Journal of
Water & Climate Change

IWA PUBLISHING

Journal of Water and Climate Change Vol 14 No 6, 1762 doi: 10.2166/wcc.2023.348

Analysis and prediction of the changes in groundwater resources under heavy precipitation and ecological water replenishment

Bowen Shi[a,b], Chao Wan[c], Weiwu Hu[b], Congchao Xu[a,b], Di Liu[a], Mingxiao Li[a], Rui Li[a,b,*], Chuanping Feng[b] and Beidou Xi[a]

[a] State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China
[b] School of Water Resources and Environment, China University of Geosciences (Beijing), Beijing 100083, China
[c] Yongding River Investment Co., LTD., Beijing, China
*Corresponding author. E-mail: 518lirui@163.com

ABSTRACT

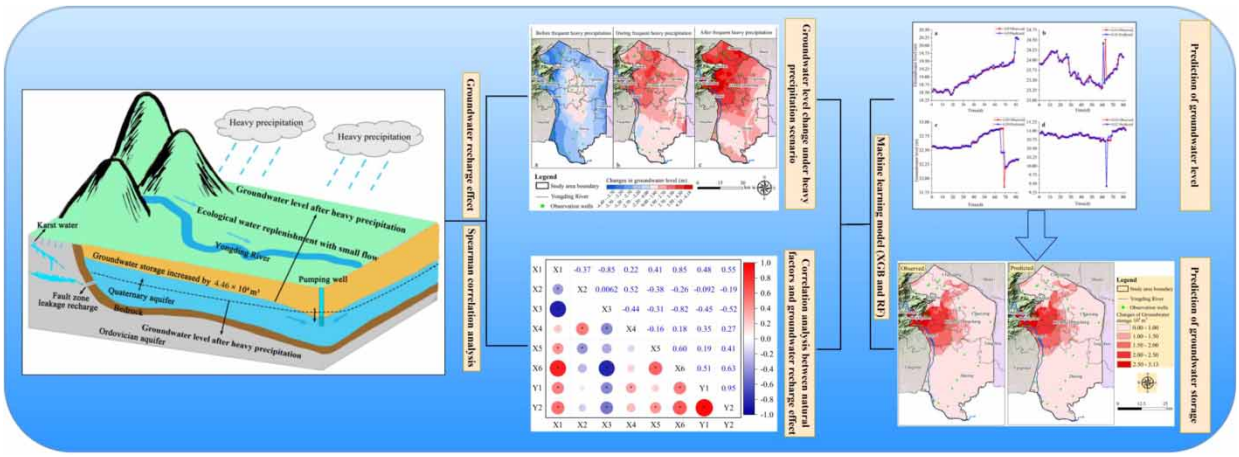
Identifying the influence of heavy precipitation and ecological water replenishment (EWR) on groundwater resources is essential for groundwater resources management and risk prevention. This study innovatively developed a groundwater resource analysis and prediction model integrated with the water level fluctuation method, correlation analysis, and machine learning method under the influence of heavy precipitation and EWR. Water level fluctuation method results showed that compared with January 1, 2021, the groundwater resources of the study area increased 4.46×10^8 m^3 on August 28. Compared with small flow of EWR, heavy precipitation was the main contributor to the rise in the groundwater level. Correlation analysis found that elevation, specific yield, and permeability coefficient show positive correlations with groundwater resource recharge. Machine learning results showed that among the water level prediction models of 35 monitoring wells, extreme gradient boosting (XGB) and random forest (RF) performed best in 30 wells and five wells, respectively. The increase in groundwater storage predicted deviated from the actual value by only 0.6×10^7 m^3 (prediction bias of 1.3%), indicating that the model prediction performance was good under the heavy precipitation condition. This study can help to better understand the change trend of groundwater resources under the conditions of heavy precipitation and EWR.

Key words: correlation analysis, groundwater resource, heavy precipitation and ecological water replenishment (EWR), machine learning, water level fluctuation method

HIGHLIGHTS

- Through hydrology, statistics, and machine learning, groundwater changes under the dual effects of heavy precipitation and ecological water replenishment are studied.
- A machine learning model is developed to predict the groundwater level and storage under heavy precipitation scenarios.
- XGB and RF models well predicted the groundwater change the next day, and the prediction deviation was only 1.3%.

GRAPHICAL ABSTRACT



1. INTRODUCTION

Groundwater refers to the water stored in rock voids below the ground (Xing *et al.* 2010), accounting for nearly 30% of the global freshwater reserves (Majumdar *et al.* 2020). Groundwater is an important source of water for drinking (supplying nearly 50% of the drinking water in the world) (Gleeson *et al.* 2019; Mohapatra *et al.* 2021), agricultural irrigation, and industrial production (Dangar *et al.* 2021; Agarwal *et al.* 2023). Changes in groundwater resources (water level and water quality changes) affect human life, industrial and agricultural production, and building safety (Schinke *et al.* 2012; Pelletier *et al.* 2015; Liu *et al.* 2019; Schreiner-McGraw *et al.* 2019; D. F. Wang *et al.* 2020). Among various factors, heavy precipitation, and ecological water replenishment (EWR) are important contributors to positive changes in groundwater resources. Specifically, heavy precipitation is induced by extreme climate, and EWR is the artificial transfer of water to water-deficient areas through water conservancy projects to replenish surface water or groundwater resources (Sun *et al.* 2023). The response of groundwater to heavy precipitation and EWR is nonlinear and complicated (Taylor *et al.* 2012; Yu & Lin 2015; Schreiner-McGraw *et al.* 2019; Pastore *et al.* 2020). The recharge of groundwater in different regions is closely related to geological structure, geographic location, and aquifer properties (van Roosmalen *et al.* 2007; Goodarzi *et al.* 2015; Fagbohun 2018; Ren *et al.* 2018; J. C. Wang *et al.* 2020). Against the background of heavy precipitation and EWR, management of groundwater resources and risk prevention measures are difficult to formulate.

Driven by the extreme climate of the North China Plain, frequent heavy precipitation events occurred in Beijing from June to August 2021. Meanwhile, the Yongding River in the territory received EWR from Guanting Reservoir. Existing studies have used numerical simulation and machine learning to analyze the impact of common precipitation and EWR (focusing on the impact of EWR) on groundwater (e.g., water level and water quality) in the Yongding River Basin (Beijing Section) (Zhang *et al.* 2018; Hu *et al.* 2019; Luo *et al.* 2019; Ji *et al.* 2021; Sun *et al.* 2021; Xu *et al.* 2022), providing references for understanding the groundwater change mechanism under human intervention. However, there is a relative lack of research on the response mechanism and changing trend of groundwater resources under the rare condition that heavy precipitation and EWR occur at the same time. Therefore, exploring the response mechanism and changing trend of groundwater under the dual influence of heavy precipitation and EWR using highly applicable, accurate, and efficient methods has high practical significance.

The changes in groundwater storage reflect the replenishment of groundwater (Li *et al.* 2015), which is a quantitative indicator of the response of groundwater to external conditions. In the water level fluctuation method, the change in groundwater storage is calculated according to the groundwater level variation (Healy & Cook 2002). This method has the advantages of simple practical application, fast calculation, and few constraints (Delottier *et al.* 2018; Gong *et al.* 2020; Z. Y. Zhang *et al.* 2020), and is widely used in calculating changes in groundwater storage (Crosbie *et al.* 2019; Yenehun *et al.* 2020). In this study, the calculation results of the water level fluctuation method were extrapolated by the Kriging interpolation analysis tool of ArcGIS, and changes in the groundwater storage of the entire study area were then obtained.

Heavy precipitation and EWR are the recharge sources of groundwater, but the recharge of groundwater may show differences across different regions, which is related to many factors (e.g., ground elevation, fault zone distribution, and aquifer

nature). Correlation analysis can provide a basis for exploring the reasons for the differences in groundwater replenishment. Specifically, Spearman correlation analysis is favorable because it has no restrictions on the distribution characteristics of the analyzed variables (Artusi *et al.* 2002; Zhu *et al.* 2023). This study used this method to analyze the relationships between complex natural factors (e.g., ground elevation, fault zone distribution, aquifer properties) and groundwater recharge. This analysis has practical significance for understanding the differences in groundwater recharge in different regions under heavy precipitation and EWR.

The numerical model (e.g., MODFLOW) is a commonly used groundwater level prediction model (Kayhomayoon *et al.* 2022; Lv *et al.* 2022), but it requires a large amount of accurate data, which can never be determined with absolute accuracy (e.g., the physical properties of the aquifer) (Chen *et al.* 2020). In addition, the establishment of the numerical model takes a lot of time and money (Jiang *et al.* 2022; Mohammed *et al.* 2023). In recent years, the machine learning model has shown good performance in groundwater level prediction (Guzman *et al.* 2019; Wei *et al.* 2019; Mohapatra *et al.* 2021; Adnan *et al.* 2023; Zhou *et al.* 2023). Unlike in the numerical model, the complex physical relationship between variables need not be considered in the machine learning model (Jing *et al.* 2023). Moreover, the machine learning model is advantageous in terms of modeling time, prediction accuracy, etc. (Chen *et al.* 2020; Natarajan & Sudheer 2020; Müller *et al.* 2021). Artificial neural network (ANN), random forest (RF), and extreme gradient boosting (XGB) have been successfully applied to groundwater level prediction under various conditions around the world (Majumdar *et al.* 2020; Yadav *et al.* 2020; Osman *et al.* 2021). Sun *et al.* (2016) used an ANN model to predict the groundwater level of a swamp forest in Singapore and achieved higher accuracy for a one-day-ahead prediction than for a seven-day-ahead prediction. X. H. Wang *et al.* (2018) used an improved RF model to predict the groundwater level of the water source of the Dagu River in Qingdao, and achieved R^2 of 0.9581. Chen *et al.* (2020) compared the groundwater level prediction accuracy of numerical simulation and machine learning models and found that the machine learning model has a higher prediction performance than the numerical simulation model; moreover, the machine learning model afforded simpler and faster practical application. El Bilali *et al.* (2021) compared the prediction performance of support vector regression (SVR), *k*-nearest neighbors algorithm (KNN), RF, and ANN on the groundwater level of seven groundwater monitoring wells in a semi-arid area (Province of Khemisset, Morocco). The results showed that except for KNN, other models showed good prediction effects. So far, the machine learning model has been rarely applied to model groundwater resources under heavy precipitation and EWR.

