

Study on wavelet multi-scale analysis and prediction of landslide groundwater

Tianlong Wang ^a, Dingmao Peng^b, Xu Wang^a, Bin Wu^c, Rui Luo^a, Zhaowei Chu^a and Hongyue Sun ^{a,*}

^a Ocean College, Zhejiang University, Zhoushan 316000, China

^b Zhejiang Institute of Communications Co., Ltd, Hangzhou 310000, China

^c ZCCC International Engineering Co., Ltd, Hangzhou 310000, China

*Corresponding author. E-mail: shy@zju.edu.cn

 TW, 0000-0002-0598-5481; HS, 0000-0002-2267-305X

ABSTRACT

Current groundwater prediction models often exhibit low accuracy and complex parameter adjustment. To tackle these limitations, a novel prediction model, called improved Aquila optimizer bi-directional long-term and short-term memory (IAO-BiLSTM) network, is proposed. IAO-BiLSTM optimizes the hyperparameters of the BiLSTM network using an IAO algorithm. IAO incorporates three novel enhancements, including population initialization, population updating, and global best individual updating, to overcome the drawbacks of current optimization algorithms. Before making predictions, the challenge posed by the highly nonlinear and non-stationary characteristics of groundwater level signals was addressed through the application of a wavelet multi-scale analysis method. Using a landslide site in Zhejiang Province as an example, a monitoring system is established, and continuous wavelet transform, cross-wavelet transform, and wavelet coherence analysis are employed to perform multi-scale feature analysis on a 2-year dataset of rainfall and groundwater depth. The findings reveal that the groundwater depth of monitoring holes exhibits similar high energy resonating periods and phase relationships, strongly correlating with rainfall. Subsequently, IAO-BiLSTM is employed to predict groundwater depth, and its results are compared with seven popular machine learning regression models. The results demonstrate that IAO-BiLSTM achieves the highest accuracy, as evidenced by its root mean squared error of 0.25.

Key words: Aquila optimizer algorithm, BiLSTM, groundwater depth, landslide, rainfall, wavelet multi-scale analysis

HIGHLIGHTS

- The wavelet multi-scale analysis method overcomes the limitation of the traditional qualitative analysis method.
- The wavelet multi-scale analysis method successfully reveals the multi-scale characteristics and the most significant influencing factors.
- An improved Aquila optimizer bi-directional long-term and short-term memory (IAO-BiLSTM) network can predict groundwater level reliably and effectively.
- IAO-BiLSTM beats the current seven hot prediction models and showed good performance.

1. INTRODUCTION

Groundwater changes can be an important contributing factor to landslides, especially in China, where 90% of these slope failures are linked to groundwater. When groundwater levels rise and fall rapidly, it compromises the shear strength of the slope rock and soil, intensifying the risk of landslides (Seifollahi-Aghmiuni *et al.* 2022). Therefore, it is of utmost importance to understand the dynamic patterns of groundwater levels in landslide-prone areas and to predict their trends efficiently and accurately.

As a complex nonlinear dynamic system, the change in the groundwater level is usually related to rainfall, reservoir water level, engineering activities, and the geological conditions of the landslide itself. It is of great significance to reveal the influencing factors of groundwater level change and to analyze the real-time monitoring and prediction of the groundwater depth potentially prone to landslides (Cao *et al.* 2020). Based on the groundwater monitoring data of the Shuping landslide in the Three Gorges Reservoir area of China, Zhao *et al.* (2016) analyzed the response of landslide groundwater, rainfall, and reservoir water level, and then constructed a classified regression tree model to predict the groundwater level. Based on the monitoring data of groundwater level, rainfall, and reservoir water level of the Baijiapu landslide in China from 2007 to

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

2013, Duan *et al.* (2019) qualitatively analyzed the characteristics, significant relationship, and lag relationship of groundwater level, concluding that rainfall is the most significant influence, which provides a basis for the prediction of groundwater level. Wei *et al.* (2020) conducted a qualitative analysis of the correlation between groundwater level, rainfall, and surface displacement in a deep landslide in China, utilizing monitoring data. They concluded that the movement of the landslide is largely dependent on the groundwater, which is, in turn, affected by rainfall. From this, they were able to predict the fluctuations in groundwater levels. Although monitoring data can be used to qualitatively assess the correlation between groundwater and its contributing factors in the aforementioned study, it is difficult to systematically quantify highly nonlinear and non-stationary signals. The lack of statistical analysis and rigor further impede the ability to provide sufficient support for predicting groundwater levels (Truong *et al.* 2021).

In recent years, researchers have increasingly employed advanced mathematical methods to investigate the dynamic changes and correlations of groundwater, among which wavelet analysis is one of the most commonly used approaches. Based on the wavelet theory, this method can analyze the characteristic and correlation relationships of two time series in the time–frequency domain. Wavelet analysis has gained wide acceptance and substantial utilization in diverse fields, including meteorology, hydrology, astronomy, and medicine (Kainthura & Sharma 2022; Su *et al.* 2023). In this study, a long-term and multi-point field monitoring approach was employed to monitor the dynamic changes of groundwater in slope areas, and the wavelet theory was used to conduct a multi-scale analysis of groundwater depth at landslide points (Sharma *et al.* 2021). By utilizing methods such as continuous wavelet transform (CWT), cross-wavelet transform (XWT), and wavelet coherence analysis (WTC), the response characteristics, periodic scale characteristics, dominant resonance period, significant periods, lag relationships, and multi-scale coherence between groundwater depth and influencing factors were revealed. The utilization of this approach enabled a comprehensive investigation of the multi-scale characteristics, evolution patterns, and influence relationships between groundwater depth and influencing factors. Such findings could potentially offer new research directions for studying the dynamic changes of groundwater level time series and serve as a foundation for selecting influential factors in predicting groundwater levels.

Groundwater levels are highly nonlinear and volatile, making accurate prediction difficult (Vadiati *et al.* 2022). Though deterministic models, such as water balance, hydrogeology analogy, analytical methods, numerical simulations, and physical simulations, can provide reliable predictions when the geological conditions are not complex and involve a single water layer, they are not well suited to mid-term and long-term forecasts (Nikvar Hassani *et al.* 2016). Likewise, traditional stochastic models like time series, gray models, random differential equations, and fuzzy mathematics present their difficulties, such as difficulty in model identification, limited ability to capture the nonlinear dependence of historical data, and a large number of parameters to be estimated (Wang *et al.* 2019). The stochastic model based on machine learning is far superior to other methods in terms of flexibility, stability, and accuracy, making it the most widely used and effective prediction technique for groundwater levels (Habibi *et al.* 2019; Li *et al.* 2023). Considering the influencing factors of rainfall and reservoir water level, Cao *et al.* (2020) leveraged the genetic algorithm-support vector machine (GA-SVM) to predict Tangjiao landslide's groundwater levels in the Three Gorges Reservoir area of China, demonstrating the GA-backpropagation (BP) neural networks as a reliable prediction method. Wei *et al.* (2020), while taking rainfall factors into account, adopted the particle swarm optimization (PSO)-SVM and the physical seepage model to assess a landslide in China and found PSO-SVM to yield higher accuracy than the physical seepage model. Furthermore, Huang *et al.* (2017) compared the prediction accuracy of PSO-SVM based on the chaos theory, PSO-SVM based on linear correlation analysis, and BP based on the chaos theory and discovered that the chaotic PSO-SVM model performed better than the other two due to the chaotic characteristics of groundwater level.

Optimization algorithms such as PSO and GA can be used to optimize the hyperparameters of machine learning models like BP and SVM and have been successfully applied to the prediction of landslide groundwater. However, these algorithms have certain drawbacks, such as a limited search space, susceptibility to local optima, overfitting, slow convergence, weak generalization performance, and limited learning ability for time series features (Bozorg-Haddad *et al.* 2017; Xiao *et al.* 2023), all of which hamper the accuracy of the model. Therefore, this study adopts an improved Aquila optimizer (IAO) algorithm to optimize the bi-directional long-term and short-term memory (IAO-BiLSTM) network model for the prediction of landslide groundwater. Compared with the Aquila optimizer (AO) algorithm, the IAO algorithm is improved in three aspects: population initialization based on Tent chaotic mapping, global optimal individual iteration based on an adaptive weight factor, and population iteration based on reverse differential evolution. It overcomes the shortcomings of the traditional intelligent optimization algorithm, such as premature convergence, susceptibility to local optima, and slow iteration, thereby

enabling the prediction model to find the optimal super-parameters. In addition, the BiLSTM network is more efficient than recurrent neural networks (RNNs) in processing groundwater sequence data due to the ability to retain important information for long-term memory and to a certain extent forget other information. In recent years, BiLSTM has been widely used in groundwater prediction (Ghasemlounia *et al.* 2021), but there are still few studies on groundwater level measurement in landslide sites. To verify the performance of the algorithm and its feasibility and scientific validity in the field of landslide groundwater prediction, this study will compare the prediction of landslide groundwater depth based on a case study of a landslide in Zhejiang, China, using IAO-BiLSTM and seven currently commonly used prediction models, and then verify the effectiveness of the method.

The remaining sections of this study are organized as illustrated in Figure 1. Section 2 introduces the basic principles of wavelet theory and the IAO-BiLSTM model. Section 3 describes the hydrogeological and monitoring system conditions of the case study. In Section 4, a wavelet multi-scale analysis is conducted for rainfall and groundwater depth in the landslide site, and the predictive factors are selected. Section 5 considers the time lag effect and utilizes rainfall to predict the groundwater depth at the water level monitoring hole of W3. Section 6 extends the prediction to other monitoring boreholes to further validate the feasibility of the model and investigates the effect of the distance between the water level monitoring boreholes and drainage tunnels on the magnitude of prediction errors. Finally, Section 7 presents the research conclusions derived from this study.

