IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 9, 2020, accepted December 21, 2020, date of publication December 28, 2020,
date of current version January 12, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3047626

Prediction of Water Inrush in Long-Lasting Shutdown Karst Tunnels Based on the HGWO-SVR Model

DUNWEN LIU, QIAN XU, YU TANG, AND YINGHUA JIAN
School of Resources and Safety Engineering, Central South University, Changsha 410083, China

Corresponding author: Yu Tang (tangyu9433@163.com)

This work was supported by the Postdoctoral Science Foundation of Central South University through the Initiation Fund for Postdoctoral Research under Grant 228697.

ABSTRACT It is very important to predict the long-term shutdown karst tunnel water inrush for preventing tunnel construction accidents. However, it is urgent to study a new prediction model to solve the problems of insufficient sample size and low prediction accuracy for long-term shutdown karst tunnel water inrush prediction. In this study, the water inrush and atmospheric rainfall in a tunnel project in China were monitored for over five months. By adopting hybrid grey wolf optimization (HGWO) algorithm and support vector regression (SVR) method, the HGWO-SVR tunnel water inrush prediction model was proposed. The atmospheric rainfall of the day and yesterday and yesterday's water inrush were considered in the HGWO-SVR model, and the model was used to predict the tunnel water inrush. The results show that the predicted water inrush value is basically consistent with the measured value. After the parameters of SVR model are optimized by HGWO algorithm, the HGWO-SVR prediction model has the advantages of high precision and less sample demand. The model is more suitable for the prediction of long-term shutdown tunnel water inrush with less measured sample. Thus, the proposed prediction model can effectively be used as a new approach for tunnel water inrush in some similar projects.

INDEX TERMS Karst tunnel, water inrush, long-lasting shutdown, hybrid gray wolf optimization, support vector regression, predictive analysis.

I. INTRODUCTION

Due to the development of construction for highway tunnels in China gradually advanced to the southwest, a large number of water inrush problems are becoming increasingly serious. Tunnel water inrush risk is throughout the whole life cycle from design, construction to operation management [1]. The uncertainty of water inrush leads to the risk of repeated stoppages and resumption. Long-time shutdown will delay the construction progress, and the huge scale of water inrush is easy to form secondary accidents such as collapse, sudden mud and even threatens the workers' life. For example, two major water inrush occurred in the Malujing Tunnel in Enshi, Hubei Province on January 21, 2006 and April 11, 2008, separately, resulted in 15 deaths and more than 2 years' delays. The Zhongjiashan Tunnel of the Jilian expressway in Jiangxi Province experienced 14 consecutive water inrush with a total volume of over 20,000 cubic meters from July to the end of October 2012. Therefore, the research on the influencing factors of water inrush and water volume prediction in long-time shutdown karst tunnel is of great practical significance and application value. Therefore, it is necessary to carry out water inrush prediction in long-time shutdown tunnel.

To control the safety risks induced by tunnel inrush, scholars have carried out many related academic studies on tunnel inrush prediction. For example, X. H. Xu proposed a deterministic mathematical model for the prediction of water inrush in a fissure enclosure tunnel under the seepage and stress coupled environment based on the theory of equivalent treatment of the permeability properties [2]. M. Jun used a dynamic design and information-based construction methods, distinguishing between the forward evolution of the design phase and the inverse method of the construction phase for the calculation of water inrush [3]. S.S. Shi revealed the mechanism that karst areas is prone to water inrush from the perspective of storage patterns. He systematically

The associate editor coordinating the review of this manuscript and approving it for publication was Jiajia Jiang.

analyzed the advantages and disadvantages of each prediction method and the applicable conditions [4]. Then, the different water-bearing geological formations of the karst zone was divided, and a more reasonable tunnel inrush prediction was proposed. R. Liu proposed a combination of experimental and numerical calculations to establish and analyze an evolutionary model of the permeability characteristics during the destruction of the surrounding rock [5]. He simulated the blasting process and water ingress in a submarine tunnel and successfully predicted the water inrush under blasting vibration. In view of the complicated mechanism of tunnel water inrush and the difficulty of quantitative prediction, X. Kang has realized the dynamic evaluation of water inrush in the gradual process of tunnel construction by various theoretical methods [6]. Fuzzy mathematical analysis was used to analyze the water inrush content of seven common ions, based on real-time monitoring data. Then, the groundwater types were classified and a hydrogeological model of the tunnel was developed to successfully predict the probability and volume of inrush. Y. Yang proposed an online discriminatory model for influx water using laser-induced fluorescence (LIF) and convolutional neural network (CNN) [7]. The model was used to automatically discriminate the influx water source. Based on this, K. Bian used the support vector machine model of cuckoo search optimization (CS-MSVM) to reconstruct the spectral signal at different stages [8]. The CEEMD signal processing method combined with LIF spectroscopy was used to effectively identify the source of influx water. H. N. Wu *et al.* proposed a new numerical method to simulate local water seepage in the lining of a shield tunnel by using a one-dimensional seepage unit in AQAQUS software, to simulate long-term seepage in the tunnel joints [9]. The method overcomes the limitations of the solid grid division of the seepage unit and reduces the calculation errors caused by the size difference between the seepage unit and the lining unit. Based on the control point source theory, X. X. Liu *et al.* proposed an anisotropic and isotropic groundwater flow prediction calculation method for tunnels in pressurized aquifers, taking various factors into consideration [10]. The method is well applied to actual projects for water inrush prediction. Y. X. Wu *et al.* established a three-dimensional flow-solid coupled finite element model to analyze the effects of the seepage of water-bearing barrier below the excavation surface [11]. Under the condition of constant flow rate and fixed pressure drop, they simulated the uniform and local weir wall leakage of the confined aquifer. This facilitates the understanding of the effects of diaphragm leakage on groundwater seepage and soil deformation.

