Journal of Groundwater Science and Engineering 14(2026) 100–122

Journal homepage: http://gwse.iheg.org.cn

Review Article

AI and ML in groundwater exploration and water resources management: Concepts, methods, applications, and future directions

Adla Andalu[1], M Gopal Naik[1], Sandeep Budde[2,3,4*]

[1] *Department of Civil Engineering, Osmania University, Hyderabad, Telangana, India.*
[2] *Bharti Institute of Public Policy, Indian School of Business, Hyderabad, India.*
[3] *Nitte School of Architecture Planning and Design, Bangalore, Nitte Deemed University, India.*
[4] *School of Urban and Regional Planning, University of Alberta, Edmonton AB - T6G 2R3, Canada.*

Abstract: The integration of Artificial Intelligence (AI) and Machine Learning (ML) into groundwater exploration and water resources management has emerged as a transformative approach to addressing global water challenges. This review explores key AI and ML concepts, methodologies, and their applications in hydrology, focusing on groundwater potential mapping, water quality prediction, and groundwater level forecasting. It discusses various data acquisition techniques, including remote sensing, geospatial analysis, and geophysical surveys, alongside preprocessing methods that are essential for enhancing model accuracy. The study highlights AI-driven solutions in water distribution, allocation optimization, and real-time resource management. Despite their advantages, the application of AI and ML in water sciences faces several challenges, including data scarcity, model reliability, and the integration of these tools with traditional water management systems. Ethical and regulatory concerns also demand careful consideration. The paper also outlines future research directions, emphasizing the need for improved data collection, interpretable models, real-time monitoring capabilities, and interdisciplinary collaboration. By leveraging AI and ML advancements, the water sector can enhance decision-making, optimize resource distribution, and support the development of sustainable water management strategies.

Keywords: Artificial intelligence; Machine learning; Groundwater exploration; Hydrological modeling; Remote sensing applications; Water resources management

Received: 22 Dec 2024/ Accepted: 28 Jul 2025/ Published: 22 Oct 2025

Introduction

Natural groundwater and Earth's water resources are integral components of the hydrological cycle, providing essential life-sustaining resources such as potable water for domestic use, irrigation for agriculture, and water for various industrial activities (Chakraborty et al. 2021). As a finite resource, water underpins the balance of ecosystems, human life, and global development. However, the growing global population, along with increased urbanization, industrialization and climate change, has collectively exerted pressure on the water resources, leading to a reduced availability of freshwater.

To meet rising demands for clean water, groundwater, the most abundant source of fresh water on Earth, is being increasingly exploited. In recent years, the groundwater ecosystem has undergone significant deterioration due to changes in climate and land use. Climate change exserts considerable stress on water resources, particularly groundwater, by influencing temperature and precipitation patterns (Wang et al. 2021). Rising temperatures impact groundwater quality by altering primary productivity and microbial activity, affecting

---

[*] Corresponding author: Sandeep Budde, *E-mail address*: sbudde@ualberta.ca

DOI: 10.26599/JGSE.2026.9280059

Andalu A, Naik M G, Budde S. 2026. AI and ML in groundwater exploration and water resources management: Concepts, methods, applications, and future directions. Journal of Groundwater Science and Engineering, 14(1): 100-122.

geochemical processes, increasing the risk of harmful algal blooms, leading to greater pesticide usage, and elevating wildfire risks (Dao et al. 2024).

Traditionally, groundwater modelling has relied on Process-Based Physical Models (PBPMs) to assess and predict groundwater quality. However, these models oten assume linear relationships between variables, limiting their ability to capture complex interactions among influencing factors. Additionally, PBPMs tends to be computationally intensive. In recent years, Machine Learning (ML) models have emerged as promising alternatives for extracting valuable insights from complex datasets. Due to their efficiency and predictive capapilities, ML techniques are increasingly being used in place of classical PBPMs for groundwater modelling.

Consequently, adaptive groundwater management strategies should incorporate Artificial Intelligence (AI) and ML technologies. Advanced FeedForward Neural Networks (FFNNs), for example, are particularly well-suited to processing extensive data (Kamyab et al. 2023). These technologies play a vital role in the efficient management of large datasets for tasks such as identifying groundwater potential zones, predicting water quality, and optimally designing water distribution systems (Huang et al. 2022; Yang et al. 2024).

This paper aims to review current AI and ML methodologies for groundwater exploration and water resources management, providing guidance for research, development, as well as policymaking aimed at sustainable societal benefit. Integrating AI/ML techniques with structural parameters can significantly improve outcomes and enhance decision-making in water resource management. Among the many approaches, techniques such as deep learning, neural networks and predictive modelling are proving especially effective in handling complex hydrological datasets (Mohammadi, 2021). These methods enhance the accuracy and reliability of groundwater resource evaluation by integrating multiple data sources (Nguyen et al. 2022).

Beyond groundwater exploration, AI and ML techniques can aid water resource management by addressing global environmental changes and uncertainties, such as the future climate change impacts on surface water-groundwater interactions. Their use extends to municipal water systems for optimizing drinking water distribution, forecasting water requirements, detecting anomalies, and improving water use efficiency (WUE) (Swain et al. 2022).

The scale and adaptability of snail-farming make it potentially deployable across a wide range of geographies and climates, playing an important role in both poverty reduction and environmental conservation (Apostolou et al. 2021). Although AI and ML have made significant advancements, they remain in early stages of a growth curve that would enable their effective implementation in groundwater exploration and water resource management.

Two fundamental bottlenecks hindering further progress are data quality and availability. Many AI/ML models perform poorly when trained on subpar or incomplete datasets. Furthermore, these models often to generalize or predict accurately due to insufficient representation of, or incorrect assumptions about, hydrological variability across multiple spatial scales (Chang and Guo, 2020).

To overcome these limitations, greater emphasis must be placed on improving the interpretability of AI and ML models through enhanced data collection techniques and robust parameterization. Integrating diverse datasets and developing interpretable, reliable models are key steps to advancing AI/ML research and applications in water resources management (Kamyab et al. 2023).

The ability of AI and ML to support real-time monitoring and predictive analytics could drive transformative improvements in the responsiveness and adaptability of water management systems (Kamyab et al. 2023). Addressing these research priorities will enhance the sustainable management of vital water resources across various spatiotemporal scales, promoting greater precision, efficiency, and effectiveness through sustainability-driven practices (Chakraborty, 2024).

Accordingly, this paper aims to offer valuable insights for researchers, practitioners, and policymakers committed to safeguarding freshwater resources. By encouraging the adoption of sustainable practices, it seeks to support the long-term availability of clean water to meet the growing demands of an expanding global population.

1 Concepts of AI and ML in water resources

1.1 Definition and principles of AI and ML

Artificial Intelligence (AI), also known as machine intelligence, originates from the field of computer science. It encompasses three primary functions: learning (gathering information and understanding how to use it), reasoning (making decisions based on learned information), and self-correction (improving performance over time). AI is a broad term that refers to the simulation of human-like behavior by machines.

Machine Learning (ML), a subset of AI, enables computers to learn from data autonomously without being explicitly programmed (Sowmya and Sathisha, 2023). ML models are capable of processing large and complex datasets to extract patterns and make predictions. Common ML methods include Neural Networks, Decision Trees, Random Forest Classifiers, Regression, and Support Vector Machines (SVM) (Sarker, 2021).

ML is generally categorized into three main types: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves training a model on input data (known as Training Data) paired with known output (Labels). The model learns from this training data, makes predictions, and adjusts its parameters if the prediction are incorrect. Algorithms such as Linear Regression, Logistic Regression, SVM, Decision Trees/Forests, and Neural Networks fall under this category. These models have been widely applied to predict groundwater levels, water quality, and potential groundwater zones (Mahamat et al. 2021). Unsupervised learning deals with unlabelled input data. The system attempts to identify patterns or underlying structures without prior knowledge or labels (Malakar et al. 2021). Techniques like clustering (e.g., k-means, hierarchical clustering), Principal Component Analysis (PCA), factor analysis, and t-SNE, are widely utilized in data science for pattern recognition and dimensionality reduction. These models are especially useful in hydrogeological studies to classify regions based on similar features within a given area (Varouchakis et al. 2023). Reinforcement learning is a type of ML in which an agent learns how to behave within an environment by performing actions and observing the resulting rewards or penalties (Agyeman et al. 2023). The agent takes actions and receives feedback based on how effectively those actions influence the environment. Through this process, the agent learns the optimal sequence of actions that yields the highest cumulative reward. This learning is typically acquired empirically through direct interaction by acting, observing outcomes, and adjusting behaviour accordingly (Wang et al. 2018). Reinforcement learning has wide-ranging applications in water resource management. It is particularly useful for optimizing the allocation and equitable distribution of limited water resources. By simulating control strategies, it can identify those that enhance overall system efficiency (Kamyab et al. 2023) (Ding and Du, 2023). These algorithms learn the most effective water distribution and consumption strategies over time, ultimately enhancing operational efficiency and supporting sustainable practices (Kamyab et al. 2023). By leveraging reinforcement learning, water resource managers can make more informed, data-driven decisions, maximizing usage efficiency and contributing to the long-term sustainability of watershed systems.

1.2 Evolution and historical development of AI and ML in water resources

Significant advancements have been made in recent years in the application of AI and ML tools to water resource management. Traditionally, hydrological data prediction relied on basic statistical models, such as regression and autoregressive techniques for time series analysis (Kamyab et al. 2023). However, with increasing computational power and the growing availability of data, the field has evolved to incorporate more sophisticated methods, including ANN, Genetic Algorithms, Decision Trees, and other ML techniques (Xin and Mou, 2022). Unlike traditional statistical models, these AI and ML approaches are capable of processing large and complex datasets, uncovering intricate patterns and relationships that would otherwise remain hidden. These sophisticated tools allow water resources practitioners to model and understand dynamic system behaviours more accurately and even enable predictive capabilities. Consequently, they support more effectivel and optimized decision-making, surpassing what conventional approaches can offer (Xu et al. 2023).

The application of AI and ML in groundwater exploration and management began to emerge in the early 1990s. Supervised learning techniques, particularly ANN (Kamyab et al. 2023), were among the first tools employed for modelling groundwater levels and quality. Unsupervised learning methods, such as clustering, were implemented through tools like Natural Classes to delineate similarity zones, each characterized by homogeneous hydrogeological properties. Early ML applications in water resources typically involved menu-driven statistical analysis and regression models to predict standard hydrogeological parameters such as water stage, flow rate, and water quality (Shi et al. 2023; Varouchakis et al. 2023). Traditional regression models, including linear and polynomial regression, were widely applied to forecast groundwater table depth using predictor variables such as temperature and precipitation

(Kumar and Saini, 2021).

