Groundwater for Sustainable Development 33 (2026) 101612

Contents lists available at ScienceDirect

Groundwater for Sustainable Development

journal homepage: www.elsevier.com/locate/gsd

Research paper

A novel spatiotemporal and machine learning framework for sustainable groundwater monitoring and management

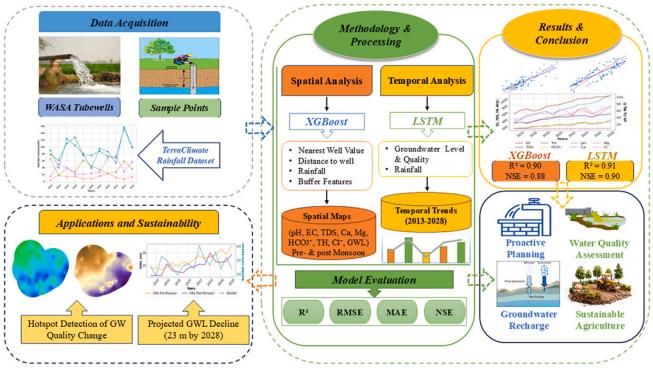
Sheraz Maqbool [a], Muhammad Imran Khan [a,*], Aamir Raza [b], Mumtaz Ali [c,e,**], Naeem Saddique [a], Qaisar Saddique [d]

[a] Department of Irrigation and Drainage, University of Agriculture, Faisalabad, 38000, Punjab, Pakistan
[b] Precision Agriculture Center, Department of Soil, Water and Climate, University of Minnesota, St. Paul, MN, 55108, USA
[c] UniSQ College, University of Southern Queensland, Toowoomba, QLD, 4350, Australia
[d] Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, Telegrafenberg, 14473, Potsdam, Germany
[e] Scientific Research Centre, Al-Ayen Iraqi University, Thi-Qar, 64001, Iraq

HIGHLIGHTS

- Integrated GIS-ML framework for groundwater quality and level assessment.
- XGBoost and LSTM models excel in spatial estimation and temporal forecasting.
- 10-year analysis tracks declining water levels and changing water quality.
- Supports scalable, low-cost decision support for semi-arid groundwater management.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Agriculture
Geographical information system (GIS)
Groundwater
Interpolation
Machine learning (ML)
Deep learning (DL)

ABSTRACT

Groundwater resources worldwide are increasingly under pressure due to over-extraction and inadequate management, highlighting the urgent need for real-time monitoring and accurate forecasting to ensure sustainable usage. Although numerous studies have addressed groundwater issues, there remains a lack of clarity regarding the spatial and temporal dynamics of groundwater levels and their relationship with key water quality indicators. This study aims to fill that gap by investigating the spatial and temporal variations in Groundwater Level (GWL) along with key water quality indicators such as pH, Electrical Conductivity (EC), Total Dissolved Solids (TDS), Calcium (Ca), Magnesium (Mg), Total Hardness (TH), Bicarbonates (HCO_3^-), and Chlorides (Cl^-). Observed data were obtained from Water and Sanitation Agency (WASA) Faisalabad observation wells, covering the period from 2013 to 2023, along with 102 uniformly distributed sample points and corresponding weather data from the TerraClimate dataset. To assess spatiotemporal variation and forecast, several machine learning (ML) and deep learning (DL) models were evaluated for their accuracy and reliability for groundwater trend

* Corresponding author.
** Corresponding author. UniSQ College, University of Southern Queensland, Toowoomba, QLD, 4350, Australia.
E-mail addresses: Imran.khan@uaf.edu.pk (M.I. Khan), mumtaz.ali@unisq.edu.au (M. Ali).

https://doi.org/10.1016/j.gsd.2026.101612
Received 6 June 2025; Received in revised form 6 March 2026; Accepted 7 March 2026
Available online 9 March 2026

projections. The eXtreme Gradient Boosting (XGBoost) model exhibited superior performance in spatial analysis, achieving the highest coefficient of determination (R^2) of 0.90 and a Nash–Sutcliffe Efficiency (NSE) of 0.88, while the Long Short-Term Memory (LSTM) model excelled in temporal forecasting with an $R^2 = 0.91$, NSE = 0.90, both accompanied by the lowest Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) values among the evaluated models. The study revealed significant seasonal variations, particularly post-monsoon increases in chemical concentrations, likely from anthropogenic and agricultural sources. These findings highlight the robustness and precision of ML & DL techniques, especially XGBoost and LSTM, in capturing complex groundwater dynamics.

1. Introduction

Groundwater is an essential resource; however anthropogenic activities and weather extreme events are deteriorating and threatening its availability (Scanlon et al., 2023). Currently, 5 billion individuals from the global population of 8 billion are experiencing water stress (Abdalla and Moubark, 2018; Arnell, 1999). In several regions globally, freshwater resources are inadequate to meet domestic, economic, and ecological requirements (Cosgrove and Loucks, 2015). Groundwater ecosystems provide services that are extremely valuable to society, the economy, and food production, including water purification and high-quality storage from decades to millennia (Gleeson et al., 2016; Griebler and Avramov, 2015; Kılıç, 2020; Saccò et al., 2024). Overuse of groundwater is a global issue with serious implications, and there are currently few effective solutions (Molle et al., 2018).

In many areas, groundwater is frequently improperly handled, which has also contributed to major agricultural development in the last 50 years, while being essential for 40% of worldwide agriculture and providing drinking water to 2 billion people (Giordano, 2009; Gleeson et al., 2020; Hao et al., 2019). Unsustainable groundwater pumping already outpaces river and precipitation recharge, which causes significant declines in groundwater levels and groundwater storage losses, particularly in heavily irrigated areas (Dalin et al., 2017; de Graaf et al., 2019). Groundwater quality investigation has become a prominent topic globally since the industrial revolution and groundwater quality concerns are becoming just as significant as groundwater quantity issues (Burri et al., 2019; Li, 2016). Climate change also contributes to the water shortage, making it more severe in areas where temperature and precipitation pattern disturbed (Singha and Navarre-Sitchler, 2022; Vaux, 2011). Climate change, along with intensified groundwater extraction driven by widespread pumping wells and groundwater development for industrial, municipal, and agricultural purposes, has contributed to rising groundwater withdrawals over the last five decades (Halder et al., 2020; Konikow and Kendy, 2005; Lashkaripour and Ghafoori, 2011; Rodell et al., 2018). Despite the necessity for management to address significant overexploitation and groundwater quality concerns, the extensive utilization of groundwater has facilitated a substantial rise in agricultural output and met the water demands of metropolitan regions (Babiker et al., 2007; Van Steenbergen and Oliemans, 2002). Chemicals, pesticides, and fertilizers used in agriculture pose a risk of contaminating groundwater (Motlagh et al., 2020).

Remote sensing and Geographic Information Systems (GIS) have become revolutionary technologies that are changing how we monitor, evaluate, and manage groundwater resources for long-term sustainability (Kumar et al., 2015; Lin et al., 2019; Shaikh and Birajdar, 2024; Shakoor et al., 2015; Zeilhofer et al., 2007). GIS-based groundwater mapping, supported by groundwater quality and level datasets and spatial analysis techniques such as Inverse Distance Weighted (IDW) interpolation, has been widely applied to characterize spatial variability in aquifer systems (Khouni et al., 2021; Ohlert et al., 2023; Zarco-Perello and Simões, 2017). Hydrochemical indicators are increasingly incorporated into groundwater quality assessments to support irrigation and drinking water suitability, reflecting advances in spatial analysis frameworks (Shaw and Sharma, 2024; Yan et al., 2024). Forecasting trends are essential for the growing population, and Artificial Intelligence (AI) provides an adaptable framework for identifying complex patterns in water resources data (Ahmed et al., 2019). Evaluating and predicting groundwater quality is essential for the proper management of water bodies, enabling timely interventions to keep contamination levels within acceptable limits (Huang et al., 2024; Jain et al., 2021; Li et al., 2014). As groundwater filtration and storage are very costly, pollution prevention, monitoring, and artificial models-based forecasting are the most dependable way to safeguard groundwater supplies (El Bilali et al., 2021; Hadžić et al., 2015; Masood et al., 2022). The application of AI increases the precision of prediction models even further, providing useful instruments for predicting variations in groundwater levels, as well as water quality (Shaikh and Birajdar, 2023). For arid and semi-arid regions to retain freshwater supplies, which is crucial for sustainable development, groundwater quality identification and management are crucial (Asadi et al., 2019). Recent studies have applied tree-based ensembles, boosting, and neural networks for effective water resources quality and quantity forecasting (Mosavi et al., 2021; Singha et al., 2021; Tao et al., 2022).

Globally, Pakistan is an important agricultural country in Asia and ranks as the 3rd largest user of groundwater, with just 27% of its land requiring irrigation from surface water supplies; whereas the remaining 73% is irrigated either directly or indirectly by groundwater, with 90% of abstraction occurring in the Punjab region (Qureshi, 2018, 2020; Qureshi et al., 2010). Agricultural development in Pakistan heavily depends on irrigation, necessitating substantial water from already stressed aquifers due to the low rainfall and fluctuating surface water sources (Awais et al., 2017; Kirby et al., 2017; Qureshi, 2015). Beyond agriculture, groundwater is an important resource for cities and residential areas, used for drinking, cleaning, and industrial applications (Abbas et al., 2023; Khaliq et al., 2021; Quddoos et al., 2024). Due to the inadequate supply of water in Pakistan's major rivers such as the Chenab, Ravi, and Sutlej, the majority of farmers use groundwater for irrigation, placing substantial pressure on groundwater resources (Ahmad et al., 2022; Ahmad et al., 2016; ul Hasan and Fatima, 2025). Essentially, a steady and suitable supply of irrigation water from the Chenab is necessary for economic security, which is declining due to urbanization and pollution (Ali et al., 2025; Kausar et al., 2019). Safe drinking water remains limited in access in many areas, and contaminated water harms infrastructure and creates health hazards (Sultana et al., 2018).

Although GIS mapping, IDW interpolation, and data-driven models have been used in groundwater studies, existing applications typically treat these components separately (e.g., mapping only, or time-series forecasting only), evaluate a small subset of variables, and rarely examine whether model suitability changes across task type (spatial estimation vs. temporal forecasting) and hydrological season (pre-vs. post-monsoon). In this study, we integrated, a season-aware, and task-specific workflow that (i) jointly analyzes groundwater level and multiple hydrochemical indicators, (ii) predicts values at unmonitored locations using hydrologically meaningful covariates (nearest-well value, distance-to-well, rainfall, and zoning/buffer features) rather than proximity-based interpolation alone, (iii) validates mapped outputs using independent field sample points, and (iv) benchmarks ML and DL models under the same splits and metrics to identify which approaches are most reliable for spatial assessment versus temporal forecasting. This study provides clearer decision support for sustainable monitoring and

management in data-limited settings than studies that apply a single method in isolation.

2. Materials and methods

2.1. Study area

Chiniot, Pakistan's agricultural hub, benefits from the Chenab River, boosting food security and economic stability. Groundwater irrigation protects indigenous habitats and provides drinking water to the 1.36 million population, highlighting the importance of water needs for this area (Ahmad et al., 2020). Japan International Cooperation Agency (JICA) and Inline Field Area (IFA) project wells data were evaluated from 2013 to 2023 to assess groundwater resources across an area of 743.5 square kilometers, utilizing 25 JICA and 29 IFA wells stations. The JICA wells were strategically distributed across the central-southern zones of the district, located between longitudes 73.070141°E to 72.948615°E and latitudes 31.623242°N to 31.550918°N. In contrast, the IFA wells were in the northeastern part of the district, ranging from 72.954420°E to 72.976381°E longitude and 31.603797°N to 31.605238°N latitude. A total of 102 groundwater sample points were examined during the pre- and post-monsoon seasons of 2021–2023 (two years of data for training and one year of data for validation) to assess spatial and seasonal fluctuations in groundwater (Fig. 1). The sampling points were selected to ensure uniform spatial coverage across the study area and across the predefined grid zones (Wadoux et al., 2021), including both near-river and far-from-river locations. Some residual sampling bias may remain due to field access constraints and land-use heterogeneity; this limitation is mitigated through spatially structured validation. From May to September, arid summers with temperatures over 45 °C transition into monsoon rains that boost agriculture, followed by mild autumn, and cold, dry winters, making the region agriculturally significant and economically contributive.

2.2. Groundwater data acquisition

Groundwater quality and GWL data were collected from WASA, Faisalabad, Pakistan covering 2013-2023. This included groundwater monitoring records from 29 Inline Field Area (IFA) wells and 25 Japan International Cooperation Agency (JICA) well stations. Additional sampling points across Chiniot were used to enhance spatial coverage for groundwater level and quality assessment, as well as for forecasting purposes. To evaluate the impact of climatic factors on groundwater dynamics, historical monsoon-season rainfall data were sourced from the TerraClimate dataset (Abatzoglou et al., 2018), which provides high-resolution gridded climate data suitable for hydrological modeling. The study integrated both hydrochemical and hydrological parameters, which are listed (Table 1), along with their respective data sources.

2.3. Data preprocessing and quality assurance

All groundwater parameters were compiled and standardized to consistent units. Duplicate records for the same well and sampling date were removed; where replicate measurements were available, the mean value was retained. Seasonal labels were assigned to each observation based on the sampling month to create consistent pre-monsoon and post-monsoon groups. Basic screening was performed to remove physically impossible values. Missing values were quantified by parameter, year, and season before modeling (Abi et al., 2024). Records with missing dependent variables were excluded from the corresponding model. Missing predictor values were imputed using a training-only procedure to avoid leakage: the imputation method was fitted on the training set

Table 1
Sources and parameters of groundwater and climatic data used in the study.

