falqi, I. & Hasan, M. A. (2019) Modeling groundwater potential zone in a semi-arid region of Aseer using fuzzy-AHP and geoinformation techniques, *Water (Switzerland)*, **11** (12), 2656. https://doi.org/10.3390/w11122656.

Mallick, J., Naikoo, M. W., Talukdar, S., Ahmed, I. A., Rahman, A., Islam, A. R. M. T., Pal, S., Ghose, B. & Shashtri, S. (2022) Developing groundwater potentiality models by coupling ensemble machine learning algorithms and statistical techniques for sustainable groundwater management, *Geocarto International*, **37** (25), 7927–7953. https://doi.org/10.1080/10106049.2021.1987535.

Marjuanto, A. A., Putranto, T. T. & Sugianto, D. N. (2019) Mapping of groundwater vulnerability index in the alluvial plain of Semarang City using the susceptibility index method, *E3S Web of Conferences*, **125** (2019), 01010. https://doi.org/10.1051/e3sconf/201912501010.

Maskooni, E. K., Naghibi, S. A., Hashemi, H. & Berndtsson, R. (2020) Application of advanced machine learning algorithms to assess groundwater potential using remote sensing-derived data, *Remote Sensing*, **12** (17), 2742. https://doi.org/10.3390/RS12172742.

Mazumder, T. & Saroar, M. M. (2025) Lightning-induced vulnerability assessment in Bangladesh using machine learning and GIS-based approach, *Progress in Disaster Science*, **25** (July 2024), 100406. https://doi.org/10.1016/j.pdisas.2025.100406.

Mekonnen, M. M. & Hoekstra, A. Y. (2016) Sustainability: four billion people facing severe water scarcity, *Science Advances*, **2** (2), 2–7. https://doi.org/10.1126/sciadv.1500323.

Melese, T. & Belay, T. (2022) Groundwater potential zone mapping using analytical hierarchy process and GIS in Muga Watershed, Abay Basin, Ethiopia, *Global Challenges*, **6** (1). https://doi.org/10.1002/gch2.202100068.

Minár, J., Evans, I. S. & Jenčo, M. (2020) A comprehensive system of definitions of land surface (topographic) curvatures, with implications for their application in geoscience modelling and prediction, *Earth-Science Reviews*, **211**, 103414. https://doi.org/10.1016/J.EARSCIREV.2020.103414.

Mohtaram, A., Shafizadeh-Moghadam, H. & Ketabchi, H. (2024) Reconstruction of total water storage anomalies from GRACE data using the LightGBM algorithm with hydroclimatic and environmental covariates, *Groundwater for Sustainable Development*, **26**, 101260. https://doi.org/10.1016/J.GSD.2024.101260.

Mokarram, M., Roshan, G. & Negahban, S. (2015) Landform classification using topography position index (case study: salt dome of Korsia-Darab plain, Iran), *Modeling Earth Systems and Environment*, **1** (4), 1–7. https://doi.org/10.1007/s40808-015-0055-9.

Mosavi, A., Sajedi Hosseini, F., Choubin, B., Goodarzi, M., Dineva, A. A. & Rafiei Sardooi, E. (2021) Ensemble boosting and bagging based machine learning models for groundwater potential prediction, *Water Resources Management*, **35** (1), 23–37. https://doi.org/10.1007/s11269-020-02704-3.

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, *Environmental Monitoring and Assessment*, **188** (1), 1–27. https://doi.org/10.1007/s10661-015-5049-6.

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, *Journal of Hydrology*, **589**, 125197. https://doi.org/10.1016/j.jhydrol.2020.125197.

Nampak, H., Pradhan, B. & Manap, M. A. (2014) Application of GIS based data driven evidential belief function model to predict groundwater potential zonation, *Journal of Hydrology*, **513**, 283–300. https://doi.org/10.1016/j.jhydrol.2014.02.053.