This study integrated the water level fluctuation method, correlation analysis, and machine learning method to establish a groundwater response prediction model considering actual conditions of heavy precipitation and EWR. The objectives of this study were to (i) study the changes in the groundwater level and groundwater storage in Yongding River Basin (Beijing Section) under extreme conditions, (ii) discuss the reasons for the differences in groundwater recharge in different regions, and (iii) predict the groundwater level and groundwater storage for the next day under the condition of heavy precipitation. This study has a practical reference value for understanding groundwater recharge and change under heavy precipitation and EWR and has guiding significance for the advanced management and risk prevention of groundwater resources.

2. MATERIALS AND METHODOLOGY

2.1. Description of the study area

The study area is located in the central and southern parts of the Beijing Plain, China. The Yongding River runs northwest to southeast (Figure 1). There are many high mountains in the northwest of the study area, but most of the area is covered by open plains. Accordingly, the Yongding River Basin (Beijing Section) can be divided into a mountain gorge area and a plain area. This study focused on the plain area. The gorge area, which is an important groundwater recharge area, is mainly composed of Ordovician limestone, and the groundwater type is karst water. The strata in the plain area are mainly quaternary alluvial deposits, and the groundwater type is loose rock pore water. Fault zones occur in the northwest of the plain area. The aquifer particles gradually become finer from northwest to southeast, and the structure transitions from a single sand–gravel layer to a multi-layer sand layer (Supplementary Material, Figure S1; Zhou *et al.* 2013; Xu *et al.* 2022). The main sources of groundwater recharge in the study area are precipitation, river seepage, and lateral inflow, and discharge occurs through artificial mining and lateral outflow.

The study area belongs to the warm temperate semi-humid and semi-arid monsoon climate, with an annual average temperature of 11.6 °C, an annual average precipitation of 556 mm, and an annual average evaporation of 1,333 mm. The climate is hot and humid in summer, and dry and windy in winter. During the period from June 1 to August 27, 2021, driven by the

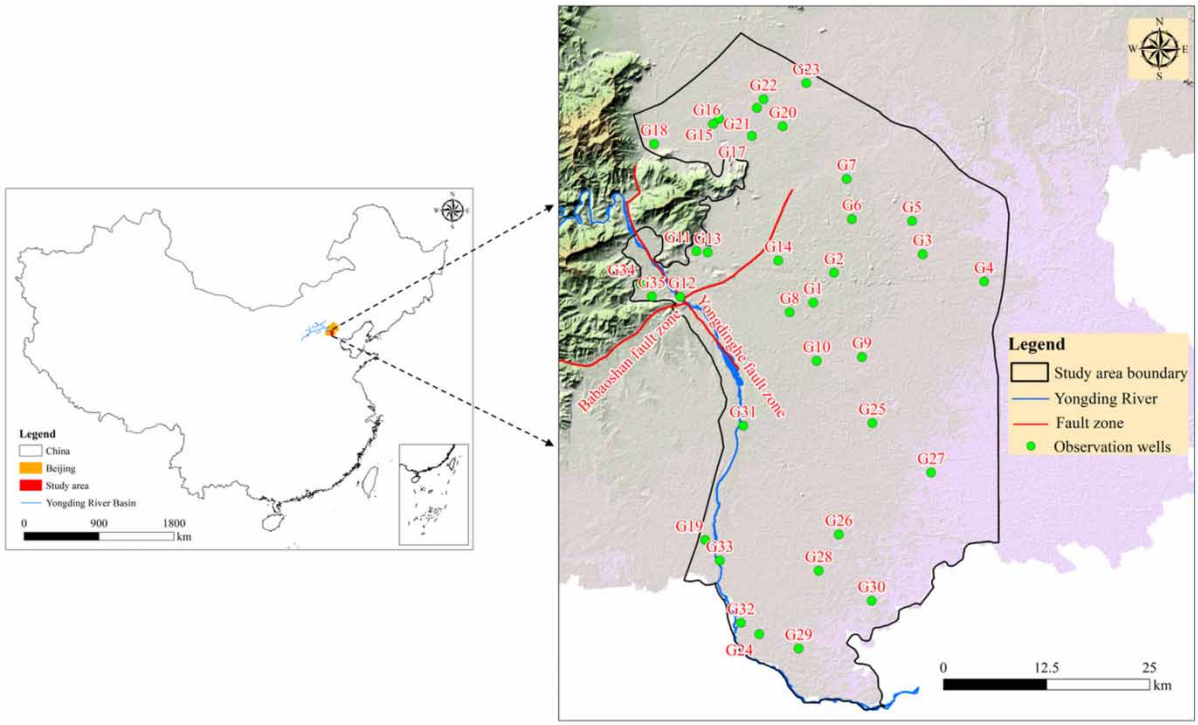


Figure 1 | Topography of the study area and location of groundwater monitoring wells.

extreme climate of the North China Plain, the precipitation in Beijing totaled 627 mm, with 77 mm in June, 405 mm in July, and 145 mm in August, which is higher than the annual precipitation (556 mm) of previous years. In order to facilitate the discussion of the impact of different precipitation conditions on the groundwater resources of the study area, the study time was divided into stages: before frequent heavy precipitation (June 1–30), frequent heavy precipitation period (July 1–31), and after frequent heavy precipitation (August 1–27). Precipitation in the study area is shown in Figure 2.

On January 1, 2021, the Beijing Municipal Government launched the Yongding River spring EWR plan, which transfers water from the Guanting Reservoir in the northwest of Beijing to the Yongding River. From January 1 to August 27, 2021, the average flow of EWR was approximately 4.687 m^3/s, the replenishment time was 239 days, and the total amount of replenishment was approximately 9.68×10^7 m^3.

2.2. Data collection

The data collected in this study include daily water level data from 35 groundwater monitoring wells (Nos G1–G35) in the plain area of the Yongding River Basin (Beijing Section) from June 1 to August 27, 2021, monitoring well surface elevation, specific yield data, precipitation data over Beijing, and EWR data. The groundwater level data and monitoring well ground elevation data were obtained from the China National Groundwater Monitoring Project (http://jcgc.cigem.cn/wechat/, visited in August 2021). Aquifer-specific yield data were extracted from the literature (Yu *et al.* 2017; Hu *et al.* 2019). The precipitation and EWR flow data were obtained from the Beijing Water Affairs Bureau (http://nsbd.swj.beijing.gov.cn:8088/uacp/pageview/bjsw/main). The distances between the monitoring wells and the Yongding River fault zone and the Babaoshan fault zone were measured using the ranging function of ArcGIS (version 10.2).

2.3. Methods

2.3.1. Numerical methods

2.3.1.1. Water level fluctuation method. According to the groundwater level variation and specific yield, the water level fluctuation method (Equation (1)) can be used to calculate the groundwater recharge within a limited range covered by a

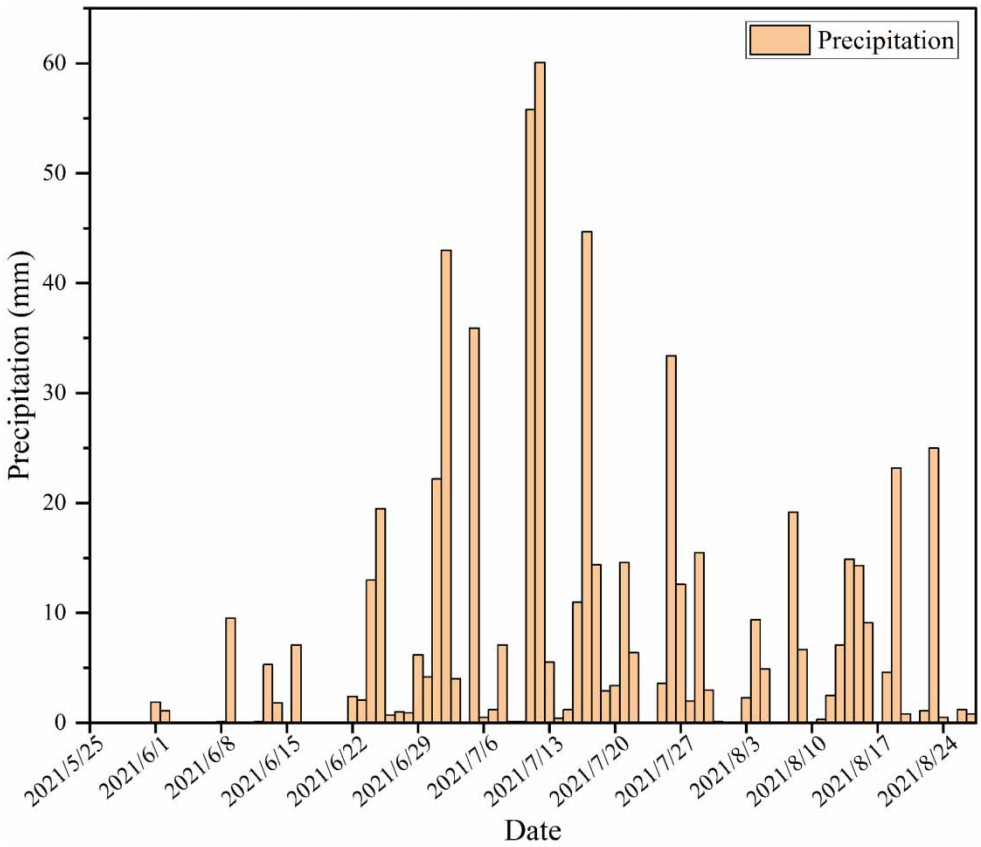


Figure 2 | Daily precipitation in Beijing from June 1 to August 27, 2021.

single monitoring point (Healy & Cook 2002).

$$R = S_y \times \frac{\Delta h}{\Delta t} \tag{1}$$

where R refers to the recharge of groundwater resources (mm/day), S_y refers to the specific yield (dimensionless), Δh refers to the groundwater level fluctuation value within a certain period of time (mm), and Δt refers to the time span (d).