Data Type	Parameters	Sources
Groundwater data	pH Electrical Conductivity (EC) Total Dissolved Solids (TDS) Calcium (Ca) Magnesium (Mg) Total Hardness (TH) Bicarbonates (HCO_3^-) Chlorides (Cl^-) Groundwater Level (GWL) Rainfall (pre-post monsoon)	Water and Sanitation Agency (WASA), Faisalabad
Climatic Data		Terra Climate Dataset https://www.climatologylab.org/terraclimate.html

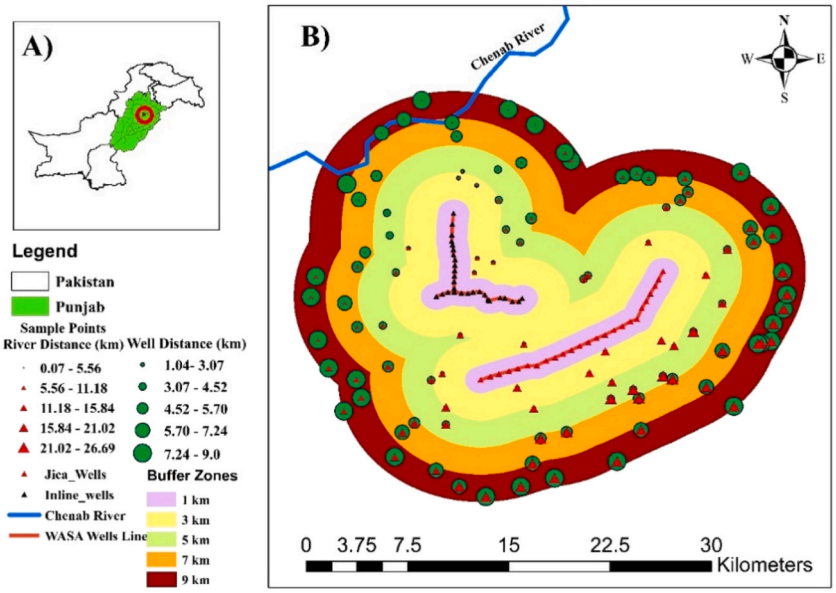


Fig. 1. Geographic distribution of the study area (a) Pakistan boundary and Punjab Province administration boundary (b) distribution of IFA, JICA Wells and sample points in buffer zones.

and then applied to validation data (Kratzert et al., 2018). For temporal modeling, short gaps in well time series were filled using linear interpolation; whereas longer gaps were not included when forming training sequences, consistent with groundwater time-series preprocessing practice (Evans et al., 2020).

2.4. Methodology

Groundwater observation data were collected from wells and constructed buffer zones around each well location. Time series data were available for these wells, but our interest was to estimate groundwater parameters at unmonitored spatial locations using existing well data. For this purpose, sample data were collected within each buffer zone and developed predictive models. These models used inputs including: (1) the nearest well's observed value, (2) the distance from the target point to that nearest well, and (3) rainfall data, as independent variables (IVs). To prevent information leakage during validation, the "nearest well value" feature for any target point was computed using training wells only, excluding any observation belonging to the validation split. The same rule was applied when generating grid predictions used for mapping. The dependent variable (DV) was the actual observed value at the target point. Several ML models were evaluated across all water quality and level parameters. Once the models were trained, we generated a 3 × 3 km spatial grid over the study area. For each grid point, its distance was calculated from the nearest observation, and the model was used to predict its parameter value. These predicted values were then spatially interpolated using the Inverse Distance Weighted (IDW) method. For model validation, sample points collected were used in the final year and compared observed values against predicted ones. Similarly, for forecasting, various models were trained and tested on the time series data from wells to project future groundwater quality and level trends. Fig. 2 presents the workflow chart and study approach.

2.4.1. ML models applied

With the expansion of hydro-environmental datasets, ML/DL approaches are increasingly used for spatial estimation and time-series forecasting of water variables (Zhu et al., 2022). Different models were selected using two-step rationale. Firstly, representative algorithm families were benchmarked from different groundwater spatial and temporal studies. Tree-based ensembles and gradient-boosting methods e.g., RF/ExtraTrees/XGBoost were consistently strong for hydro-environmental predictors used in spatial estimation and potential mapping, because they capture non-linear interactions, handle multi-collinearity, and are robust with limited-to-moderate sample sizes (Pourmorad et al., 2024). Based on this evidence, different models were included (i) representative linear baselines (Ridge/Lasso) to provide interpretable reference performance and control collinearity (Arashi et al., 2021; Lee et al., 2022), (ii) widely used non-linear ML baselines (KNN/SVM) as common benchmarks in groundwater prediction for fast processing of geographical datasets (Afrifa et al., 2022; Arabgol et al., 2016; Suthaharan and Suthaharan, 2016; Zendehboudi et al., 2018), (iii) ensemble learners (RF, Extra Trees, AdaBoost, and XGBoost) effectively model complex predictor interactions in spatial/tabular datasets: RF/Extra Trees improve generalization via feature randomization, while boosting approaches reduce overfitting through sequential learning and regularization (Acosta et al., 2020; Chen et al., 2015; Kasahun et al., 2025; Sun et al., 2024), and (iv) DL time-series models (ANN/CNN/LSTM) for swiftly evaluating and anticipating hydro parameters temporally (Chebud et al., 2012; Gupta et al., 2019; Kouadri et al., 2022; Palani et al., 2008; Patra and Chu, 2023; Ubah et al., 2021). Sequence models especially LSTM-family networks were well-suited for groundwater level and quality forecasting because they learn lagged and long-memory dependencies that are common in hydroclimatic time series (Gilbert et al., 2025). Given their capacity to model sequential and temporal data, Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) were selected since their proven efficacy in forecasting time series variations in water parameters (Baek et al., 2020; Hu et al., 2019; Liu et al., 2019). Second, to avoid subjective selection, all selected models were systematically benchmarked on the same train and test splits using R^2, RMSE, NSE and MAE, and the best-performing

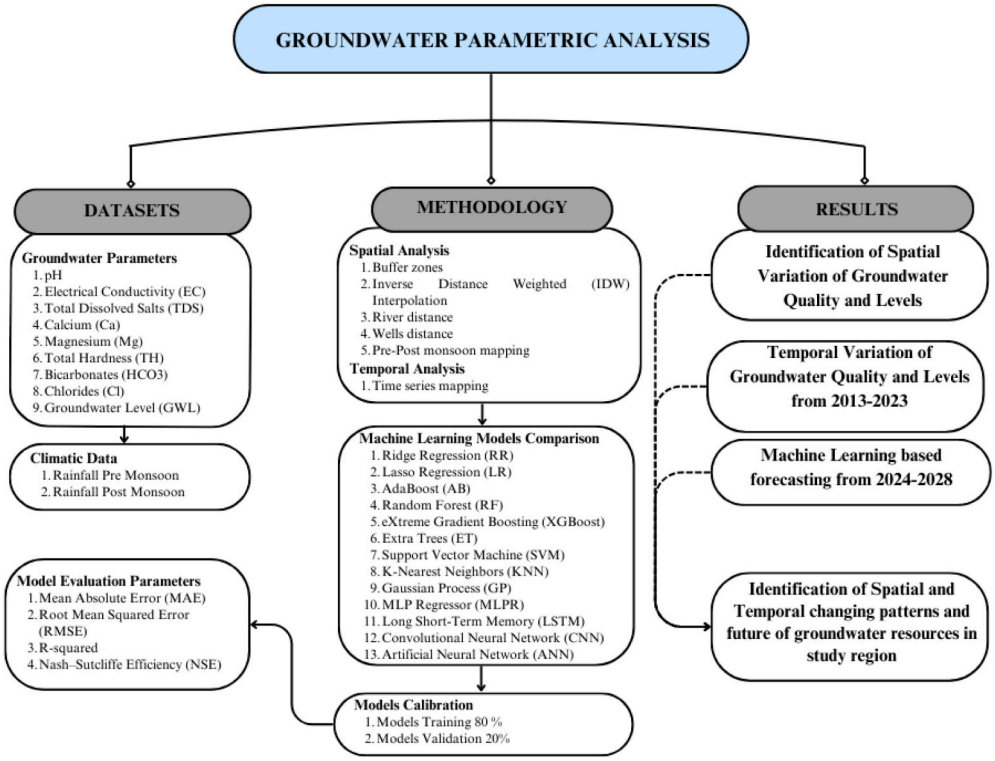


Fig. 2. Methodology workflow for spatiotemporal forecasting of groundwater quality and levels using ML & DL.

model for each parameter and season and for forecasting was included. Table 2 summarizes the suitability model for each data type and rationale for selection.

Based on groundwater hydrochemical and level data recorded from wells sampled point wells (denoted as P) and WASA-operated wells (denoted as W); a series of box plots was prepared to assess differences in groundwater parameters. Two separate seasonal periods, pre-monsoon and post-monsoon, parametric changes were examined. Every value was analyzed for important water quality markers including pH, EC, TDS, Ca, Mg, TH, HCO_3^-, and Cl^-. Four comparable categories, P Pre-Monsoon, P Post-Monsoon, W Pre-Monsoon, and W Post-Monsoon, were formed out of the data. Box plots were created for every parameter across the four categories, central tendency, variability, and possible outliers.

2.4.2. Evaluation of models

The dataset was split into two different parts; 80 % of the data were used for model training and 20% for model testing (Vabalas et al., 2019). The comparison was made between models to assess each model in terms of accuracy and different parameters evaluation under different scenarios. Groundwater spatial maps were validated against actual maps prepared from historical data of 102 sample points. A similar division was used for temporal graphs of average yearly changes in water quality and levels of 54 wells with 10 years of historical data.

Assessing models' accuracy is important and error metrics offer crucial standards for model calibration, verification, and validation for machine learning models (Hodson et al., 2021; Li, 2017). An error set of metrics, including the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE) and Nash–Sutcliffe efficiency (NSE) was used to evaluate the models. R^2, RMSE, MAE, and NSE were computed from Equations (1)–(4), respectively (Chicco et al., 2021). In the following formulas, X_i is the predicted value ith value, and Y_i is the actual value ith value, m total observations, and $\overline{Y}$ average of actual value.

$$R^2 = 1 - \frac{\sum_{i=1}^{m} (X_i - Y_i)^2}{\sum_{i=1}^{m} (\overline{Y} - Y_i)^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{m}(X_i - Y_i)^2} \quad (2)$$

$$MAE = \frac{1}{m} \sum_{i=1}^{m} |X_i - Y_i| \quad (3)$$

Table 2
ML & DL models, data types, and benchmarking rationale.

Model	Data Type	Benchmarking Rationale
Ridge/ Lasso	Spatial/tabular	Interpretable linear baseline; collinearity control
KNN/SVM	Spatial/tabular	Common benchmarks in groundwater prediction, fast processing of geographical datasets
DT	Spatial/tabular	Interpretable non-linear baseline
RF/Extra Trees	Spatial/tabular (mixed hydro-env predictors)	Robust on tabular hydro-env predictors, capacity against lower variance and manage geographical data
AdaBoost	Spatial/tabular	Boosting benchmark vs ensembles
XGBoost	Spatial/tabular	Strong non-linear tabular learning
ANN	Spatial/tabular; temporal via lag features	Standard DL baseline for water variables
CNN	Temporal/spatial grids	Learning local temporal patterns
LSTM	Time sequences	Sequence model for time-series; captures long-term dependencies; common benchmark for groundwater forecasting

$$NSE = 1 - \frac{\sum_{i=1}^{m} (Y_i - X_i)^2}{\sum_{i=1}^{m} (Y_i - \overline{Y})^2} \quad (4)$$

2.4.3. Spatial cross-validation and overfitting mitigation

In spatial datasets, random train–test partitioning fails to account for spatial autocorrelation, leading to biased and typically overestimated model generalization performance. So, spatially structured cross-validation was used on the training period data to obtain a more realistic estimate of model generalization and robustness and to tune the hyperparameters (Ploton et al., 2020; Tziachris et al., 2023). In each fold, all the samples were held out related to a certain buffer zone and used models trained on the remaining groups. This strategy reduces inflation of predictive skill that occurs when training and testing samples are spatially proximate (Roberts et al., 2017). To prevent feature leakage, proximity-based predictors (e.g., nearest well value and distance-to-well) were computed using training wells only within each fold and during grid prediction. Final model analysis used a temporally independent holdout year that was not used for tuning. Overfitting was further controlled using regularization and model-complexity constraints (e.g., tree depth/min samples) and early stopping where applicable (Meyer and Pebesma, 2022).

2.4.4. Spatial Mapping

Groundwater data of collectively randomly distributed points of 2022-2023 were used for the comparative analysis of machine learning models (Fig. 3). Historical data of all groundwater parameters were collected and arranged properly. For 2022 and 2023, groundwater data of pre- and post-monsoon were used for models training. After getting the best model results, the model was validated with 2024 data. Based on those values, spatial maps were prepared for pre-monsoon and post-monsoon season using Geographical Information System software by ESRI (ArcGIS 10.8). GIS assists in identifying pollution sources and resolving associated issues by identifying specific regions with significant alterations in water quality and supply (Chabuk et al., 2020; Shaikh and Birajdar, 2024). Buffer zones were created according to specific distances from JICA and IFA wells to check groundwater changes over specific distances. Grid zones were prepared in the study area. All points were interpolated by using ArcGIS by considering conditioning variables. Inverse Distance Weighted (IDW) interpolation was used to generate continuous seasonal groundwater surfaces because it is a transparent deterministic method that assumes nearby observations are more similar than distant ones, assigning higher weights to closer points; an assumption commonly adopted for environmental mapping in GIS workflows (Childs, 2004; Li, 2021; Pro, 2021; Singh and Verma, 2019; Yasin et al., 2024). In this framework, spatial values are first estimated at unmonitored grid locations using machine-learning predictors; IDW is then applied as a consistent raster surface-generation step for seasonal visualization and hotspot screening. Kriging requires defining and fitting a valid variogram model under stationarity/intrinsic assumptions, and mapping results can be sensitive to variogram uncertainty and data support (Oliver and Webster, 2014; Webster and Oliver, 1993). Seasonal splitting (pre-/post-monsoon) can further reduce effective sample support for stable variogram estimation, so IDW was used to maintain consistency across parameters while emphasizing independent validation of mapped outputs. The first phase involved using machine learning models to predict parameter values at grid locations throughout the study region. Then, IDW was used on these grid values to make continuous raster surfaces that could be used in GIS. Spatial validation was conducted by extracting raster-estimated values from withheld validation sampling locations and comparing them to observed data using R^2, RMSE, NSE, and MAE. The IDW technique created spatial maps of all groundwater parameters before and after the monsoon season for all extracted values.