Natekin, A. & Knoll, A. (2013) Gradient boosting machines, a tutorial, *Frontiers in Neurorobotics*, **7** (Dec). https://doi.org/10.3389/fnbot.2013.00021.

Newman, C. P., Russell, C. A., Kisfalusi, Z. D. & Paschke, S. S. (2025) Groundwater hydrology, groundwater and surface-water interactions, water quality, and groundwater-flow simulations for the Wet Mountain Valley alluvial aquifer, Custer and Fremont Counties, Colorado, 2017–19, *Scientific Investigations Report*, **2025**. https://doi.org/10.3133/SIR20245105.

Nguyen, H. D., Nguyen, V. H., Du, Q. V. V., Nguyen, C. T., Dang, D. K., Truong, Q. H., Dang, N. B. T., Tran, Q. T., Nguyen, Q. H. & Bui, Q. T. (2024) Application of hybrid model-based machine learning for groundwater potential prediction in the north central of Vietnam, *Earth Science Informatics*, **17** (2), 1569–1589. https://doi.org/10.1007/s12145-023-01209-y.

Nowreen, S., Taylor, R. G., Shamsudduha, M., Salehin, M., Zahid, A. & Ahmed, K. M. (2020) Groundwater recharge processes in an Asian mega-delta: hydrometric evidence from Bangladesh, *Hydrogeology Journal*, **28** (8), 2917–2932. https://doi.org/10.1007/s10040-020-02238-3.

Nugroho, W. H. A., Nurwatik, N. & Widya, L. K. (2024) Groundwater potential mapping using random forest and extreme gradient boosting algorithms, *IOP Conference Series: Earth and Environmental Science*, **1418** (1), 012035. https://doi.org/10.1088/1755-1315/1418/1/012035.

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.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Ok, E. & Emmanuel, M. (2025) *Understanding the Gradient Boosting Algorithm in XGBoost*. Available at: https://www.researchgate.net/publication/390137877_Understanding_the_Gradient_Boosting_Algorithm_in_XGBoost.

O'Neill, B. C., Kriegler, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., Mathur, R. & van Vuuren, D. P. (2014) A new scenario framework for climate change research: the concept of shared socioeconomic pathways, *Climatic Change*, **122** (3), 387–400. https://doi.org/10.1007/S10584-013-0905-2.

Panahi, M., Sadhasivam, N., Pourghasemi, H. R., Rezaie, F. & Lee, S. (2020) Spatial prediction of groundwater potential mapping based on convolutional neural network (CNN) and support vector regression (SVR), *Journal of Hydrology*, **588** (May), 125033. https://doi.org/10.1016/j.jhydrol.2020.125033.

Park, S. & Kim, J. (2021) The predictive capability of a novel ensemble tree-based algorithm for assessing groundwater potential, *Sustainability (Switzerland)*, **13** (5), 1–19. https://doi.org/10.3390/su13052459.

Porun, M. M. S., Tauhidur, M. & Hassan, F. (2025) Mapping groundwater potential zone by robust machine learning algorithms & remote sensing techniques in agriculture dominated area, Bangladesh, *Cleaner Water*, **3** (January), 100064. https://doi.org/10.1016/j.clwat.2025.100064.

Pradhan, B. (2009) Groundwater potential zonation for basaltic watersheds using satellite remote sensing data and GIS techniques, *Central European Journal of Geosciences*, **1** (1), 120–129. https://doi.org/10.2478/v10085-009-0008-5.

Prasad, P., Loveson, V. J., Kotha, M. & Yadav, R. (2020) Application of machine learning techniques in groundwater potential mapping along the west coast of India, *GIScience and Remote Sensing*, **57**, 735–752. https://doi.org/10.1080/15481603.2020.1794104.