In this study, the ordinary Kriging spatial interpolation was combined with the water level fluctuation method to calculate the variation of groundwater storage within the region. First, the calculation area is divided into several grids of equal area. Second, the water level change value and specific yield of the groundwater monitoring well are distributed in each grid by ordinary Kriging (the semi-variogram model is the spherical semi-variogram model). Finally, Equation (2) is used to calculate and superimpose the changes in groundwater storage of each grid to obtain the change in the groundwater storage of the entire area (Xu *et al.* 2022):

$$Q = \sum_{i=1}^{n} \Delta h_i \times S_y \times S_i \tag{2}$$

where Q refers to the variation in groundwater storage in a specific period (m^3), S_y refers to the specific yield (dimensionless), S_i refers to the area of each grid (m^2), Δh_i refers to the variation of groundwater levels of the ith grid in a specific period (m), and n refers to the number of grids.

According to Equation (2), the change value (Q) of groundwater storage in the plain area of the Yongding River Basin (Beijing Section) from June 1 to August 27, 2021, was calculated. The number of grids in the study area was 59,675, and the size of

a single grid was 200 m × 200 m. The specific yield distribution is shown in Supplementary Material, Figure S2. In this study, the 'Spatial Analyst' tool in ArcGIS (version 10.2) was used to implement ordinary Kriging interpolation.

2.3.1.2. Spearman correlation analysis. The replenishment of groundwater resources in different regions may be different, which is attributable to differences in natural and anthropogenic factors. Natural factors mainly include the ground elevation (X1), distance from the monitoring well to the Yongding River fault zone (X2), distance from the monitoring well to the Babaoshan fault zone (X3), precipitation (X4), specific yield (X5), and permeability coefficient (X6). Anthropogenic factors are mainly groundwater exploitation and utilization (Xu *et al.* 2022). In this study, the data of groundwater exploitation were excluded, and only the influence of natural factors on the replenishment of groundwater resources was explored. Since the above variables do not have normal distribution characteristics, the Pearson correlation method is not applicable (Hazra & Gogtay 2016). The Spearman correlation method does not require variable distribution characteristics. Therefore, Spearman correlation analysis was conducted between natural factors, groundwater level changes (Y1), and groundwater storage change (Y2).

2.3.2. Machine learning methods

2.3.2.1. Construction of the prediction model. When predicting the water level, a large prediction step leads to a decrease in prediction accuracy (Sun *et al.* 2016). Therefore, the prediction step of this study is 1, that is, the groundwater level for the next day is predicted. The water level data, rainfall data, and EWR flow data for 88 days (from June 1 to August 27, 2021) were used to establish the prediction model of 35 groundwater monitoring wells, and predict the groundwater level for the next day in the study area (i.e. the output variable of the model) under the heavy rainfall and EWR conditions. The 'train_test_split' function in Python (version No. 3.8.3) was used to divide the dataset into two sub-datasets, including a training set with 70% data and a test set with 30% data.

First, the input variables of the machine learning model were determined. The correlation between the input variable and the output variable is very important for prediction accuracy (Huang *et al.* 2017; Amaranto *et al.* 2019; Wu *et al.* 2021). The input variables of the model were set by referring to the results of the Spearman correlation analysis; natural factors having a strong correlation with the groundwater level were selected as input variables. As the historical groundwater level significantly affects the prediction accuracy of the next-day groundwater level (Alizamir *et al.* 2018), the groundwater level data of day t and day $t-1$ were also selected as input variables.

Subsequently, a machine learning model was established, and the accuracy of its prediction was assessed. Numerous studies have shown that the RF and XGB models based on decision trees have better performance in predicting groundwater levels than ANN, SVM, ANFIS, and other models (Miro *et al.* 2021; Osman *et al.* 2021; Ruidas *et al.* 2021; Pham *et al.* 2022; Rao *et al.* 2022). Therefore, this study selected RF and XGB as the groundwater level prediction models. After selecting the input variables, the prediction model between the input variable and the output variable was established through the RF algorithm and XGB algorithm. In this study, the RF model and the XGB model were established for each monitoring well (35 in total), and the model with the highest performance was selected as the groundwater level prediction model for each monitoring well after comparison. In order to assess the predictive ability of the model clearly and comprehensively, the 35 monitoring wells were grouped under four categories according to the similarity of data changes through systematic cluster analysis. One representative monitoring well was selected from each category to evaluate the model prediction accuracy in detail, and the model was then extended to all monitoring wells. The machine learning models used in this study were all built with Python (version No. 3.8.3).

Additionally, the model predicted value was used to draw the change chart of groundwater storage on August 27. The water level fluctuation method was used to convert the groundwater level data predicted by the model into the change in groundwater storage. A graph of changes in the groundwater storage of the study area on August 27 relative to January 1 was mapped. The charts of the predicted and actual groundwater storage were then compared to further evaluate the model prediction accuracy.

2.3.2.2. Random forest. RF is a machine learning algorithm based on decision trees, which can be applied to classification and regression problems (Breiman 2001). RF comprises multiple decision trees, and no relationship exists between each decision tree in the forest. The final output of the model is jointly determined by each decision tree in the forest. When dealing with classification problems, each decision tree in the forest gives the final category, and the category of the test

sample is determined by voting; when dealing with regression problems, the mean output of each decision tree is used as the final result (Breiman 2001). The performance of the RF model is closely related to parameters such as ntree (quantity of decision trees), mtry (number of variables selected on each decision tree node), and max depth (maximum number of divisions of decision tree nodes) (Biau & Scornet 2016; Brédy *et al.* 2020; Rahman *et al.* 2020). In this study, the ntree values of the RF model were set as 10, 20, 50, 100, and 200, mtry values were set as 1, 2, and 3, and max depth values were set as 3, 5, 7, and 9. The optimal parameter value was selected by the RF model algorithm after automatic calculation. The modeling steps are described in Section 3.3.1.

2.3.2.3. Extreme gradient boosting. XGB is one of the best-performing machine learning algorithms (for both classification and regression problems), with higher computation speed than other algorithms (Osman *et al.* 2021). As a powerful prediction tool, XGB is a representative algorithm of the lifting method (Boosting) in ensemble algorithms. Through the ensemble algorithm, XGB models can build multiple decision trees, and then summarize the modeling results of all decision trees to obtain better regression than the single-tree model (Fan *et al.* 2021). Unlike RF, the XGB model accumulates new tree models on the basis of the single-tree model and gradually integrates many trees to form a model with strong predictive performance. Moreover, the XGB model can prevent overfitting of the model through regularization, and the sum of the weighted contributions of all decision tree predictions is provided as the final output (Zanotti *et al.* 2019; Park & Kim 2021). In this study, the ntree values of the XGB model were set as 1,000 and 2,000 and the remaining parameters (e.g., max depth) were automatically selected by XGB. The modeling steps are described in Section 3.3.1.

2.3.3. Assessment of model performance

In this study, the model was evaluated in terms of the coefficient of determination (R^2) and root mean square error (RMSE), which are the two most used indicators for evaluating the accuracy of prediction models (Moriasi *et al.* 2007; Evans *et al.* 2020; Nguyen 2022). R^2 ranges from 0 to 1, and the larger the value of R^2, the better the model fitting; on the contrary, RMSE represents the deviation between the predicted value and the observed value, and the smaller the RMSE, the better the model prediction. The calculation methods of R^2 and RMSE are shown in the following equations:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - m_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \tag{3}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y_i - m_i)^2}{N}} \tag{4}$$

where y_i refers to the observed value (the actual value), $\bar{y}$ refers to the mean of the observed value, m_i refers to the predicted value, $\overline{m_i}$ refers to the mean of the predicted value, and N refers to the total number of samples.

3. RESULTS AND DISCUSSION

3.1. Analysis of groundwater resource changes

3.1.1. Analysis of the groundwater level change

In this study, three days (June 30, July 31, and August 27) were selected from the three phases, to represent stages before, during, and after frequent heavy precipitation, and the spatial distribution diagram of the groundwater level change (Figure 3) was drawn using the ordinary Kriging interpolation analysis tool of ArcGIS. Supplementary Material, Table S1 shows the semi-variogram error analysis of ordinary Kriging interpolation. The best semi-variograms for interpolation analysis of groundwater level changes on June 30, July 31, and August 27 were exponential function, spherical function, and exponential function, respectively (Supplementary Material, Table S1).

3.1.1.1. Before frequent heavy precipitation. With reference to January 1, the overall study area was in a state of decline on June 30 (Figure 3(a)), except for the groundwater level in the central part of the study area, where the groundwater level

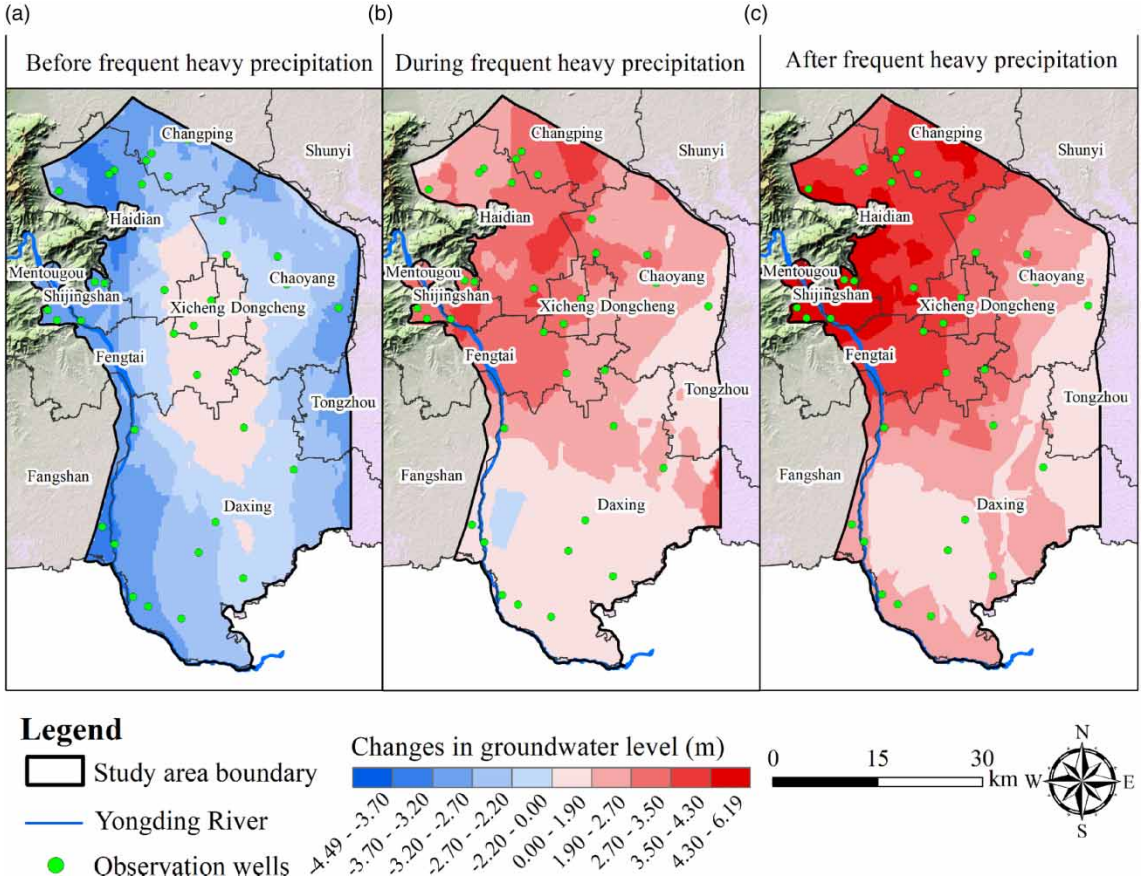


Figure 3 | Temporal change in the groundwater levels during the EWR and heavy precipitation period in 2021 (a, b, and c are the groundwater level changes on June 30, July 31, and August 27, respectively, as compared with that on January 1 when EWR was not conducted).