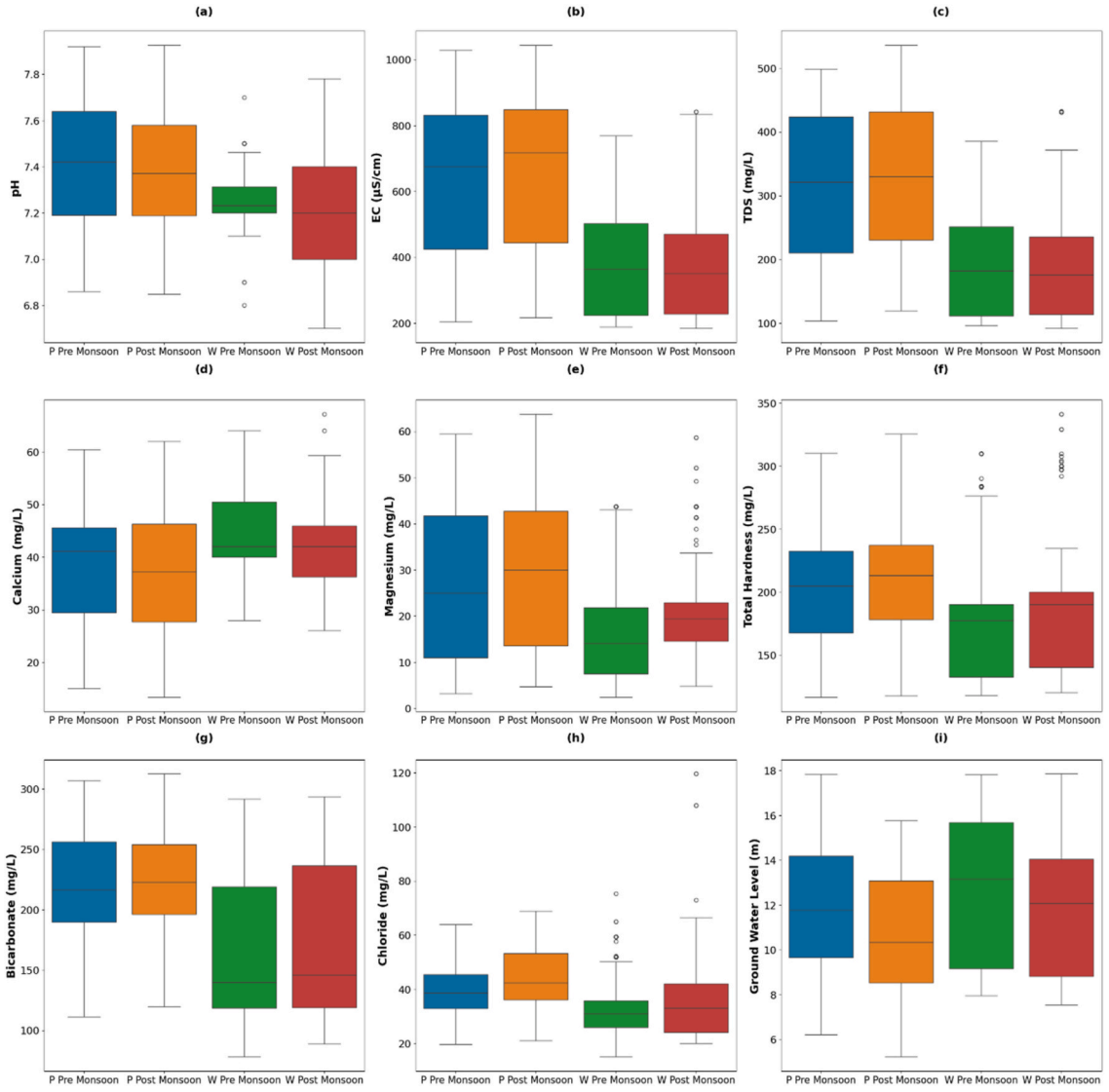


Fig. 3. Seasonal variation in groundwater quality parameters (pre- and post-monsoon) across different locations W indicating for WASA wells and P for the sample points.

2.4.5. Temporal mapping

Groundwater samples were obtained from several well locations over two separate time intervals: (pre-monsoon and post-monsoon) to evaluate seasonal fluctuations in water quality. Machine learning models were used for training (80%) and testing (20%) of the data for temporal analysis of groundwater patterns. The average value was used to study seasonal variations. Based on analysis of historical data using machine learning, the best model was used for forecasting groundwater level fluctuations and its chemical parameters. After training and testing the data, the temporal graphs were prepared to monitor seasonal fluctuations of groundwater parameters. Data was statistically analyzed and shown using box plots to assess trends, variability, and possible changes in groundwater quality over time. Forecasts were extended to 2028 to support near-term groundwater management planning and early-warning decisions, providing actionable lead time while remaining within a conservative horizon relative to the 10-year historical record (2013–2023). Because future pumping policies, land-use change, and industrial/agricultural inputs were not explicitly modelled, the 2024–2028 projections were interpreted as baseline continuation of recent system behavior by generating an uncertainty envelope (lower–upper bounds) rather than deterministic long-range predictions under changing boundary conditions (Abbas et al., 2024; Ejaz et al., 2023). Where rainfall was used as a predictor, future rainfall inputs for 2024–2028 were treated using a baseline assumption based on mean seasonal rainfall and were additionally offset in sensitivity tests to represent realistic dry and wet conditions (Boo et al., 2024; Chu et al., 2022; Nourani et al., 2022).

3. Results

3.1. Comparative analysis of model's evaluation for spatial distribution

The model's accuracy ranges from 0 to 1 for R^2. Similarly, the root mean square error and mean absolute error should be close to zero. So, after acquiring results from different machine learning models, XGBoost and RF (Table 3) models showed best accuracy. These results were obtained by comparing ten different models to check their accuracy and to handle large data. In case of EC and Cl^-., gradient boosting model has shown best results. The maximum $R^2 = 0.90$, NSE = 0.88 for XGBoost model in case of EC for pre-monsoon season. In case of post-monsoon season, XGBoost has also shown accuracy with $R^2 = 0.90$, and NSE = 0.87, MAE = 2.26 (mg/L), and RMSE = 3.42 (mg/L) in case of chloride. The same trend follows in case of groundwater level where in pre-monsoon season, the RF model has shown maximum accuracy when

Table 3
Comparative analysis of best models results for each groundwater parameter pre and post monsoon season; pH, EC (μS/cm), TDS (mg/L), Ca (mg/L), Mg (mg/L), TH (mg/L), HCO_3^-(mg/L), Cl^- (mg/L), GWL (m).

Groundwater Parameter	Season	Model	RMSE	MAE	NSE	R^2
pH	Pre-Monsoon	Ridge Regression	0.11	0.08	0.78	0.77
	Post-Monsoon	Ridge Regression	0.11	0.09	0.76	0.76
Electrical Conductivity	Pre-Monsoon	XGBoost	80.42	57.74	0.88	0.90
	Post-Monsoon	XGBoost	88.87	66.90	0.85	0.87
Total Dissolved Solids	Pre-Monsoon	XGBoost	47.77	36.88	0.80	0.81
	Post-Monsoon	Random Forest	51.54	36.66	0.73	0.77
Calcium	Pre-Monsoon	Lasso Regression	5.36	8.34	0.66	0.67
	Post-Monsoon	Ridge Regression	4.46	6.39	0.70	0.72
Magnesium	Pre-Monsoon	XGBoost	8.68	5.92	0.78	0.74
	Post-Monsoon	Random Forest	5.157	6.45	0.76	0.76
Total Hardness	Pre-Monsoon	Random Forest	21.51	19.11	0.63	0.63
	Post-Monsoon	Extra Trees	29.03	13.51	0.65	0.67
Bicarbonates	Pre-Monsoon	Random Forest	24.0	23.03	0.65	0.65
	Post-Monsoon	AdaBoost	27.10	19.12	0.66	0.67
Chlorides	Pre-Monsoon	XGBoost	3.42	2.26	0.89	0.90
	Post-Monsoon	XGBoost	6.26	5.86	0.87	0.89
Groundwater Level	Pre-Monsoon	XGBoost	3.09	1.93	0.85	0.86
	Post-Monsoon	Extra trees	3.51	2.91	0.80	0.80

compared with other models. As most of these values are above 0.75 in case of R^2 showing model results as good fit. Due to several variables like river distance, rainfall shifting patterns and groundwater pre- and post-monsoon trends, these values can vary.

In the comparison of ML models applied to groundwater data during pre- and post-monsoon seasons, XGBoost consistently achieved high accuracy for many critical metrics, including EC, TDS, Cl^-, and GWL, especially during the pre-monsoon season. The model attained R^2 values of 0.86–0.90, with NSE >0.8, accompanied by comparatively low RMSE, and MAE scores, signifying robust reliability and generalization capacity. XGBoost outperforms other models in forecasting groundwater fluctuations due to its superior handling of non-linear relationships, outliers, and variable importance across diverse environmental conditions.

3.2. Groundwater validation

3.2.1. Pre-monsoon validation

Model evaluation of groundwater parameters for the pre-monsoon season is shown (Fig. 4), comparing predicted values against actual field measurements. Each of the nine plots corresponds to a different parameter, including pH, EC, TDS, Ca, Mg, TH, HCO_3^-, Cl^-, and GWL were prepared to examine predicted and actual value trends. The pH model (Fig. 4a) produced a R^2 of 0.72, signifying a robust connection between predicted and actual values; however, the regression slope (0.78) indicated a propensity to slightly underpredict at elevated pH levels. The forecast of Electrical Conductivity (EC) (Fig. 4b) exhibited the best precision among all parameters, with a R^2 of 0.87 and a regression slope of 0.90, closely conforming to the 1:1 reference line. TDS (Fig. 4c) had a strong correlation with a R^2 of 0.76, although it displayed little underestimate at elevated concentrations. The model exhibited good performance for Ca and Mg, with R^2 values of 0.69 and 0.70, respectively (Fig. 4d & e), with regression lines indicating a slight tendency towards underprediction TH (Fig. 4f) had the least favorable model fit, with a R^2 of 0.61, suggesting more variability in its behavior during the pre-monsoon period, which may be influenced by changing subsurface conditions or seasonal recharge patterns. The model for HCO_3^- (Fig. 4g) yielded a R^2 of 0.63, with data points moderately aligned along the best-fit line, indicating a persistent underestimation. Conversely, Cl^- (Fig. 4h) was forecasted with considerable precision (R^2 = 0.85) and negligible divergence from the ideal line (slope = 0.88), affirming robust predictive reliability. The model demonstrated strong performance in forecasting GWL (Fig. 4i), with a R^2 of 0.76; however, the slope of 0.72 suggested a minor inclination to underestimate deeper water levels.

3.2.2. Post-monsoon validation

For post-monsoon season different models showed different results for individual parameters. The performance was evaluated for the post-monsoon season (Fig. 5) by scatter plots comparing predicted and observed values for essential groundwater quality metrics and groundwater levels. The pH predictions (Fig. 5a) demonstrated a robust association with actual values (R^2 = 0.74), while the slope (0.72) suggests a minor underestimation at elevated pH levels. The model's prediction for EC (Fig. 5b) attained great accuracy with a R^2 of 0.85 and a regression slope of 0.75, while it marginally underestimated values over 800 μS/cm. Likewise, TDS (Fig. 5c) were well predicted (R^2 = 0.79), with a slope of 0.82 indicating a steady trend capture with little underprediction. The model demonstrated modest predictive accuracy for Ca (Fig. 5d) with a R^2 of 0.61 and a regression slope of 0.62, indicating bigger residuals and less concordance at elevated values. Predictions for Mg (Fig. 5e) exhibited more reliability (R^2 = 0.73), but a slope of 0.70 indicates a tendency for underestimation. Predictions of TH (Fig. 5f) had the lowest performance (R^2 = 0.60), because of compounding variability in component ions. HCO_3^- concentrations (Fig. 5g) were predicted with reasonable accuracy, evidenced by a R^2 of 0.67, while a regression slope of 0.79 suggested a favorable approximation with minor bias. The model for Cl^- (Fig. 5h) attained a R^2 of 0.78 and a slope of 0.83, indicating dependable predictive capability with little underestimation. Ultimately, GWL (Fig. 5i) demonstrated a strong correlation between observed and predicted values, with a R^2 of 0.70 and a slope of 0.71. The model's post-monsoon forecasting ability was good across all parameters, especially for EC, TDS, and Cl^-.