2. METHODS

2.1. Wavelet theory

2.1.1. Continuous wavelet transform

For a given time series $Y(t)$, its CWT is defined in Equation (1). In Equation (1), τ and s are the position and scale parameters, respectively, which determine the exact locality of the mother wavelet and its scaling factor, and $\bar{\varphi}_{\tau,s}$ is a cluster of wavelet functions obtained by stretching and translation. There are many types of mother wavelets, such as the Morlet wavelet, Haar wavelet, Daubechies wavelet, Mexican Hat wavelet, Meyer wavelet, and so on. Among them, the Morlet wavelet is widely used because of its easy-to-interpret output results; thus, this study uses the Morlet wavelet as the mother wavelet (Di *et al.* 2021).

$$W_Y(\tau, s) = \int_{-\infty}^{+\infty} Y(t) \cdot \bar{\varphi}_{\tau,s}(t) \cdot dt \quad (1)$$

2.1.2. Cross-wavelet transform

Cross-wavelet transform (XWT) combines wavelet transform with cross-spectrum analysis to study the correlation between two time series in the time–frequency domain from multiple time scales. The XWT of two time series X and Y is defined as Equation (2). In Equation (2), $*$ is a complex conjugate, and the corresponding crossover wavelet power spectral density is W_n^{xy} . The larger its value is, the more evidence there is of a shared area of high energy between the two time series and a strong correlation between them.

$$W_n^{xy} = W_n^x \cdot W_n^{y*} \quad (2)$$

2.1.3. Wavelet coherence analysis

WTC is a useful tool for detecting the covariance intensity of two time series in time–frequency space, providing insight into the correlation of its low energy region. Wavelet coherence is defined as Equation (3). In Equation (3), S is the smoothing operator. The significance of the wavelet coherence spectrum is tested by the Monte Carlo method. The range of wavelet correlation value R^2 is 0–1, which can be interpreted as the local correlation coefficient in the time–frequency domain. The larger the value is, the more strongly the correlation is established (Grinsted *et al.* 2004).

$$R_n^2(s) = \frac{|S(s^{-1}W_n^{xy}(s))|^2}{S(s^{-1}|W_n^x(s)|^2) \cdot S(s^{-1}|W_n^y(s)|^2)} \quad (3)$$

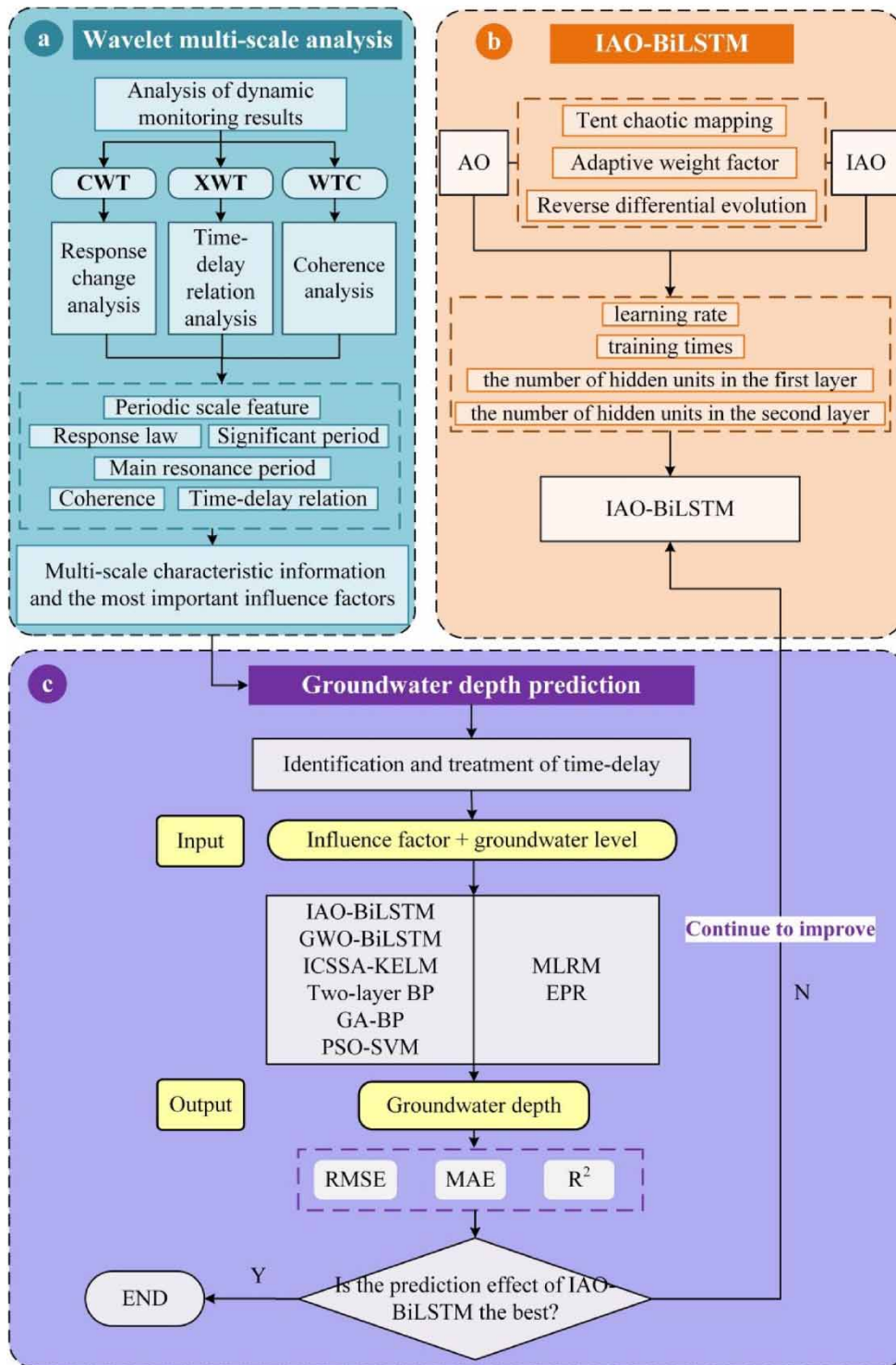


Figure 1 | Flow chart of this study.

2.2. IAO-BiLSTM

LSTM proposed by Hochreiter and Schmidhuber is an advanced version of the RNN. It retains RNN’s capability to remember previous information while replacing its looped hidden layer with a ‘memory block’ to filter out noise, reduce memory load, and address the issue of RNN’s vanishing gradients. It is widely used in the field of time series prediction. The

computation steps of LSTM are as follows: first, the forget gate is calculated using Equation (4). Then, the input gate is computed using Equation (5). Next, the cell state is updated according to Equation (6). Finally, the output gate is computed using Equation (7). In Equations (4)–(7), σ is the sigmoid activation function; W_f , W_i , and W_o are the weights of the sigmoid activation function of the forgetting, input, and output gates, respectively; b_f , b_i , and b_o are the bias of the activation function of the forgetting, input, and output gates, respectively; x_t and x_{t-1} are the memory activation values of the moments t and $t-1$, respectively; h_{t-1} is the activation function of the cellular output of the moments $t-1$; and C_t and C'_t are the new cellular and the cellular memory status, respectively.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (4)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (5)$$

$$C_t = f_t \times C_{t-1} + i_t \times C'_t \quad (6)$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_{t-1}] + b_o) \quad (7)$$

BiLSTM is proposed based on the LSTM network; each memory block of BiLSTM contains two LSTM layers, the input time sequence X_1, X_2, \dots, X_t at each moment, the two hidden layer states h_1, h_2, \dots, h_t and h'_1, h'_2, \dots, h'_t obtained from the forward LSTM layer, and the backward LSTM layer which is opposite to the two time sequences, and then the two hidden layer states are connected to get the same outputs y_1, y_2, \dots, y_t , the forward LSTM and the backward LSTM layers can get the past information and the future information of the input sequence, respectively. BiLSTM uncovers the relationship between the current data and the preceding as well as succeeding data points. This characteristic enhances the model's accuracy and improves data utilization efficiency. Compared to LSTM, BiLSTM is particularly suitable for handling data with significant temporal features, such as groundwater data.

The AO algorithm simulates the different hunting methods of different prey, namely, selecting the search space (X_1) by using vertical bending, exploring the search space (X_2) through the contour flight of short gliding attack, exploring the convergent search space (X_3) by the low-altitude flight of slow descent attack, and diving by walking and grabbing prey (X_4). Due to space limitations, the specific principles and computational steps of the AO algorithm can be found in the article of [Abualigah et al. \(2021\)](#).

AO is characterized by strong global exploration ability, high search efficiency, and fast convergence speed. It can be used to optimize the hyperparameters in the BiLSTM model. However, its local development ability is insufficient, and as a result, it is easy to fall into local optimization ([Brociek et al. 2022](#)). In this study, an improved AO (IAO) algorithm was utilized to enhance the research methodology. IAO employs population initialization based on Tent chaotic mapping, global optimal individual iterative update based on adaptive weight primer, and population iterative update with reverse differential evolution to increase the diversity of the population, improve the activity range of the population, and enhance the optimization efficiency and robustness of the algorithm, thus minimizing the chances of premature convergence and local optimization. The implementation steps for optimizing the BiLSTM model using the IAO algorithm are outlined below, and the detailed process is depicted in [Figure 2](#). The key mathematical notations used in Equations (8)–(14) are listed in [Table 1](#).

Step 1: Initialize parameters, such as setting the population size to 30 and the maximum number of iterations to 100.