With advancements in data acquisition technologies, AI and ML models became capable of handling increasingly large and complex datasets using more sophisticated algorithms (Xin and Mou, 2022). Over the past decade, deep learning techniques have gained momentum in water resource management. The expansion of remote sensing and Geographic Information Systems (GIS) has enabled the integration of diverse data sources, offering a more comprehensive understanding of groundwater systems. For instance, combining remote sensing data with ground-based observations and historical records enhance the accuracy of groundwater assessments (Quan and Wang, 2020).

AI models such as Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) have been deployed for tasks including groundwater level prediction, water quality assessment, and irrigation optimization. These models, originally developed for image processing, have also been successfully adapted for classifying satellite imagery and identifying groundwater potential zones (Li et al. 2023). Furthermore, hybrids approaches that combine traditional hydrological models with AI/ML techniques have been explored to improve predictive accuracy and model robustness (Li et al. 2023).

1.3 Importance of AI and ML in hydrology and water resource management

In the field of hydrology, AI and ML have significantly transformed to the analysis of large and complex datasets. These tools enable the development of intelligent predictive models that can capture the inherent complexity of hydrological systems, for example, by forecasting groundwater recharge rates or detecting anomalies in water quality. They are also increasingly applied to solve large-scale optimization problems, such as improving the efficiency of water distribution networks (Shi et al. 2023). In groundwater exploration, AI models integrate meteorological, geological, and land use data to estimate recharge rates, offering more accurate and data-driven insights into subsurface water dynamics. These models can also analyze environmental factors correlated with recharge variabilities (Kamyab et al. 2023). Furthermore, ML algorithms can be utilized for early warning systems, detecting potential pollution or contamination events by identifying patterns and trends in time series water quality data (Sahour et al. 2023). Such predictive capabilities are critical for maintaining sustainable water management, particularly in regions prone to drought or historically burdened by poor quality.

AI and ML applications in water resource management extend far beyond groundwater estimation. They play a pivotal role in optimizing distribution systems, modelling river basins, and scheduling irrigation. Predictive models developed from historical and real-time data, for example, those forecasting water demand or detecting system anomalies (Feng et al. 2024), are key for supporting economic activities. These tools not only enable efficient and sustainable use of limited water resources but also contribute to broader environmental sustainability goals. One of the key advantages of AI and ML solutions is their adaptability across diverse geographic and climatic regions, making them ideal tools for promoting sustainable water management practices.

When applied appropriately, AI- and ML-based technologies empower water resource managers to make informed, data-driven decisions, enhancing water use efficiency and promoting the long-term sustainability of available resources (Feng et al. 2024) (Kamyab et al. 2023). However, despite their potential, the application of AI and ML in groundwater exploration and management remains underutilized, primarily due to issues related to data quality and availability. Moreover, hydrological systems vary significantly across regions, and addressing spatial variabilities requires high-quality input data to ensure robust model training and performance (Varouchakis et al. 2023).

Advancements are still needed in data integration methods, model accuracy, and interpretability, along with cross-disciplinary collaborations to unlock the full potential of AI and ML in water resource management. Enhancing real-time monitoring and responsiveness could further transform the dynamic adaptability of water systems.

In general, the application of AI and ML technologies suggests a promising shift towards more accurate, efficient, and environmentally sustainable water management practices. Fig. 1 highlights key research areas and emerging trends in AI applications within hydrology (Biazar et al. 2025), showing a marked increase in AI utilization over recent years. For example, Abu et al. (2024) investigated groundwater characterization and quality forecasting at the boundary between the Savanna region and Upper West region of Ghana using multivariate statistics and ML techniques. They evaluated model performance based on Mean Absolute Error, Mean Squared Error, and Root

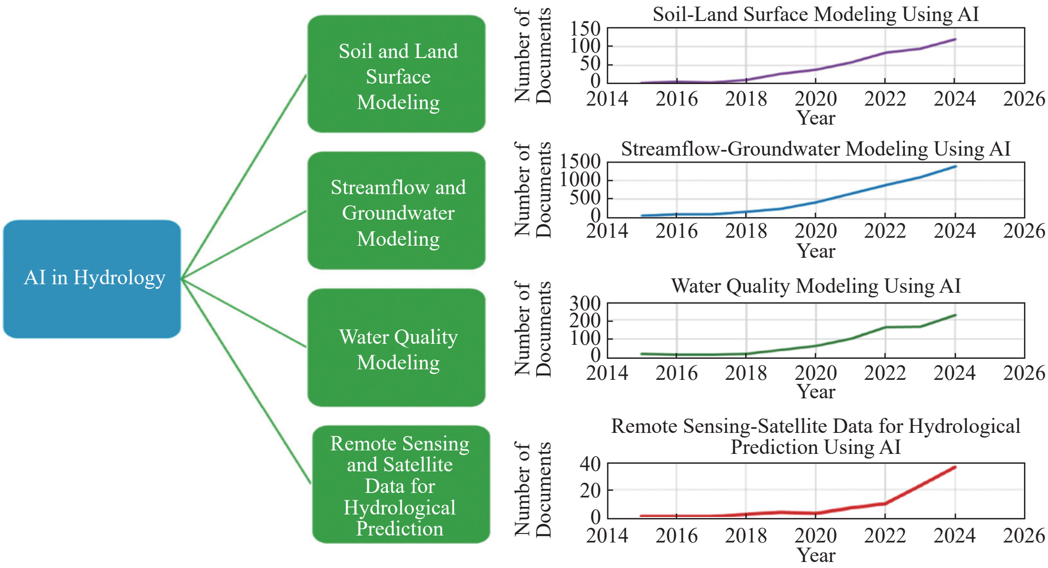

Fig. 1 Key research areas in the application of AI in Hydrology (Biazar et al. 2025)

Mean Squared Error. Their results ranked the models for Water Quality Index forecasting in the following order: Artificial Neural Network (ANN) >Random Forest Regressor >Decision Tree Regressor >Multiple Linear Regression.

Fig. 2 shows the deployment of various AI models for groundwater level prediction from 2001 to 2023 (Pourmorad et al. 2024). In a study by Mohammad et al. (2020), ANN was applied to predict groundwater level in the Diyala river basin, Iraq, with the model effectively identifying optimal well-drilling locations to maximize groundwater availability. Similarly, Bui et al. (2017) affirmed that ANN models provide reliable and highly accurate groundwater level forecasts.

Moreover, research evaluating ML techniques for predicting groundwater fluctuations in arid and semi-arid regions using data from the Gravity Recovery and Climate Experiment (GRACE) satellite mission highlighted the potential of Decision Tree models. When combined with GRACE data, Decision Trees accurately captured the complex relationships between GRACE data and groundwater dynamics, offering reliable predictions and valuable insights to support sustainable groundwater management strategies (Eftekhari and Khashei-Suiki, 2025).

2 Methods of AI and ML in groundwater exploration

2.1 Data collection techniques

Data collection techniques have significantly advanced to accommodate the complex three-dimensional nature of groundwater systems, encompassing both traditional field-based meth-

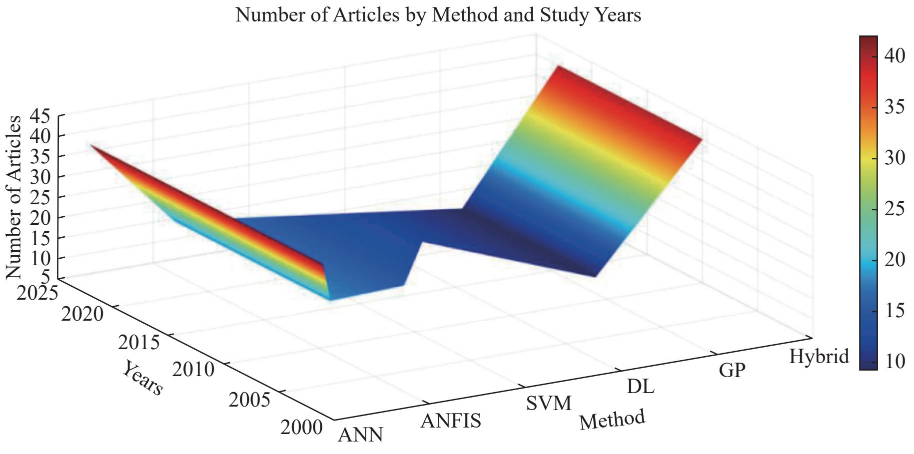

Fig. 2 Application of various AI methods for groundwater level prediction from 2001 to 2023 (Pourmorad et al. 2024). Note: ANN – Artificial Neural Network, ANFIS - Adaptive Neuro-Fuzzy Inference Systems, SVM – Support vector Machines, DL- Deep Learning, GP- Genetic Programming

ods and modern remote sensing approaches. Traditional hydrogeological data collection methods include field surveys, well logs, measurements of water table levels, discharge rates, and other in-situ observations. These point-based methods provide detailed insights into the geological characteristics and hydraulic behaviour of aquifers. The growing use of remotely sensed data from satellite images and airborne geophysical surveys has become an important complement to traditional field-based groundwater exploration techniques.

Remote sensing technologies are increasingly favoured in groundwater exploration due to their extensive spatial coverage and ability to provide frequent, repeatable observations, capabilities that are difficult to replicate through conventional means. These technologies offer synoptic, computer-aided assessments of both surface and subsurface conditions, serving as powerful tools for rapid, regional-scale mapping at the landscape level. Methods such as Synthetic Aperture Radar (SAR) and Light Detection and Ranging (LiDAR) enable the detection of geological structures and soil characteristics that directly influence groundwater potential (Song and Jung, 2023). Collectively, these remote sensing approaches provide a comprehensive view of surface and subsurface conditions that would otherwise be extremely difficult, if not impossible, for hydrologists and water resource managers to obtain through conventional field-based methods alone.

High-resolution satellite imagery, such as from Landsat and Sentinel, provide valuable insights into land surface dynamics that can assist in identifying potential groundwater recharge zones and monitoring the effects of land use changes on aquifers. When integrated into GIS to produce thematic maps, these remote sensing datasets offer a powerful means of delineating areas characterized by greater aquifer depth and thickness at relatively shallow depths, which are key indicators of groundwater potential (Derdour et al. 2022).

2.2 Geospatial Data Integration and Analysis

The delineation of groundwater potential zones using geospatial technology requires the integration of data obtained through various methods, including field surveys and remote sensing, particularly at the catchment scale. GIS serve as essential tools for integrating these diverse datasets and conducting multi-dimensional spatial analyses to identify correlations with hydrological phenomena, an important step in groundwater exploration. Incorporating large-scale models with geospatial layers such as geology, hydrogeology, topography, and land use and land cover can significantly enhance the prediction of groundwater availability and quality.