Ptr, A. F. L., Siregar, M. M. & Daniel, I. (2024) Analysis of gradient boosting, XGBoost, and CatBoost on mobile phone classification, *Journal of Computer Networks, Architecture and High Performance Computing*, **6** (2), 661–670. https://doi.org/10.47709/CNAHPC.V6I2.3790.

Qadir, J., Bhat, M. S., Alam, A. & Rashid, I. (2020) Mapping groundwater potential zones using remote sensing and GIS approach in Jammu Himalaya, Jammu and Kashmir, *GeoJournal*, **85** (2), 487–504. https://doi.org/10.1007/s10708-019-09981-5.

Quino Lima, I., Ormachea Muñoz, M., Ramos Ramos, O. E., Quintanilla Aguirre, J., Maity, J. P., Ahmad, A. & Bhattacharya, P. (2021) Hydrogeochemical contrasts in the shallow aquifer systems of the Lower Katari Basin and Southern Poopó Basin, Bolivian Altiplano, *Journal of South American Earth Sciences*, **105**, 102914. https://doi.org/10.1016/j.jsames.2020.102914.

Rahman, M. M., Althobiani, F., Shahid, S., Virdis, S. G. P., Kamruzzaman, M., Rahaman, H., Momin, M. A., Hossain, M. B. & Ghandourah, E. I. (2022) GIS and remote sensing-based multi-criteria analysis for delineation of groundwater potential zones: a case study for industrial zones in Bangladesh, *Sustainability (Switzerland)*, **14** (11), 6667. https://doi.org/10.3390/su14116667.

Rahman, M., Islam, M. M., Kim, H. J., Alam, M., Sadiq, S., Rahman, M. K., Sadir Hossan, M., Islam, M. T., Raju, M. R., Alam, M. S., Ahmad, S. I. & Dewan, A. (2024) Optimizing urban water sustainability: integrating deep learning, genetic algorithm, and CMIP6 GCM for groundwater potential zone prediction within a social-ecological-technological framework, *Advances in Space Research*, **73** (12), 5925–5948. https://doi.org/10.1016/J.ASR.2024.03.033.

Rana, M. M. S. P., Hossain, M. A. & Nasher, N. M. R. (2022) Identification of groundwater potential zone using geospatial techniques of agriculture dominated area in Dinajpur district, Bangladesh, *Environmental Challenges*, **7** (June 2021), 100475. https://doi.org/10.1016/j.envc.2022.100475.

Rao, P., Wang, Y., Liu, Y., Wang, X., Hou, Y., Pan, S., Wang, F. & Zhu, D. (2022) A comparison of multiple methods for mapping groundwater levels in the Mu Us Sandy Land, China, *Journal of Hydrology: Regional Studies*, **43**, 101189. https://doi.org/10.1016/j.ejrh.2022.101189.

Rasool, U., Yin, X., Xu, Z., Rasool, M. A., Senapathi, V., Hussain, M., Siddique, J. & Trabucco, J. C. (2022) Mapping of groundwater productivity potential with machine learning algorithms: a case study in the provincial capital of Baluchistan, Pakistan, *Chemosphere*, **303** (May), 135265. https://doi.org/10.1016/j.chemosphere.2022.135265.

Razavi-Termeh, S. V., Sadeghi-Niaraki, A., Abba, S. I., Ali, F. & Choi, S. M. (2024) Enhancing spatial prediction of groundwater-prone areas through optimization of a boosting algorithm with bio-inspired metaheuristic algorithms, *Applied Water Science*, **14** (11), 1–25. https://doi.org/10.1007/s13201-024-02301-4.

Riehl, K., Neunteufel, M. & Hemberg, M. (2023) Hierarchical confusion matrix for classification performance evaluation, *Journal of the Royal Statistical Society Series C: Applied Statistics*, **72** (5), 1394–1412. https://doi.org/10.1093/jrsssc/qlad057.

Roshani, A. & Hamidi, M. (2022) Groundwater level fluctuations in coastal aquifer: using artificial neural networks to predict the impacts of Climatical CMIP6 scenarios, *Water Resources Management*, **36** (11), 3981–4001. https://doi.org/10.1007/s11269-022-03204-2.