Step 2: Initialize the population based on the Tent chaotic mapping. The population $P(N)$ is initialized using Equations (8)–(10). Equation (9) represents the transformation of the value X_{i+1} , generated by Equation (8), into the value range $[LB, UB]$ of the variable being optimized. Specifically, in the j -dimensional space, X_{i+2j} denotes the newly generated value. Additionally, Equation (10) introduces chaotic perturbation.

$$X_{i+1,j} = \begin{cases} 2x_{i,j}, & 0 \leq x_{i,j} \leq \frac{1}{2}, \\ 2(1 - x_{i,j}), & \frac{1}{2} < x_{i,j} \leq 1, \end{cases} \quad i = 0, 1, 2, \dots, N; j = 1, 2, \dots, D \quad (8)$$

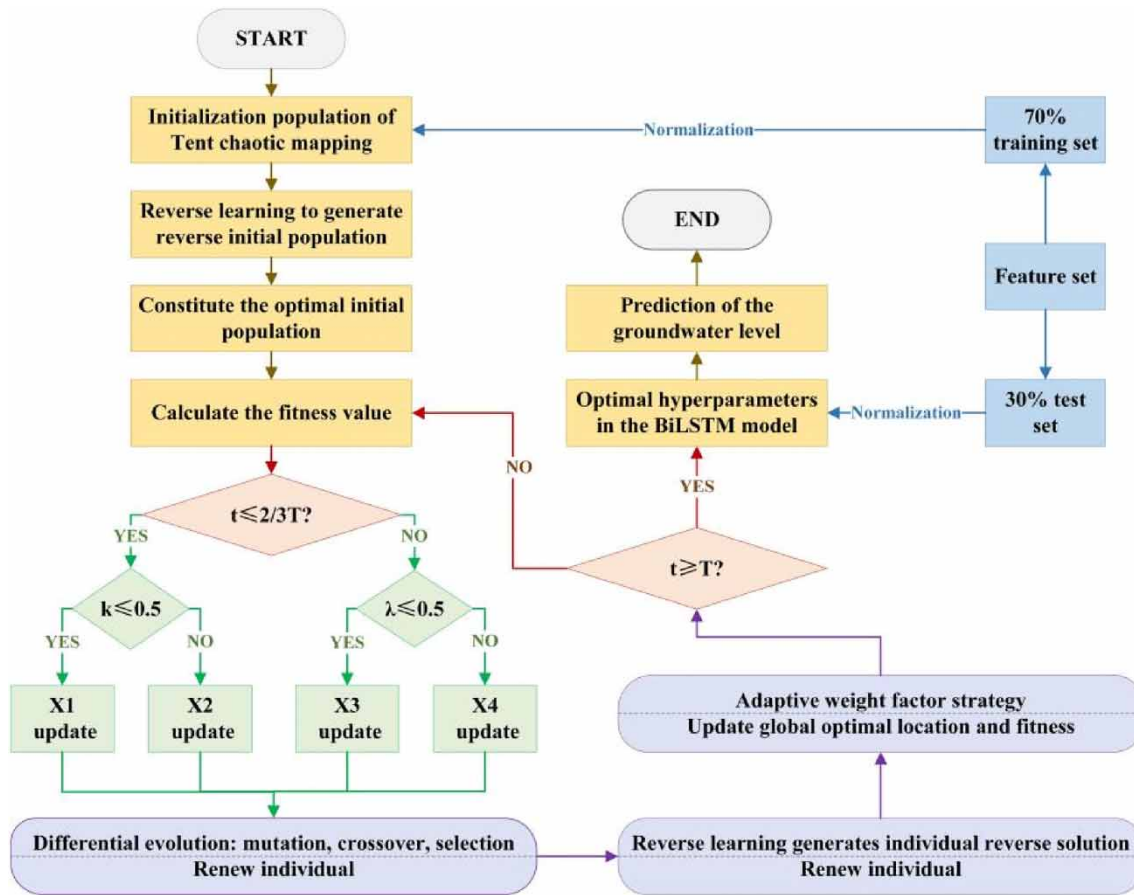


Figure 2 | IAO-BiLSTM model flow chart.

$$X_{i+2,j} = X_{i+1,j} \times (UB - LB) + LB, \quad i = 0, 1, 2, \dots, N; j = 1, 2, \dots, D \tag{9}$$

$$\text{new}X' = (X' + X)/2 \tag{10}$$

Step 3: Calculate the opposite population based on the opposition-based learning (OBL) strategy by the following equation:

$$O_{ij} = L_j + U_j - P_{ij}, \quad i = 1, 2, \dots, N; j = 1, 2, \dots, D \tag{11}$$

Step 4: Calculate the fitness of the initial population and the opposite population separately. N individuals with the lowest fitness are selected from the combined sets of both populations to form the optimal initial population.

Step 5: Calculate the fitness value of the optimal initial population and subsequently evaluate against predetermined criteria. If the criteria are met, the position of the Aquila is updated using the optimization process X_1 or X_2 . On the other hand, if the criteria are not met, the position of the Aquila is updated using the optimization process X_3 or X_4 . Additionally, the global optimal position and its corresponding fitness are updated accordingly.

Step 6: Generate new solutions based on the opposition-based differential evolution strategy after each individual updates their position. This involves applying the mutate, crossover, and select steps from the GA to the original solutions. Subsequently, new solutions are selected along with the best individual from the original solutions for the next iteration. Simultaneously, the global optimal position and its corresponding fitness are updated.

Step 7: Generate individual opposite solutions by applying Equation (12) based on the OBL strategy. The best individual from the original solutions is then selected to proceed to the next iteration along with these opposite solutions.

Table 1 | Notations used in Equations (8)–(14)

Symbol	Description
$X_{i+1,j}$	The position of the $(i + 1)$ th Aquila individual in the j th dimension after Tent chaotic mapping
$x_{i,j}$	The position of the i th Aquila individual in the same dimension
N	The population size
D	The problem dimension
LB	The lower bounds of the tuned-parameter optimization problem
UB	The upper bounds of the tuned-parameter optimization problem
new X'	The individual after the perturbation
X'	The individual to be chaotically perturbed
X	The generated chaotic perturbation variable
$P_{i,j}$	The i th individual in the j th dimension of the population
$O_{i,j}$	The i th individual in the j th dimension of the inverse population
L_j	The lower bounds of the problem in the j th dimension
U_j	The upper bounds of the problem in the j th dimension
ε	A random number in the interval [0,1]
$X(t + 1)$	The solution at the $(t + 1)$ th iteration
$X_{\text{best}}(t)$	The optimal solution at the t th iteration
t	The current iteration
T	The maximum number of iterations
e	The base of the natural logarithm

Simultaneously, the global optimal position and its corresponding fitness are updated.

$$\begin{cases} X_{i,j}^*(t) = k(\text{UB}_j(t) + \text{LB}_j(t)) - X_{i,j}(t) \\ \text{UB}_j(t) = \max(X_{i,j}(t)), \text{LB}_j(t) = \min(X_{i,j}(t)) \\ i = 1, 2, \dots, N, j = 1, 2, \dots, D \end{cases} \quad (12)$$

Step 8: Generate the new global optimal location and its corresponding fitness value by Equations (13) and (14) based on the adaptive weight strategy. The best individual from the original values is then selected for updating along with this new global optimal location.

$$X_1(t + 1) = \eta X_{\text{best}}(t) \quad (13)$$

$$\eta = 1 - e^{1 + \tan(\pi(\varepsilon - 0.5)) - ((T+t)/(T-t))} \quad (14)$$

Step 9: Determine whether the current iteration count has reached the maximum predefined number of iterations. If it has, the optimal parameters of the model at the current iteration count are outputted. Otherwise, the algorithm proceeds to Step 5 and continues iterating until the maximum iteration count is reached.

Step 10: Assign the optimal parameters produced in Step 9 to the BiLSTM model. Subsequently, the predictive factors are inputted, and the model is executed to generate the corresponding output predictions.

3. STUDY AREA AND MONITORING SYSTEM

3.1. Research background of the study area

The study area is a landslide located at the junction of Xinqiao Township, Changshan County, and Kecheng District, Quzhou City, Zhejiang Province, China (Figure 3(a) and 3(b)). It is situated in the middle and lower mountain areas of western Zhejiang, with the highest elevation being 1,050 m. The length of the landslide is about 350 m, with a 140 m wide leading edge

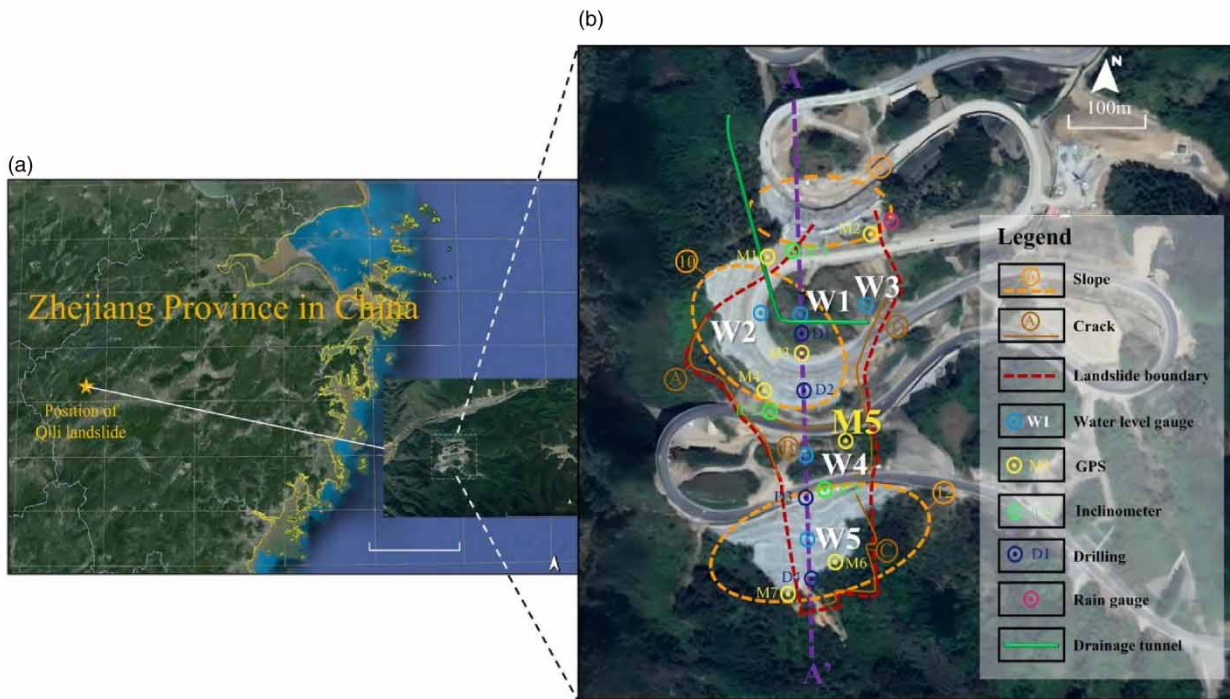


Figure 3 | Study area plan and monitoring system profile.