The diagnostic capabilities of large-scale data derived from remote sensing technologies such as SAR and LiDAR significantly enhance groundwater exploration mapping. These methods enable the identification of geological structures, soil moisture levels, and vegetation patterns associated with underlying groundwater systems. SAR is particularly effective in detecting subtle surface deformations, while LiDAR provides high-resolution, three-dimensional surface data. Together, they contribute to a comprehensive understanding of subsurface control mechanisms that influence regional groundwater flow and yield. When integrated into GIS, these datasets facilitate the creation of detailed thematic maps, which are instrumental for identifying groundwater potential zones as regional scales (Dabas et al. 2022).

Satellite platforms such as Landsat and Sentinel offer high-resolution imagery that captures land surface dynamics over time, including the distribution and variation of surface water bodies such as lakes, rivers, and wetlands. This information not only supports surface water assessments but also aids in identifying favourable zones for groundwater recharge, as well as detecting human-induced impacts on aquifer systems (Cheng et al. 2022). Mapping potential recharge zones through changes in land use further strengthens the hydrogeological conceptual models needed for sustainable groundwater management.

The need to monitor continuous changes in groundwater levels over time has prompted researchers to develop remote sensing-based empirical methods for estimating the total volume of water flowing through hydrological systems. The use of very high-resolution imagery for thematic mapping in GIS has facilitated the interpretation and evaluation of groundwater exploration areas, even at considerable depths (Pandey et al. 2020). Regional-scale maps derived from remote sensing data offer improved accuracy in identifying groundwater origin points and estimating aquifer depths, thus aiding satellite-based hydrogeological assessments (Derdour et al. 2022).

Further, assessing groundwater contamination through indicators such as surface temperature anomalies, vegetation stress, and soil moisture is critically important. These factors can be effectively monitored using remote sensing technologies designed for environmental assessment and monitoring (Surinaidu et al. 2021). This process involves detecting abnormal patterns in features derived from connectivity and spatial relationships, an area where ML models are well suited (Surinaidu et al. 2021). In addition to point-source contamination detection, thermal infrared (TIR) imagery provides valuable insights by identifying thermal anomalies on the surface, which often correspond to groundwater contamination hotspots. TIR technology is more sensitive than conventional remote sensing methods and can detect even subtle changes in surface temperature, making it a powerful tool for the ongoing monitoring and early forecasting of groundwater contamination events.

The maps provide water resource managers with the ability to zoom into specific areas of interest, locate surface water bodies by name or features (like rivers), and identify potential pollutant sources along with their associated impacts. This functionality supports public health practitioners in analyzing environmental interventions and can help reduce costs in protecting environmental resources, thereby positively affecting human welfare. Early detection during groundwater harvesting is critical for the sustainable conservation of water resources (Dahan, 2020).

ML has been a breakthrough in this field, dramatically enhancing prediction accuracy, especially in groundwater potential mapping through the use of CNNs. Deep learning algorithms, such as spatial CNNs, can process large-scale satellite imagery to learn spatial patterns from single or multiple scenes. Researchers use full-image classification and attribution techniques to map optimal sites for water storage, including outputs tailored to stakeholder needs. The fully automated and efficient nature of CNN-based models has made them computationally valuable tools for groundwater potential mapping, providing crucial insights for designing targeted exploration plans that support sustainable water resources management (Liu et al. 2022). Remote sensing combined with ML has proven highly effective in identifying promising groundwater prospects (Jaafarzadeh et al. 2021).

Geophysical surveys employ a variety of non-invasive techniques, such as Electrical Resistivity Imaging (ERI), Ground-Penetrating Radar (GPR), and Seismic Refraction, to generate comprehensive images of subsurface conditions that are often inaccessible through conventional methods. ERI has become particularly important in groundwater exploration, as it provides real-time measurements of the electrical resistivity of subsurface layers. Since water-saturated layers typically exhibit lower resistivity than dry materials, ERI effectively aids in identifying aquifers. This technique delivers precise information on the distribution and depth of groundwater-bearing formations, which is crucial for targeted exploration and sustainable management (Alghamdi et al. 2023).

GPR uses high-frequency electromagnetic waves emitted into the ground, with antennas measuring the returning signals. This method is valuable for detecting water-saturated layers, bedrock configurations, and other underground structures such as voids, pipes, and soil horizons. Being non-invasive, GPR provides detailed information essential for aquifer identification and mapping (Abdullah et al. 2022).

Seismic refraction involves measuring the travel time of artificially induced seismic waves through the subsurface. By analysing propagation paths and reflections of these waves at subsurface boundaries, high-resolution images of the ground can be produced. This technique is well suited for precisely locating the water table and understanding the configuration of aquifer systems. It supports the creation of 3D, high-resolution images of water-bearing layers, thereby enhancing groundwater exploration and management efforts (Zhu et al. 2023).

The integration of geophysical methods with remote sensing and GIS tools significantly improves the resolution, efficiency, and reliability of subsurface surveys, yielding more reliable results. Moreover, the application of AI and ML to large geophysical datasets, commonly referred to as Big GeoData, has considerably enhanced interpretation accuracy. This advancement is particularly critical in groundwater exploration.

2.3 Data processing and preprocessing methods

Fast and consistent data processing solutions are essential for the effective management of groundwater resources, particularly through the intelligent extraction of valuable information from diverse geospatial datasets, including large volumes of ground-based and remote sensing data. In the application of AI and ML to groundwater exploration, two of the most crucial steps are data processing and preprocessing. Data preparation involves transforming and unifying raw data into a format suitable for further analysis. This includes tasks such as handling missing values, scaling input variables, and engineering relevant features for machine learning models (Ahmed et al. 2023).

Effective data preprocessing improves the performance and accuracy of AI/ML models, thereby enhancing the reliability of predictions related to groundwater availability and quality. Key preprocessing steps include imputing missing values, normalizing data and selecting meaningful features. Techniques such as data imputation, interpolation, and normalization ensure consistent scaling of variables, which contributes to better model performance and more accurate estimates.

To achieve accurate predictions and reliable results in groundwater investigation and analysis, AI/ML models should focus on essential data preprocessing steps. These include handling missing values through techniques such as mean or median imputation, as well as more advanced approaches like multiple imputations; scaling input variables using standardization; and selecting only the most influential features affecting groundwater dynamics to improve both model interpretability and performance (Sahour et al. 2023; Varouchakis et al. 2023). Environmental predictors are typically processed using mathematical transformations to ensure balanced representation, such as through expression space normalization (Appling et al. 2022). Feature selection and extraction techniques identify the most relevant variables for modeling, thereby enhancing model robustness while maintaining transparency in human decision-making processes (Suwadi et al. 2022).

Advanced data fusion techniques, integrating satellite imagery with various geophysical surveys and in-situ measurements, are essential for effective groundwater exploration. These comprehensive datasets provide detailed information on subsurface strata, enabling more accurate identification of potential groundwater zones. The combination of multiple spatially explicit data sources deepens our understanding of complex hydrogeological systems, leading to improved groundwater management and exploration practices (Varouchakis et al. 2023; Li et al. 2021). For example, remote sensing data (e.g., satellite images or aerial photographs), when combined with field measurements from wells and ground-based sensors, provides multi-dimensional insights into groundwater systems by linking present observations with historical records and archives.

Integrating datasets across different spatial, temporal and thematic dimensions enhances our understanding of regional hydrogeological conditions, crucial for both researchers and water resource managers. Therefore, robust data integration and synchronization from diverse sources are crucial for validation, cross-verification and corroboration of results, ultimately strengthening scientific knowledge of groundwater systems. This integrated approach serves as a reliable and effective tool for developing targeted groundwater exploration strategies and advancing sustainable water resource management (Derdour et al. 2022).

2.4 AI and ML algorithms used

2.4.1 Supervised learning

Supervised learning refers to the training of models using labelled datasets, enabling them to make predictions or classifications based on input-output mappings. Common supervised learning algorithms include regression models, decision trees, and SVM. These methods have been extensively applied in groundwater studies, particularly for predicting groundwater yield, estimating water table depths, and assessing aqufier productivity (Appling et al. 2022).

In classification tasks, decision trees and SVMs are widely favoured due to their ability to model data relationships and provide interpretable outcomes based on attribute-based decision boundaries (Nalevanková et al. 2023; Appling et al. 2022). Regression techniques, both linear and non-linear (e.g., quadratic regression), are used to predict variables such as Wavelet Thresholding Denoising (WTD) in relation to climatic parameters. These models effectively integrate diverse data types, including soil moisture and hydraulic head, thereby enhancing the predictive capabilities crucial for informed water resource management (Nalevanková et al. 2023).

Decision trees are typically valuable for delineating groundwater potential zones, providing transparent and interpretable results. They can also be combined with other machine learning techniques to improve overall model accuracy/robustness. SVMs have been effectively employed in groundwater system modelling, particularly where the input features include complex variables like soil properties, climatic conditions, and other factors influencing hydrological behaviours (Sun et al. 2022; Varouchakis et al. 2023). In many groundwater detection applications, input data often consists of diverse types, whether explicitly numerical or transformed from categorical forms, resulting in high or even infinite-dimensional feature spaces. In such cases, SVM excels due to their capacity to handle high-dimensional data.

2.4.2 Unsupervised learning

Unsupervised learning is employed to extract meaningful insights from unlabeled datasets by identifying hidden patterns or underlying structures within the data. Clustering algorithms, such as k-means and hierarchical clustering, are among the most widely used techniques in this context they group data points based on similar properties. These clustering approaches enable the delineation of groundwater potential zones and regional segregation based on hydrostratigraphic similarity. Such classification is essential for accurately characterizing complex groundwater systems and understanding their spatial variability. The unsupervised classification of discrete groundwater zones also facilitates region-wide exploration, monitoring, and management strategies, thereby enhancing the efficient utilization of water resources (Sahour et al. 2023; Varouchakis et al. 2023).

Additionally, dimensionality reduction techniques such as PCA and t-SNE are instrumental in addressing the heterogeneity of groundwater datasets. PCA reduces the dimensionality of multivariate data by transforming the original variables into a smaller set of principal components that remain most of the variance, thereby simplifying model inputs while preserving key structural relationships. On the other hand, t-SNE is a nonlinear method that maps high-dimensional data to a lower-dimensional space for visualization, effectively revealing subtle clustering patterns that might not be apparent in higher dimensions. By preserving local data relationships, t-SNE excels at uncovering intricate groupings and interdependencies. Together, these dimensionality reduction methods enhance the efficiency, interpretability and predictive performance ML models for groundwater exploration, particularly in identifying key drivers and simplifying complex feature spaces (Sahour et al. 2023; He et al. 2022).