rebounded slightly. The groundwater level increased by 0.04–1.86 m, with an average of 0.84 m, in Haidian District, Xicheng District, Chaoyang District, Fengtai District, and Daxing District, representing 11 monitoring wells. The area of groundwater level rise covered approximately 419 km^2, accounting for 17.5% of the total area of the study area; the water level declined by 0.01–4.49 m, with an average of 1.20 m, in Changping District, Haidian District, Mentougou District, Shijingshan District, Chaoyang District, Fengtai District, and Daxing District, representing 24 monitoring wells. The area of the groundwater level decline covered approximately 1,971 km^2, accounting for 82.5% of the total area of the study area.

From January 1 to June 30 (low precipitation in six months, about 140 mm in total), the flow of EWR was approximately 4.687 m^3/s, and the amount of water replenishment was approximately 7.33×10^7 m^3. However, only 17.5% of the study area showed a slight increase in the groundwater level (Figure 3(a)), which indicates that the continuous small-flow water replenishment method has a relatively insignificant contribution to the recovery of the groundwater level. The small recovery of the groundwater level in the central part of the study area may be related to the small amount of precipitation (about 76.9 mm) in Beijing during June 1–30. Groundwater mining for agricultural irrigation in spring may be the reason for the decline of groundwater levels in the study area (Xu *et al.* 2022). The plain area of the Yongding River Basin (Beijing Section) includes a 14 km^2 area of winter wheat plantation, which is often irrigated from May to June (L. M. Wang *et al.* 2018).

3.1.1.2. Frequent heavy precipitation period. With reference to January 1, the groundwater level showed relative increases in most parts of the study area on July 31 (Figure 3(b)). The water level increased by 0.04–4.33 m (average of 1.29 m) in Changping District, Haidian District, Mentougou District, Shijingshan District, Xicheng District, Chaoyang District, Fengtai District, Fangshan District, and Daxing District, representing 25 monitoring wells. The area of groundwater level rise covered 2,362 km^2, accounting for 98.8% of the total area of the study area (2,390 km^2); the water level declined (mainly concentrated in Daxing District) by 0.02–2.11 m (average of 0.75 m) in Haidian District, Chaoyang District, and

Daxing District, representing ten monitoring wells. The area of groundwater level decline covered approximately 28 km^2, accounting for 1.2% of the total area of the study area. Comparing the changes in groundwater levels on June 30 and July 31, the groundwater level sharply rose during the frequent heavy precipitation period, and the groundwater recharge was prominent.

3.1.1.3. After frequent heavy precipitation. With reference to January 1, the overall groundwater level of the study area increased by 0.09–6.19 m (average of 1.92 m) on August 27 (Figure 3(c)). The area of the groundwater level rise covered approximately 2,390 km^2, accounting for 100% of the total area (2,390 km^2). On average, the groundwater level increased by 0.45 m between July 31 and June 30 and by 0.63 m between August 27 and July 31. The comparison results show that the groundwater level exhibited lower recovery during the frequent heavy precipitation period (July) than after the frequent heavy precipitation period (August), and the response of the groundwater level to the frequent heavy precipitation has a hysteresis effect (Qi *et al.* 2018). The increase in the groundwater level of the study area was the largest in the northwest and gradually decreased from northwest to southeast (Figure 3(b) and 3(c)). The existence of the Babaoshan fault zone in the northwest (Figure 1) and the difference in regional aquifer permeability may be the reasons for this phenomenon. The precipitation recharges the groundwater in the plain area through the Babaoshan fault zone (Wang *et al.* 2010), and monitoring wells (recharge point) farther from the fault zone receive less replenishment. From the northwest to the southeast, the aquifer sediments in the study area change from coarse to fine, and the permeability decreases. Therefore, the groundwater recharged by precipitation is more prominent in the northwest than in the southeast. In addition, the karst groundwater in the mountainous area preferentially recharges the piedmont groundwater in the northwest, which is also an important reason for the difference in the increase in the groundwater level of the study area.

3.1.2. Analysis of groundwater resource storage

From January 1 to August 27, 2021, the groundwater storage of the study area increased by 4.46×10^8 m^3, with obvious recovery (Figure 4). The best semi-variogram selected for ordinary Kriging interpolation was the exponential function (Supplementary Material, Table S1).

Changes in groundwater storage were more distinct in Changping District, Haidian District, Shijingshan District, Mentougou District, Fengtai District, and Xicheng District than in other districts (Figure 4). These six districts cover an area of approximately 804 km^2, accounting for only 33.6% of the total study area (2,390 km^2), but the increase in groundwater resources amounted to 2.92×10^8 m^3, accounting for 65.5% of the total increase (4.46×10^8 m^3). These six districts, which constitute the 'western suburban groundwater reservoir' in the plain area of Beijing (Li *et al.* 2010), are located in the upper part of the Yongding River alluvial fan, close to the Babaoshan fault zone, and have good, water-rich aquifers (consisting of single and double layers of sand and pebbles). Under the action of frequent heavy precipitation, they are more likely to store a large amount of groundwater resources.

The increases in groundwater storage in Dongcheng District, Chaoyang District, Fangshan District, Daxing District, and Tongzhou District amounted to 1.54×10^8 m^3, accounting for 34.5% of the total increase (4.46×10^8 m^3), and the increases were mainly concentrated in Daxing District (7.7×10^7 m^3) and Chaoyang District (5.4×10^7 m^3). Daxing District has a relatively small area of impervious roads, which is conducive to groundwater recharge by precipitation and surface runoff infiltration. Chaoyang District has a high level of urbanization, and anti-seepage infrastructure such as buildings and roads reduce the recharge efficiency of precipitation and surface runoff (Han *et al.* 2017; Chen *et al.* 2018; L. L. Zhang *et al.* 2020). However, districts close to the water-rich area on the west side (Haidian District, Xicheng District, and Dongcheng District) received lateral recharge from the groundwater on the west side.

3.2. Factors influencing the groundwater resource recharge

The correlation between different natural factors (X1–X6), groundwater level changes (Y1), and change in groundwater storage (Y2) is shown in Figure 5. Ground elevation (X1) was significantly positively correlated with the groundwater level change (Y1) and groundwater storage change (Y2). The topography of the study area is high in the northwest and low in the southeast. The northwest is closer to the upper part of the Yongding River alluvial fan than the southeast and the aquifer permeability coefficient is larger, which is conducive to groundwater conservation and storage (Supplementary Material, Figure S1; Ochoa *et al.* 2013). Therefore, groundwater recharge is more significant in the northwest.

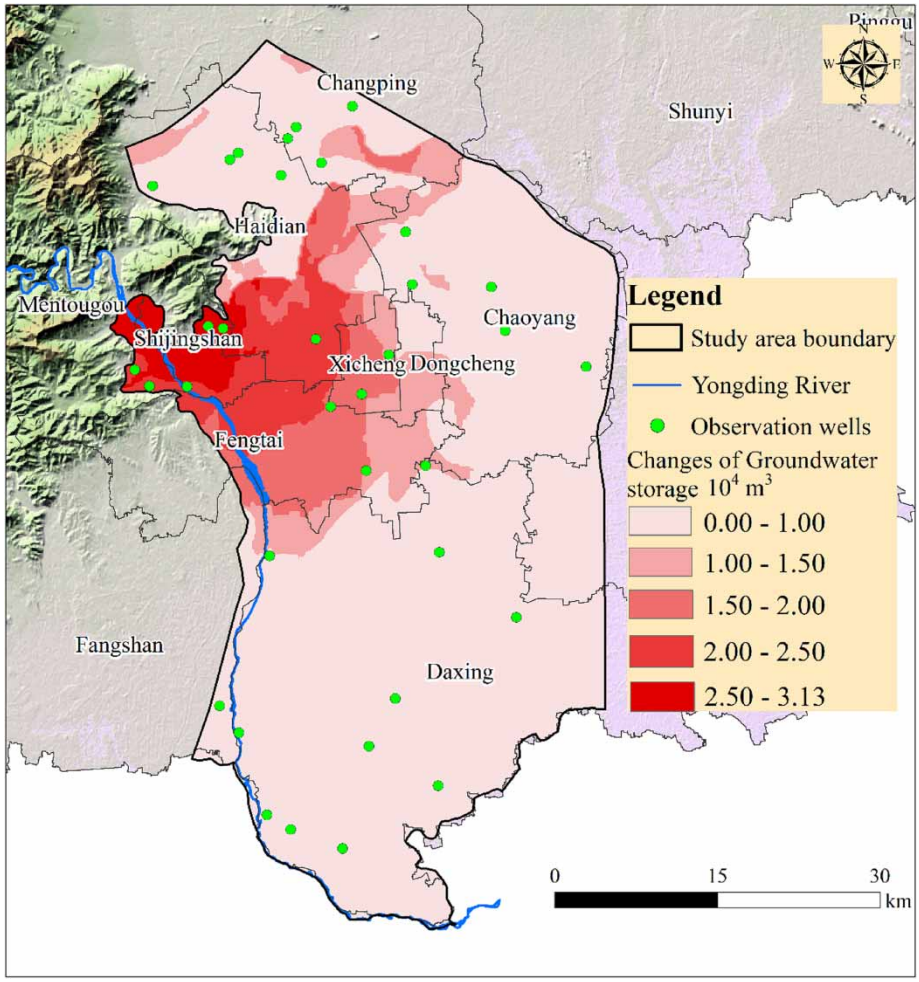


Figure 4 | Changes in groundwater storage on August 27, 2021.