and a 55 m wide trailing edge. It is characterized by a fault zone and four main fractures, which are near the Jurassic and Ordovician stratigraphic boundary and fault zone.

The landslide strata are mainly composed of quaternary residual soil formation (Q), upper Jurassic Huangjian formation (J_{5n}), lower Ordovician Ningguo formation (O_{1n}), and lower Ordovician Yinzhubu formation (O_{1y}). The landslide area belongs to the subtropical monsoon climate zone with a prevailing northeasterly wind throughout the year. The rainy season extends from February to mid-July, particularly from May to June and early July, while the dry season lasts from late July to January the following year. From August to October, it is occasionally affected by typhoons, and there are heavy rains, which often lead to a drought in autumn. The average annual rainfall in the study area is 1,804 mm and the evaporation is 1,424 mm. Groundwater in landslide areas can be mainly divided into two types: pore water in loose rocks and fracture water in bedrock. The thickness of the water-bearing layer of pore water in loose rocks is about 2.0–6.1 m, and it is replenished by atmospheric precipitation and irrigation return flow and is discharged in the form of springs or lateral recharge. Fracture water in bedrock can be divided into weathering fracture water and structural fracture water according to the occurrence conditions of groundwater. The thickness of the water-bearing layer of weathering fracture water is relatively small and mostly within 10 m, while structural fracture water mainly occurs in the structural fractures of the Ordovician strata. Both types of fracture water are mainly replenished by atmospheric precipitation and high mountain condensation water and are discharged in the form of springs.

There were several cracks in the study area between August 2014 and November 2015, which led to landslides at the site in early 2016. Through detailed mechanism analysis, Sun *et al.* (2019) proved that the landslide is a deep one formed by the reactivation of the local ancient landslide under the influence of groundwater and construction activities. To ensure the stability of the groundwater depth in the area, a 305 m-long drainage hole was constructed from August 2016 to November 2016, which was able to lower the groundwater level in the landslide.

Groundwater in landslide areas mainly comes from atmospheric precipitation and high mountain condensation water and is also influenced by engineering activities such as drainage tunnels. As a highly nonlinear and non-stationary signal (Kousali *et al.* 2022), the dynamic changes of groundwater have complex relationships that require the selection of important influencing factors for groundwater depth prediction in conjunction with monitoring systems. Additionally, it is necessary to confirm whether rainfall can be a significant predictor for groundwater depth prediction.

3.2. Monitoring system

Systematic monitoring was conducted in a landslide area from January 1, 2016 to December 31, 2017 to ensure the safety of construction. The monitoring content mainly includes three parts: rainfall, groundwater depth, and surface displacement (Figure 3(b)). The specific monitoring methods and instruments are as follows: (1) for rainfall monitoring, automatic tipping bucket rain gauge real-time monitoring is used to monitor the rainfall in the region. (2) For groundwater depth monitoring, five water level monitoring holes (W1–W5) are arranged in the landslide area, and then water level gauges are arranged in the water level monitoring holes to monitor the change of groundwater depth in the landslide. It should be noted that the monitored groundwater level refers to the groundwater depth. (3) For surface displacement, seven GPS stations (M1–M7) are arranged in the slope area, and then the displacement of the monitoring point is obtained by GPS satellite positioning measurement.

4. WAVELET MULTI-SCALE ANALYSIS

4.1. Analysis of dynamic monitoring results

To explore the dynamic monitoring results and then provide an analysis basis for the subsequent multi-scale analysis between rainfall and groundwater depth, the rainfall, surface displacement, surface displacement rate (SDR), and the groundwater depth of five water level monitoring holes (W1–W5) were selected to analyze the monitoring data from January 5, 2016 to December 31, 2017. The monitoring results of surface displacement and SDR are derived specifically from the M5 site, which is the area with the greatest landslide displacement during the monitoring period and can fully represent the landslide's movement (Wei *et al.* 2019).

Figure 4 presents the variation curves of rainfall, groundwater depth, surface displacement, and displacement rate during the monitoring process. From Figure 4, it can be seen that there is a good response relationship between rainfall (RL) and groundwater depth (W1–W5) as well as SDR. For example, when the rainfall increases, the groundwater depth decreases while the SDR increases, which is in line with the objective law. At the same time, the changes in monitoring parameters before and after the construction of the drainage tunnel can be observed. In November 2016, after the tunnel was constructed, the water level in the buried depth became higher as a whole, that is, the groundwater level decreased, and the change of the water level was no longer as drastic as it used to be before the tunnel was built, which implies that the tunnel

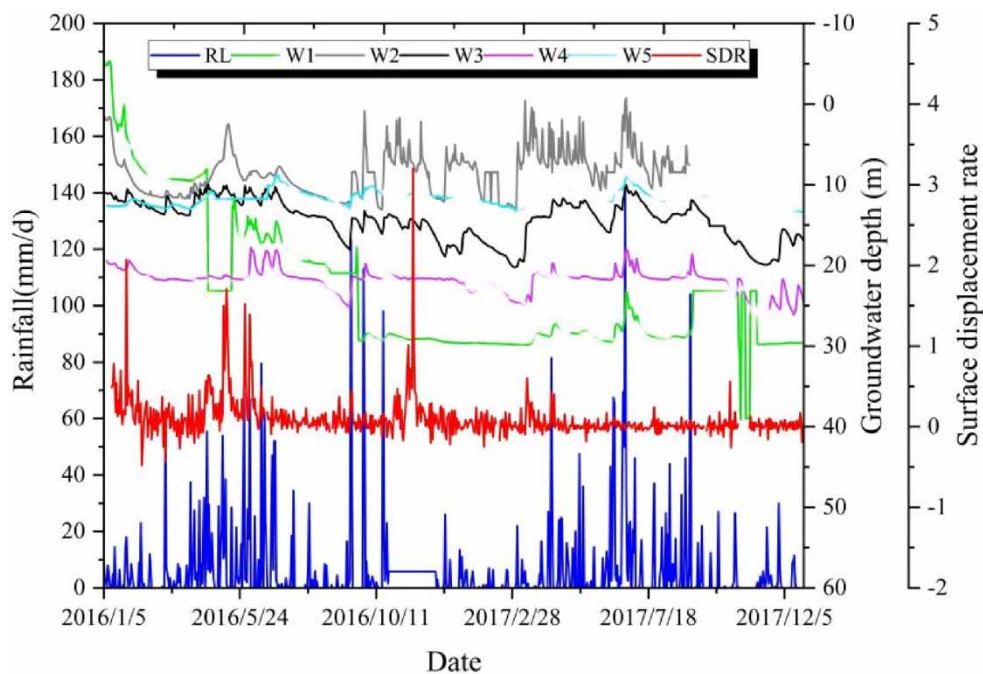


Figure 4 | Dynamic monitoring data in the landslide site.

effectively lowers the groundwater depth and keeps it in a comparatively steady state. Likewise, after the tunnel was constructed, the increasing rate of surface displacement decreased in comparison with the situation before construction, which reveals that the tunnel successfully enhances the stability of the slope and decreases the deformation rate of the slope.

In addition, it can be seen from Figure 4 that after the construction of the drainage tunnel, the groundwater depth of W1 changed most greatly, rising from 15.04 mm/day on July 2, 2016 to 29.07 mm/day on November 10, 2016. W1 is the nearest water level monitoring hole among the five groundwater depth monitoring holes, indicating the significant role of the drainage hole in drainage.

4.2. Response change analysis

To analyze whether rainfall is the most important influencing factor for groundwater in landslide sites, CWT (Figure 5) was conducted on the daily rainfall data from January 5, 2016 to December 30, 2017 and the daily average groundwater depth of five water level monitoring holes, which can further explain their periodic scale characteristics.