Various clusters of groundwater or hydrogeological settings can be identified using clustering methods such as k-means and hierarchical clustering, which group data points based on similarity among their attributes (Agyemang, 2022; Varouchakis et al. 2023). These techniques help uncover hidden patterns in complex datasets. Several dimensionality reduction algorithms, such as PCA or t-SNE, are often applied to simplify complex groundwater datasets. These approaches enhance model performance by filtering out noise and retaining only the most informative features, thereby improving interpretability and reducing computational cost. In the context of groundwater exploration, these methods facilitate the analysis of large, multivariate datasets where geological complexity or data sparsity would otherwise hinder pattern recognition and classification.

Furthermore, Hierarchical Cluster Analyses (HCA) and PCA have been employed to assess groundwater quality, contributing to the development of more effective monitoring systems. HCA groups samples with similar hydrodynamic and physicochemical characteristics, allowing for the classification of distinct groundwater types or zones. In contrast, PCA identifies the key variables driving observed patterns in water quality, helping to highlight the most influential parameters (Agyemang, 2022). Together, these unsupervised learning methods provide valuable insights into the intricate dynamics of groundwater systems and simplify the analysis of large, multivariate datasets. This supports water managers in designing cost-effective and site-specific monitoring programs. The combination of HCA and PCA not only aids in classifying groundwater samples but also improves understanding of water quality control points (Abdelaziz et al. 2020), Enabling more targeted protection strategies and sustainable resource use.

2.4.3 Deep learning

As a subset of ML and AI, deep learning uses multi-layered ANNs to capture complex patterns and representations in data. By stacking multiple layers, deep learning models can extract increasingly abstract and high-level features, enabling superior performance in tasks such as image recognition, natural language processing, and predictive analytics. Their ability to handle non-linear, multidimensional relationships allow them to outperform traditional ML algorithms, especially in complex datasets (Chrysos et al. 2021).

These techniques have been widely used in groundwater exploration for tasks such as image classification, time series forecasting, and anomaly detection. CNNs are particularly effective for analyzing satellite imagery. They can automatically extract relevant features, such as vegetation patterns, surface temperature, and soil moisture, that are indicative of groundwater availability. CNNs are increasingly used to identify potential groundwater zones, improving accuracy in groundwater prospecting by leveraginghigh-resolution satellite images (Wang et al. 2022).

RNNs, particularly Long Short-Term Memory (LSTM) architectures are well suited for modelling temporal dependencies in hydrological time series. These models have demonstrated strong performance in forecasting groundwater levels and streamflow dynamics, enabling more efficient irrigation scheduling and water usage planning. Accurate LSTM-based forecasts support long-term sustainability of water resources (Shrestha and Pradhanang, 2023), especially in water-scarce regions.

The advancement of groundwater exploration research is accelerating with the integration of deep learning and ML algorithms applied to diverse datasets. Hybrid models, such as those combining CNNs with conventional ML techniques, have shown improved accuracy in generating groundwater potential maps (Pandey et al. 2014). These hybrid approaches leverage the complementary strengths of each method. Furthermore, incorporating deep learning into real-time monitoring and predictive analytics holds great promise for enhancing groundwater management by improving decision-making, increasing water-use efficiency, and ultimately contribute to more sustainable resource utilization.

Predicting groundwater levels and quality using deep learning models in real-time allows water managers to respond within actionable timeframes, enabling the implementation of adaptive response strategies. Deep learning techniques can automate groundwater monitoring by processing data from sensor networks and acquisition systems, thereby generating continuous, real-time feedback on the state of water resources. This integration facilitates data-driven and timely decision-making, helping to optimize groundwater utilization, prevent depletion, and support the long-term sustainability of water resources (Mokua et al. 2021).

Groundwater exploration typically involves a sequence of data collection, information extraction, and application of various modelling approaches, including AI and ML techniques. Data collection primarily relies on remote sensing and geophysical surveys to assess subsurface characteristics. However, the reliability and performance of AI and ML models are critically dependent on advanced preprocessing techniques. These preprocessing steps are essential to ensure data quality, consistency, and relevance for model training. When effectively implemented, AI and ML approaches can significantly enhance groundwater exploration efforts, support more sustainable resource management and aligning with regulatory standards and ministry quidelines (Sahour et al. 2023).

3 Applications of AI and ML in groundwater exploration

3.1 Identification of groundwater potential zones

The identification of groundwater potential zones is a critical step in sustainable groundwater resource management. With the advancement of AI and ML technologies, this process has become significantly more accurate and efficient. Traditional approaches relied heavily on manual techniques and fieldwork, where spatial distribution of features were interpreted from physical samples, borehole data and existing maps. These methods were labour-intensive and often constrained by limited spatial coverage and interpretative subjectivity.

The integration of remote sensing and geophysical surveys with AI and ML algorithms has revolutionized this process. Machine learning models can now process large-scale hydro-geospatial datasets to identify and predict groundwater potential zones with high precision (Sahour et al. 2023). Numerous studies have employed classical ML algorithms such as decision trees, random forests, and SVM to classify and predict groundwater potential areas at both regional and global scales, often compensating for the scarcity or inconsistency of input data. For example, layers representing soil type, land use/land cover, precipitation, and elevation (for topography) have been used to train random forest models to generate groundwater potential maps. These models have demonstrated high predictive accuracy and frequently outperform traditional empirical methods (Jaafarzadeh et al. 2021).

Additionally, deep learning models have been increasingly applied to process remote sensing images and extract meaningful features indicative of groundwater potential. CNNs, in particular, are highly effective at learning spatial patterns from satellite imagery, such as vegetation health status, surface temperature, and soil moisture, factors closely correlated with groundwater availability (Wang et al. 2022). These models are typically trained on large datasets, which helps to minimise bias and improve adaptability, thereby enhancing prediction accuracy across diverse geographic regions. The ability to integrate deep learning with publicly available datasets, such as remote sensing imagery, geophysical surveys, and hydrogeological information, has allowed for the development of more advanced and scalable groundwater detection systems (Clark et al. 2022). Data fusion techniques, which combine remote sensing data with outcomes from methods like Electrical Resistivity Imaging (ERI) and Ground-Penetrating Radar (GPR), have proven particularly valuable. These integrated approaches provide richer insights into subsurface conditions, such as soil structure and aquifer accessibility, thereby improving the reliability of groundwater potential assessments (Ding et al. 2020). Prasad et al. (2020) conducted a comparative study using three ML models, namely, random forest, boosted regression tree and random forest assisted SVM. Their results demonstrated that the random forest model achieved the highest accuracy (94%), followed closely by the random forest assisted SVM (93.4%), while the boosted regression tree model yielded slightly lower accuracy (89.8%).

3.2 Groundwater quality prediction

Groundwater quality is critical not only for ensuring safe drinking water but also for maintaining overall ecosystem health. A growing number of studies have demonstrated the successful application of AI & ML techniques to predict key groundwater quality parameters, including pH, TDS, and contaminant concentrations. These models are particularly effective in capturing complex, nonlinear relationships between multiple variables by using both historical dataset and real-time monitoring data to deliver highly precise forecasts.

Common supervised learning algorithms used in groundwater quality prediction include linear regression, decision trees, and neural networks. These models can be trained on historical water quality records to identify patterns and trends, thereby improving their ability to forecast future contaminant levels. For example, neural networks have shown strong performance in predicting nitrate concentrations in groundwater, particularly due to their ability to model complex, nonlinear relationships between environmental inputs and water quality outcomes (Xin and Mou, 2022).

Ensemble learning techniques, which combine multiple base models to improve overall prediction performance, have also gain popularity in groundwater quality assessments. Algorithms such as random forests and gradient boosting combine the predictive capabilities of multiple learners to produce more robust and accurate models. These approaches have demonstrated superior performance compared to individual models in various groundwater quality prediction studies (Sahour et al. 2023).

Time series prediction of groundwater quality parameters has been effectively achieved using LSTM-based deep learning models. LSTM networks are designed to retain and learn from sequential data, making them particularly suitable for capturing temporal trends and fluctuations in water quality over time. For instance, LSTM models have been applied to forecast salinity levels in coastal aquifers, thereby supporting informed decision-making in water resource management and aiding the development of more sustainable practices in vulnerable coastal regions (Xu et al. 2023).

3.3 Groundwater level forecasting

Groundwater level forecasting plays a crucial role in effective water resources management and planning. AI and ML-based models have become increasingly popular for predicting groundwater levels due to their ability to process complex and nonlinear relationships among diverse input variables, include precipitation, temperature, land use, and historical water levels. These models have demonstrated strong performance in generating reliable forecasts, thereby enabling water managers to make informed decisions on resource allocation, land conservation strategies, and long-term sustainability planning.

Supervised machine learning algorithms, such as linear regression, SVM, and decision rules, are commonly used to predict groundwater levels. These models are trained to identify patterns from input variables and use the pattern to predict future levels. Recent studies also suggested that such models perform particularly well when optimized with proper configurations and enhanced through preprocessing techniques such as normalization and feature selection. These advancements contribute to improved prediction accuracy.

Deep learning models, particularly RNNs and their variants such as LSTM networks, have also shown strong potential in groundwater level forecasting. These models are designed to capture temporal dependencies in sequential data, making them well-suited for time series prediction. LSTM networks, in particular, can model long-term relationships and complex nonlinear interactions among input variables, which have been proved through various studies (Guo et al. 2023).

Several hybrid models combining AI/ML methods with conventional hydrological models have been developed to enhance groundwater level forecasting. These hybrid models integrate the data-driven strength of AI/ML with the physics-based hydrological models, resulting in more accurate and robust predictions, thereby improving forecasting performance and deepening our understanding of groundwater system dynamics (Xu et al. 2023).

4 AI and ML in water resource management

4.1 AI Applications in water distribution systems

Water distribution systems are critical infrastructures that supply potable water for municipal, domestic and industrial uses. Due to their complexity and need for high operational efficiency, they are well-suited for an AI or ML solutions. Municipalities around the world are increasingly applying AI/ML technologies to optimize water delivery systems, particularly for tasks such as leak detection and demand forecasting, ultimately enhancing operational performance (Kamyab et al. 2023). Leakage in water distribution system is a critical problem, conventionally addressed through physical inspections and acoustic sensors, which are often costly and time-consuming. Modern systems rely on pressure sensors, flow meters, and acoustic devices to monitor anomalies. AI and ML algorithms can analyze the data collected from these devices to enhance leak detection and monitoring (Choudhary et al. 2021). Various ML models, including SVMs, neural networks, and clustering algorithms, have been developed to identify specific patterns associated with leaks (Vahldiek et al. 2022). These models assist utilities in systematically managing water loss and optimizing operations to enhance the overall efficiency of water distribution systems.