The distance between the recharge point and the Yongding River fault zone (X2) showed no distinct correlation with the groundwater level change (Y1) and groundwater storage change (Y2). Before heavy precipitation in 2021 (January to June) the flow of EWR was 4.687 m^3/s. The EWR mode with continuous small flow did not significantly promote the recovery of the groundwater level on both sides of the Yongding River fault zone (Figure 3(a)). During the heavy precipitation period from June to August 2021, the groundwater level in the study area rebounded greatly, but the groundwater recharge was not significantly related to the distance between the recharge point and the Yongding River fault zone (Figures 3(b) and 3(c)). Under heavy precipitation, precipitation recharges the groundwater in the plain area through the Babaoshan fault zone, while the Yongding River fault zone may have mainly acted as a flood discharge channel. Therefore, the phenomenon of precipitation recharging groundwater in the plain area through the Yongding River fault zone was not distinct.

The distance between the recharge point and the Babaoshan fault zone (X3) was significantly negatively correlated with the groundwater level change (Y1) and groundwater storage change (Y2). The Babaoshan fault zone is distributed in the northwest of the study area, and its strike is from northeast to southwest (Figure 1). During heavy precipitation, the fault zone can hold precipitation and surface runoff, recharging the groundwater on both sides of the fault zone (Wang *et al.* 2010). Moreover, aquifers on both sides of the fault zone are rich in water, and groundwater replenishment is prominent. With the increasing distance between the recharge point and the Babaoshan fault zone, groundwater recharge becomes relatively insignificant.

Precipitation (X4) was significantly positively correlated with the groundwater level change (Y1). The heavy precipitation from June to August 2021 is the key factor in the groundwater level rise. Precipitation not only recharges the Quaternary groundwater in the plain through runoff infiltration but also through the Babaoshan fault zone in the northwest of the

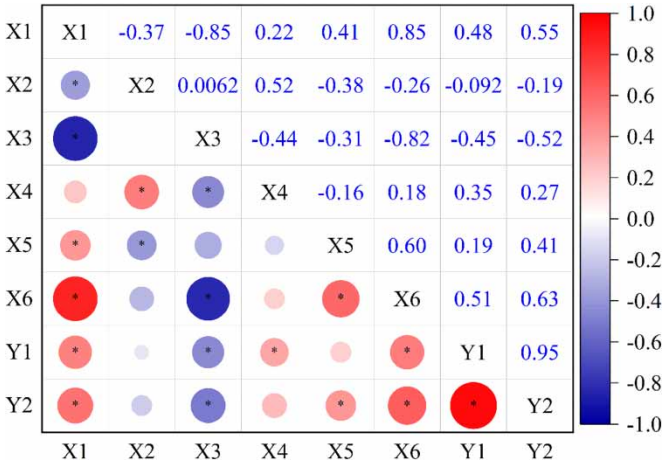


Figure 5 | Correlation between various natural factors and groundwater resource recharge (X1 is ground elevation, X2 is the distance between the recharge point and the Yongding River fault zone, X3 is the distance between the recharge point and the Babaoshan fault zone, X4 is precipitation, X5 is specific yield, X6 is aquifer permeability coefficient, Y1 is the groundwater level change, Y2 is the groundwater storage change, $n = 35$, '*' means p-value ≤ 0.05).

region (Gudmundsson 2000; Apaydin 2010). However, no significant correlation was found between precipitation (X4) and groundwater storage change (Y2). As shown in Equation (2), the change in the water level and the specific yield jointly determine the change in groundwater storage. Various increases in groundwater levels and different specific yields (Supplementary Material, Figure S2 and Figure 3) of the aquifer may lead to changes in groundwater resource storage in areas with high precipitation that is less or greater than that in areas with a low precipitation. Consequently, the calculated change in groundwater storage did not show any significant correlation with precipitation.

Specific yield (X5) showed a significant positive correlation with the groundwater storage change (Y2). Areas with large aquifer-specific yields showed relatively large changes in groundwater storage (Supplementary Material, Figure S2 and Figure 4). A large specific yield indicates that the aquifer particles are coarse and uniform in size, which is beneficial to the storage of groundwater resources (Healy & Cook 2002). No significant correlation was found between specific yield (X5) and the groundwater level change (Y1). Precipitation is the main source of groundwater recharge in the study area. In this study, the amount of precipitation corresponding to a large specific yield is uncertain, resulting in the lack of a significant correlation between specific yield and groundwater level changes.

The aquifer permeability coefficient (X6) showed a significant positive correlation with the groundwater level change (Y1) and groundwater storage change (Y2). In the study area, aquifer sediments gradually become finer from northwest to southeast, implying lower permeability. Accordingly, the infiltration of precipitation and surface runoff decreased gradually. During precipitation, aquifers with higher permeability will generate preferential flow, further promoting recharge (Hu *et al.* 2017).

In summary, heavy precipitation from June to August served as the main source of groundwater recharge in the study area during the study period. Ground elevation, aquifer-specific yield, and permeability coefficient showed a significant positive correlation with the recharge of groundwater resources. Heavy precipitation recharges the groundwater in the plain area through the Babaoshan fault zone. The distance from the recharge point to the Babaoshan fault zone is significantly negatively correlated with the groundwater recharge.

3.3. Analysis of groundwater resource prediction under heavy precipitation

In this study, the RF and XGB models were established for each monitoring well (35 in total) under the heavy precipitation from June to August, and the model with the best performance was selected as the groundwater level prediction model for each monitoring well. Referring to the results of the correlation analysis (Figure 5), precipitation, which was strongly correlated with the groundwater level, was selected as one of the input variables. EWR was excluded because its contribution to the recovery of the groundwater level was relatively insignificant (Section 3.1.1). Therefore, this study assumes that precipitation was the most important factor causing changes in groundwater levels, and the impact of the surface water system (Yongding River) on groundwater was negligible. The input variable of each model was selected in combinations of

precipitation, $t-1$ day groundwater level, and t day groundwater level. The results of systematic clustering analysis are shown in Supplementary Material, Table S2: 35 monitoring wells grouped into I, II, III, and IV categories according to the changes in water levels, and G9, G16, G20, and G32, respectively, were selected as representative monitoring wells.

3.3.1. Evaluation of the groundwater level model

The groundwater level prediction of the four representative monitoring wells is shown in Table 1 and Figure 6. For G9, G16, G20, and G32, XGB, XGB, RF, and XGB, respectively, provided the best results for the next day (i.e., $t+1$ day) groundwater

Table 1 | Model prediction of four representative monitoring wells

Well	Model	Input variables	Ratio of test set to training set	R^2	RMSE
G9	XGB	Rain, GWL ($t-1$), GWL (t)	3:7	0.993	0.037
G16	XGB	Rain, GWL ($t-1$), GWL (t)	3:7	0.851	0.117
G20	RF	Rain, GWL (t)	3:7	0.724	0.105
G32	XGB	Rain, GWL (t)	3:7	0.977	0.028

Note: R^2 and RMSE values represent the model's performance on the overall dataset.

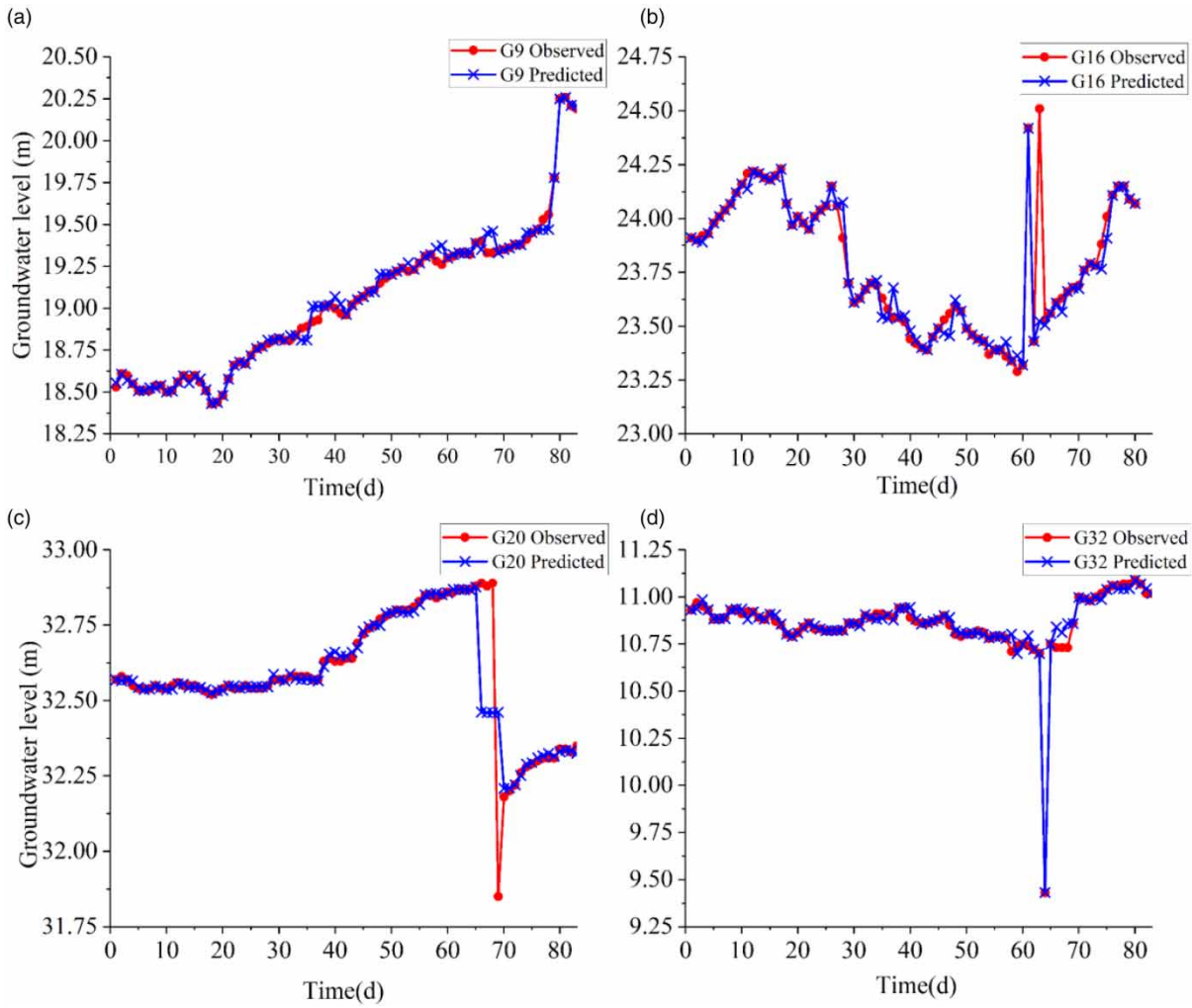


Figure 6 | Comparison of four representative monitoring wells between the actual (red line) and model-predicted (blue line) values of groundwater levels on $t+1$ day. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/wcc.2023.348.