Roy, S. K., Hasan, M. M., Mondal, I., Akhter, J., Roy, S. K., Talukder, S., Islam, A. K. M. S., Rahman, A. & Karuppannan, S. (2024) Empowered machine learning algorithm to identify sustainable groundwater potential zone map in Jashore District, Bangladesh, *Groundwater for Sustainable Development*, **25**, 101168. https://doi.org/10.1016/J.GSD.2024.101168.

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.

Sadeak, S., Al Amin, M., Chowdhury, T., Mia, M. B., Alam, M. J., Ahmed, K. M. & Khan, M. R. (2023) Comparison of the groundwater recharge estimations of the highly exploited aquifers in Bangladesh and their sustainability, *Groundwater for Sustainable Development*, **20**, 100896. https://doi.org/10.1016/J.GSD.2022.100896.

Saha, R., Dey, N. C., Rahman, M., Bhattacharya, P. & Rabbani, G. H. (2019) Geogenic arsenic and microbial contamination in drinking water sources: exposure risks to the coastal population in Bangladesh, *Frontiers in Environmental Science*, **7** (May), 1–12. https://doi.org/10.3389/fenvs.2019.00057.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Saini, A. (2024) *Guide on AdaBoost Algorithm*. Analytics Vidhya. Available at: https://www.analyticsvidhya.com/blog/2021/09/adaboost-algorithm-a-complete-guide-for-beginners.

Salim, M. Z., Choudhari, N., Kafy, A. A., Nath, H., Alsulamy, S., Rahaman, Z. A., Aldosary, A. S., Rahmand, M. T. & Al-Ramadan, B. (2024) A comprehensive review of navigating urbanization induced climate change complexities for sustainable groundwater resources management in the Indian subcontinent, *Groundwater for Sustainable Development*, **25**, 101115. https://doi.org/10.1016/J.GSD.2024.101115.

Salvadore, E., Bronders, J. & Batelaan, O. (2015) Hydrological modelling of urbanized catchments: a review and future directions, *Journal of Hydrology*, **529** (P1), 62–81. https://doi.org/10.1016/j.jhydrol.2015.06.028.

Sánchez-Gómez, A., Schürz, C., Molina-Navarro, E. & Bieger, K. (2024) Groundwater modelling in SWAT + : considerations for a realistic baseflow simulation, *Groundwater for Sustainable Development*, **26** (May), 101275. https://doi.org/10.1016/j.gsd.2024.101275.

Sarkar, S. K., Esraz-Ul-Zannat, M., Das, P. C. & Mohiuddin Ekram, K. M. (2022a) Delineating the groundwater potential zones in Bangladesh, *Water Supply*, **22** (4), 4500–4516. https://doi.org/10.2166/ws.2022.113.

Sarkar, S. K., Talukdar, S., Rahman, A., Shahfahad & Roy, S. K. (2022b) Groundwater potentiality mapping using ensemble machine learning algorithms for sustainable groundwater management, *Frontiers in Engineering and Built Environment*, **2** (1), 43–54. https://doi.org/10.1108/febe-09-2021-0044.

Sarkar, S. K., Alshehri, F., Shahfahad, Rahman, A., Pradhan, B. & Shahab, M. (2024a) Mapping groundwater potentiality by using hybrid machine learning models under the scenario of climate variability: a national level study of Bangladesh, *Environment, Development and Sustainability*, **27** (8), 0123456789. https://doi.org/10.1007/s10668-024-04687-2.

Sarkar, S. K., Rudra, R. R., Talukdar, S., Das, P. C., Nur, M. S., Alam, E., Islam, M. K. & Islam, A. R. M. T. (2024b) Future groundwater potential mapping using machine learning algorithms and climate change scenarios in Bangladesh, *Scientific Reports*, **14** (1), 1–17. https://doi.org/10.1038/s41598-024-60560-2.