AI and ML technologies have also enhanced demand forecasting in water distribution systems. Water utilities can tailor their operations and predictions to meet demand on an hourly basis by accurately predicting water usage, thereby optimizing supply with respect to cost and energy efficiency (Kamyab et al. 2023). State-of-the-art univariate statistical models such as Autoregressive Integrated Moving Average (ARIMA), neural network, and ensemble learning methods can predict water demands based on historical usage data. In addition, weather-related factors and social elements play a significant role in this predictive endeavour (Bata et al. 2020). These models allow utilities to anticipate both short-term and long-term demand, thereby supporting effective resource planning and infrastructure development.

Furthermore, optimization algorithms for pump scheduling, pressure management, and energy utilization, as part of AI and ML frameworks, enable operational efficiency in water distribution systems. Models developed using genetic algorithms and particle swarm optimization aim to minimize energy consumption while maintaining adequate pressure levels across the network (Kamyab et al. 2023). These solutions are designed to be cost-effective and easy-to-implement, providing utilities with a reliable and comprehensive approach to water system management.

4.2 Water resource allocation optimization

Water is an important commodity, essential for efficient distribution across agriculture, industry and domestic sectors. AI/ML-based tools have been developed to optimize water allocation by considering factors such as availability (Kamyab et al. 2023), and minimizing the environmental impact of allocation patterns. In general, water resource allocation optimization involves balancing conflicting objectives, making multi-objective algorithms suitable for solving such problems. These algorithms aim to balance supply accuracy and efficiency with reduced costs and lower environmental footprints. Models for determining optimal allocation strategies based on various input parameters have been developed using techniques such as genetic algorithms, simulated annealing, and multi-objective particle swarm optimization (Wang et al. 2022).

In addition, to develop long-term predictive models that support the design of water resource allocation strategies, both unsupervised and supervised ML techniques have been enforced. Many neural network- and SVM based models incorporate sector-specific information, such as residential vs. industrial use, population growth, economic activity, and climate conditions, to explain water demand. These models assist water managers in making informed decisions on resource allocation, addressing supply-demand imbalances, and reducing waste generation (Mustafa et al. 2021).

AI and ML technologies are also used to develop decision support systems for water resource allocation. By integrating data from satellites, in situ observations, and historical records, researchers can quantify water availability relative to demand. Optimization algorithms and predictive models are then applied to assess both water allocation options and policy impacts. These tools help identify opportunities for informed decision-making, particularly in sectors where resource use has minimal environmental impact.

4.3 Case studies and practical applications

The potential of AI and ML technologies has been explored through various practical applications in water resource management, demonstrating added advantage in accuracy, decision-making, and sustainability. Numerous real-world case studies show how these technologies have been successfully employed across diverse sectors.

In urban water management, AI/ML models have optimized drinking water distribution networks, improved leak detection in aging pipelines, and predicted future demands. For instance: in Barcelona, Spain, an ML algorithm connected to pressure sensors and flow meters successfully reduced water loss by 20% (Daniel and Cominola, 2023). Similarly, in Singapore, neural networks were applied to predict water demand using historical usage and weather data, helping utility companies achieve greater energy efficiency and operational optimization (Hsia et al. 2021).

In agriculture water management, AI/ML technologies have improved irrigation scheduling, crop water requirements predictions, and applied water efficiency. One approach involved a ML-based system that triangulated real-time soil moisture data and crop needs to reduce water use significantly (Ding and Du, 2023). A recent study in India leveraged remotely sensed data and neural networks to estimate peak crop water requirements, improving irrigation management and Water Use Efficiency (WUE).

In flood forecasting, countries like the Netherlands have used AI/ML-based models combing data from river gauges, weather stations, and remote sensing systems to deliver flood predictions up to three days in advance, enabling timely evacuations (Goh, 2021). A project (Global South East Asia project) in Southeast Asia showed the ability of ML models to predict major flood events and mitigate risks for vulnerable populations (Huynh and Kiang, 2023). In Bangladesh, a suite of AI algorithms, trained on historical data and current observations, has been integrated into local government strategies to manage floods more effectively (Kamyab et al. 2023).

ML has also been used to predict nitrate concentrations in groundwater by analyzing agricultural land use, soil types, and precipitation data by using neural networks (Lara et al. 2023) for better water quality management. Another study applied a random forest model to analyze partial heavy metal concentrations and derived geochemical properties predict contamination sources and guide remediation strategies (Liu et al. 2022; Yu and Yang, 2021).

Case studies like these illustrate the wide-ranging applicability of AI and ML technologies in advancing water resource management across contexts. By leveraging big data, predictive analytics, and advanced modelling techniques, water managers can make more informed decisions, enhance operational efficiency, and alleviate pressure on natural resources, driving transformational change toward sustainable water conservation and management.

The findings of this study offer significant practical applications in groundwater exploration and water resource management, particularly in addressing challenges related to data scarcity, prediction accuracy, and sustainable water distribution. The integration of AI and ML techniques enhances the ability to analyze vast datasets, identify groundwater potential zones, predict water quality variations, and forecast groundwater levels with improved precision. One of the key applications of this research lies in AI-driven groundwater potential mapping, where machine learning models can delineate high-yield aquifers more effectively than traditional methods. This can assist policymakers and hydrologists in optimizing groundwater extraction and preventing overexploitation. Additionally, real-time water quality prediction using machine learning models can aid in the early detection of contamination, ensuring safer drinking water supplies and supporting environmental conservation efforts. Furthermore, AI-based water resource allocation models can optimize the distribution of available water based on demand, climatic conditions, and hydrological constraints. This is particularly valuable in regions facing water scarcity, where efficient management is crucial for agriculture, industry, and domestic use. By leveraging predictive analytics, water authorities can implement proactive measures to mitigate drought impacts, reduce wastage, and enhance overall water sustainability.

5 Challenges and limitations

5.1 Data quality and availability issues

Data quality and availability represent some of the most significant challenges in applying AI and ML to groundwater exploration and management. These models rely heavily on high-quality, comprehensive datasets to produce accurate and reliable results. However, such data is often unavailable thorough local web-based services, particularly in developing countries, where information on critical domains like health service access and infrastructure is frequently incomplete, inconsistent, or outdated (Agweyu et al. 2023).

Moreover, data sources are highly heterogeneous. Groundwater data may be derived from remote sensing, geophysical surveys, and in-situ measurements, each with varying spatial and temporal resolutions, formats, and accuracy levels. Integrating such diverse datasets is complex and demands robust data processing techniques to ensure uniformity and reliability (Aldoseri et al. 2023).

Another key issue is the lack of long-term data. Groundwater systems are influenced by climatic, geological, and anthropogenic factors over extended periods. Short-term datasets may fail to capture these trends, leading to inaccurate or biased model predictions. Hence, continuous monitoring and long-term data collection are essential for enhancing the performance and resilience of AI/ML models (Chakraborty, 2024).

5.2 Model accuracy and reliability

Developing specialized AI/ML models for groundwater and water resource management hosts great promise, but optimizing their performance remains a key challenge. A primary concern is overfitting, where model perform well on training data but fail to generalize to unseen data. This issue can be addressed using techniques such as cross-validation and regularization, while ensemble learning has also proven effective in improving model robustness during training (Bradshaw et al. 2023).

Another key challenge is interpretability. Many AI/ML models, especially those based on deep learning, are often regarded as "black boxes", making it difficult to understand how or why they produce certain predictions. This lack of transparency can limit their adoption, particularly in water resource management, where stakeholders not only require high predictive accuracy but also the ability to trust and interpret model outputs. In response, efforts such as eXplainable AI (eXAI) are being used to develop more interpretable models (Kobayashi and Alam, 2023).

Moreover, the performance of AI/ML models is closely tied to the quality and relevance of the data used during training. Models developed for one region often cannot be directly applied to another due to differences in climate, geospatial characteristics and socio-economic conditions. This necessitates the need for region-specific models, which must also be routinely updated with new data to maintain predictive accuracy (Fan et al. 2022).

5.3 Integration with existing water management systems

Overcoming the challenge of technical and infrastructural bottlenecks that emerge when AI/ML technologies are integrated into water management systems presents strong opportunities for collaboration between emerging technology startups and industrial conglomerates. Most existing water management systems rely on conventional approaches and are not readily compatible with new AI/ML solutions. These systems require rubust infrastructure, skilled personnel, and ongoing maintenance, which may act as barriers, especially in resource-limited settings (Rathor and Kumari, 2021). One significant challenge is the fragmentation of data sources. Water resource management practices aim to provide integrated data from diverse sources, including remote sensing, geophysical surveys, and in-situ measurements. These datasets often vary widely in formats, scales, and sensor accuracy, requiring re-integration into an unified AI/ML model. Standardizing data formats and protocols can help address this issue, although challenges remain in scaling these solutions globally (Behar et al. 2023).

Moreover, effectively integrating AI and ML requires interdisciplinary collaboration in the field of water management. This calls for integrated teams of hydrologists, data scientists, and engineers to co-develop AI/ML solutions aligned with policymaker needs. Achieving this requires substantial communication and coordination across disciplines, which can be challenging due to differing expertise, terminologies, and perspectives (Zowghi and Rimini, 2023).

5.4 Ethical and regulatory aspects

Ethical and regulatory issues are increasingly coming to the forefront as stakeholders in water resource management begin to integrate AI/ML into their operations. With critical concerns such as water quality data, user consumption patterns, and infrastructure details, often handled by private companies that collect, analyse, and operate on this information, data privacy becomes a major concern. Since this data is highly sensitive, safeguarding it is essential to ensure privacy and prevent misuse (Slavković and Seeman, 2023).

Although these considerations are general, some fundamental ethical principles for best practice in the field can be deduced (Siva Barathi et al. 2024). Transparency and accountability are key: Decision-makers should have at least a high-level understanding of how AI/ML models function and arrive at their predictions. This underscores the need for explainable AI and transparent reporting. In addition, there must be mechanisms to attribute responsibility for AI/ML model outputs to the developers and users of these systems (Ryan et al. 2024).

The powerful capabilities of AI and ML technologies also demand updated regulations that can ensure those tolls are used responsibly and effectively for public benefit. Current regulatory frameworks may fall short in addressing challenges such as data ownership, algorithmic bias, and automated decision-making. Policymakers are therefore urged to introduce new legislation that both safeguards the ethical use of AI/ML in water resource management and supports ongoing innovation and development (Han et al. 2023).