Based on the analysis of the wavelet coefficients' time-frequency characteristics, in conjunction with Figure 5(a), the red noise test at the 95% confidence level indicates that the rainfall during the study period exhibits prominent oscillation periods primarily at time scales of 10–25 and 365 days. The strong rainfall behavior occurred on the 250th day and 550th day in this area, which indicated that the scale of the 365 day cycle was significant. Furthermore, Figure 5(b)–5(f) shows that W1–W5 all have the main oscillation period time scale of 25–45 days. During the period of rainfall, the groundwater depth of W1–W5 has a strong signal, which indicates that groundwater depth is significantly affected by rainfall. The small scale covers the time-domain discontinuity and large span, while the large scale covers time-domain continuity and small span. It is worth noting that the periodic time scale, oscillation intensity, and response of the groundwater depth of W2–W5 are similar, indicating that the groundwater depth of W2–W5 is greatly affected by rainfall, while W1 weakens the response relationship due to the influence of the drainage tunnel. In summary, due to the limitations of topography and hydroclimatic conditions, the groundwater depth in the landslide site has no other main sources of supply except rainfall. Therefore, it can be preliminarily

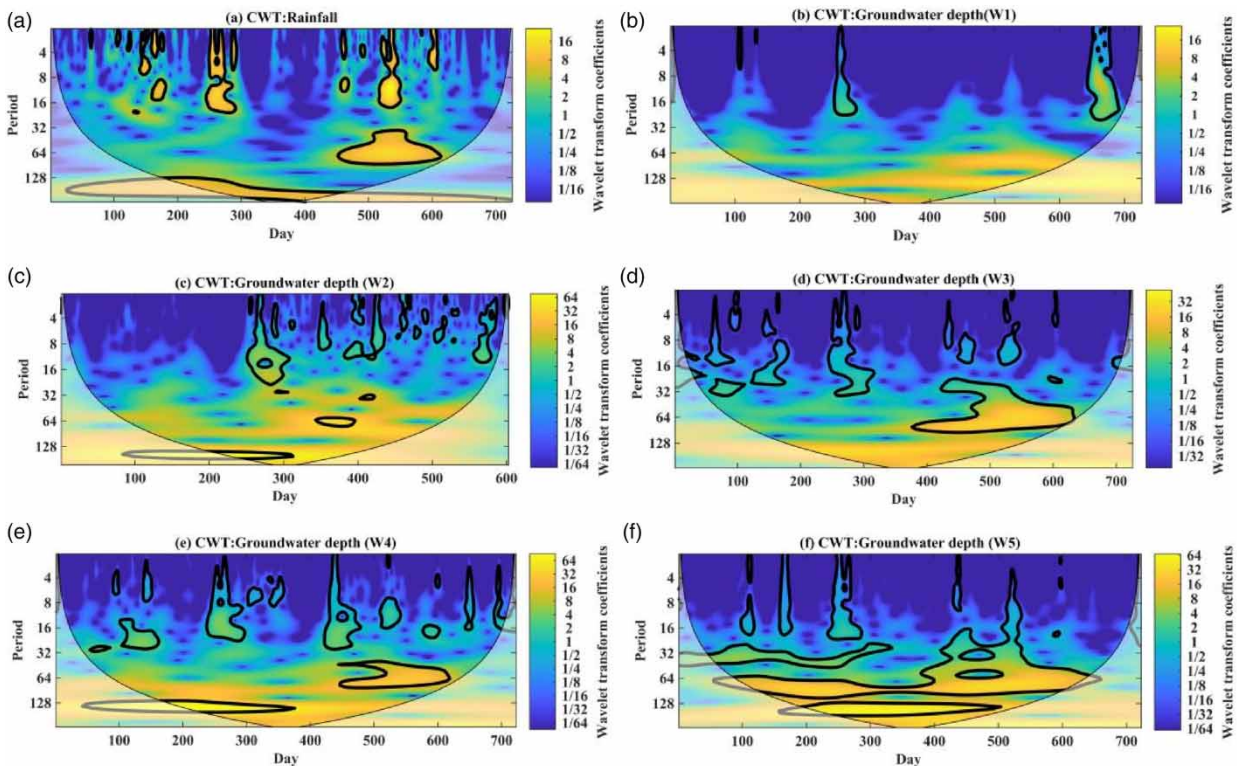


Figure 5 | CWT power spectrum of rainfall and groundwater depth; the objects of CWT power spectrum of (a)–(f) are successively the rainfall and groundwater depth of W1–W5.

concluded that rainfall is the main driving factor leading to the significant change of groundwater depth during the study period.

4.3. Time lag relation analysis

Based on the similar periodic scale between rainfall and groundwater depth of W2–W5, the XWT (Figure 6) is used to analyze the main resonance period, significant period, and time lag effects of rainfall and groundwater depth. Figure 6 shows that W2–W5 have a similar high energy resonating period and phase relationship, but the distribution is scattered in the time–frequency domain. On the 365 day time scale, the phase relationship between rainfall and groundwater depth is consistent and the arrow is pointing to the upper left, which implies a negative correlation between the two factors, and a time lag effect of rainfall leading to the change in the groundwater depth. This also indirectly reflects the process of the transformation from rainfall to groundwater depth in the study site and confirms the preliminary conclusion in Section 4.2 that ‘rainfall is the main driving factor leading to the significant change of groundwater depth during the study period’. In addition, the resonating signal of W2–W5 tends to become increasingly stronger, indicating that the correlation between the groundwater depth and rainfall is gradually enhanced, suggesting that the drainage tunnel has a significant influence on the correlation between the groundwater depth and rainfall, since W2 and W3 are closer to the drainage tunnel at the leading edge of the slope, while W4 and W5 are farther away.

4.4. Coherence analysis

WTC was applied to the groundwater depth of W2–W5 and rainfall to gain further insights into their multi-scale coherence characteristics and understand their correlation. From the results of Figure 7, it can be seen that there is a strong correlation between rainfall and groundwater depth of W2–W5, with a correlation coefficient greater than 0.8 in the whole monitoring

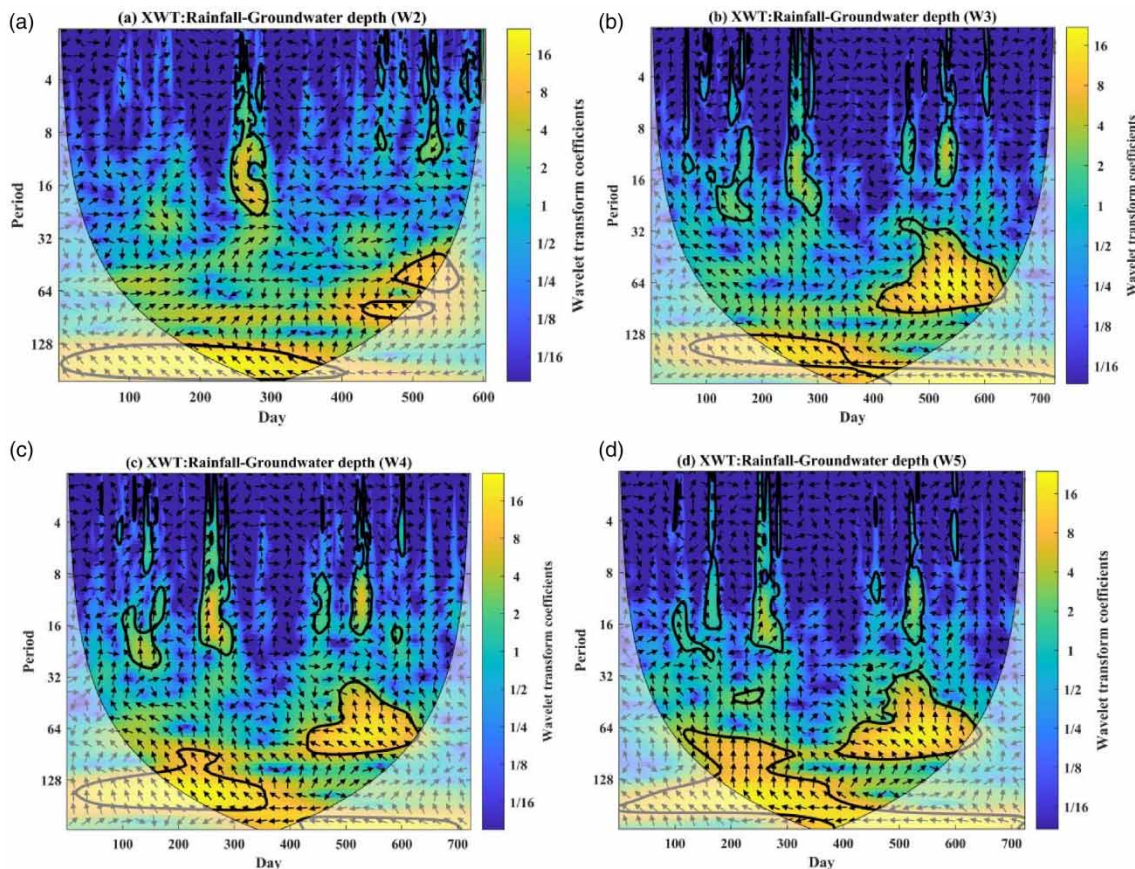


Figure 6 | Cross-wavelet transform power spectrum of rainfall and groundwater depth: (a)–(d) the crossover wavelet transform power spectrum of the groundwater depth of W2–W5 and rainfall in turn.