level prediction. It is noteworthy that the input variables of each model were not exactly the same (Table 1). The model showed higher prediction ability for the groundwater level of G9 and G32, with higher R^2 (0.993 and 0.977 for G9 and G32, respectively) and lower RMSE (0.037 and 0.028 m for G9 and G32, respectively) (Table 1, Figure 6(a) and 6(d)). For G16 and G20, the models showed relatively general performances, with R^2 of 0.851 and 0.724, and RMSE of 0.117 and 0.105 m, respectively (Table 1, Figure 6(b) and 6(c)), which is related to the existence of continuous outliers or cliff-like anomalies in the water level data. For instance, the groundwater level of G16 showed continuous large changes from August 7 to August 11 (23.32, 24.42, 23.43, 24.51, and 23.53 m, respectively), and the groundwater level of G20 changed from 32.89 m on August 12 to 31.85 m on August 13, followed by an upward trend. The abnormal water level data may be attributable to errors of the water level monitoring instrument (Gribovszki *et al.* 2013). Additionally, the machine learning model was applied to all 35 monitoring wells and the predictive performance of the model was discussed.

The predicted $t+1$ day groundwater levels of the 35 monitoring wells are shown in Supplementary Material, Table S3. Overall, the R^2 of 30 models was greater than 0.9, and the R^2 of 33 models was greater than 0.8. The machine learning model (XGB and RF) developed in this study has high groundwater-level prediction accuracy under heavy precipitation conditions. Among the water level prediction models of the 35 monitoring wells, XGB showed the best performance in 30 wells (G1–G3, G5–G9, G13–G19, G21–G35), and RF showed the best performance in five wells (G4, G10–G12, G20). The XGB model showed higher prediction performance than the RF model. The R^2 value of the groundwater level model for G3, G16, G17, G20, and G27 remained below 0.9 at 0.873, 0.851, 0.759, 0.724, and 0.818, respectively. There are a few abnormal values or cliff-type abnormal values in the groundwater level data of these five monitoring wells. The low R^2 of some models is due to the water level data outliers, which may be attributable to the failure of water level monitoring instruments (Gribovszki *et al.* 2013).

3.3.2. Prediction of groundwater storage

The observed and predicted values of the groundwater level on August 27, 2021, are shown in Supplementary Material, Table S3. Among the 35 groups of actual and predicted data, the maximum prediction error was 8.9 cm, the minimum prediction error was 0.1 cm, and the average prediction error was 1.78 cm. The prediction error was less than 1 cm for 17 groups (accounting for 48.6%), less than 3 cm for 29 groups (accounting for 82.9%), and greater than 3 cm for only six groups

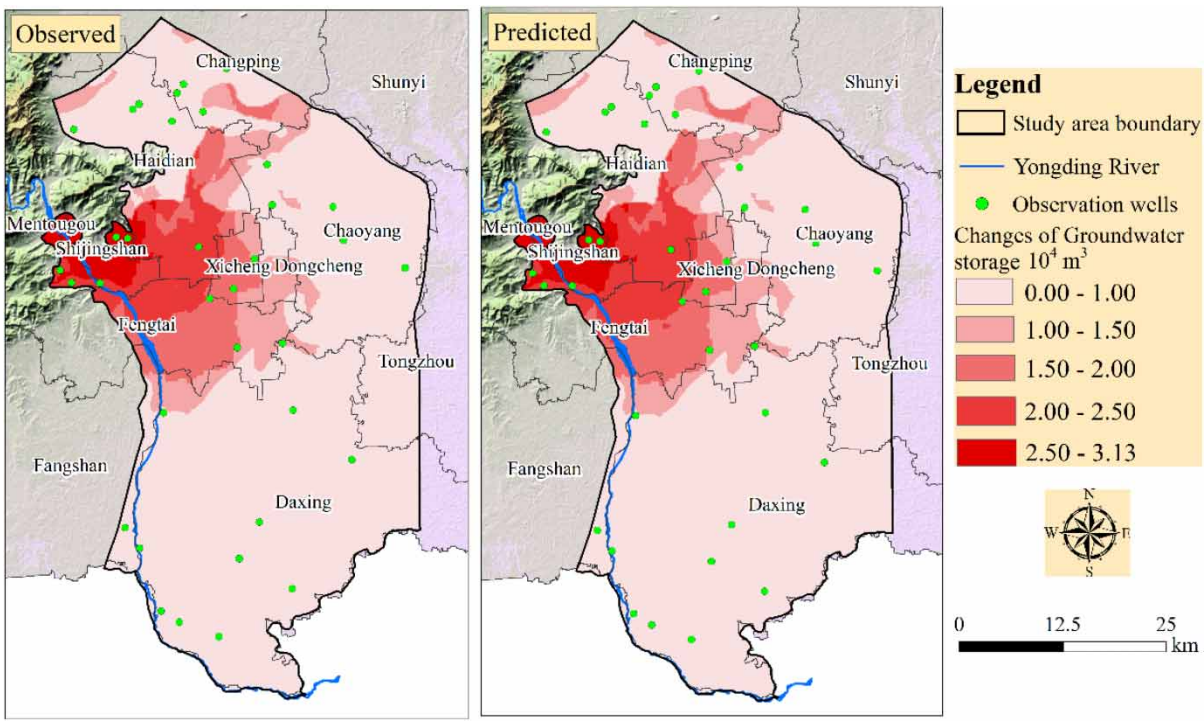


Figure 7 | Comparison of actual and predicted changes in the groundwater storage on August 27, 2021.

(accounting for 17.1%). The water level fluctuation method was used to convert the groundwater level change predicted by the model (the difference between the groundwater levels on August 27 and January 1) into the groundwater storage change. The change in groundwater storage based on the predicted value of the model was highly consistent with the actual change in groundwater storage on August 27 (Figure 7). The increase in groundwater storage calculated using the model predicted value was 4.52×10^8 m^3, deviating from the actual calculated value (4.46×10^8 m^3) by 0.6×10^7 m^3. Accordingly, the prediction bias is only 1.3% (within the acceptable range). The machine learning model developed in this study has high prediction accuracy and can be used as a tool to predict changes in groundwater resources under heavy precipitation.

4. CONCLUSIONS

In this study, a coupling model integrated with the water level fluctuation method, correlation analysis, and machine learning method was first proposed to analyze the changing trend of groundwater under the influence of heavy precipitation and EWR. The results of the groundwater level analysis showed that the EWR with the continuous small flow (4.687 m^3/s) did not significantly contribute to the recovery of the groundwater level in the study area, whereas heavy precipitation significantly contributed to the recovery of the groundwater level. The calculation results of the water level fluctuation method showed that the groundwater storage increased by 4.46×10^8 m^3 after heavy precipitation. Correlation analysis found that ground elevation, aquifer-specific yield, and the permeability coefficient showed a significant positive correlation with groundwater resource replenishment. The distance from the recharge point to the Babaoshan fault zone after heavy precipitation was significantly negatively correlated with the recharge of groundwater resources, indicating that the fault zone may serve as a migration channel for precipitation. The performance assessment results of the machine learning models showed that XGB had better predictive performance than RF in most cases. The groundwater resource prediction model developed based on XGB and RF had a prediction error of 0.6×10^7 m^3 (prediction bias of 1.3%), and the model performance is relatively good.

The more groundwater monitoring wells that are involved in Kriging interpolation, the more real the regional groundwater changes will be. Therefore, we recommend that the local government increase the number of groundwater monitoring wells in areas with large changes in the groundwater level and reserves to grasp more accurate groundwater changes in extreme conditions (e.g., heavy precipitation and EWR conditions). In addition, it is also recommended to use the groundwater level data with a longer time-range to establish a groundwater prediction model based on XGB and RF, so as to obtain the changing trend in groundwater in extreme cases in advance and to make timely management decisions. This study can help to better understand the changing trend in groundwater resources under the influence of heavy precipitation and ecological replenishment and has guiding significance for the advanced management and risk prevention of groundwater resources.

ACKNOWLEDGEMENTS

This work was supported by the Beijing Municipal Science and Technology Commission Project (Z191100006919001) and the Fundamental Research Business Special Project of the Central Public Welfare Scientific Research Institutes (JY-2013YQ06072101), and the National Key R&D Plan of the In Situ Real Time Online Monitoring Technology and Equipment for Typical Organic Pollutants in Groundwater (SQ2022YFC3700182). The authors are grateful for the valuable comments and suggestions given by the editors and the anonymous reviewers.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Adnan, R. M., Dai, H.-L., Mostafa, R. R., Islam, A. M. T., Kisi, O., Heddam, S. & Zounemat-Kermani, M. 2023 Modelling groundwater level fluctuations by ELM merged advanced metaheuristic algorithms using hydroclimatic data. *Geocarto International* **38** (1), 2158951.