3.3. Groundwater spatial maps

3.3.1. Pre-monsoon spatial variability

The pH map (Fig. 6a) exhibits a predominantly uniform distribution, with values ranging from 6.88 to 7.87, consistently within neutral range across the research area, while somewhat more acidic regions are seen in the western side. The EC map (Fig. 6b) exhibits significant heterogeneity, with elevated EC values (>1000 μS/cm) concentrated in the central region, suggesting potential contamination or mineral dissolution in these areas, while lower EC values are found in the southern and eastern areas. The geographical distribution of TDS (Fig. 6c) mirrors that of EC, with heightened concentrations reaching 496.4 mg/L in central and northeastern regions, indicating areas of increased salinity or solute buildup. Ca concentrations (Fig. 6d) are predominantly elevated in dispersed areas, particularly in the southeastern region, with values from 21.84 to 106.42 mg/L. Mg (Fig. 6e) exhibits significant hotspots in central regions, with concentrations reaching a maximum of 57.80 mg/L. The distribution of HCO_3^- (Fig. 6f) has a distinct north-to-south gradient, with elevated concentrations (>290 mg/L) in southern regions, indicative of carbonate weathering or organic decomposition

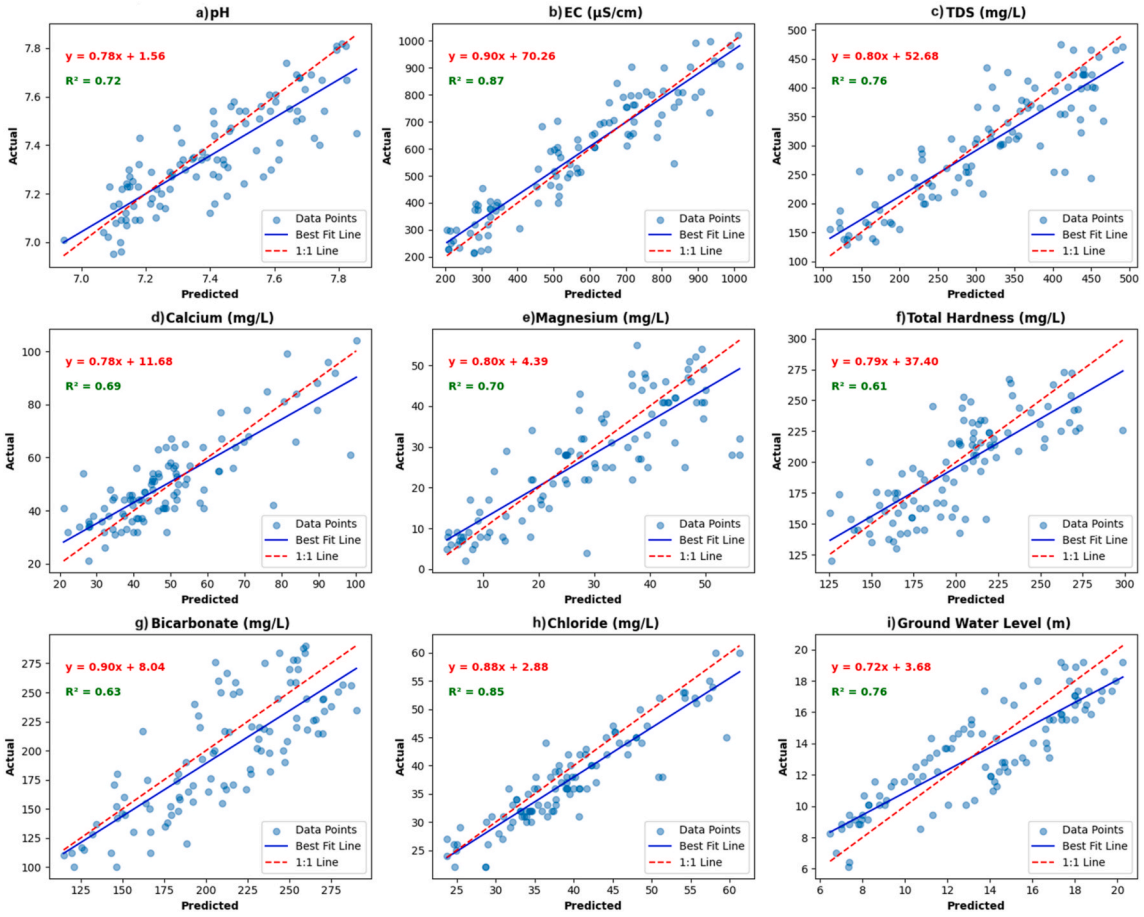


Fig. 4. Predicted vs actual values of groundwater parameters during the pre-monsoon season with regression lines and R^2 values indicating model performance.

influences. TH (Fig. 6g) has a heterogeneous distribution, with values between 121.32 and 304.78 mg/L, reaching a maximum in regions with high amounts of calcium and magnesium. The Cl^- map (Fig. 6h) reveals localized elevated values (>60 mg/L), especially in the northern and western sectors of the area, perhaps signifying pollution by domestic wastewater or agricultural runoff. The GWL map (Fig. 6i) illustrates depth fluctuations between 8.26 and 19.9 m, with greater depths before the monsoon season.

3.3.2. *Post-monsoon spatial variability*

After seasonal rainfall recharge, the post-monsoon geographical distribution maps show the hydrochemical response of the aquifer system exhibiting partial dilution, with a little acidic tendency in the northern sector and more neutral conditions in the southern part. The pH values (Fig. 7a) ranged from 6.89 to 7.81, displaying essentially near-neutral conditions across the region. Remaining higher in the central and northern zones, EC (Fig. 7b) peaked at 1006.64 µS/cm, this shows that there were localized areas with higher ionic strength. The most important sign of post-monsoon dilution is in Ca and some sections of TDS. The fact that high EC stays the same and Cl^- (Fig. 7h) levels rise in certain areas shows that there are localized flushing or inputs such as runoff or return flow and near-surface salt mobilization instead of uniform dilution across the aquifer. With increased concentrations centrally and reduced values farther southeast, TDS (Fig. 7c) ranged from 128.05 to 513.42 mg/L, showing partial dilution from monsoonal recharge. With a consistent trend of decreasing concentration throughout much of the area, the distribution of Ca (Fig. 7d) ranged from 20.07 to 101.84 mg/L, with isolated peaks. With southern areas showing values over 60 mg/L, maybe due to lithological impacts or subsurface interactions after the rise in the water table, magnesium (Mg) (Fig. 7e) showed a clear geographical gradient. With increasing levels in the southern and eastern regions, HCO_3^- concentrations (Fig. 7f) displayed a broad range from 125.23 to 295.59 mg/L, in line with increased carbonate weathering during recharging. With a range from 121.28 to 316.93 mg/L, TH (Fig. 7g) showed a distribution like that of Ca and Mg, suggesting that these two ions essentially control hardness. With concentrations ranging from 23.23 to 67.89 mg/L, Cl^- levels showed notable fluctuation; increased levels localized in the northwest region, most likely influenced by human sources such as domestic waste or agricultural runoff. With a shallower water table in the northern parts near the river (~7.27 m) relative to the deeper levels in the southern areas (~16.55 m), the GWL map (Fig. 7i) suggested a possible recharge gradient from north to south following the monsoon. The geographical distributions taken together show the impact of seasonal recharge on groundwater quality, pointing to areas of concern including EC and TDS hotspots as well as the aquifer's buffering capacity for pH and HCO_3^- stability.

3.4. *Temporal variability of model results*

A thorough comparison was performed to assess the performance of several prediction models in estimating groundwater quality indicators and levels using temporal data from 2013 to 2023. The evaluated models comprised SVM, KNN, RF, XGBoost, ANN, CNN, and LSTM networks , as depicted in Table 4 below. Their performance was evaluated using four established statistical metrics: the R^2, NSE, RMSE, and MAE. The results revealed that LSTM consistently outperformed all other models across the evaluated parameters, including pH, EC, TDS, Ca, Mg, TH, HCO_3^-, Cl^-, GWL and the relationship with rainfall. In the case of pH prediction, LSTM achieved the highest R^2 value of 0.91 with good fit NSE, along with the lowest RMSE and MAE, indicating superior accuracy and

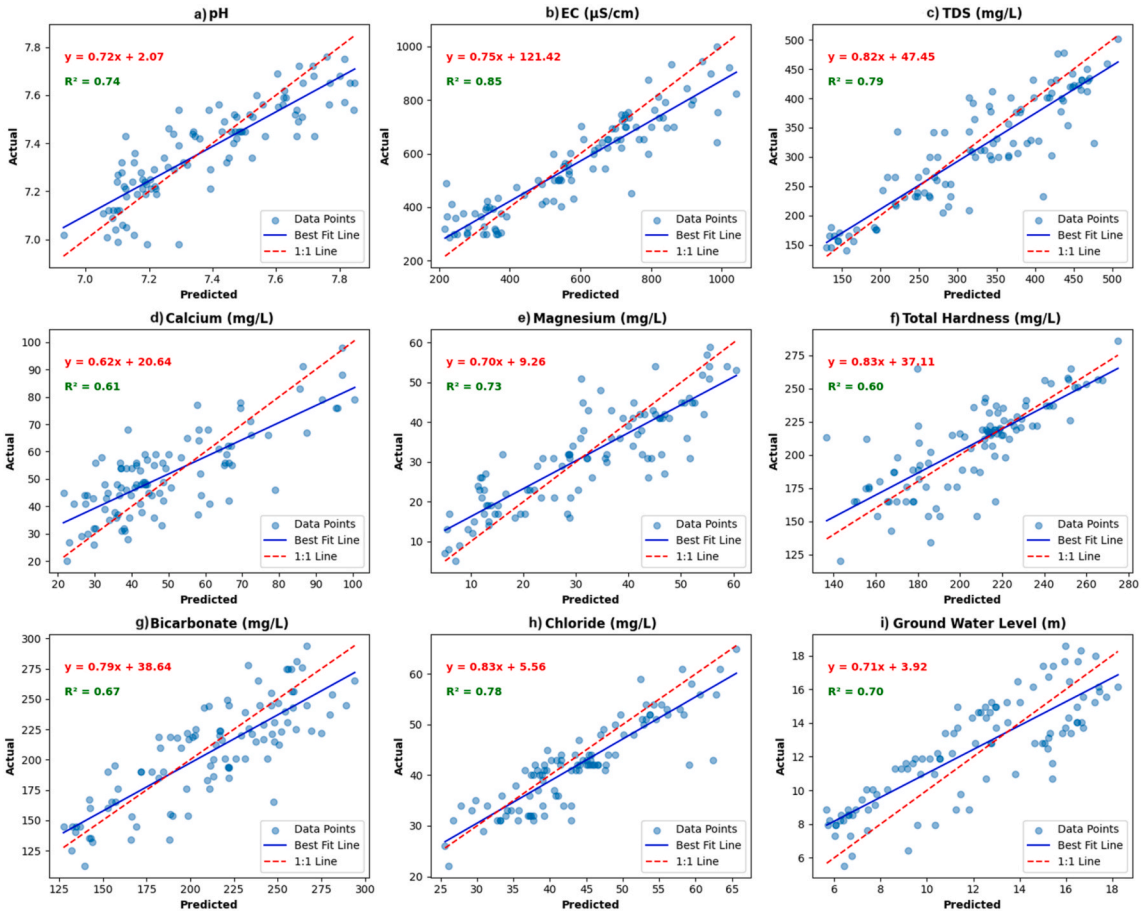


Fig. 5. Predicted vs actual values of groundwater parameters during the post-monsoon season with regression lines and R^2 values indicating model performance.

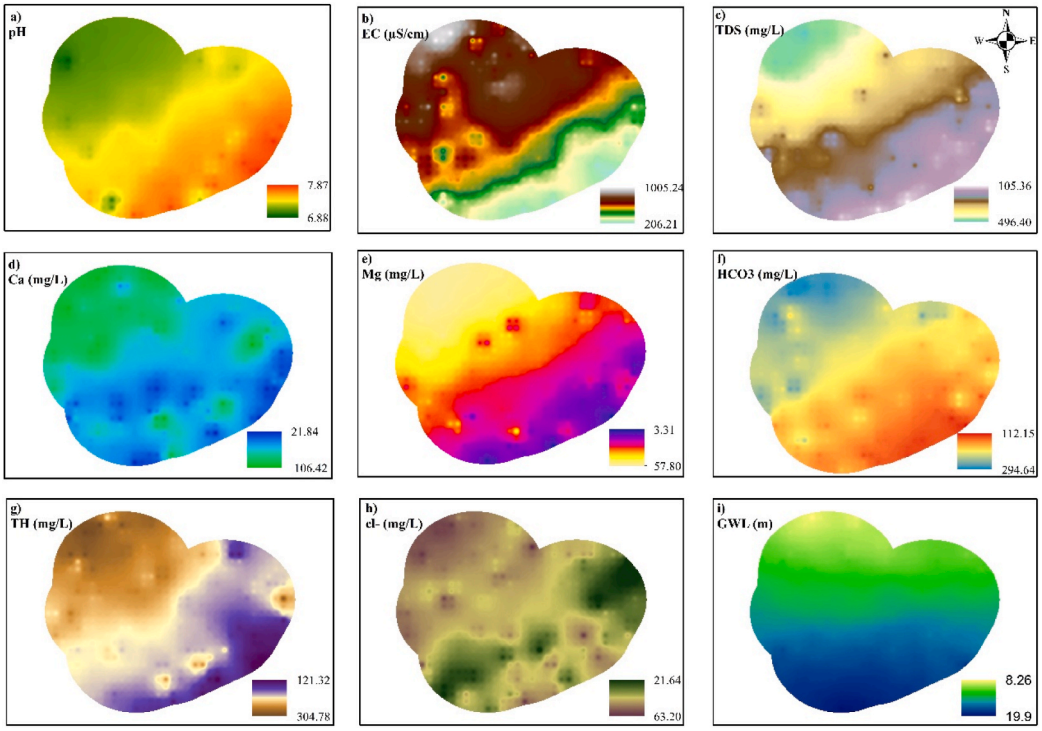


Fig. 6. Comparative investigation of groundwater variability: a comparative analysis-2024; pre-monsoon season (a) pH (b) EC (c) TDS (d) Ca (e) Mg (f) HCO_3^- (g) TH (h) Cl^- (i) GWL.