Sarker, S. (2022) Fundamentals of climatology for engineers: lecture note, *Eng*, **3** (4), 573–595. https://doi.org/10.3390/eng3040040.

Sarker, S. & Leta, O. T. (2025) Review of watershed hydrology and mathematical models, *Eng*, **6** (6), 1–48. https://doi.org/10.3390/eng6060129.

Schäuble, H., Marinoni, O. & Hinderer, M. (2008) A GIS-based method to calculate flow accumulation by considering dams and their specific operation time, *Computers and Geosciences*, **34** (6), 635–646. https://doi.org/10.1016/j.cageo.2007.05.023.

Shahid, S. & Hazarika, M. K. (2010) Groundwater drought in the northwestern districts of Bangladesh, *Water Resources Management*, **24** (10), 1989–2006. https://doi.org/10.1007/s11269-009-9534-y.

Shrestha, N. (2020) Detecting multicollinearity in regression analysis, *American Journal of Applied Mathematics and Statistics*, **8** (2), 39–42. https://doi.org/10.12691/AJAMS-8-2-1.

Siddik, M. S., Tulip, S. S., Rahman, A., Islam, M. N., Haghighi, A. T. & Mustafa, S. M. T. (2022) The impact of land use and land cover change on groundwater recharge in northwestern Bangladesh, *Journal of Environmental Management*, **315** (August), 115130. https://doi.org/10.1016/j.jenvman.2022.115130.

Sidle, R. C. & Onda, Y. (2004) Hydrogeomorphology: overview of an emerging science, *Hydrological Processes*, **18** (4), 597–602. https://doi.org/10.1002/hyp.1360.

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** (2), e24308. https://doi.org/10.1016/j.heliyon.2024.e24308.

Singhal, A., Jaseem, M., Divya, Sarker, S., Prajapati, P., Singh, A. & Jha, S. K. (2024) Identifying potential locations of hydrologic monitoring stations based on topographical and hydrological information, *Water Resources Management*, **38** (1), 369–384. https://doi.org/10.1007/s11269-023-03675-x.

Sivanandam, C., Perumal, V. M. & Mohan, J. (2024) A novel light GBM-optimized long short-term memory for enhancing quality and security in web service recommendation system, *J Supercomput*, **80**, 2428–2460. https://doi.org/10.1007/s11227-023-05552-1.

Sobaga, A., Habets, F., Beaudoin, N., Léonard, J. & Decharme, B. (2024) Decreasing trend of groundwater recharge with limited impact of intense precipitation: evidence from long-term lysimeter data, *Journal of Hydrology*, **637** (April), 131340. https://doi.org/10.1016/j.jhydrol.2024.131340.

Sonego, P., Kocsor, A. & Pongor, S. (2008) ROC analysis: applications to the classification of biological sequences and 3D structures, *Briefings in Bioinformatics*, **9** (3), 198–209. https://doi.org/10.1093/BIB/BBM064.

Sresto, M. A., Siddika, S., Haque, M. N. & Saroar, M. (2021) Application of fuzzy analytic hierarchy process and geospatial technology to identify groundwater potential zones in north-west region of Bangladesh, *Environmental Challenges*, **5** (July), 100214. https://doi.org/10.1016/j.envc.2021.100214.

Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M., Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M. F. P., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., Gurdak, J. J., Allen, D. M., Shamsudduha, M., Hiscock, K., Yeh, P. J.-F., Holman, I. & Treidel, H. (2013) Ground water and climate change, *Nature Climate Change*, **3** (4), 322–329. https://doi.org/10.1038/nclimate1744.

Tekin, O., Cetin, M., Varol, T., Ozel, H. B., Sevik, H. & Zeren Cetin, I. (2022) Altitudinal migration of species of fir (*Abies* spp.) in adaptation to climate change, *Water, Air, and Soil Pollution*, **233** (9). https://doi.org/10.1007/s11270-022-05851-y.