Agarwal, V., Akyilmaz, O., Shum, C. K., Feng, W., Yang, T.-Y., Forootan, E., Syed, T. H., Haritashya, U. K. & Uz, M. 2023 Machine learning based downscaling of GRACE-estimated groundwater in Central Valley, California. *Science of the Total Environment* **865**, 161138.

Alizamir, M., Kisi, O. & Zounemat-Kermani, M. 2018 Modelling long-term groundwater fluctuations by extreme learning machine using hydro-climatic data. *Hydrological Sciences Journal* **63** (1), 63–73.

Amaranto, A., Munoz-Arriola, F., Solomatine, D. P. & Corzo, G. 2019 A spatially enhanced data-driven multimodel to improve semiseasonal groundwater forecasts in the High Plains Aquifer, USA. *Water Resources Research* **55** (7), 5941–5961.

Apaydin, A. 2010 Relation of tectonic structure to groundwater flow in the Beypazari region, NW Anatolia, Turkey. *Hydrogeology Journal* **18** (6), 1343–1356.

Artusi, R., Verderio, P. & Marubini, E. 2002 Bravais-Pearson and Spearman correlation coefficients: meaning, test of hypothesis and confidence interval. *International Journal of Biological Markers* **17** (2), 148–151.

Biau, G. & Scornet, E. 2016 A random forest guided tour. *TEST* **25** (2), 197–227.

Brédy, J., Gallichand, J., Celicourt, P. & Gumiere, S. J. 2020 Water table depth forecasting in cranberry fields using two decision-tree-modeling approaches. *Agricultural Water Management* **233**, 106090.

Breiman, L. 2001 Random forests. *Machine Learning* **45** (1), 5–32.

Chen, Z., Jiang, W. G., Wang, W. J., Deng, Y., He, B. & Jia, K. 2018 The impact of precipitation deficit and urbanization on variations in water storage in the Beijing-Tianjin-Hebei urban agglomeration. *Remote Sensing* **10** (1), 4.

Chen, C., He, W., Zhou, H., Xue, Y. R. & Zhu, M. D. 2020 A comparative study among machine learning and numerical models for simulating groundwater dynamics in the Heihe River Basin, northwestern China. *Scientific Reports* **10** (1), 3904.

Crosbie, R. S., Doble, R. C., Turnadge, C. & Taylor, A. R. 2019 Constraining the magnitude and uncertainty of specific yield for use in the water table fluctuation method of estimating recharge. *Water Resources Research* **55** (8), 7343–7361.

Dangar, S., Asoka, A. & Mishra, V. 2021 Causes and implications of groundwater depletion in India: a review. *Journal of Hydrology* **596**, 126103.

Delottier, H., Pryet, A., Lemieux, J. M. & Dupuy, A. 2018 Estimating groundwater recharge uncertainty from joint application of an aquifer test and the water-table fluctuation method. *Hydrogeology Journal* **26** (7), 2495–2505.

El Bilali, A., Taleb, A. & Brouziyne, Y. 2021 Comparing four machine learning model performances in forecasting the alluvial aquifer level in a semi-arid region. *Journal of African Earth Sciences* **181**, 104244.

Evans, S., Williams, G. P., Jones, N. L., Ames, D. P. & Nelson, E. J. 2020 Exploiting Earth observation data to impute groundwater level measurements with an extreme learning machine. *Remote Sensing* **12** (12), 2044.

Fagbohun, B. J. 2018 Integrating GIS and multi-influencing factor technique for delineation of potential groundwater recharge zones in parts of Ilesha schist belt, southwestern Nigeria. *Environmental Earth Sciences* **77** (3), 69.

Fan, J. L., Zheng, J., Wu, L. F. & Zhang, F. C. 2021 Estimation of daily maize transpiration using support vector machines, extreme gradient boosting, artificial and deep neural networks models. *Agricultural Water Management* **245**, 106547.

Gleeson, T., Villholth, K., Taylor, R., Perrone, D. & Hyndman, D. 2019 Groundwater: a call to action. *Nature* **576** (7786), 213–213.

Gong, C. C., Wang, W. K., Zhang, Z. Y., Wang, H., Luo, J. & Brunner, P. 2020 Comparison of field methods for estimating evaporation from bare soil using lysimeters in a semi-arid area. *Journal of Hydrology* **590**, 125334.

Goodarzi, M., Abedi-Koupai, J., Heidarpour, M. & Safavi, H. R. 2015 Evaluation of the effects of climate change on groundwater recharge using a hybrid method. *Water Resources Management* **30** (1), 133–148.

Gribovszki, Z., Kalicz, P. & Szilágyi, J. 2013 Does the accuracy of fine-scale water level measurements by vented pressure transducers permit for diurnal evapotranspiration estimation? *Journal of Hydrology* **488**, 166–169.

Gudmundsson, A. 2000 Active fault zones and groundwater flow. *Geophysical Research Letters* **27** (18), 2993–2996.

Guzman, S. M., Paz, J. O., Tagert, M. L. M. & Mercer, A. E. 2019 Evaluation of seasonally classified inputs for the prediction of daily groundwater levels: NARX networks vs support vector machines. *Environmental Modeling & Assessment* **24** (2), 223–234.

Han, D., Currell, M. J., Cao, G. & Hall, B. 2017 Alterations to groundwater recharge due to anthropogenic landscape change. *Journal of Hydrology* **554**, 545–557.

Hazra, A. & Gogtay, N. 2016 Biostatistics series module 6: correlation and linear regression. *Indian Journal of Dermatology* **61** (6), 593–601.

Healy, R. W. & Cook, P. G. 2002 Using groundwater levels to estimate recharge. *Hydrogeology Journal* **10** (1), 91–109.

Hu, K. X., Awange, J. L., Khandu, Forootan, E., Goncalves, R. M. & Fleming, K. 2017 Hydrogeological characterisation of groundwater over Brazil using remotely sensed and model products. *Science of the Total Environment* **599–600**, 372–386.

Hu, H. Z., Mao, X. M. & Yang, Q. 2019 Development of a groundwater flow and reactive solute transport model in the Yongding River alluvial fan, China. *Frontiers of Earth Science* **13** (2), 371–384.

Huang, F. M., Huang, J. S., Jiang, S. H. & Zhou, C. B. 2017 Prediction of groundwater levels using evidence of chaos and support vector machine. *Journal of Hydroinformatics* **19** (4), 586–606.

Ji, Z., Cui, Y., Zhang, S., Chao, W. & Shao, J. 2021 Evaluation of the impact of ecological water supplement on groundwater restoration based on numerical simulation: a case study in the section of Yongding River, Beijing Plain. *Water* **13** (21), 3059.

Jiang, Z. W., Yang, S. H., Liu, Z. Y., Xu, Y., Shen, T., Qi, S. T., Pang, Q. Q., Xu, J. Z., Liu, F. P. & Xu, T. 2022 Can ensemble machine learning be used to predict the groundwater level dynamics of farmland under future climate: a 10-year study on Huaibei Plain. *Environmental Science and Pollution Research* **29** (29), 44653–44667.

Jing, H., He, X., Tian, Y., Lancia, M., Cao, G. L., Crivellari, A., Guo, Z. L. & Zheng, C. M. 2023 Comparison and interpretation of data-driven models for simulating site-specific human-impacted groundwater dynamics in the North China Plain. *Journal of Hydrology* **616**, 128751.

Kayhomayoon, Z., Ghordoyee-Milan, S., Jaafari, A., Arya-Azar, N., Melesse, A. M. & Moghaddam, H. K. 2022 How does a combination of numerical modeling, clustering, artificial intelligence, and evolutionary algorithms perform to predict regional groundwater levels? *Computers and Electronics in Agriculture* **203**, 107482.

Li, Y., Shao, J., Ye, C., Xing, G. & Cui, Y. 2010 A discussion on the patterns of groundwater reservoir in the west suburb of Beijing. *Earth Science Frontiers* **17** (6), 192–199.

Li, B., Rodell, M. & Famiglietti, J. S. 2015 Groundwater variability across temporal and spatial scales in the central and northeastern US. *Journal of Hydrology* **525**, 769–780.

Liu, Y. C., Fei, Y. H., Meng, S. H. & Cui, X. X. 2019 Hydrochemical evolution of groundwater and soils in the water-level-fluctuation zone. *Environmental Earth Sciences* **78** (22), 647.

Luo, Z. R., Zhao, S. Q., Wu, J., Zhang, Y. X., Liu, P. B. & Jia, R. T. 2019 The influence of ecological restoration projects on groundwater in Yongding River Basin in Beijing China. *Water Supply* **19** (8), 2391–2399.

Lv, Y. X., Jiang, J., Chen, L., Hu, W. & Jiang, Y. J. 2022 Elaborate simulation and predication of the tunnel drainage effect on karst groundwater field and discharge based on Visual MODFLOW. *Journal of Hydrology* **612**, 128023.

Majumdar, S., Smith, R., Butler, J. J. & Lakshmi, V. 2020 Groundwater withdrawal prediction using integrated multitemporal remote sensing data sets and machine learning. *Water Resources Research* **56** (11), e2020WR028059.

Miro, M. E., Groves, D., Tincher, B., Syme, J., Tanverakul, S. & Catt, D. 2021 Adaptive water management in the face of uncertainty: integrating machine learning, groundwater modeling and robust decision making. *Climate Risk Management* **34**, 100383.

Mohammed, K. S., Shabanlou, S., Rajabi, A., Yosefvand, F. & Izadbakhsh, M. A. 2023 Prediction of groundwater level fluctuations using artificial intelligence-based models and GMS. *Applied Water Science* **13** (2), 54.

Mohapatra, J. B., Jha, P. Y., Jha, M. K. & Biswal, S. 2021 Efficacy of machine learning techniques in predicting groundwater fluctuations in agro-ecological zones of India. *Science of the Total Environment* **785**, 147319.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. L. 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE* **50** (3), 885–900.

Müller, J., Park, J., Sahu, R., Varadharajan, C., Arora, B., Faybishenko, B. & Agarwal, D. 2021 Surrogate optimization of deep neural networks for groundwater predictions. *Journal of Global Optimization* **81** (1), 203–231.

Natarajan, N. & Sudheer, C. 2020 Groundwater level forecasting using soft computing techniques. *Neural Computing & Applications* **32** (12), 7691–7708.