6 Future research directions

6.1 Improving data collection and integration methods

For future research, the approach should address issues related to data quality and availability concerns to strengthen conventional methods in groundwater exploration applications as well as in water resources management. With the advancement of powerful remote sensing technologies (i.e., high-resolution satellites and drones), we can now monitor changes in land surface dynamics longitudinally and with increasing accuracy, including evaluations of vegetation health, soil moisture, and even human disturbances (Guo et al. 2023). This highlights the need to enhance ML-based approaches by integrating them with conventional techniques, such as geophysical surveys and in-situ measurements, to establish a comprehensive dataset for groundwater systems, which are often composed of heterogeneous subsets.

Another key building block is the technology for data fusion, which must handle input data of varying spatial and temporal scales that require merging (Varouchakis et al. 2023). Deep learning can be used to combine remote sensing data, geophysical survey results, and historical records on groundwater conditions to enhance the accuracy of estimates. Additionally, systems can leverage Internet of Things (IoT) devices for real-time monitoring of water usage, groundwater levels, and quality, providing fresh data at frequent intervals and enabling proactive, dynamically responsive systems.

Standardizing data formats across various sources and frameworks is essential for seamless integration and interoperability. Ensuring consistency enables efficient data exchange between databases, fostering greater collaboration among researchers, practitioners, and policymakers. This, in turn, facilitates the implementation of advanced AI/ML algorithms for improved decision-making in groundwater and water resource management (Ahmed et al. 2023).

6.2 Developing robust and interpretable models

Developing AI models that are both robust and interpretable remains a critical direction for future research. Addressing challenges such as overfitting, bias, and limited generalizability is essential to improve model accuracy and reliability. Techniques such as cross-validation, regularization, and ensemble learning can help mitigate overfitting and enhance generalization. Additionally, incorporating Bayesian methods for uncertainty quantification and integrating human priors into model structures can lead to more explainable AI/ML models. For instance, combining AI/ML with conventional hydrological modelling, such as the calibration of groundwater dynamics and water resources (Chatterjee et al. 2023), can leverage the strength of both approaches, resulting in more accurate and reliable hybrid models.

Decision-makers must have a clear understanding of AI-driven water resource models. To enhance transparency, explainable AI techniques and model-agnostic approaches, such as interpretable neural networks, have been increasingly adopted (Alshehri and Rahman, 2023). These methods allow users to understand the reasoning behind each prediction, enabling the identification of potential errors and improving model reliability. Transparency is essential for effective real-world implementations, helping water resource managers understand why a particular policy action is recommended. This understanding fosters greater trust and confidence, encouraging the regular, or even daily, use of AI/ML tools in water management. Moreover, interpretability can help detect mistakes or biases that might be overlooked by black-box models, further improving model reliability and credibility. However, a major challenge in fully integrating AI into the water resource management sector lies in developing models that are both robust and interpretable (Kamyab et al. 2023). A promising direction is the creation of hybrid models that combines the strengths of both transparent and high-performing approaches, aiming to improve overall prediction accuracy and robustness. Such models can offer valuable insights into system behaviour, supporting more efficient and informed water resource management (Mohammed et al. 2022).

6.3 Real-time monitoring and predictive analytics

Real-time monitoring and predictive analytics are essential for enhancing the responsiveness and effectiveness of water management systems. One key application is the use of IoT devices to monitor groundwater levels and quality. These capabilities can be further strengthened by integrating conventional remote sensing techniques with AI/ML algorithms to provide utimely data for informed decision-making (Kamyab et al. 2023). Advancements in sensor technology, data transmission, and processing are needed to develop miniaturized, real-time monitoring systems. High-resolution water sensors, when paired with IoT devices, can deliver valuable data on groundwater conditions. To ensure this data is transmitted swiftly and reliably, next-generation wireless communication technologies such as 5G are crucial. This high-frequency data stream can then be processed in real-time by AI/ML algorithms, enabling predictive models to support decision-making based on historical trends (e.g., past climate data) and generate new insights for water resource management. Spatio-temporal analyses of groundwater quality derived from real-time monitoring allow for proactive adaptation to changes in groundwater dynamics, quality, and usage patterns, as demonstrated in monitored catchments, thereby improving the resilience, efficiency, and sustainability of managed ecosystems (El - Aassar et al. 2023; Li et al. 2026).

Water managers can use predictive analytics to forecast changes in groundwater levels, quality, and usage, enabling timely and informed responses. ML models, such as neural networks and ensemble-learning methods, analyze historical records, real-time sensor data, and usage patterns to identify trends and predict future conditions. These insights support proactive water resource management, allowing for pre-emptive actions that enhance the effectiveness, resilience, and long-term sustainability of water management strategies (Agyeman et al. 2023).

6.4 Cross-disciplinary research and collaboration

Water resource management is a complex and multifaceted challenge that requires cross-disciplinary research and collaboration among hydrologists, data scientists, engineers, policymakers, and other stakeholders. Interdisciplinary approaches are key in the development of holistic tools and well-considered solutions that address water management issues. These approaches are suitable for a broad audience, allowing different insights to be drawn from various fields of expertise. Such efforts will be important in future research to facilitate collaboration and communication across disciplines, leveraging a wide range of knowledge, skills, and perspectives. By pooling diverse ideas and expertise, more integrated and effective strategies for addressing water resource challenges can emerge. We argue that interdisciplinary studies offer a promising pathway for system-level representation of complex water-related challenges and the design of integrated solutions.

Collaborating in research activities, such as international projects or participation in global networks, not only promotes knowledge disseminate but also helps build capacity. These initiatives create valuable opportunities for researchers to co-develop AI/ML solutions relevant to water management. In addition, collaborations among academia, industry, and government can accelerate the translation of research outcomes into practical applications and foster the adoption of innovative technologies in real-world settings. Organized interdisciplinary initiatives involving experts from hydrology, computer science, engineering and related fields can generate powerful synergies. Through such collaboration, more robust and creative solutions to the numerous challenges in water resource management can be developed. These efforts will also help realize the full global potential of AI/ML technologies in water-related fields and contribute to a shared base of knowledge and experiences (Li et al. 2023) .

6.5 Open science and data sharing

We share this belief, and we think open science and data sharing can help accelerate the development of AI/ML models for groundwater exploration and water resource management. This model of open access model is particularly valuable, as it fosters innovation by making data, algorithms, and research discoveries accessible for collaboration, verification, and transparency. With available data, methodologies, and research findings, we can build upon existing work to validate results and develop integrated approaches using diverse tools. This ensures that progress benefits both individual researchers and the broader scientific community. By leveraging multiple data sources and documenting algorithms, scholars can create more robust and reliable AI/ML models.

Interdisciplinary collaboration among researchers in hydrology, computer sciences, or environmental sciences can also be facilitated by open science and data sharing practices. Such cooperation across regions and river basins in Mexico, and potentially with communities throughout Central America, represents a significant step forward in addressing complex water resource challenges. Ultimately, the major driver of progress in AI/ML-based groundwater exploration and water resource management will be the widespread adoption of open science and data sharing principles (Kamyab et al. 2023).

Open data platforms and repositories make it easy for researchers and practitioners to access high-quality datasets for developing and validating AI/ML models. Decisions regarding specific standards for data formats and protocols used on these platforms, enabling common understanding among datasets collected from various sources, must also be addressed at this stage. In another context, the promotion of open-source software-based solutions might encourage broader adoption of AI/ML interventions in water resources management (Chakraborty, 2024; Kamyab et al. 2023; Cheng et al. 2025). Policies that favour open science and data, while respecting ethical considerations and privacy, should be established. This is crucial for safegarding privacy, civil liberties, and maintaining public trust. Policymakers must develop regulations that require data sharing and partnerships, with strong privacy and security safeguards (Toward a 21st Century National Data Infrastructure: Managing Privacy and Confidentiality Risks with Blended Data, 2024; McGraw and Mandl, 2021).

7 Conclusions

(1) The integration of AI and ML is transforming water resource management by enabling advanced data analysis, pattern recognition, and predictive modelling. From hydrology to groundwater exploration, these technologies enhance efficiency, accuracy, and decision-making. These technologies play a crucial role in identifying groundwater potential zones, predicting water quality, and optimizing water distribution systems, contributing to the sustainable management of this vital resource.

(2) Techniques such as remote sensing, geospatial analysis, and deep learning have significantly advanced groundwater exploration. By processing large datasets and revealing hidden patterns, AI models enhance the prediction of groundwater levels and quality. The integration of geophysical and hydrological data enables more precise identification of potential water sources, reducing the reliance on traditional, trial-and-error methods.

(3) Despite these advantages, challenges persist, including data quality, limited real-time data access, and the need for high computational power. Model reliability and transparency are also critical, as inaccurate predictions can lead to ineffective or even harmful management strategies. Addressing these challenges is critical for the successful implementation of AI technologies.

(4) Ethical and regulatory considerations must also be addressed to ensure responsible AI deployment. Transparent data use, unbiased model training, and adherence to environmental regulations are key to fostering stakeholder trust. Successful integration further requires close collaboration between policymakers, engineers, and AI experts to develop scalable and adaptable solutions.

(5) Future research should focus on improving data collection methods, developing robust AI models, and enhancing real-time monitoring capabilities. Cross-disciplinary collaboration, especially among hydrologists, data scientists, and policymakers, is essential to advancing the field. Promoting open science and data-sharing practices can accelerate innovations and lead to more reliable and sustainable water management solutions.

References

Abdelaziz S, Gad MI, El Tahan AHMH. 2020. Groundwater quality index based on PCA: Wadi el-natrun, Egypt. Journal of African Earth Sciences, 172: 103964. DOI：10.1016/j.jafrearsci.2020.103964.

Abdullah MS, Karim HH, Samueel ZW. 2022. Investigation structural settlement by ground penetrating radar (case study). IOP Conference Series: Earth and Environmental Science, 961(1): 012037. DOI：10.1088/1755-1315/961/1/012037.

Abu M, Musah R, Zango MS. 2024. A combination of multivariate statistics and machine learning techniques in groundwater characterization and quality forecasting. Geosystems and Geoenvironment, 3(2): 100261. DOI：10.1016/j.geogeo.2024.100261.

Agweyu A, Hill K, Diaz T, et al. 2023. Regular measurement is essential but insufficient to improve quality of healthcare. BMJ, 380: e073412. DOI：10.1136/bmj-2022-073412.

Agyeman BT, Naouri M, Appels WM, et al. 2023. Integrating machine learning paradigms and mixed-integer model predictive control for irrigation scheduling. Cornell University. DOI:10.48550/arxiv.2306.08715

Agyemang VO. 2022. Application of geostatistical techniques in the assessment of groundwater contamination in the Afigya Kwabre District of Ghana. Applied Water Science, 12(3): 53. DOI：10.1007/s13201-022-01582-x.