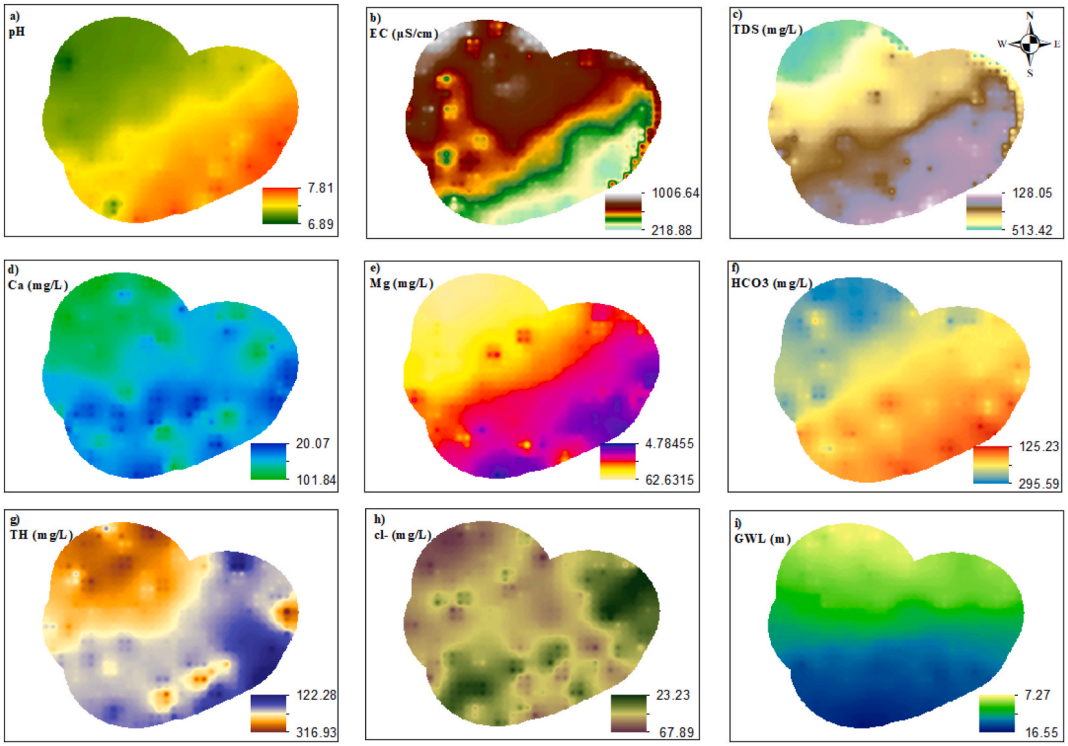


Fig. 7. Comparative investigation of groundwater variability: a comparative analysis-2024; post-monsoon season (a) pH (b) EC (c) TDS (d) Ca (e) Mg (f) HCO_3^- (g) TH (h) Cl^- (i) GWL.

Table 4
Comparative evaluation of ML & DL models for groundwater quality and level prediction using R^2, RMSE, and MAE; pH, EC (µS/cm), TDS (mg/L), Ca (mg/L), Mg (mg/L), TH (mg/L), HCO_3^- (mg/L), Cl^- (mg/L), GWL (m).

Evaluation Parameter	SVM	KNN	RF	XGBoost	ANN	CNN	LSTM
pH (R^2)	0.82	0.78	0.85	0.87	0.84	0.86	0.89
pH (RMSE)	0.35	0.42	0.31	0.29	0.33	0.31	0.27
pH (MAE)	0.28	0.34	0.25	0.23	0.26	0.24	0.22
pH (NSE)	0.83	0.77	0.84	0.86	0.85	0.87	0.86
EC (R^2)	0.79	0.75	0.83	0.86	0.82	0.84	0.88
EC (RMSE)	113.01	115.38	98.77	92.47	105.22	96.82	88.69
EC (MAE)	89.44	102.11	78.51	72.30	83.74	75.28	68.90
EC (NSE)	0.80	0.76	0.82	0.82	0.84	0.86	0.89
TDS (R^2)	0.81	0.77	0.84	0.88	0.83	0.85	0.89
TDS (RMSE)	55.22	54.62	46.38	40.19	60.57	54.82	37.27
TDS (MAE)	58.79	54.46	41.25	56.83	54.32	39.59	33.12
TDS (NSE)	0.80	0.74	0.81	0.87	0.85	0.87	0.90
Calcium (R^2)	0.76	0.72	0.8	0.83	0.79	0.81	0.85
Calcium (RMSE)	6.43	5.22	8.89	9.66	7.51	6.12	5.72
Calcium (MAE)	9.82	7.34	7.50	7.68	8.12	8.06	6.37
Calcium (NSE)	0.78	0.70	0.76	0.82	0.80	0.80	0.86
Magnesium (R^2)	0.74	0.75	0.78	0.81	0.77	0.79	0.83
Magnesium (RMSE)	8.66	9.86	7.58	6.93	8.17	7.34	6.40
Magnesium (MAE)	6.95	7.87	6.04	5.57	6.49	5.83	5.12
Magnesium (NSE)	0.75	0.73	0.78	0.80	0.79	0.81	0.83
Total Hardness (R^2)	0.81	0.76	0.84	0.87	0.82	0.85	0.89
Total Hardness (RMSE)	25.72	30.25	20.57	18.19	23.66	19.80	16.42
Total Hardness (MAE)	17.42	15.13	19.37	12.50	21.84	23.77	17.09
Total Hardness (NSE)	0.81	0.76	0.85	0.86	0.83	0.86	0.90
Bicarbonate (R^2)	0.75	0.71	0.79	0.82	0.78	0.80	0.84
Bicarbonate (RMSE)	14.51	18.32	11.83	10.69	13.15	15.44	14.22
Bicarbonate (MAE)	13.62	12.79	9.42	8.79	11.52	12.31	10.55
Bicarbonate (NSE)	0.76	0.70	0.80	0.80	0.79	0.78	0.82
Chloride (R^2)	0.77	0.73	0.81	0.84	0.80	0.82	0.86
Chloride (RMSE)	15.33	13.14	10.25	17.56	13.66	19.01	15.28
Chloride (MAE)	16.21	10.08	12.19	10	14.87	11.27	18.12
Chloride (NSE)	0.78	0.71	0.80	0.82	0.82	0.80	0.80
Rainfall & GWL (R^2)	0.83	0.79	0.86	0.89	0.85	0.87	0.91
Rainfall & GWL (RMSE)	1.31	1.41	0.92	0.99	0.55	0.61	0.46
Rainfall & GWL (MAE)	0.68	0.84	0.54	0.34	0.56	0.93	0.29
Rainfall & GWL (NSE)	0.84	0.80	0.87	0.90	0.84	0.88	0.90

reliability. Similarly, for EC and TDS, LSTM demonstrated R^2 values of 0.88 and 0.89, respectively, with the lowest RMSE (88.6 μS/cm for EC and 37.2 mg/L for TDS) and MAE (68.9 μS/cm for EC and 33.1 mg/L for TDS), highlighting its capability to handle the temporal variability inherent in groundwater quality data.

For other parameters such as Ca, Mg, TH, HCO_3^-, and Cl^-, LSTM maintained the highest predictive accuracy with minimal error values. Notably, in modeling the relationship between rainfall and groundwater level; a critical aspect of groundwater management.

LSTM excelled with $R^2 = 0.91$, NSE = 0.90, RMSE = 0.46 m, and MAE = 0.29 m. This confirms the model's strength in capturing long-term dependencies and complex temporal interactions between climatic and hydrogeological factors. Among traditional ML methods, RF and XGBoost showed relatively better performance compared to SVM and KNN, which yielded lower R^2 and higher error values across most parameters. DL models, particularly LSTM, benefited from their ability to process sequential data and identify nonlinear trends over time, especially in groundwater level monitoring. These findings underscore the effectiveness of LSTM in predicting groundwater quality and levels.

3.5. *Temporal variation in groundwater during pre- and post-monsoon periods (2013–2028)*

Because future climate and land-use conditions are uncertain and are not explicitly represented in the model inputs, the 2024–2028 projections were treated as baseline forecasts conditional on recent system behavior. Climate sensitivity was assessed by scaling seasonal rainfall by ±10% (dry/wet scenarios). The resulting forecast quantifies how GWL and salinity indicators respond to realistic climatic variability and indicates that forecast uncertainty increases with lead time and is greatest

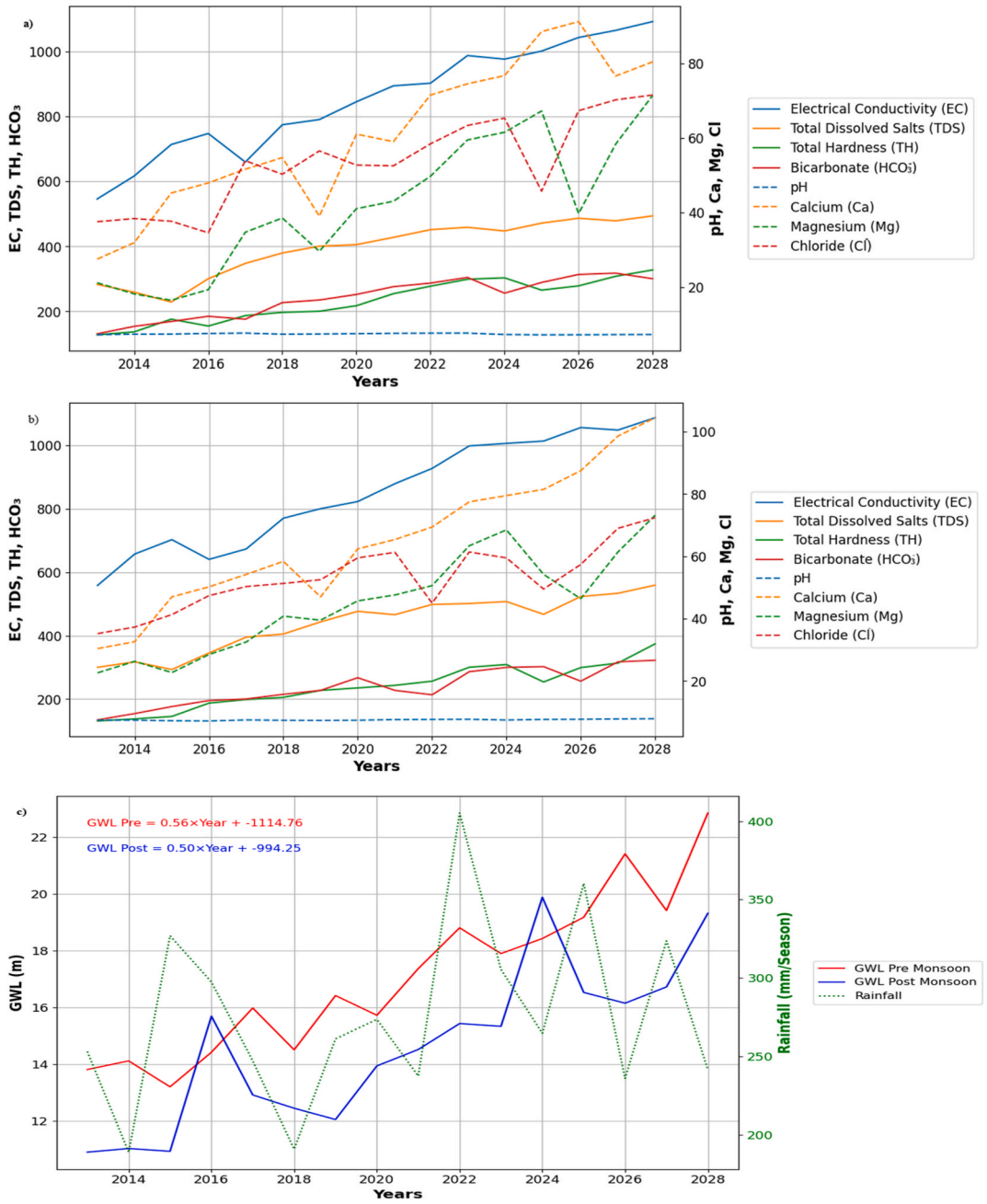


Fig. 8. Temporal variability in the study area from 2013 to 2028: a comparative analysis of (a) pre-monsoon water quality (b) post-monsoon water quality (c) groundwater level.

under dry-season assumptions. Under these scenarios, salinity-related indicators exhibited the greatest sensitivity and widening uncertainty with forecast lead time. By 2028, EC is projected to increase to approximately 1050–1100 μS/cm at some points, with a plausible envelope of 1000–1150 μS/cm under wet and dry conditions, respectively. Similarly, TDS increases to around 500–560 mg/L, ranging from 480 to 600 mg/L across rainfall scenarios. Cl^- demonstrates moderate sensitivity, reaching approximately 70–80 mg/L under baseline conditions and varying between 65 and 85 mg/L under wet and dry forcing. In contrast, pH remains stable (7.2–7.8) across the forecast horizon, reflecting buffering capacity in the aquifer system. Under wet conditions (+10% rainfall), projected depths remain slightly shallower, reaching approximately 20.8 m, while under dry forcing (−10% rainfall), groundwater levels decline more, reaching approximately 22.7 m by 2028. Similar sensitivity is observed for post-monsoon conditions, where the baseline depth is projected to reach 19.3 m, with an uncertainty envelope of approximately 21.1 m under dry rainfall forcing. Overall, the uncertainty range expands with increasing lead time, and the largest climatic sensitivity occurs for EC, TDS, and Cl^-. These scenario bounds support robust interpretation of near-term trends while acknowledging increasing climate-driven uncertainty with time.