Nguyen, H. D. 2022 Spatial modeling of flood hazard using machine learning and GIS in Ha Tinh province, Vietnam. *Journal of Water & Climate Change* **14** (1), 200–222.

Ochoa, C. G., Fernald, A. G., Guldan, S. J., Tidwell, V. C. & Shukla, M. K. 2013 Shallow aquifer recharge from irrigation in a semiarid agricultural valley in New Mexico. *Journal of Hydrologic Engineering* **18** (10), 1219–1230.

Osman, A. I. A., Ahmed, A. N., Chow, M. F., Huang, Y. F. & El-Shafie, A. 2021 Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Engineering Journal* **12** (2), 1545–1556.

Park, S. & Kim, J. 2021 The predictive capability of a novel ensemble tree-based algorithm for assessing groundwater potential. *Sustainability* **13** (5), 2459.

Pastore, N., Cherubini, C., Doglioni, A., Giasi, C. I. & Simeone, V. 2020 Modelling of the complex groundwater level dynamics during episodic rainfall events of a surficial aquifer in southern Italy. *Water* **12** (10), 2196.

Pelletier, V., Gallichand, J., Gumiere, S., Pepin, S. & Caron, J. 2015 Water table control for increasing yield and saving water in cranberry production. *Sustainability* **7** (8), 10602–10619.

Pham, Q. B., Kumar, M., Di Nunno, F., Elbeltagi, A., Granata, F., Islam, A. R. M. T., Talukdar, S., Nguyen, X. C., Ahmed, A. N. & Anh, D. T. 2022 Groundwater level prediction using machine learning algorithms in a drought-prone area. *Neural Computing & Applications* **34** (13), 10751–10773.

Qi, P., Zhang, G. X., Xu, Y. J., Wang, L., Ding, C. C. & Cheng, C. Y. 2018 Assessing the influence of precipitation on shallow groundwater table response using a combination of singular value decomposition and cross-wavelet approaches. *Water* **10** (5), 598.

Rahman, A. T. M. S., Hosono, T., Quilty, J. M., Das, J. & Basak, A. 2020 Multiscale groundwater level forecasting: coupling new machine learning approaches with wavelet transforms. *Advances in Water Resources* **141**, 103595.

Rao, P. Z., Wang, Y. C., Liu, Y., Wang, X. Y., Hou, Y. K., Pan, S. B., Wang, F. & Zhu, D. S. 2022 A comparison of multiple methods for mapping groundwater levels in the Mu Us Sandy Land, China. *Journal of Hydrology: Regional Studies* **43**, 101189.

Ren, X., Zhu, B., Liu, M., Zhang, Y., He, Z. & Rioual, P. 2018 Mechanism of groundwater recharge in the middle-latitude desert of eastern Hunshandake, China: diffuse or focused recharge? *Hydrogeology Journal* **27** (2), 761–783.

Ruidas, D., Pal, S. C., Islam, A. R. M. T. & Saha, A. 2021 Characterization of groundwater potential zones in water-scarce hardrock regions using data driven model. *Environmental Earth Sciences* **80** (24), 809.

Schinke, R., Neubert, M., Hennersdorf, J., Stodolny, U., Sommer, T. & Naumann, T. 2012 Damage estimation of subterranean building constructions due to groundwater inundation – the GIS-based model approach GRUWAD. *Natural Hazards and Earth System Sciences* **12** (9), 2865–2877.

Schreiner-McGraw, A. P., Ajami, H. & Vivoni, E. R. 2019 Extreme weather events and transmission losses in arid streams. *Environmental Research Letters* **14** (8), 084002.

Sun, Y. B., Wendi, D., Kim, D. E. & Liong, S. Y. 2016 Technical note: Application of artificial neural networks in groundwater table forecasting – a case study in a Singapore swamp forest. *Hydrology and Earth System Sciences* **20** (4), 1405–1412.

Sun, K. N., Hu, L. T., Guo, J. L., Yang, Z. Q., Zhai, Y. Z. & Zhang, S. Q. 2021 Enhancing the understanding of hydrological responses induced by ecological water replenishment using improved machine learning models: a case study in Yongding River. *Science of the Total Environment* **768**, 145489.

Sun, K. N., Hu, L. T., Sun, J. C., Zhai, Y. Z., Zhang, S. Q. & Cao, X. Y. 2023 Quantifying the contribution of ecological water replenishment on aquifer recovery using a refined groundwater model. *Science of the Total Environment* **857**, 159216.

Taylor, R. G., Todd, M. C., Kongola, L., Maurice, L., Nahozya, E., Sanga, H. & MacDonald, A. M. 2012 Evidence of the dependence of groundwater resources on extreme rainfall in East Africa. *Nature Climate Change* **3** (4), 374–378.

van Roosmalen, L., Christensen, B. S. B. & Sonnenborg, T. O. 2007 Regional differences in climate change impacts on groundwater and stream discharge in Denmark. *Vadose Zone Journal* **6** (3), 554–571.

Wang, X. S., Wang, G. C. & Dong, J. N. 2010 Hydro-mechanical coupling model of shallow rocks in the Babaoshan Fault. *Earth Science Frontiers* **17** (6), 141–146.

Wang, L. M., Liu, J., Yang, F. G., Yang, L. B., Yao, B. M. & Wang, X. L. 2018 Acquisition of winter wheat area in the Beijing-Tianjin-Hebei Region with GF-1 Satellite Data. *Acta Agronomica Sinica* **44** (5), 762–773.

Wang, X. H., Liu, T. L., Zheng, X. L., Peng, H., Xin, J. & Zhang, B. 2018 Short-term prediction of groundwater level using improved random forest regression with a combination of random features. *Applied Water Science* **8** (5), 125.

Wang, D. F., Xu, H. D., Wang, L., Wu, X. & Sun, H. Y. 2020 Statistical analyses of the effect of a drainage tunnel on landslide hydrogeological characteristics. *Hydrological Processes* **34** (11), 2418–2432.

Wang, J. C., Huo, A. D., Zhang, X. Z. & Lu, Y. D. 2020 Prediction of the response of groundwater recharge to climate changes in Heihe River basin, China. *Environmental Earth Sciences* **79** (1), 13.

Wei, Z. L., Lu, Q., Sun, H. Y. & Shang, Y. Q. 2019 Estimating the rainfall threshold of a deep-seated landslide by integrating models for predicting the groundwater level and stability analysis of the slope. *Engineering Geology* **253**, 14–26.

Wu, C. C., Zhang, X. Q., Wang, W. J., Lu, C. P., Zhang, Y., Qin, W., Tick, G. R., Liu, B. & Shu, L. C. 2021 Groundwater level modeling framework by combining the wavelet transform with a long short-term memory data-driven model. *Science of the Total Environment* **783**, 146948.

Xing, L. T., Wu, Q. A., Ye, C. H. & Ye, N. 2010 Groundwater environmental capacity and its evaluation index. *Environmental Monitoring and Assessment* **169** (1–4), 217–227.

Xu, C. C., Sun, Y., Shi, B. W., Wang, X. J., Li, R., Li, M. X., Xi, B. D. & Feng, C. P. 2022 Study on the processes influencing and importance of ecological water replenishment for groundwater resources: a case study in Yongding River. *Water* **14** (5), 828.

Yadav, B., Gupta, P. K., Patidar, N. & Himanshu, S. K. 2020 Ensemble modelling framework for groundwater level prediction in urban areas of India. *Science of the Total Environment* **712**, 135539.

Yenehun, A., Nigate, F., Belay, A. S., Desta, M. T., Van Camp, M. & Walraevens, K. 2020 Groundwater recharge and water table response to changing conditions for aquifers at different physiography: the case of a semi-humid river catchment, northwestern highlands of Ethiopia. *Science of the Total Environment* **748**, 142243.

Yu, H. L. & Lin, Y. C. 2015 Analysis of space–time non-stationary patterns of rainfall–groundwater interactions by integrating empirical orthogonal function and cross wavelet transform methods. *Journal of Hydrology* **525**, 585–597.

Yu, Y. L., Ma, M. Y., Zheng, F. D., Liu, L. C., Zhao, N. N., Li, X. X., Yang, Y. M. & Guo, J. 2017 Spatio-temporal variation and controlling factors of water quality in Yongding River replenished by reclaimed water in Beijing, North China. *Water* **9** (7), 453.

Zanotti, C., Rotiroti, M., Sterlacchini, S., Cappellini, G., Fumagalli, L., Stefania, G. A., Nannucci, M. S., Leoni, B. & Bonomi, T. 2019 Choosing between linear and nonlinear models and avoiding overfitting for short and long term groundwater level forecasting in a linear system. *Journal of Hydrology* **578**, 124015.

Zhang, M. L., Hu, L. T., Yao, L. L. & Yin, W. J. 2018 Numerical studies on the influences of the South-to-North Water Transfer Project on groundwater level changes in the Beijing Plain, China. *Hydrological Processes* **32** (12), 1858–1873.

Zhang, L. L., Wang, C. Y., Liang, G. X., Cui, Y. L. & Zhang, Q. L. 2020 Influence of land use change on hydrological cycle: application of SWAT to Su-Mi-Huai Area in Beijing, China. *Water* **12** (11), 3164.

Zhang, Z. Y., Wang, W. K., Gong, C. C. & Zhang, M. 2020 A comparison of methods to estimate groundwater recharge from bare soil based on data observed by a large-scale lysimeter. *Hydrological Processes* **34** (13), 2987–2999.

Zhou, Y. X., Dong, D. W., Liu, J. R. & Li, W. P. 2013 Upgrading a regional groundwater level monitoring network for Beijing Plain, China. *Geoscience Frontiers* **4** (1), 127–138.

Zhou, T., Wen, X. H., Feng, Q., Yu, H. J. & Xi, H. Y. 2023 Bayesian model averaging ensemble approach for multi-time-ahead groundwater level prediction combining the GRACE, GLEAM, and GLDAS data in arid areas. *Remote Sensing* **15** (1), 188.

Zhu, Z. W., Zhou, M. R., Hu, F., Wang, S. H., Ma, J. H., Gao, B., Bian, K. & Lai, W. H. 2023 A day-ahead industrial load forecasting model using load change rate features and combining FA-ELM and the AdaBoost algorithm. *Energy Reports* **9**, 971–981.

First received 23 August 2022; accepted in revised form 4 May 2023. Available online 22 May 2023