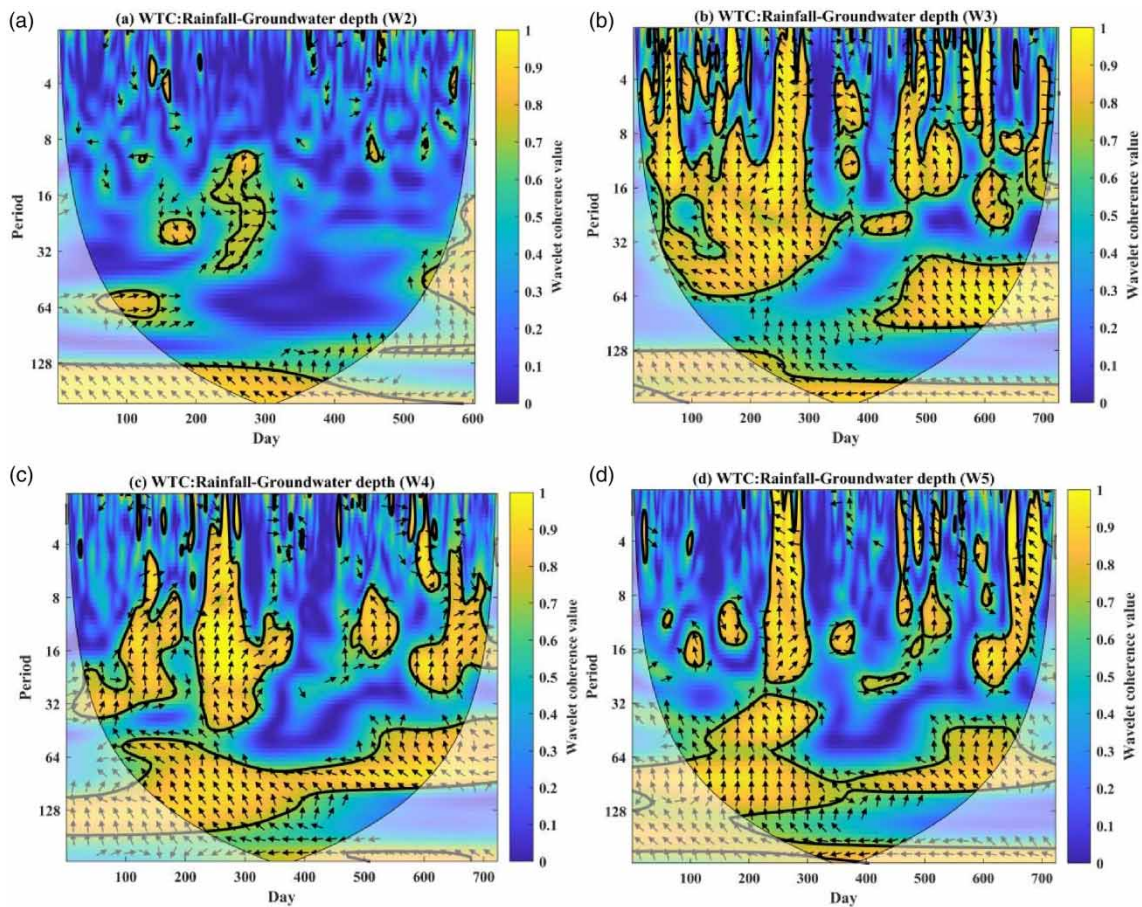


Figure 7 | Wavelet coherence spectrum of rainfall and groundwater depth: (a)–(d) the wavelet coherence spectrum of the groundwater depth of W2–W5 and rainfall in turn.

period and a 365-day cycle. Coherence between rainfall and monitoring hole W2–W5 is also strong in the 400–650 day time domain and the 35–45 day cycle scale, with a correlation coefficient greater than 0.8. However, the coherence difference can be seen in other scales. The percentage of significant coherence area can further quantify this coherence, with the result being: W5 (37.17%) > W3 (37.00%) > W4 (32.00%) > W2 (23.62%). This indicates a close relationship between rainfall and groundwater depth, although W2 is closer to the drainage tunnel, and thus has a lower percentage of coherence as it is more affected by the drainage effect.

Based on the wavelet multi-scale analysis system mentioned earlier, combined with the time–domain and frequency–domain multi-scale features analysis of the main oscillation period, main resonance period, significant period, time lag effect, coherence, and differences between groundwater depth and rainfall, it can be concluded that rainfall is the most important influencing and driving factor for groundwater depth prediction at the Qili landslide site. Therefore, this study selects rainfall as the input factor for predicting groundwater depth at the Qili landslide site.

5. PREDICTION OF GROUNDWATER DEPTH

5.1. Identification of lag period

Data collected from January 5, 2016 to December 25, 2017 are utilized, and the missing values of less than 5% in the data are replaced by the adjacent position mean processing method. Based on the calculation using the Spearman correlation coefficient, it is determined that the water level monitoring hole W3, situated at the center of the landslide, exhibits the highest correlation (–0.576) with the GPS station M5. Consequently, it is considered the most suitable choice for representing the

hydrological behavior of the landslide. Hence, this study will focus on the prediction of the groundwater depth of the monitoring hole W3.

Before making predictions, it is necessary to consider the time lag effect between rainfall and groundwater depth to make the prediction system more accurate and close to reality. Given the discreteness of the monitoring data, the Spelman correlation coefficient is used to identify the lag days between rainfall and groundwater depth at a 99% confidence level, and the results show a significant correlation, as can be seen in Figure 8.

In Figure 8, $T = 0$ shows the correlation coefficient between the rainfall of the day and the groundwater depth of the water level detection hole W3–W5 on the same day, $T = 1$ represents the correlation coefficient between the rainfall of the day and the groundwater depth of the second-day water level detection hole W3–W5, $T = 2$ represents the correlation coefficient between the rainfall of the previous day and the groundwater depth of the third-day water level detection hole W3–W5, and so on. The analysis demonstrates that as T increases from 1 to 12, the correlation coefficient initially rises and then declines. As T approaches the lag days, the correlation coefficient consistently increases. When T equals the lag days, the correlation coefficient reaches its peak, indicating the lag days between groundwater depth and rainfall. As T gradually exceeds the lag days, the influence of rainfall on groundwater depth diminishes, leading to a decreasing correlation coefficient. The highest correlation coefficient of W2–W5 corresponds to T , which is 5, 4, 3, and 2 days, respectively. This indicates that the groundwater depth fluctuations at W2–W5 lag behind rainfall by 5, 4, 3, and 2 days, respectively. It is also important to note that the correlation coefficient between rainfall and groundwater depth has been weakened due to the significant impact of engineering activities such as excavation and the construction of the drainage tunnel on the slope.

5.2. Prediction and analysis

To alleviate the impact of lag effects on the prediction of groundwater depth of W3, considering the existing lag relationship of 4 days, the rainfall data for $T = 0-4$ are utilized as the input factors for the prediction model. Subsequently, the IAO-BiLSTM model is employed to forecast the groundwater depth in the landslide area. In the process of optimization, the IAO algorithm optimizes four hyperparameters that have a great influence on the accuracy of feature information recognition in BiLSTM, namely, the number of hidden units in the first layer, the number of hidden units in the second layer, training times and learning rate. To assess the plausibility of the model, the gray wolf optimization algorithm (GWO) is used to optimize the BiLSTM network (Liu & Yang 2022) and an improved circle sparrow search algorithm to optimize kernel extreme

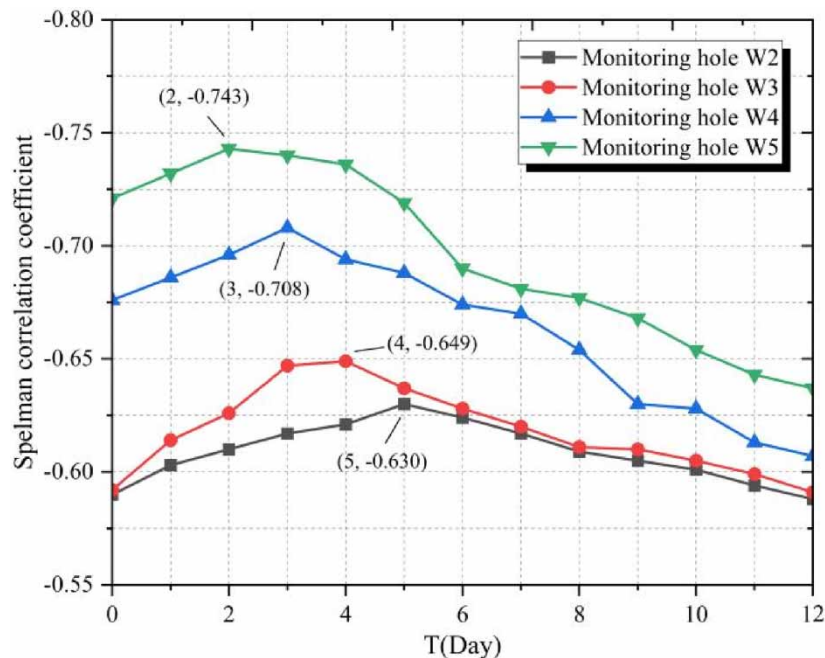


Figure 8 | Identification of the time lag effect of groundwater depth of W2–W5.

learning machine (improved circle sparrow search algorithm (ICSSA)-KELM) (Parida *et al.* 2021), two-layer BP (Yesilnacar & Sahinkaya 2012), multiple linear regression with ridge regression (MLRM), and evolutionary polynomial regression (EPR) (Doglioni *et al.* 2012), GA to optimize BP neural network, and PSO-SVM (Zhu *et al.* 2020) for comparative prediction and analysis. It is important to note that all models used in this experimental test belong to the field of machine learning and are regression models. Their utilization in this predictive case is both permissible and appropriate. In addition, to avoid computational randomness, each model is set with a population of 30 and an iteration of 100, and their average results after running 50 times are used as the prediction results (Ahmadianfar *et al.* 2021).