An analysis of groundwater quality metrics from 2013 to 2028 is shown (Fig. 8), which underscores notable trends and variations. EC exhibited significant variation, measuring around 550 μS/cm in 2013 and thereafter increasing over time. TDS had the same pattern, reaching around 250 mg/L in 2013, declining to 220 mg/L in 2015, and then increasing again (Fig. 8a). These fluctuations indicate potential sources of pollution and alterations in concentration. Total hardness exhibited minimal variations, remaining relatively constant at an average of 150-200 mg/L in the early years, with an increase observed in subsequent years, indicating stable concentrations of calcium and magnesium ions, which are vital for maintaining water quality. pH values consistently ranged from 6.9 to 7.9 (Fig. 8a), indicating stable acidity/basicity, whereas bicarbonate levels, which influence alkalinity, varied from 100 to 340 mg/L at some periods, reaching a maximum value. Magnesium levels exhibited considerable variability, while calcium levels varied somewhat, averaging 70 mg/L and peaking at 108.54 mg/L (Fig. 8a).

The concentration of Cl^- in the water decreased to 48 mg/L in 2023, increasing up to 74 mg/L in 2028, based on model estimates (Fig. 8a). These patterns emphasize how crucial it is to preserve water quality and deal with possible sources of pollution through ongoing monitoring and efficient management.

Fig. 8b illustrates the temporal variation of several water quality parameters throughout the post-monsoon period from 2013 to 2028. EC values are variable; they are projected to rise to around 1045 μS/cm by 2028, after a decrease from 650 μS/cm in 2015 to roughly 1000 μS/cm in 2023 (Fig. 8b). Elevated EC indicates salinity, thereby impacting soil health and agricultural output. The levels of TDS influence the taste and suitability of water for consumption and irrigation. The levels surged significantly from around 250 mg/L in 2013 to an estimated 500 mg/L by 2028 (Fig. 8b). TH exhibited minor fluctuations, increasing from around 28 mg/L Ca in 2013 to about 38 mg/L Ca by 2019 (Fig. 8b). The pH scale is projected to rise from an average of 6.89 in 2013 to 7.8 by 2028, with peaks reaching 7.95 at certain intervals (Fig. 8b). Mg concentrations fluctuate between 30 and 50 mg/L, with elevated values at some intervals, whereas calcium levels typically range from 40 to 80 mg/L, contributing to water hardness. From 2022 to 2025, Cl^- concentrations decrease from about 48 mg/L (Fig. 8b) and then stabilize.

The temporal study of GWL (Fig. 8c), from 2014 to 2028 demonstrates a marked and continuous decrease in groundwater availability. Pre-monsoon groundwater levels demonstrated a steady decline, increasing from roughly 14 m in 2014 to nearly 22 m by 2028, as indicated by the below regression equations (5) and (6).

$$\text{GWL Pre} = 0.56\,x - 1114.76 \quad (5)$$

$$\text{GWL Post} = 0.50\,x - 994.25 \quad (6)$$

where x is the year. The reduced groundwater levels noted after the monsoon, compared to those before, indicate replenishment from seasonal precipitation events. Seasonal precipitation exhibited significant variability throughout the years, often ranging from 200 to 400 mm every season, but without a consistent long-term upward or downward trend.

4. Discussion

4.1. Comparison with existing studies

XGBoost, LSTM, and IDW were used in previous groundwater studies and are often employed separately (Osman et al., 2021; Thakur et al., 2025). As IDW is used primarily for mapping and models such as XGBoost/LSTM primarily for prediction or forecasting, often with limited comparison across ML and DL methods and without examining seasonal or task-specific suitability. The contribution of this study is a unified workflow that (i) predicts individual groundwater parameters at unmonitored locations using conditioning variables rather than proximity alone, (ii) generates pre- and post-monsoon GIS surfaces for multiple hydrochemical indicators and GWL, and (iii) produces time-series forecasts to support early warning and management decisions within one reproducible pipeline. This study found that XGBoost and RF models as demonstrated in early studies (Guo et al., 2023; Osman et al., 2021; Wang et al., 2018) were the most effective models for predicting groundwater parameters, with XGBoost achieving $R^2 = 0.90$ for electrical conductivity and 0.90 for chloride during the pre-monsoon season. In our study, validation plots showed strong correlations between predicted and actual values; for instance, EC had $R^2 = 0.87$ pre-monsoon and 0.85 post-monsoon. Spatial analysis revealed EC values exceeding 1000 μS/cm and TDS concentrations up to 496.4 mg/L in central regions, indicating potential salinity issues. The LSTM model outperformed all others, achieving $R^2 = 0.91$ and NSE 0.90 for rainfall-GWL predictions, with the lowest RMSE and MAE across all metrics. Temporally, groundwater levels declined from 14 m in 2014 to nearly 23 m by 2028, while Cl^- levels rose from 48 mg/L to 74 mg/L, highlighting growing pollution and water scarcity. The spatial mapping of groundwater revealed significant fluctuations in groundwater level and quality parameters due to pre- and post-monsoon season as well as near and far from the river. Temporal analyses from 2013 to 2028 revealed rising trends in electrical conductivity and total dissolved solids (Rao et al., 2022), signaling worsening salinity conditions, while groundwater levels showed a significant long-term decline despite seasonal recharge, with pre-monsoon depths projected to reach approximately 23 m by 2028. Overall, the model demonstrates reliable performance in predicting most groundwater quality and level indicators for the pre-monsoon season (Sahour et al., 2020). The findings demonstrate that the model successfully generalizes the groundwater dynamics during post-monsoon recharge settings impacting parameters (Jaafarzadeh et al., 2021; Mohan et al., 2018). The regional distribution of groundwater parameters (Fig. 6) measured in the pre-monsoon season exhibits notable heterogeneity driven by land use, hydrogeological conditions, and probable human activities, consistent with recent findings from spatially explicit and interpretable groundwater studies (Mo et al., 2025; Novický et al., 2010; Ranjan et al., 2006). This trend indicates shallower water levels near the river side (Yang et al., 2018), while deeper water levels in other areas either over-extraction or less recharge in particular areas. These geographical maps collectively identify areas of concern for groundwater quality in the pre-monsoon season and can establish a baseline for focused management and monitoring efforts (Oseke et al., 2021). This suggests that groundwater levels are likely to become more variable in upcoming years due to changes in seasonal rainfall, shifting cropping patterns and associated

changes in groundwater recharge and abstraction dynamics (Fishman, 2018; Smith et al., 2024; Somorowska, 2017). The values for water quality parameters must adhere to specified ranges: pH should range from 6.5 to 8.5, EC should be between 500 and 1500 μS/cm, and TDS should not exceed 1000 mg/L. The acceptable limits for Ca are 75 mg/L, Mg must not surpass 50 mg/L, HCO_3^- should be below 400 mg/L, Cl^- levels must not exceed 250 mg/L, and TH < 500 mg/L (Edition, 2011; Meride and Ayenew, 2016). Although most of the parameters are in acceptable range, these values are increasing with possible risks for future if not addressed, characterized by notable seasonal variations affected by precipitation patterns (Venkatesan et al., 2021). Significantly, years of increased precipitation were associated with recharge, highlighting the importance of recharging (Visser et al., 2018). Nonetheless, despite interannual fluctuations in rainfall, both pre- and post-monsoon groundwater levels exhibited consistent upward trends, indicating persistent groundwater depletion primarily caused by withdrawals due to increasing groundwater dependence that exceed natural recharge capacity (Sufyan et al., 2024; Turner et al., 2019). The correlation between groundwater levels and precipitation indicates that current rainfall is inadequate to offset rising groundwater extraction rates, underscoring the critical necessity for sustainable groundwater management and recharge enhancement strategies in the region.

4.2. Implications

These findings highlight the need for targeted groundwater management strategies focusing on high-risk zones, proactive salinity control measures, and improved recharge practices. Practically, this methodology and results can be applied in the study area for (i) risk-tiered seasonal monitoring with minimal pre-/post-monsoon sampling, (ii) forecast-based triggers for management by restricting new tubewell permissions and shifting from supply-driven to demand-based management due to limited surface supply in future climatic conditions, and (iii) targeted pollution control in hotspot buffers where chemical concentrations remain elevated post-monsoon, consistent with localized anthropogenic inputs and return-flow/leaching effects. This includes enforcing stricter controls on industrial and agricultural effluent discharge, promoting land use changes that protect recharge zones, implementing community-level water budgeting, and encouraging rainwater harvesting and aquifer recharge interventions (Kashiwar et al., 2016; Patra et al., 2018). Moreover, the adoption of ML and DL models like XGBoost and LSTM in groundwater prediction frameworks can enhance early warning systems, optimize water resource allocation (Ahmad and Zhang, 2022), and support sustainable groundwater governance under future climatic and anthropogenic stressors. This framework can be employed to provide large-scale data analysis, i.e., for provincial and national groundwater management policies to monitor groundwater quality and storage.

4.3. Research limitations and future work

There were some limitations recognized in this study that can be addressed in future research. A primary problem faced was the restricted access to long-term and high-resolution data, and non-uniform datasets, which limited the analytical scope and may have affected model precision. Future research should prioritize the acquisition of extensive seasonal groundwater observations with other interpolation techniques and continuous datasets from various geographical locations. This would improve understanding of regional groundwater dynamics and enhance the predictive performance of models. Furthermore, although the study examined essential hydrological parameters, several other significant factors, such as alterations in land use, agricultural intensification, population increase, and policy implementation were not comprehensively incorporated into the existing model. Forecasts for 2024–2028 are baseline projections assuming near-term stationarity of 2013–2023 relationships; this assumption is reasonable given the limited expected changes in the study area over a five-year period. Integrating these socioeconomic and environmental factors in future research may yield a more comprehensive knowledge of groundwater resource sustainability. Long-term monitoring, interdisciplinary data integration, and the utilization of modern modeling approaches, such as artificial intelligence and remote sensing, are crucial for enhancing water management systems and assuring the sustainable utilization of groundwater resources. Future research should focus on multi-parameter integration, linking groundwater data with climate projections, land-use change, and socioeconomic drivers. Expanding the model framework to include real-time monitoring, IoT-based sensors, and remote sensing inputs could further enhance predictive capability and water resource resilience in Pakistan's semi-arid zones.

5. Conclusions

This research presented an integrated, data-driven framework for assessing groundwater quality and level fluctuations through spatial and temporal analyses, employing advanced ML models for precise forecasting. The results underscore the growing threat to groundwater sustainability in the Chiniot and Faisalabad regions, likely influenced by increased tubewell installations, rapid urbanization, and agricultural expansion. Model evaluation using pre-monsoon data revealed high predictive performance across key groundwater parameters. Notably, EC demonstrated the strongest correlation (R^2 = 0.90, NSE = 0.88), followed by Cl^- (R^2 = 0.90) and GWL (R^2 = 0.87), indicating the model's ability to capture ionic variability in the groundwater. TH showed the weakest predictive power (R^2 = 0.61), likely due to seasonal recharge dynamics and subsurface variability. Post-monsoon predictions further confirmed model accuracy, especially for EC (R^2 = 0.87), and Cl^- (R^2 = 0.89), consistent with recharge-related changes after monsoon rainfall. Spatial maps for both seasons demonstrated notable regional variability, with elevated EC and TDS levels in the central and northern zones, suggesting possible anthropogenic pollution and mineral dissolution. Groundwater levels ranged from 8.26 to 19.9 m pre-monsoon and 7.27 to 16.55 m post-monsoon, highlighting seasonal recharge yet indicating long-term decline trends. The temporal model comparison (2013–2023) revealed that Long Short-Term Memory (LSTM) networks consistently outperformed traditional models across all metrics. For instance, LSTM achieved an R^2 of 0.91, NSE = 0.90, in predicting pH and groundwater levels, with lowest RMSE (0.46 m) and MAE (0.29 mg/L) for rainfall-GWL relationships. In contrast, models such as SVM and KNN yielded inferior performance, highlighting the superiority of deep learning in handling nonlinear, time-dependent hydrochemical data. Projected trends from 2013 to 2028 indicate a progressive increase in EC from ~550 to 1045 μS/cm at some locations and TDS from ~250 to ~500 mg/L, suggesting rising salinity levels and potential impacts on irrigation and drinking water suitability. These projections represent baseline continuation of recent conditions and should be interpreted with increasing uncertainty over time. Notably, groundwater levels are projected to decline from 14 m to over 23 m pre-monsoon by 2028, underscoring critical aquifer stress and insufficient recharge despite monsoonal influence. These findings offer crucial insights for adaptive groundwater governance. By accurately modeling future groundwater dynamics, tools like XGBoost for spatial prediction and LSTM for temporal trends enable proactive planning. The integration of GIS mapping, seasonal recharge analysis, and deep learning facilitate precise identification of contamination hotspots, informs early warning systems, and supports the design of site-specific mitigation strategies. Beyond prediction accuracy, the proposed workflow can be operationalized as a decision-support tool by updating hotspot maps and short-horizon forecasts annually to guide targeted monitoring, permitting, and mitigation. This directly supports adaptive groundwater governance under limited monitoring capacity.

Funding

This research received no external funding.

Disclaimer

The findings and conclusions in this publication are those of the authors and should not be construed to represent any official government determination or policy.