Ahmed MI, Spooner B, Isherwood J, et al. 2023. A systematic review of the barriers to the implementation of artificial intelligence in healthcare. Cureus, 15(10): e46454. DOI：10.7759/cureus.46454.

Aldoseri A, Al-Khalifa KN, Hamouda AM. 2023. Re-thinking data strategy and integration for artificial intelligence: Concepts, opportunities, and challenges. Applied Sciences, 13(12): 7082. DOI：10.3390/app13127082.

Alghamdi AG, Aly AA, Majrashi MA, et al. 2023. Impact of climate change on hydrochemical properties and quality of groundwater for domestic and irrigation purposes in arid environment: A case study of Al-Baha region, Saudi Arabia. Environmental Earth Sciences, 82(1): 39. DOI：10.1007/s12665-022-10731-z.

Alshehri F, Rahman A. 2023. Coupling machine and deep learning with explainable artificial intelligence for improving prediction of groundwater quality and decision-making in arid region, Saudi Arabia. Water, 15(12): 2298. DOI：10.3390/w15122298.

Apostolou K, Staikou A, Sotiraki S, et al. 2021. An assessment of snail-farm systems based on land use and farm components. Animals, 11(2): 272. DOI：10.3390/ani11020272.

Appling AP, Oliver SK, Read JS, et al. 2022. Machine learning for understanding inland water quantity, quality, and ecology. Encyclopedia of Inland Waters. Amsterdam: Elsevier: 585−606. DOI: 10.1016/b978-0-12-819166-8.00121-3.

Biazar SM, Golmohammadi G, Nedhunuri RR, et al. 2025. Artificial intelligence in hydrology: Advancements in soil, water resource management, and sustainable development. Sustainability, 17(5): 2250. DOI：10.3390/su17052250.

Bata M, Carriveau R, Ting DSK. 2020. Short-term water demand forecasting using hybrid supervised and unsupervised machine learning model. Smart Water, 5(1): 2. DOI：10.1186/s40713-020-00020-y.

Behar JA, Levy J, Celi LA. 2023. Generalization in medical AI: a perspective on developing scalable models. Cornell University. DOI:10.48550/arXiv.2311.

Bradshaw TJ, Huemann Z, Hu JJ, et al. 2023. A guide to cross-validation for artificial intelligence in medical imaging. Radiology: Artificial Intelligence, 5(4): e220232. DOI：10.1148/ryai.220232.

Bui DD, Nguyen NC, Bui NT, et al. 2017. Climate change and groundwater resources in Mekong Delta, Vietnam. Journal of Groundwater Science and Engineering, 5(1): 76−90. DOI：10.26599/jgse.2017.9280008.

Chakraborty M, Tejankar A, Ayyamperumal R. 2021. Site suitability analysis for artificial groundwater recharge potential zone using a GIS approach in Basaltic terrain, Buldhana District, Maharashtra, India. Research Square (United States). DOI：10.21203/rs.3.rs-405615/v1.

Chakraborty S. 2024. Towards a comprehensive assessment of AI's environmental impact. Cornell University. DOI：10.48550/arxiv.2405.14004.

Chang FJ, Guo SL. 2020. Advances in hydrologic forecasts and water resources management. Water, 12(6): 1819. DOI：10.3390/w12061819.

Chatterjee SS, Ghosh R, Renganathan A, et al. 2023. Uncertainty quantification in inverse models in hydrology. Cornell University. Doi:10.48550/arxiv.2310.02193

Cheng H, Hong W, Zhang ZK, et al. 2025. Impacts of random negative training datasets on machine learning-based geologic hazard susceptibility assessment. China Geology, 8(4): 676−690. DOI：10.31035/cg2024094.

Cheng CY, Zhang F, Shi JC, et al. 2022. What is the relationship between land use and surface water quality? A review and prospects from remote sensing perspective. Environmental Science and Pollution Research, 29(38): 56887−56907. DOI：10.1007/s11356-022-21348-x.

Choudhary P, Modi A, Botre BA, et al. 2021. Leak detection in smart water distribution network. American Institute of Physics. DOI：10.1063/5.0044005.

Chrysos G, Georgopoulos M, Deng J, et al. 2021. Augmenting deep classifiers with polynomial neural networks. Cornell University. DOI: 10.48550/arxiv.2104.07916

Clark SR, Pagendam D, Ryan L. 2022. Forecasting multiple groundwater time series with local and global deep learning networks. International Journal of Environmental Research and Public Health, 19(9): 5091. DOI：10.3390/ijerph19095091.

Dabas J, Sarah S, Mondal NC, et al. 2022. Geostatistical spatial projection of geophysical parameters for practical aquifer mapping. Scientific Reports, 12: 4641. DOI：10.1038/s41598-022-08494-5.

Dahan O. 2020. Vadose zone monitoring as a key to groundwater protection. Frontiers in Water, 2: 599569. DOI：10.3389/frwa.2020.599569.

Daniel I, Cominola A. 2023. Estimating irregular water demands with physics-informed machine learning to inform leakage detection. Cornell University. DOI：10.48550/arxiv.2309.02935.

Dao PU, Heuzard AG, Le TXH, et al. 2024. The impacts of climate change on groundwater quality: A review. Science of The Total Environment, 912: 169241. DOI：10.1016/j.scitotenv.2023.169241.

de Lara A, Mieno T, Luck JD, et al. 2023. Predicting site-specific economic optimal nitrogen rate using machine learning methods and on-farm precision experimentation. Precision Agriculture, 24(5): 1792−1812. DOI：10.1007/s11119-023-10018-8.

Derdour A, Benkaddour Y, Bendahou B. 2022. Application of remote sensing and GIS to assess groundwater potential in the transboundary watershed of the Chott-El-Gharbi (Algerian–Moroccan border). Applied Water Science, 12(6): 136. DOI：10.1007/s13201-022-01663-x.

Ding JL, Yang ST, Shi Q, et al. 2020. Using apparent electrical conductivity as indicator for investigating potential spatial variation of soil salinity across seven oases along Tarim River in southern Xinjiang, China. Remote Sensing, 12(16): 2601. DOI：10.3390/rs12162601.

Ding X, Du W. 2023. Optimizing irrigation efficiency using deep reinforcement learning in the field. Cornell University. DOI：10.48550/arxiv.2304.01435.

Eftekhari M, Khashei-Siuki A. 2025. Evaluating machine learning methods for predicting groundwater fluctuations using GRACE satellite in arid and semi-arid regions. Journal of Groundwater Science and Engineering, 13(1): 5−21. DOI：10.26599/jgse.2025.9280035.

El-Aassar AH, Hagagg K, Hussien R, et al. 2023. Integration of groundwater vulnerability with contaminants transport modeling in unsaturated zone, case study El-Sharqia, Egypt. Environmental Monitoring and Assessment, 195(6): 722. DOI：10.1007/s10661-023-11298-3.

Fan J, Bai JW, Li ZY, et al. 2022. A GNN-RNN approach for harnessing geospatial and temporal information: Application to crop yield prediction. Proceedings of the AAAI Conference on Artificial Intelligence, 36(11): 11873−11881. DOI：10.1609/aaai.v36i11.21444.

Feng F, Chen Z, Ni J, et al. 2024. Machine learning to access and ensure safe drinking water supply: A systematic review. DOI：10.26434/chemrxiv-2024-cc4jd.

Goh H. 2021. Artificial intelligence in achieving sustainable development goals. Cornell University. DOI：10.48550/arXiv.2107.

Guo BL, Zhang SC, Liu K, et al. 2023. Prediction of groundwater level under the influence of groundwater exploitation using a data-driven method with the combination of time series analysis and long short-term memory: A case study of a coastal aquifer in Rizhao City, Northern China. Frontiers in Environmental Science, 11: 1253949. DOI：10.3389/fenvs.2023.1253949.

Guo Y, Xing NC, Gan FP, et al. 2023. Evaluating the hydrological components contributions to terrestrial water storage changes in Inner Mongolia with multiple datasets. Sensors, 23(14): 6452. DOI：10.3390/s23146452.

Han Y, Chen JH, Dou MT, et al. 2023. The impact of artificial intelligence on the financial services industry. Academic Journal of Management and Social Sciences, 2(3): 83−85. DOI：10.54097/ajmss.v2i3.8741.

He T, Chang H, Zhang D. 2022. Identification of physical processes and unknown parameters of 3D groundwater contaminant problems via theory-guided U-net. Cornell University. DOI：10.48550/arxiv. 2205.00134.

Hsia SC, Wang SH, Hsu SW. 2021. Smart water-meter wireless transmission system for smart cities. IEEE Consumer Electronics Magazine, 10(6): 83−88. DOI：10.1109/MCE.2020.3043997.

Huang YK, Wang XY, Xiang WJ, et al. 2022. Forward-looking roadmaps for long-term continuous water quality monitoring: Bottlenecks, innovations, and prospects in a critical review. Environmental Science & Technology, 56(9): 5334−5354. DOI：10.1021/acs.est.1c07857.

Huynh BQ, Kiang MV. 2023. AI for anticipatory action: Moving beyond climate forecasting. Cornell University. DOI:10.48550/arxiv.2307.15727

Jaafarzadeh MS, Tahmasebipour N, Haghizadeh A, et al. 2021. Groundwater recharge potential zonation using an ensemble of machine learning and bivariate statistical models. Scientific Reports, 11: 5587. DOI：10.1038/s41598-021-85205-6.

Kamyab H, Khademi T, Chelliapan S, et al. 2023. The latest innovative avenues for the utilization of artificial Intelligence and big data analytics in water resource management. Results in Engineering, 20: 101566. DOI：10.1016/j.rineng.2023.101566.

Kobayashi K, Alam SB. 2023. Explainable, interpretable & trustworthy AI for intelligent digital twin: Case study on remaining useful life. Cornell University. DOI：10.48550/arxiv.2301.06676.

Kumar K, Saini R. 2021. Application of artificial intelligence for the optimization of hydropower energy generation. EAI Endorsed Transactions on Industrial Networks and Intelligent Systems, 8(28): 170560. DOI：10.4108/eai.6-8-2021.170560.

Li J, Wang WK, Cheng DW, et al. 2021. Hydrogeological structure modelling based on an integrated approach using multi-source data. Journal of Hydrology, 600: 126435. DOI：10.1016/j.jhydrol.2021.126435.

Li P, Yang J, Islam MA, et al. 2023. Making AI less "Thirsty": Uncovering and addressing the secret water footprint of AI models. Cornell University. DOI：10.48550/arXiv.2304.

Li WB, Wang XY, He L, et al. 2026. Reservoir fluid type identification method based on deep learning: A case study of the Chang 1 Formation in the Jiyuan oilfield of the Ordos basin, China. China Geology. DOI：10.31035/cg2025010.