The predictive performance of each model is evaluated using three commonly used evaluation metrics in regression studies: root mean squared error (RMSE), mean absolute error (MAE), and Nash–Sutcliffe efficiency (NSE) coefficient. RMSE and MAE primarily assess the model's predictive error performance, while NSE primarily evaluates the quality of the regression model. The prediction results of each model are shown in Table 2. According to Table 2, the prediction performance of IAO-BiLSTM is the best, and its RMSE, MAE, and NSE are the best among the eight algorithms, which are 0.25, 0.21, and 0.99, respectively. In addition, after considering the time lag effect, the prediction error of each model is reduced. For instance, the RMSE of GWO-BiLSTM decreased from 5.11 to 3.36, resulting in a precision improvement of 34.30%. Table 3 is the hyperparameter result of fitness and optimization of IAO-BiLSTM before and after considering the time lag effect. From Table 3, the fitness after considering the time lag effect is lower (0.0827) when compared with the situation without considering it, which once again proves that the lag periods' identification and processing strategy adopted in this study is feasible and effective. Furthermore, the prediction accuracy is higher after considering the time lag effect, but the required optimal parameters' values are reduced as a whole, implying that the efficiency of the algorithm is higher, which may be due to the increase of the input parameters from 1 to 5 after considering the time lag effect. The training and test performance of IAO-BiLSTM is shown in Figure 9. As can be seen from Figure 9, on the whole, the fitting results and prediction results of IAO-BiLSTM are close to the actual values, which are basically in line with the prediction trend and have relatively stable and accurate prediction performance.

6. DISCUSSION

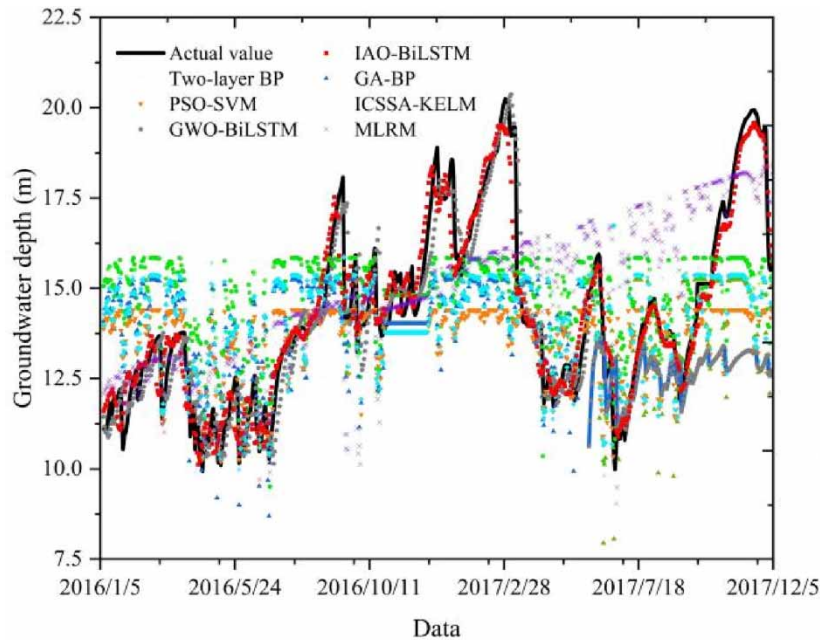
To mitigate computational randomness and validate the effectiveness of IAO-BiLSTM in predicting groundwater depth, as well as to assess the scientific validity of the time lag effect identification and resolution strategy, additional predictions are conducted for water level monitoring holes within the landslide area. However, W1's groundwater depth correlation with rainfall is weak due to its proximity to the drainage tunnel, thus making it impossible to accurately predict the

Table 2 | Prediction results of groundwater depth of W3

Model	Time lag	RMSE	MAE	NSE	Precision improvement (%)
IAO-BiLSTM	Before	0.26	0.23	0.99	3.16
	After	0.25	0.21	0.99	
GWO-BiLSTM	Before	5.11	4.81	-3.04	34.30
	After	3.36	2.64	-0.75	
ICSSA-KELM	Before	3.23	2.34	-0.61	14.77
	After	2.75	2.15	-0.17	
GA-BP	Before	2.87	2.29	-0.28	2.40
	After	2.80	2.14	-0.21	
Two-layer BP	Before	2.53	2.09	-7.54	13.67
	After	2.19	1.74	-1.51	
PSO-SVM	Before	2.98	2.35	-0.38	5.53
	After	2.82	2.27	-0.23	
MLRM	Before	2.91	2.50	-0.31	12.10
	After	2.55	2.20	-0.01	
EPR	Before	2.31	1.19	0.15	6.02
	After	2.17	1.12	0.27	

Table 3 | Fitness and hyperparameters before and after considering the time lag effect

Time lag	Fitness value	Hyperparameters			Training times	Learning rate
		The number of hidden units in the first layer	The number of hidden units in the second layer			
Before	0.0932	100	29.6711	43.2858	0.0100	
After	0.0827	23.1126	33.3668	36.7120	0.0076	

**Figure 9** | Training and test performance of groundwater depth of W3.

groundwater depth via rainfall data. Therefore, W1 is excluded here. Incorporating the time lag relationships between groundwater depth and rainfall (Figure 8), the IAO-BiLSTM model was utilized to predict the groundwater depth of W2, W4, and W5. Specific prediction results are presented in Table 4. Overall, the prediction performance of W2, W4, and W5 is relatively good, with RMSE values of 0.63, 0.22, and 0.15, respectively, which meet the prediction requirements. This further confirms the feasibility and applicability of the IAO-BiLSTM prediction model.

Figure 10 demonstrates the relationship between the prediction accuracy of groundwater depth for each monitoring hole and its respective distance from the drainage tunnel. As the distance between W2–W5 and the drainage tunnel increases sequentially, the corresponding prediction error decreases accordingly. This occurs because the drainage tunnel attenuates the correlation between groundwater depth and rainfall, thereby impacting prediction accuracy. Consequently, when

Table 4 | Prediction results of groundwater depth of W2, W4, and W5

Monitoring hole number	Time lag	RMSE	MAE	NSE	Fitness value	Precision improvement (%)
W2	Before	0.65	0.54	0.84	0.0507	3.96
	After	0.63	0.52	0.86	0.0424	
W4	Before	0.26	0.23	0.96	0.0195	15.38
	After	0.22	0.17	0.99	0.0182	
W5	Before	0.18	0.15	0.97	0.0283	13.52
	After	0.15	0.13	0.99	0.0221	

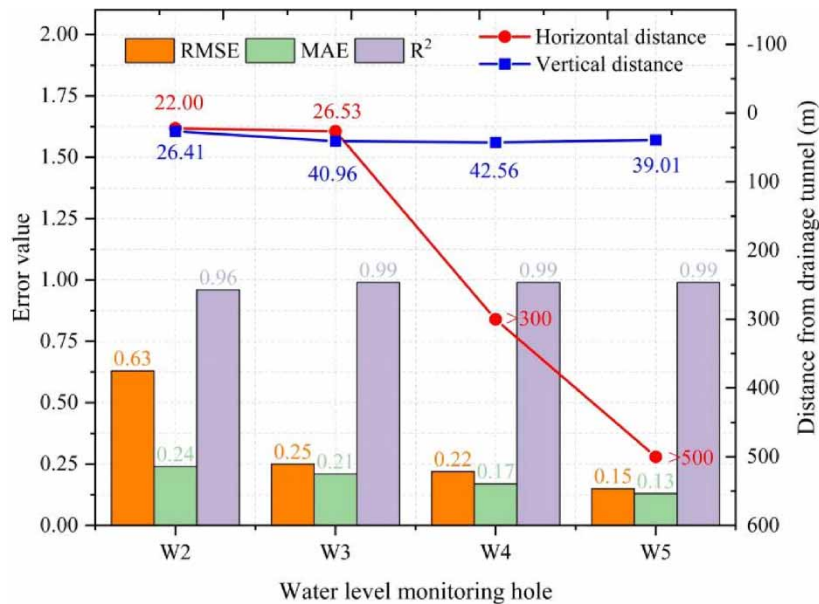


Figure 10 | Influence of distance between the water level monitoring hole and the drainage tunnel on prediction.

making predictions for groundwater depth of water level monitoring holes situated in proximity to and notably influenced by the drainage tunnel, factors related to the tunnel should be considered to enhance prediction precision.

7. CONCLUSIONS

Using rainfall and groundwater depth monitoring data from a landslide site in Zhejiang Province, China, over 2 years, this study applies a wavelet multi-scale analysis method to investigate and identify the multi-scale characteristics of groundwater and finds the most important influencing factors for predicting groundwater depth. Subsequently, the time lag effect was taken into consideration, and the IAO-BiLSTM model was applied to predict groundwater depth in the landslide site. The main conclusions are as follows: (1) the groundwater depth of the water level monitoring hole (W2–W5) possesses a remarkable similarity in an energy resonating period and a phase relationship and demonstrates a strong correlation with rainfall, as indicated by a correlation coefficient exceeding 0.8. Rainfall is the most important influence and driving factor in predicting groundwater depth within landslide sites. (2) A negative correlation between rainfall and the groundwater depth is observed, with the level lagging behind the rainfall. The Spelman correlation coefficient is applied to measure the lag period of the W2–W5 water level monitoring holes, which were 5, 4, 3, and 2 days, respectively. (3) The construction of a drainage tunnel effectively reduces the groundwater depth and SDR and improves the stability of the slope. In addition, the closer the water level monitoring hole is to the drainage hole, the greater the weakening of the correlation and response between the water level and rainfall. (4) Considering the time lag effect and after comparing the results of GWO-BiLSTM, ICSSA-KELM, GA-BP, two-layer BP, PSO-SVM, MLRM, and EPR, it can be seen that IAO-BiLSTM used in this study has higher accuracy, and its RMSE, MAE, and R^2 are 0.25, 0.21, and 0.99, respectively. It is proven that IAO-BiLSTM can effectively predict the groundwater level of landslides and is a reliable prediction model. This study provides a new approach to revealing the dynamic response relationship between groundwater levels and their influencing factors and offers a novel method for predicting the dynamic changes in groundwater levels of landslides.