CRediT authorship contribution statement

Sheraz Maqbool: Writing – original draft, Visualization, Data curation, Formal analysis, Investigation, Methodology, Conceptualization. **Muhammad Imran Khan:** Writing – review & editing, Formal analysis, Supervision. **Aamir Raza:** Writing – review & editing, Validation. **Mumtaz Ali:** Writing – review & editing, Formal analysis. **Naeem Saddique:** Writing – review & editing. **Qaisar Saddique:** Formal analysis.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Sci. Data 5 (1), 1–12.
Abbas, M., Arshad, M., Shahid, M.A., 2023. Zoning of groundwater level using innovative trend analysis: case study at Rechna Doab, Pakistan. Water Resources and Irrigation Management-WRIM 12 (1-3), 64–80.
Abbas, S.A., Bailey, R.T., White, J.T., Arnold, J.G., White, M.J., Čerkasova, N., Gao, J., 2024. A framework for parameter estimation, sensitivity analysis, and uncertainty analysis for holistic hydrologic modeling using SWAT+. Hydrol. Earth Syst. Sci. 28 (1), 21–48.
Abdalla, F., Moubark, K., 2018. Assessment of well performance criteria and aquifer characteristics using step-drawdown tests and hydrogeochemical data, west of Qena area, Egypt. J. Afr. Earth Sci. 138, 336–347.
Abi, A., Walter, J., Chesnaux, R., Saeidi, A., 2024. Groundwater level monitoring using exploited domestic wells: outlier removal and imputation of missing values. Hydrogeol. J. 32 (3), 723–737.
Acosta, M.R.C., Ahmed, S., Garcia, C.E., Koo, I., 2020. Extremely randomized trees-based scheme for stealthy cyber-attack detection in smart grid networks. IEEE Access 8, 19921–19933.
Afrifa, S., Zhang, T., Appiahene, P., Varadarajan, V., 2022. Mathematical and machine learning models for groundwater level changes: a systematic review and bibliographic analysis. Future Internet 14 (9), 259.
Ahmad, I., Zhang, F., 2022. Optimal agricultural water allocation for the sustainable development of surface and groundwater resources. Water Resour. Manag. 36 (11), 4219–4236.
Ahmad, M.N., Sultana, R., Din, M.S.U., Syed, N.A., Parvaiz, R.A., Ahmad, M., Ahmad, T., 2022. Water quality mapping of district Chiniot, Pakistan by using GIS. FUUAST *Journal of Biology*, 12 (1), 1–8.
Ahmad, M.N., Sultana, R., Salahuddin, M., Ahmad, J.S., 2016. Assessment of groundwater resources in Kirana hills region, Rabwah, district Chiniot, Pakistan. International Journal of Economic and Environmental Geology 7 (2), 54–58.
Ahmad, M.N., Sultana, R., Yoshida, M., Salahuddin, M., 2020. Groundwater contamination issues in Chiniot area, Punjab, Pakistan. Int. J. Environ. Sci. Dev 11, 123–127.
Ahmed, A.N., Othman, F.B., Afan, H.A., Ibrahim, R.K., Fai, C.M., Hossain, M.S., Ehteram, M., Elshafie, A., 2019. Machine learning methods for better water quality prediction. J. Hydrol. 578, 124084.
Ali, M., Ishaq, A., Warsi, A.Z., Mohsin, A., Akbar, M., Khalil, M.N., 2025. A geographical view of Pakistan's river systems and its effects on climate change. Contemporary Journal of Social Science Review 3 (1), 289–305.
Arabgol, R., Sartaj, M., Asghari, K., 2016. Predicting nitrate concentration and its spatial distribution in groundwater resources using support vector machines (SVMs) model. Environ. Model. Assess. 21, 71–82.
Arashi, M., Roozbeh, M., Hamzah, N.A., Gasparini, M., 2021. Ridge regression and its applications in genetic studies. PLoS One 16 (4), e0245376.
Arnell, N.W., 1999. Climate change and global water resources. Glob. Environ. Change 9, S31–S49.
Asadi, E., Isazadeh, M., Samadianfard, S., Ramli, M.F., Mosavi, A., Nabipour, N., Shamshirband, S., Hajnal, E., Chau, K.-W., 2019. Groundwater quality assessment for sustainable drinking and irrigation. Sustainability 12 (1), 177.
Awais, M., Arshad, M., Shah, S.H.H., Anwar-ul-Haq, M., 2017. Evaluating groundwater quality for irrigated agriculture: spatio-temporal investigations using GIS and geostatistics in Punjab, Pakistan. Arabian J. Geosci. 10, 1–15.
Babiker, I.S., Mohamed, M.A., Hiyama, T., 2007. Assessing groundwater quality using GIS. Water Resour. Manag. 21, 699–715.
Baek, S.-S., Pyo, J., Chun, J.A., 2020. Prediction of water level and water quality using a CNN-LSTM combined deep learning approach. Water 12 (12), 3399.
Boo, K.B.W., El-Shafie, A., Othman, F., Khan, M.M.H., Birima, A.H., Ahmed, A.N., 2024. Groundwater level forecasting with machine learning models: a review. Water Res. 252, 121249.
Burri, N.M., Weatherl, R., Moeck, C., Schirmer, M., 2019. A review of threats to groundwater quality in the anthropocene. Sci. Total Environ. 684, 136–154.
Chabuk, A., Al-Madhlom, Q., Al-Maliki, A., Al-Ansari, N., Hussain, H.M., Laue, J., 2020. Water quality assessment along Tigris River (Iraq) using water quality index (WQI) and GIS software. Arabian J. Geosci. 13, 1–23.
Chebud, Y., Naja, G.M., Rivero, R.G., Melesse, A.M., 2012. Water quality monitoring using remote sensing and an artificial neural network. Water, Air, Soil Pollut. 223, 4875–4887.
Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., 2015. Xgboost: extreme gradient boosting. R package version 0.4-2 1 (4), 1–4.
Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput. Sci. 7, e623.
Childs, C., 2004. Interpolating surfaces in ArcGIS spatial analyst. ArcUser. July-September 3235 (569), 32–35.
Chu, H., Bian, J., Lang, Q., Sun, X., Wang, Z., 2022. Daily groundwater level prediction and uncertainty using LSTM coupled with PMI and bootstrap incorporating teleconnection patterns information. Sustainability 14 (18), 11598.
Cosgrove, W.J., Loucks, D.P., 2015. Water management: current and future challenges and research directions. Water Resour. Res. 51 (6), 4823–4839.
Dalin, C., Wada, Y., Kastner, T., Puma, M.J., 2017. Groundwater depletion embedded in international food trade. Nature 543 (7647), 700–704.
de Graaf, I.E., Gleeson, T., Van Beek, L., Sutanudjaja, E.H., Bierkens, M.F., 2019. Environmental flow limits to global groundwater pumping. Nature 574 (7776), 90–94.
Edition, F., 2011. Guidelines for drinking-water quality. WHO Chron. 38 (4), 104–108.
Ejaz, F., Guthke, A., Wöhling, T., Nowak, W., 2023. Comprehensive uncertainty analysis for surface water and groundwater projections under climate change based on a lumped geo-hydrological model. J. Hydrol. 626, 130323.
El Bilali, A., Taleb, A., Brouziyne, Y., 2021. Groundwater quality forecasting using machine learning algorithms for irrigation purposes. Agric. Water Manag. 245, 106625.
Evans, S., Williams, G.P., Jones, N.L., Ames, D.P., Nelson, E.J., 2020. Exploiting earth observation data to impute groundwater level measurements with an extreme learning machine. Remote Sens. 12 (12), 2044.
Fishman, R., 2018. Groundwater depletion limits the scope for adaptation to increased rainfall variability in India. Clim. Change 147, 195–209.
Gilbert, J., Boateng, C.D., Aryee, J.N., Osei, M.A., Wemegah, D.D., Gidigasu, S.S., Britwum, A., Afful, S.K., Touré, H., Mensah, V., 2025. A systematic review of machine learning models for groundwater level prediction. Applied Computing and Geosciences, 100303.
Giordano, M., 2009. Global groundwater? Issues and solutions. Annu. Rev. Environ. Resour. 34 (1), 153–178.
Gleeson, T., Befus, K.M., Jasechko, S., Luijendijk, E., Cardenas, M.B., 2016. The global volume and distribution of modern groundwater. Nat. Geosci. 9 (2), 161–167.
Gleeson, T., Cuthbert, M., Ferguson, G., Perrone, D., 2020. Global groundwater sustainability, resources, and systems in the Anthropocene. Annu. Rev. Earth Planet Sci. 48 (1), 431–463.
Griebler, C., Avramov, M., 2015. Groundwater ecosystem services: a review. Freshw. Sci. 34 (1), 355–367.
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. J. Hydrol. 621, 129599.
Gupta, R., Singh, A., Singhal, A., 2019. Application of ANN for water quality index. Int. J. Mach. Learn. Comput 9 (5), 688–693.
Hadžić, E., Lazović, N., Mulaomerović-Šeta, A., 2015. The importance of groundwater vulnerability maps in the protection of groundwater sources. Key Study: sarajevsko Polje. Procedia Environ. Sci. 25, 104–111.
Halder, S., Roy, M.B., Roy, P.K., 2020. Analysis of groundwater level trend and groundwater drought using Standard Groundwater Level Index: a case study of an eastern river basin of West Bengal, India. SN Appl. Sci. 2 (3), 507.
Hao, Z., Zhao, H., Zhang, C., Zhou, H., Zhao, H., Wang, H., 2019. Correlation analysis between groundwater decline trend and human-induced factors in Bashang region. Water 11 (3), 473.
Hodson, T.O., Over, T.M., Foks, S.S., 2021. Mean squared error, deconstructed. J. Adv. Model. Earth Syst. 13 (12), e2021MS002681.