Liu X, Zhang JQ, Huang XL, et al. 2022. Heavy metal distribution and bioaccumulation combined with ecological and human health risk evaluation in a typical urban plateau lake, Southwest China. Frontiers in Environmental Science, 10: 814678. DOI：10.3389/fenvs.2022.814678.

Liu YD, Yan DD, Zheng KX. 2022. Design of a comprehensive assessment model for the stability and engineering geology of slope based on improved convolutional neural network. Computational Intelligence and Neuroscience, 2022: 1639311. DOI：10.1155/2022/1639311.

Mahamat AA, Boukar MM, Ibrahim NM, et al. 2021. Machine learning approaches for prediction of the compressive strength of alkali activated termite mound soil. Applied Sciences, 11(11): 4754. DOI：10.3390/app11114754.

Malakar P, Sarkar S, Mukherjee A, et al. 2021. Use of machine learning and deep learning methods in groundwater. Elsevier BV, 545-557. DOI：10.1016/b978-0-12-818172-0.00040-2.

McGraw D, Mandl KD. 2021. Privacy protections to encourage use of health-relevant digital data in a learning health system. NPJ Digital Medicine, 4: 2. DOI：10.1038/s41746-020-00362-8.

Mohammadi B. 2021. A review on the applications of machine learning for runoff modeling. Sustainable Water Resources Management, 7(6): 98. DOI：10.1007/s40899-021-00584-y.

Mohammed SJ, Zubaidi SL, Ortega-Martorell S, et al. 2022. Application of hybrid machine learning models and data pre-processing to predict water level of watersheds: Recent trends and future perspective. Cogent Engineering, 9(1). DOI：10.1080/23311916.2022.2143051.

Mohammad AT, Jalut QH, Abbas NL. 2020. Predicting groundwater level of wells in the Diyala river Basin in eastern Iraq using articial neural network. Journal of Groundwater Science and Engineering, 8(1): 87−96. DOI：10.19637/j.cnki.2305-7068.2020.01.009.

Mokua N, Maina CW, Kiragu H. 2021. A raw water quality monitoring system using wireless sensor networks. International Journal of Computer Applications, 174(21): 35−42. DOI：10.5120/ijca2021921113.

Mustafa HM, Mustapha A, Hayder G, et al. 2021. Applications of IoT and artificial intelligence in water quality monitoring and prediction: A review. 2021 6th International Conference on Inventive Computation Technologies (ICICT). January 20-22, 2021, Coimbatore, India. IEEE: 968−975. DOI：10.1109/ICICT50816.2021.9358675.

Nalevanková P, Fleischer P, Mukarram M, et al. 2023. Comparative assessment of sap flow modeling techniques in European beech trees: Can linear models compete with random forest, extreme gradient boosting, and neural networks? Water, 15(14): 2525. DOI：10.3390/w15142525.

Nguyen AD, Le Nguyen P, Vu VH, et al. 2022. Accurate discharge and water level forecasting using ensemble learning with genetic algorithm and singular spectrum analysis-based denoising. Scientific Reports, 12: 19870. DOI：10.1038/s41598-022-22057-8.

Pandey NK, Shukla AK, Shukla S, et al. 2014. Assessment of underground water potential zones using modern geomatics technologies in Jhansi district, Uttar Pradesh, India. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL−8: 377−381. DOI：10.5194/isprsarchives-xl-8-377-2014.

Pandey S, Tripathi S, Singh A. 2020. Delineation of ground water potential zone using remote sensing, GIS and GPS, in mauranipur block, Jhansi district (UP), India. International Journal of Current Microbiology and Applied Sciences, 9(6): 2138−2145. DOI：10.20546/ijcmas.2020.906.261.

Pourmorad S, Kabolizade M, Dimuccio LA. 2024. Artificial intelligence advancements for accurate groundwater level modelling: An updated synthesis and review. Applied Sciences, 14(16): 7358. DOI：10.3390/app14167358.

Prasad P, Loveson VJ, Kotha M, et al. 2020. Application of machine learning techniques in groundwater potential mapping along the west coast of India. GIScience & Remote Sensing, 57(6): 735−752. DOI：10.1080/15481603.2020.1794104.

Quan HB, Wang X. 2020. Research on application of GIS technology in water environment planning of basin. Journal of Physics: Conference Series, 1649(1): 012006. DOI：10.1088/1742-6596/1649/1/012006.

Rathor S, Kumari S. 2021. A social application of artificial intelligence & IoT for water conservation. IOP Conference Series: Materials Science and Engineering, 1116(1): 012191. DOI：10.1088/1757-899x/1116/1/012191.

Ryan P, Porter Z, Al-Qaddoumi J, et al. 2024. What's my role? modelling responsibility for AI-based safety-critical systems. Cornell University. DOI：10.48550/arxiv.2401.09459.

Sahour S, Khanbeyki M, Gholami V, et al. 2023. Evaluation of machine learning algorithms for groundwater quality modeling. Environmental Science and Pollution Research, 30(16): 46004−46021. DOI：10.1007/s11356-023-25596-3.

Sarker IH. 2021. Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2(3): 160. DOI：10.1007/s42979-021-00592-x.

Siva Barathi A, Manapragada NVSK, Rai PK, et al. 2024. Artificial intelligence and machine learning-based building solutions: Pathways to ensure occupant comfort and energy efficiency with climate change. Big Data, Artificial Intelligence, and Data Analytics in Climate Change Research. Singapore: Springer Nature Singapore: 57−79. DOI：10.1007/978-981-97-1685-2_4.

Shi J, Yin Z, Myana R, et al. 2023. Deep learning models for flood predictions in South Florida. Cornell University. DOI：10.48550/arxiv.2306.15907.

Shrestha SG, Pradhanang SM. 2023. Performance of LSTM over SWAT in rainfall-runoff modeling in a small, forested watershed: A case study of cork brook, RI. Water, 15(23): 4194. DOI：10.3390/w15234194.

Slavković A, Seeman J. 2023. Statistical data privacy: A song of privacy and utility. Annual Review of Statistics and Its Application, 10: 189−218. DOI：10.1146/annurev-statistics-033121-112921.

Song H, Jung J. 2023. Scalable surface water mapping up to fine-scale using geometric features of water from topographic airborne LiDAR data. Cornell University. DOI：10.48550/arxiv.2301.06567.

Sowmya GS, Sathisha HK. 2023. Detecting financial fraud in the digital age: The AI and ML revolution. International Journal for Multidisciplinary Research, 5(5): 6139. DOI：10.36948/ijfmr.2023.v05i05.6139.

Sun N, Zhang S, Peng T, et al. 2022. Multi-variables-driven model based on random forest and Gaussian process regression for monthly streamflow forecasting. Water, 14(11): 1828. DOI：10.3390/w14111828.

Surinaidu L, Nandan MJ, Sahadevan DK, et al. 2021. Source identification and management of perennial contaminated groundwater seepage in the highly industrial watershed, south India. Environmental Pollution, 269: 116165. DOI：10.1016/j.envpol.2020.116165.

Suwadi NA, Derbali M, Sani NS, et al. 2022. An optimized approach for predicting water quality features based on machine learning. Wireless Communications and Mobile Computing, 2022: 3397972. DOI：10.1155/2022/3397972.

Swain S, Mishra SK, Pandey A, et al. 2022. Inclusion of groundwater and socio-economic factors for assessing comprehensive drought vulnerability over Narmada River Basin, India: A geospatial approach. Applied Water Science, 12(2): 14. DOI：10.1007/s13201-021-01529-8.

Toward a 21st Century National Data Infrastructure: Managing privacy and confidentiality risks with blended data. 2024. DOI：10.17226/27335.

Vahldiek K, Rüger B, Klawonn F. 2022. Leakages in district heating networks—Model-based data set quality assessment and localization. Sensors, 22(14): 5300. DOI：10.3390/s22145300.

Varouchakis EA, Solomatine D, Perez GAC, et al. 2023. Combination of geostatistics and self-organizing maps for the spatial analysis of groundwater level variations in complex hydrogeological systems. Stochastic Environmental Research and Risk Assessment, 37(8): 3009−3020. DOI：10.1007/s00477-023-02436-x.

Wang HL, Shen Y, Liang L, et al. 2022. River extraction from remote sensing images in cold and arid regions based on attention mechanism. Wireless Communications and Mobile Computing, 2022: 9410381. DOI：10.1155/2022/9410381.

Wang JB, Sun T, Wang XY. 2022. Research on the application of water resources optimal allocation model based on fuzzy optimization theory. Polish Journal of Environmental Studies, 31(6): 5241−5251. DOI：10.15244/pjoes/150046.

Wang YJ, Gu XC, Yang G, et al. 2021. Impacts of climate change and human activities on water resources in the Ebinur Lake Basin, Northwest China. Journal of Arid Land, 13(6): 581−598. DOI：10.1007/s40333-021-0067-4.

Wang X, Xiong W, Wang H, et al. 2018. Look before you leap: Bridging model-free and model-based reinforcement learning for planned-ahead vision-and-language navigation. Cornell University. DOI：10.48550/arXiv.1803.

Xin L, Mou TY. 2022. Research on the application of multimodal-based machine learning algorithms to water quality classification. Wireless Communications and Mobile

Computing, 2022: 9555790. DOI：10.1155/2022/9555790.

Xu H, Lv B, Chen J, et al. 2023. Research on a prediction model of water quality parameters in a marine ranch based on LSTM-BP. Water, 15(15): 2760. DOI：10.3390/w15152760.

Xu Q, Shi Y, Bamber JL, et al. 2023. Physics-aware machine learning revolutionizes scientific paradigm for machine learning and process-based hydrology. Cornell University. DOI：10.48550/arxiv.2310.05227.

Yang L, Cheng YP, Wen X-R, et al. 2024. Development, hotspots and trend directions of groundwater numerical simulation: A bibliometric and visualization analysis. Journal of Groundwater Science and Engineering, 12(4): 411−427. DOI：10.26599/JGSE.2024.9280031.

Yu J, Yang R. 2021. Study on the predictive algorithm of plant restoration under heavy metals. Scientific Programming, 2021: 6193182. DOI：10.1155/2021/6193182.

Zhu W, Wang WG, Wang DY, et al. 2023. Application of the electromagnetic method to the spatial distribution of subsurface saline and fresh water in the coastal mudflat area of Jiangsu Province. Sensors, 23(14): 6405. DOI：10.3390/s23146405.

Zowghi D, Rimini FD. 2023. Diversity and inclusion in artificial intelligence. Cornell University. DOI：10.48550/arXiv.2305.