Tests, D. (1978) Basic principles of ROC analysis, *Seminars in Nuclear Medicine*, **VIII** (4), 283–298.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

Uddin, M. J., Hu, J., Islam, A. R. M. T., Eibek, K. U. & Nasrin, Z. M. (2020) A comprehensive statistical assessment of drought indices to monitor drought status in Bangladesh, *Arabian Journal of Geosciences*, **13** (9). https://doi.org/10.1007/s12517-020-05302-0.

Uliasz-Misiak, B., Winid, B., Lewandowska-Śmierzchalska, J. & Matuła, R. (2022) Impact of road transport on groundwater quality, *Science of The Total Environment*, **824**, 153804. https://doi.org/10.1016/J.SCITOTENV.2022.153804.

Varol, T., Canturk, U., Cetin, M., Ozel, H. B., Sevik, H. & Zeren Cetin, I. (2022) Identifying the suitable habitats for Anatolian boxwood (*Buxus sempervirens* L.) for the future regarding the climate change, *Theoretical and Applied Climatology*, **150** (1–2), 637–647. https://doi.org/10.1007/s00704-022-04179-1.

Vogeti, R. K., Srinivasa Raju, K., Nagesh Kumar, D., Rajesh, A. M., Somanath Kumar, S. V. & Jha, Y. S. K. (2023) Application of hydrological models in climate change framework for a river basin in India, *Journal of Water and Climate Change*, **14** (9), 3150–3165. https://doi.org/10.2166/wcc.2023.188.

Vörösmarty, G. & Dobos, I. (2020) Green purchasing frameworks considering firm size: a multicollinearity analysis using variance inflation factor, *Supply Chain Forum*, **21** (4), 290–301. https://doi.org/10.1080/16258312.2020.1776090.

Wagavkar, S. (2023) *Correlation Matrix*. IEEE Transactions on Computers. Chicago, IL, United States. Available at: https://builtin.com/data-science/correlation-matrix#.

Waikar, M. L. & Nilawar, A. P. (2014) Identification of groundwater potential zone using remote sensing and GIS technique, *International Journal of Innovative Research in Science, Engineering and Technology*, **3** (5), 12163–12174.

Wang, Z., Wang, J., Yu, D. & Chen, K. (2023) Groundwater potential assessment using GIS-based ensemble learning models in Guanzhong Basin, China, *Environmental Monitoring and Assessment*, **195** (6). https://doi.org/10.1007/s10661-023-11388-2.

Wang-Erlandsson, L., Tobian, A., van der Ent, R. J., Fetzer, I., te Wierik, S., Porkka, M., Staal, A., Jaramillo, F., Dahlmann, H., Singh, C., Greve, P., Gerten, D., Keys, P. W., Gleeson, T., Cornell, S. E., Steffen, W., Bai, X. & Rockström, J. (2022) A planetary boundary for green water, *Nature Reviews Earth and Environment*, **3** (6), 380–392. https://doi.org/10.1038/S43017-022-00287-8.

WBG (2022) *Water resources management overview: Development news, research, data*. Washington, D.C.: World Bank. Retrieved from: https://www.worldbank.org/en/topic/waterresourcesmanagement.

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

Wei, L., Yang, M., Li, Z., Shao, J., Li, L., Chen, P., Li, S. & Zhao, R. (2022b) Experimental investigation of relationship between infiltration rate and soil moisture under rainfall conditions, *Water (Switzerland)*, **14** (9), 1347. https://doi.org/10.3390/W14091347.

Wischmeier, W. H. & Smith, D. D. (1978) Predicting rainfall erosion losses: A guide to conservation planning. In: *USDA Agricultural Handbook No. 537*. Washington, D.C.: U.S. Department of Agriculture, pp. 285–291.