Hu, Z., Zhang, Y., Zhao, Y., Xie, M., Zhong, J., Tu, Z., Liu, J., 2019. A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture. Sensors 19 (6), 1420.
Huang, Y., Wang, C., Wang, Y., Lyu, G., Lin, S., Liu, W., Niu, H., Hu, Q., 2024. Application of machine learning models in groundwater quality assessment and prediction: progress and challenges. Front. Environ. Sci. Eng. 18 (3), 29.
Jaafarzadeh, M.S., Tahmasebipour, N., Haghizadeh, A., Pourghasemi, H.R., Rouhani, H., 2021. Groundwater recharge potential zonation using an ensemble of machine learning and bivariate statistical models. Sci. Rep. 11 (1), 5587.
Jain, C.K., Sharma, S.K., Singh, S., 2021. Assessment of groundwater quality and determination of hydrochemical evolution of groundwater in Shillong, Meghalaya (India). SN Appl. Sci. 3 (1), 33.
Kasahun, M., Diriba, D., Kawo, N.S., Karuppannan, S., Tedla, H.Z., 2025. Data-Driven insights into factors controlling groundwater potential using explainable and ensemble machine learning models. Earth Systems and Environment 1–27.
Kashiwar, S.R., Dongarwar, U.R., Mondal, B., Kundu, M.C., 2016. An overview on the ground water recharge by rain water harvesting. Journal of Energy Research and Environmental Technology 3 (2), 146–148.
Kausar, F., Qadir, A., Ahmad, S.R., Baqar, M., 2019. Evaluation of surface water quality on spatiotemporal gradient using multivariate statistical techniques: a case study of River Chenab, Pakistan. Pol. J. Environ. Stud. 28 (4), 2645–2657.
Khaliq, A., Maqbool, A., Tansar, H., Bakhsh, A., Saeed, M., Sarwar, M.T., Hui, W., 2021. Groundwater pumping modeling for the sustainable management of urban water supply in Faisalabad city, Pakistan. Arabian J. Geosci. 14, 1–14.
Khouni, I., Louhichi, G., Ghrabi, A., 2021. Use of GIS based Inverse Distance Weighted interpolation to assess surface water quality: case of Wadi El Bey, Tunisia. Environ. Technol. Innovat. 24, 101892.
Kılıç, Z., 2020. The importance of water and conscious use of water. International Journal of Hydrology 4 (5), 239–241.
Kirby, M., Mainuddin, M., Khaliq, T., Cheema, M.J.M., 2017. Agricultural production, water use and food availability in Pakistan: historical trends, and projections to 2050. Agric. Water Manag. 179, 34–46.
Konikow, L.F., Kendy, E., 2005. Groundwater depletion: a global problem. Hydrogeol. J. 13, 317–320.
Kouadri, S., Pande, C.B., Panneerselvam, B., Moharir, K.N., Elbeltagi, A., 2022. Prediction of irrigation groundwater quality parameters using ANN, LSTM, and MLR models. Environ. Sci. Pollut. Control Ser. 29 (14), 21067–21091.
Kratzert, F., Klotz, D., Brenner, C., Schulz, K., Herrnegger, M., 2018. Rainfall–runoff modelling using long short-term memory (LSTM) networks. Hydrol. Earth Syst. Sci. 22 (11), 6005–6022.
Kumar, N., Yamaç, S.S., Velmurugan, A., 2015. Applications of remote sensing and GIS in natural resource management. Journal of the Andaman Science Association 20 (1), 1–6.
Lashkaripour, G.R., Ghafoori, M., 2011. The effects of water table decline on the groundwater quality in aquifer of Torbat Jam Plain, Northeast Iran. Int. J. Emerg. Sci. 1 (2), 153.
Lee, J.H., Shi, Z., Gao, Z., 2022. On LASSO for predictive regression. J. Econom. 229 (2), 322–349.
Li, F., Qiao, J., Zhao, Y., Zhang, W., 2014. Risk assessment of groundwater and its application. Part II: using a groundwater risk maps to determine control levels of the groundwater. Water Resour. Manag. 28, 4875–4893.
Li, J., 2017. Assessing the accuracy of predictive models for numerical data: not r nor r2, why not? Then what? PLoS One 12 (8), e0183250.
Li, P., 2016. Groundwater quality in western China: challenges and paths forward for groundwater quality research in western China. Expo. Health 8 (3), 305–310.
Li, Z., 2021. An enhanced dual IDW method for high-quality geospatial interpolation. Sci. Rep. 11 (1), 9903.
Lin, M., Biswas, A., Bennett, E.M., 2019. Spatio-temporal dynamics of groundwater storage changes in the Yellow River Basin. J. Environ. Manag. 235, 84–95.
Liu, P., Wang, J., Sangaiah, A.K., Xie, Y., Yin, X., 2019. Analysis and prediction of water quality using LSTM deep neural networks in IoT environment. Sustainability 11 (7), 2058.
Masood, A., Tariq, M.A.U.R., Hashmi, M.Z.U.R., Waseem, M., Sarwar, M.K., Ali, W., Farooq, R., Almazroui, M., Ng, A.W., 2022. An overview of groundwater monitoring through point-to satellite-based techniques. Water 14 (4), 565.
Meride, Y., Ayenew, B., 2016. Drinking water quality assessment and its effects on residents health in Wondo genet campus, Ethiopia. Environ. Syst. Res. 5 (1), 1–7.
Meyer, H., Pebesma, E., 2022. Machine learning-based global maps of ecological variables and the challenge of assessing them. Nat. Commun. 13 (1), 2208.
Mo, Y., Xu, J., Zhu, S., Xu, B., Wu, J., Jin, G., Wang, Y.-G., Li, L., 2025. Spatial heterogeneity of groundwater depths in coastal cities and their responses to multiple factors interactions by interpretable machine learning models. Geoscience *Frontiers*, 16 (3), 102033.
Mohan, C., Western, A.W., Wei, Y., Saft, M., 2018. Predicting groundwater recharge for varying land cover and climate conditions–a global meta-study. Hydrol. Earth Syst. Sci. 22 (5), 2689–2703.
Molle, F., López-Gunn, E., Van Steenbergen, F., 2018. The local and national politics of groundwater overexploitation. Water Altern. (WaA) 11 (3).
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 Resour. Manag. 35, 23–37.
Motlagh, A.M., Yang, Z., Saba, H., 2020. Groundwater quality. Water Environ. Res. 92 (10), 1649–1658.
Nourani, V., Khodkar, K., Gebremichael, M., 2022. Uncertainty assessment of LSTM based groundwater level predictions. Hydrol. Sci. J. 67 (5), 773–790.
Novický, O., Kašpárek, L., Uhlik, J., 2010. Vulnerability of groundwater resources in different hydrogeological conditions to climate change. In: Taniguchi, M., Holman, I. P. (Eds.), Groundwater Response to Changing Climate. cRc Press/Balkema, Leiden, Netherlands, pp. 1–10.
Ohlert, P.L., Bach, M., Breuer, L., 2023. Accuracy assessment of inverse distance weighting interpolation of groundwater nitrate concentrations in Bavaria (Germany). Environ. Sci. Pollut. Control Ser. 30 (4), 9445–9455.
Oliver, M., Webster, R., 2014. A tutorial guide to geostatistics: computing and modelling variograms and kriging. Catena 113, 56–69.
Oseke, F.I., Anornu, G.K., Adjei, K.A., Eduvie, M.O., 2021. Assessment of water quality using GIS techniques and water quality index in reservoirs affected by water diversion. Water Energy Nexus 4, 25–34.
Osman, A.I.A., Ahmed, A.N., Chow, M.F., Huang, Y.F., El-Shafie, A., 2021. Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. Ain Shams Eng. J. 12 (2), 1545–1556.
Palani, S., Liong, S.-Y., Tkalich, P., 2008. An ANN application for water quality forecasting. Mar. Pollut. Bull. 56 (9), 1586–1597.
Patra, S., Sahoo, S., Mishra, P., Mahapatra, S.C., 2018. Impacts of urbanization on land use/cover changes and its probable implications on local climate and groundwater level. Journal of urban management 7 (2), 70–84.
Patra, S.R., Chu, H.-J., 2023. Regional groundwater sequential forecasting using global and local LSTM models. Journal of Hydrology. Reg. Stud. 47, 101442.
Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., 2020. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. Nat. Commun. 11 (1), 4540.
Pourmorad, S., Kabolizade, M., Dimuccio, L.A., 2024. Artificial intelligence advancements for accurate groundwater level modelling: an updated synthesis and review. Appl. Sci. 14 (16), 7358.
Pro, A., 2021. How inverse distance weighted interpolation works. Pristupljeno: https://pro.arcgis.com/en/pro-app/help/analysis/geostatistical-analyst/how-inversedistance-weighted-interpolation-works.htm, 17.
Quddoos, A., Muhmood, K., Naz, I., Aslam, R.W., Usman, S.Y., 2024. Geospatial insights into groundwater contamination from urban and industrial effluents in Faisalabad. Discover Water 4 (1), 50.
Qureshi, A.S., 2015. Improving food security and livelihood resilience through groundwater management in Pakistan. Glob. Adv. Res. J. Agric. Sci 4, 687–710.
Qureshi, A.S., 2018. Challenges and opportunities of groundwater management in Pakistan. Groundwater of South Asia 735–757.
Qureshi, A.S., 2020. Groundwater governance in Pakistan: from colossal development to neglected management. Water 12 (11), 3017.
Qureshi, A.S., Gill, M.A., Sarwar, A., 2010. Sustainable groundwater management in Pakistan: challenges and opportunities. Irrigation and drainage. The Journal of the International Commission on Irrigation and Drainage 59 (2), 107–116.
Ranjan, S.P., Kazama, S., Sawamoto, M., 2006. Effects of climate and land use changes on groundwater resources in coastal aquifers. J. Environ. Manag. 80 (1), 25–35.
Rao, N.S., Sunitha, B., Das, R., Kumar, B.A., 2022. Monitoring the causes of pollution using groundwater quality and chemistry before and after the monsoon. Phys. Chem. Earth, Parts A/B/C 128, 103228.
Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40 (8), 913–929.
Rodell, M., Famiglietti, J.S., Wiese, D.N., Reager, J., Beaudoing, H.K., Landerer, F.W., Lo, M.-H., 2018. Emerging trends in global freshwater availability. Nature 557 (7707), 651–659.
Saccò, M., Mammola, S., Altermatt, F., Alther, R., Bolpagni, R., Brancelj, A., Brankovits, D., Fišer, C., Gerovasileiou, V., Griebler, C., 2024. Groundwater is a hidden global keystone ecosystem. Glob. Change Biol. 30 (1), e17066.
Sahour, H., Gholami, V., Vazifedan, M., 2020. A comparative analysis of statistical and machine learning techniques for mapping the spatial distribution of groundwater salinity in a coastal aquifer. J. Hydrol. 591, 125321.
Scanlon, B.R., Fakhreddine, S., Rateb, A., de Graaf, I., Famiglietti, J., Gleeson, T., Grafton, R.Q., Jobbagy, E., Kebede, S., Kolusu, S.R., 2023. Global water resources and the role of groundwater in a resilient water future. Nat. Rev. Earth Environ. 4 (2), 87–101.
Shaikh, M., Birajdar, F., 2023. Groundwater management and sustainable farming practices: a socioeconomic analysis of their interplay in rural agriculture-a case study of Solapur, Maharashtra. International Journal for Innovative Science Research Trends and Innovation 8 (9).
Shaikh, M., Birajdar, F., 2024. Advancements in remote sensing and GIS for sustainable groundwater monitoring: applications, challenges, and future directions. International Journal of Research in Engineering, Science and Management 7 (3), 16–24.
Shakoor, A., Arshad, M., Bakhsh, A., Ahmed, R., 2015. GIS based assessment and delineation of groundwater quality zones and its impact on agricultural productivity. Pak. J. Agri. Sci 52 (3), 837–843.
Shaw, S.K., Sharma, A., 2024. Assessment of groundwater quality and suitability for irrigation purpose using irrigation indices, remote sensing and GIS approach. Groundw. Sustain. Dev. 26, 101297.
Singh, P., Verma, P., 2019. A comparative study of spatial interpolation technique (IDW and Kriging) for determining groundwater quality. GIS and geostatistical techniques for groundwater science 43–56.
Singha, K., Navarre-Sitchler, A., 2022. The importance of groundwater in critical zone science. Groundwater 60 (1), 27–34.

Singha, S., Pasupuleti, S., Singha, S.S., Singh, R., Kumar, S., 2021. Prediction of groundwater quality using efficient machine learning technique. Chemosphere 276, 130265.
Smith, J.A., Baeck, M.L., Miller, A.J., Claggett, E.L., 2024. Rainfall frequency analysis based on long-term high-resolution radar rainfall fields: spatial heterogeneities and temporal nonstationarities. Water Resour. Res. 60 (3), e2023WR035640.
Somorowska, U., 2017. Climate-driven changes to streamflow patterns in a groundwater-dominated catchment. Acta Geophysica 65 (4), 789–798.
Sufyan, M., Martelli, G., Teatini, P., Cherubini, C., Goi, D., 2024. Managed aquifer recharge for sustainable groundwater management: new developments, challenges, and future prospects. Water 16 (22), 1–28.
Sultana, R., Salahuddin, M., Ahmad, M.N., 2018. Economic impact assessment of brackish groundwater in Kirana Hills region, District Chiniot, Pakistan. Int J Econ Environ Geol 9 (3), 19–24.
Sun, Y., Zhang, J., Zhang, Y., 2024. Adaboost algorithm combined multiple random forest models (Adaboost-RF) is employed for fluid prediction using well logging data. Phys. Fluids 36 (1).
Suthaharan, S., Suthaharan, S., 2016. Support Vector Machine. Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning, pp. 207–235.
Tao, H., Hameed, M.M., Marhoon, H.A., Zounemat-Kermani, M., Heddam, S., Kim, S., Sulaiman, S.O., Tan, M.L., Sa'adi, Z., Mehr, A.D., 2022. Groundwater level prediction using machine learning models: a comprehensive review. Neurocomputing 489, 271–308.
Thakur, A., Chandel, A., Shankar, V., 2025. Prediction of groundwater levels using a long short-term memory (LSTM) technique. J. Hydroinform. 27 (1), 51–68.
Turner, S.W., Hejazi, M., Yonkofski, C., Kim, S.H., Kyle, P., 2019. Influence of groundwater extraction costs and resource depletion limits on simulated global nonrenewable water withdrawals over the twenty-first century. Earths Future 7 (2), 123–135.
Tziachris, P., Nikou, M., Aschonitis, V., Kallioras, A., Sachsamanoglou, K., Fidelibus, M. D., Tziritis, E., 2023. Spatial or random cross-validation? The effect of resampling methods in predicting groundwater salinity with machine learning in mediterranean region. Water 15 (12), 2278.
Ubah, J., Orakwe, L., Ogbu, K., Awu, J., Ahaneku, I., Chukwuma, E., 2021. Forecasting water quality parameters using artificial neural network for irrigation purposes. Sci. Rep. 11 (1), 24438.
ul Hasan, F., Fatima, B., 2025. A review of drivers contributing to unsustainable groundwater consumption in Pakistan. Groundw. Sustain. Dev., 101414
Vabalas, A., Gowen, E., Poliakoff, E., Casson, A.J., 2019. Machine learning algorithm validation with a limited sample size. PLoS One 14 (11), e0224365.
Van Steenbergen, F., Oliemans, W., 2002. A review of policies in groundwater management in Pakistan 1950–2000. Water Policy 4 (4), 323–344.
Vaux, H., 2011. Groundwater under stress: the importance of management. Environ. Earth Sci. 62, 19–23.
Venkatesan, G., Subramani, T., Karunanidhi, D., Sathya, U., Li, P., 2021. Impact of precipitation disparity on groundwater fluctuation in a semi-arid region (Vellore district) of southern India using geospatial techniques. Environ. Sci. Pollut. Control Ser. 28, 18539–18551.
Visser, A., Moran, J.E., Singleton, M.J., Esser, B.K., 2018. Importance of river water recharge to the San Joaquin Valley groundwater system. Hydrol. Process. 32 (9), 1202–1213.
Wadoux, A.M.-C., Heuvelink, G.B., De Bruin, S., Brus, D.J., 2021. Spatial cross-validation is not the right way to evaluate map accuracy. Ecol. Model. 457, 109692.
Wang, W., Zhang, Z., Duan, L., Wang, Z., Zhao, Y., Zhang, Q., Dai, M., Liu, H., Zheng, X., Sun, Y., 2018. Response of the groundwater system in the Guanzhong Basin (central China) to climate change and human activities. Hydrogeol. J. 26 (5), 1429–1441.
Webster, R., Oliver, M., 1993. How large a sample is needed to estimate the regional variogram adequately?. In: Geostatistics Tróia'92, vol. 1. Springer, pp. 155–166.
Yan, Y., Zhang, Y., Yao, R., Wei, C., Luo, M., Yang, C., Chen, S., Huang, X., 2024. Groundwater suitability assessment for irrigation and drinking purposes by integrating spatial analysis, machine learning, water quality index, and health risk model. Environ. Sci. Pollut. Control Ser. 31 (27), 39155–39176.
Yang, Y., Zhang, M., Sun, Z., Han, J., Wang, J., 2018. The relationship between water level change and river channel geometry adjustment in the downstream of the Three Gorges Dam. J. Geogr. Sci. 28 (12), 1975–1993.
Yasin, K.H., Gelete, T.B., Iguala, A.D., Kebede, E., 2024. Optimal interpolation approach for groundwater depth estimation. MethodsX 13, 102916.
Zarco-Perello, S., Simões, N., 2017. Ordinary kriging vs inverse distance weighting: spatial interpolation of the sessile community of Madagascar reef, Gulf of Mexico. PeerJ 5, e4078.
Zeilhofer, P., Zeilhofer, L.V.A.C., Hardoim, E.L., Lima, Z.M.d., Oliveira, C.S., 2007. GIS applications for mapping and spatial modeling of urban-use water quality: a case study in District of Cuiabá, Mato Grosso, Brazil. Cad. Saúde Pública 23 (4), 875–884.
Zendehboudi, A., Baseer, M.A., Saidur, R., 2018. Application of support vector machine models for forecasting solar and wind energy resources: a review. J. Clean. Prod. 199, 272–285.
Zhu, M., Wang, J., Yang, X., Zhang, Y., Zhang, L., Ren, H., Wu, B., Ye, L., 2022. A review of the application of machine learning in water quality evaluation. Eco-Environ. Health 1 (2), 107–116.