ACKNOWLEDGEMENTS

The authors would like to thank the project of the National Natural Science Foundation of China for its support as well as the editors and reviewers for their useful comments and suggestions.

FUNDING

This study was supported by the National Natural Science Foundation of China (Key Project) (No. 42230702).

DATA AVAILABILITY STATEMENT

All relevant data are available from an online repository or repositories: <https://doi.org/10.5061/dryad.djh9w0w6d>.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Abualigah, L., Yousri, D., Abd Elaziz, M., Ewees, A. A., Al-qaness, M. A. & Gandomi, A. H. 2021 *Aquila optimizer: A novel meta-heuristic optimization algorithm*. *Computers & Industrial Engineering* **157**, 107250. doi:10.1016/j.cie.2021.107250.
- Ahmadianfar, I., Noshadian, S., Elagib, N. A. & Salarijazi, M. 2021 *Robust diversity-based sine-cosine algorithm for optimizing hydropower multi-reservoir systems*. *Water Resources Management* **35**, 3513–3538. doi:10.1007/s11269-021-02903-6.
- Bozorg-Haddad, O., Azarnivand, A., Hosseini-Moghari, S.-M. & Loáiciga, H. A. 2017 *Optimal operation of reservoir systems with the symbiotic organisms search (SOS) algorithm*. *Journal of Hydroinformatics* **19**, 507–521. doi:10.2166/hydro.2017.085.
- Brociek, R., Pleszczyński, M., Zielonka, A., Wajda, A., Coco, S., Lo Sciuto, G. & Napoli, C. 2022 *Application of heuristic algorithms in the tomography problem for pre-mining anomaly detection in coal seams*. *Sensors (Basel, Switzerland)* **22**, 7297–7319. doi:10.3390/s22197297.
- Cao, Y., Yin, K., Zhou, C. & Ahmed, B. 2020 *Establishment of landslide groundwater level prediction model based on GA-SVM and influencing factor analysis*. *Sensors (Basel, Switzerland)* **20**, 845–866. doi:10.3390/s20030845.
- Di, Y., An, X., Zhong, W., Liu, S. & Ming, D. 2021 *The time-robustness analysis of individual identification based on resting-state EEG*. *Frontiers in Human Neuroscience* **15**, 672946. doi:10.3389/fnhum.2021.672946.
- Dogliani, A., Fiorillo, F., Guadagno, F. M. & Simeone, V. 2012 *Evolutionary polynomial regression to alert rainfall-triggered landslide reactivation*. *Landslides* **9**, 53–62. doi:10.1007/s10346-011-0274-8.
- Duan, G., Chen, D. & Niu, R. 2019 *Forecasting groundwater level for soil landslide based on a dynamic model and landslide evolution pattern*. *Water* **11**, 2163. doi:10.3390/w11102163.
- Ghasemlounia, R., Gharehbaghi, A., Ahmadi, F. & Saadatnejadgharahassanlou, H. 2021 *Developing a novel framework for forecasting groundwater level fluctuations using bi-directional long short-term memory (BiLSTM) deep neural network*. *Computers and Electronics in Agriculture* **191**, 106568. doi:10.1016/j.compag.2021.106568.
- Grinsted, A., Moore, J. C. & Jevrejeva, S. 2004 *Application of the cross wavelet transform and wavelet coherence to geophysical time series*. *Nonlinear Processes in Geophysics* **11**, 561–566. doi:10.5194/npg-11-561-2004.
- Habibi, V., Ahmadi, H., Jafari, M. & Moeini, A. 2019 *Application of nonlinear models and groundwater index to predict desertification case study: Sharifabad watershed*. *Natural Hazards* **99**, 715–733. doi:10.1007/s11069-019-03769-z.
- Huang, F., Huang, J., Jiang, S.-H. & Zhou, C. 2017 *Prediction of groundwater levels using evidence of chaos and support vector machine*. *Journal of Hydroinformatics* **19**, 586–606. doi:10.2166/hydro.2017.102.
- Kainthura, P. & Sharma, N. 2022 *Machine learning driven landslide susceptibility prediction for the Uttarkashi region of Uttarakhand in India*. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* **16**, 570–583. doi:10.1080/17499518.2021.1957484.
- Kousali, M., Salarijazi, M. & Ghorbani, K. 2022 *Estimation of non-stationary behavior in annual and seasonal surface freshwater volume discharged into the Gorgan Bay, Iran*. *Natural Resources Research* **31**, 835–847. doi:10.1007/s11053-022-10010-5.
- Li, F., Liu, G., Tao, Q. & Zhai, M. 2023 *Land subsidence prediction model based on its influencing factors and machine learning methods*. *Natural Hazards* **116**, 3015–3041. doi:10.1007/s11069-022-05796-9.
- Liu, Z. & Yang, J. 2022 *Research on short-term load forecasting based on GWO-BILSTM*. *Journal of Physics: Conference Series* **2290**, 12100. doi:10.1088/1742-6596/2290/1/012100.
- Nikvar Hassani, A., Katibeh, H. & Farhadian, H. 2016 *Numerical analysis of steady-state groundwater inflow into Tabriz line 2 metro tunnel, northwestern Iran, with special consideration of model dimensions*. *Bulletin of Engineering Geology and the Environment* **75**, 1617–1627. doi:10.1007/s10064-015-0802-1.
- Parida, N., Mishra, D., Das, K. & Rout, N. K. 2021 *Development and performance evaluation of hybrid KELM models for forecasting of agro-commodity price*. *Evolutionary Intelligence* **14**, 529–544. doi:10.1007/s12065-019-00295-6.
- Seifollahi-Aghmiuni, S., Kalantari, Z., Egidi, G., Gaburova, L. & Salvati, L. 2022 *Urbanisation-driven land degradation and socioeconomic challenges in peri-urban areas: Insights from Southern Europe*. *Ambio* **51**, 1446–1458. doi:10.1007/s13280-022-01701-7.
- Sharma, G. D., Tiwari, A. K., Jain, M., Yadav, A. & Erkut, B. 2021 *Unconditional and conditional analysis between COVID-19 cases, temperature, exchange rate and stock markets using wavelet coherence and wavelet partial coherence approaches*. *Heliyon* **7**, e06181. doi:10.1016/j.heliyon.2021.e06181.
- Su, C., Wang, B., Lv, Y., Zhang, M., Peng, D., Bate, B. & Zhang, S. 2023 *Improved landslide susceptibility mapping using unsupervised and supervised collaborative machine learning models*. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* **17**, 387–405. doi:10.1080/17499518.2022.2088802.

- Sun, H., Wu, X., Wang, D., Xu, H., Liang, X. & Shang, Y. 2019 Analysis of deformation mechanism of landslide in complex geological conditions. *Bulletin of Engineering Geology and the Environment* **78**, 4311–4323. doi:10.1007/s10064-018-1406-3.
- Truong, A. Y.-Q., Saway, B. F., Bouzaher, M. H., Rasheed, M. N., Monjazebe, S., Everest, S. D., Giampalmo, S. L., Hartman, D., Hartman, C., Kablinger, A. S. & Trestman, R. L. 2021 Systematic content analysis of patient evaluations of START NOW psychotherapy reveals practical strategies for improving the treatment of opioid use disorder. *BMC Psychiatry* **21**, 23–36. doi:10.1186/s12888-020-03024-x.
- Vadiati, M., Rajabi Yami, Z., Eskandari, E., Nakhaei, M. & Kisi, O. 2022 Application of artificial intelligence models for prediction of groundwater level fluctuations: Case study (Tehran-Karaj alluvial aquifer). *Environmental Monitoring and Assessment* **194**, 619. doi:10.1007/s10661-022-10277-4.
- Wang, Y.-Q., Wang, Z.-F. & Cheng, W.-C. 2019 A review on land subsidence caused by groundwater withdrawal in Xi'an, China. *Bulletin of Engineering Geology and the Environment* **78**, 2851–2863. doi:10.1007/s10064-018-1278-6.
- Wei, Z., Shang, Y., Sun, H., Xu, H. & Wang, D. 2019 The effectiveness of a drainage tunnel in increasing the rainfall threshold of a deep-seated landslide. *Landslides* **16**, 1731–1744. doi:10.1007/s10346-019-01241-4.
- Wei, Z., Wang, D., Sun, H. & Yan, X. 2020 Comparison of a physical model and phenomenological model to forecast groundwater levels in a rainfall-induced deep-seated landslide. *Journal of Hydrology* **586**, 124894. doi:10.1016/j.jhydrol.2020.124894.
- Xiao, Z., Liang, Z., Wang, J., Li, B., Hu, Y. & Wang, J. 2023 An improved butterfly optimization algorithm and its application in cascade hydropower generation operation. *Journal of Hydroinformatics* **25**, 1121–1138. doi:10.2166/hydro.2023.026.
- Yesilnacar, M. I. & Sahinkaya, E. 2012 Artificial neural network prediction of sulfate and SAR in an unconfined aquifer in southeastern Turkey. *Environmental Earth Sciences* **67**, 1111–1119. doi:10.1007/s12665-012-1555-9.
- Zhao, Y., Li, Y., Zhang, L. & Wang, Q. 2016 Groundwater level prediction of landslide based on classification and regression tree. *Geodesy and Geodynamics* **7**, 348–355. doi:10.1016/j.geog.2016.07.005.
- Zhu, C., Zhang, J., Liu, Y., Ma, D., Li, M. & Xiang, B. 2020 Comparison of GA-BP and PSO-BP neural network models with initial BP model for rainfall-induced landslides risk assessment in regional scale: A case study in Sichuan, China. *Nature Hazards* **100**, 173–204. doi:10.1007/s11069-019-03806-x.

First received 17 August 2023; accepted in revised form 28 October 2023. Available online 20 November 2023