Wu, H., Ye, X., Du, X., Wang, W., Li, H. & Dong, W. (2025) Assessing groundwater level variability in response to climate change: a case study of large plain areas, *Journal of Hydrology: Regional Studies*, **57** (December 2024), 102180. https://doi.org/10.1016/j.ejrh.2025.102180.

Xie, X., Wang, Y., Duan, M. & Li, J. (2013) Provenance and paleoenvironment impact on arsenic accumulation in aquifer sediments from the Datong Basin, China: implications from element geochemistry, *Procedia Earth and Planetary Science*, **7** (October 2015), 904–907. https://doi.org/10.1016/j.proeps.2013.03.058.

Xiong, H., Guo, X., Wang, Y., Xiong, R., Gui, X., Hu, X., Li, Y., Qiu, Y., Tan, J. & Ma, C. (2023) Spatial prediction of groundwater potential by various novel boosting-based ensemble learning models in mountainous areas, *Geocarto International*, **38** (1). https://doi.org/10.1080/10106049.2023.2274870.

Xiong, H., Yang, S., Tan, J., Wang, Y., Guo, X. & Ma, C. (2024) Effects of DEM resolution and application of solely DEM-derived indicators on groundwater potential mapping in the mountainous area, *Journal of Hydrology*, **636**, 131349. https://doi.org/10.1016/J.JHYDROL.2024.131349.

Yen, H. P. H., Pham, B. T., Phong, T. V., Ha, D. H., Costache, R., Le, H. V., Nguyen, H. D., Amiri, M., Tao, N. V. & Prakash, I. (2021) Locally weighted learning based hybrid intelligence models for groundwater potential mapping and modeling: a case study at Gia Lai province, Vietnam, *Geoscience Frontiers*, **12** (5), 1–14. https://doi.org/10.1016/j.gsf.2021.101154.

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, *Environmental Earth Sciences*, **75** (8), 1–19. https://doi.org/10.1007/s12665-016-5424-9.

Zanotti, C., Rotiroti, M., Sterlacchini, S., Cappellini, G., Fumagalli, L., Stefania, G. A., Nannucci, M. S., Leoni, B. & Bonomi, T. (2019) Choosing between linear and nonlinear models and avoiding overfitting for short and long term groundwater level forecasting in a linear system, *Journal of Hydrology*, **578** (September), 124015. https://doi.org/10.1016/j.jhydrol.2019.124015.

Zegaar, A., Ounoki, S. & Telli, A. (2024) Machine learning for groundwater quality classification: a step towards economic and sustainable groundwater quality assessment process, *Water Resources Management*, **38** (2), 621–637. https://doi.org/10.1007/s11269-023-03690-y.

Zeydalinejad, N., Javadi, A. A. & Webber, J. L. (2024) Global perspectives on groundwater infiltration to sewer networks: a threat to urban sustainability, *Water Research*, **262**, 122098. https://doi.org/10.1016/J.WATRES.2024.122098.

Zhang, C., Zhang, Y., Shi, X., Almpanidis, G., Fan, G. & Shen, X. (2019) On incremental learning for gradient boosting decision trees, *Neural Processing Letters*, **50** (1), 957–987. https://doi.org/10.1007/S11063-019-09999-3.

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest

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** (1), 48–64. https://doi.org/10.1016/J.ASR.2024.03.038.

Zzaman, R. U., Nowreen, S., Khan, I. M., Islam, M. R., Ibtehaz, N., Rahman, M. S., Zahid, A., Farzana, D., Sharmin, A. & Rahman, M. S. (2022) A machine learning-based approach for groundwater mapping, *Natural Resources Research*, **31** (1), 281–299. https://doi.org/10.1007/s11053-021-09977-4.

First received 14 July 2025; accepted in revised form 23 September 2025. Available online 16 October 2025

Downloaded from http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2025.216/1599047/jwc2025216.pdf
by guest