Many scholars have also started to apply the machine learning algorithms to predict time series data. For example, the bidirectional long short-term memory (Bi-LSTM) network proposed by Shen S.L. *et al.*, predicts the future series of data based on the observational learning before the model, and performs its learning process to integrate the time sequence [12]. The proposed model can effectively predict the diameter of jet grouted columns in soft soil in real time, which is helpful for engineering construction. Based on the transitional Markov chain and Monte Carlo method, Mnasri S. *et al.* improved the Bayesian method for evaluating the generalization ability of prediction models and model selection [13]. The method effectively helps researchers to find the most appropriate model in terms of predictive ability, generalization ability, and model complexity. In addition, they propose a multi-step automatic model selection EPR technique based on multi-objective optimization [14]. The technique consists of two stages: intelligent rough model selection and model fineness identification. The technique can effectively address the difficulties of existing multi-objective evolutionary polynomial regression methods for optimal evolutionary polynomial model decision making. However, the above-mentioned machine learning research methods are applicable to large numbers of samples. When the number of study samples is limited, accuracy is difficult to guarantee. This article focuses on the problem of water inrush prediction under long-lasting shutdown, which has rarely been addressed by scholars. In addition, the actual data obtained from the project is limited in time span, which is not conducive to accurate prediction. Therefore, it is necessary to conduct further research on the prediction of water inrush in karst tunnels under long-lasting shutdown.

In this article, a mixed gray wolf optimization support vector regression (HGWO-SVR) prediction model is proposed, choosing the three factors of yesterday's rainfall, today's rainfall and yesterday's water inrush. Considering that the parameter establishment of the support vector machine is the key to its model building, the artificial intelligence optimization algorithm can better help the support vector machine to establish the model parameters compared to other search methods such as traditional mesh search [15], [16]. In this article, we use the global search optimization of differential evolution mixed with the local search of the gray wolf algorithm, the hybrid gray wolf optimization algorithm (HGWO), to optimize the parameters of the time series prediction supporting vector regression model. The prediction algorithm effectively avoids the inefficiency and subjectivity of traditional parameter search and other intelligent algorithm selections, and improves the training speed and scalability of the prediction model. The model is applied to the inrush prediction of a long-lasting downtime tunnel in Guizhou province, which provides a powerful reference for the scientific resumption and drainage organization.

II. THEORETICAL BASIS

A. HYBRID GRAY WOLF OPTIMIZATION ALGORITHM

The gray wolf algorithm is a global stochastic optimization method, inspired by the leadership hierarchy and hunting mechanism elements of the gray wolf in the natural environment. Through the testing of 29 known functions, it has been documented that the algorithm performs better than other heuristics (e.g., PSO, GSA, DE, EP, ES) in terms of spatial exploration, local optimization avoidance, and convergence

performance [17]. It exhibits excellent performance in single-peak, multimodal, and composite functions. It has been proven to work well in areas such as power dispatching and multiple input/output system optimization [18], [19]. The gray wolf's attacking behavior in the gray wolf algorithm is similar to local search, and deviating from the current prey to find better prey is similar to global search. Therefore, when the gray wolf attacks the prey, the gray wolf optimization (GWO) can easily come to a standstill. The DE algorithm was first proposed by Storn for solving global optimization problems [20]. Since the DE algorithm is a simple and powerful stochastic algorithm with few control parameters, it has been widely used in various engineering fields.

Hybrid gray wolf optimization (HGWO) is a combination of the gray wolf algorithm and differential evolution proposed by Zhu in 2015 [21], to artificially reduce the limitations of the gray wolf optimization (GWO). Differential evolution is used to start from a randomly generated population and generate the next generation of populations through mutation, crossover and selection operations. It ensures that the gray wolf algorithm jumps out of the local optimum and improves the convergence speed of the optimization process [22].

The principle was used to establish a wolf pack rank system and hunting process [23], with the following steps.

First, initialize the wolf pack via equation (1). The individuals are uniformly distributed in the space defined by the objective function.

$$X_{k,p}(0) = X_p^{min} + \text{rand}(0, 1) \times (X_p^{max} - X_p^{min}) \tag{1}$$

where $X_{k,p}$ (0) is the p-dimensional position of the search space for the kth wolf in the initial pack, rand(0,1) is a random constant between [0,1], X_p^{max} and X_p^{min} are the upper and lower limits of the p-dimensional search space respectively.

Based on the above steps, an initialized wolf pack is formed. And a social hierarchy model of the wolf pack is developed based on the adaptation ranking of each wolf. Appoint wolf X_α as the largest ranking, wolf X_β as the second ranking, and wolf X_δ as the third ranking.

Second, use equation (2) - equation (4) to update each wolf position.

$$\begin{cases} \vec{D}_\alpha = |\vec{C}_\alpha \cdot \vec{X}_\alpha - \vec{X}(t)| \\ \vec{D}_\beta = |\vec{C}_\beta \cdot \vec{X}_\beta - \vec{X}(t)| \\ \vec{D}_\delta = |\vec{C}_\delta \cdot \vec{X}_\delta - \vec{X}(t)| \end{cases} \tag{2}$$

$$\begin{cases} \vec{X}_1(t+1) = \vec{X}_\alpha(t) - \vec{A}_1 \cdot \vec{D}_\alpha \\ \vec{X}_2(t+1) = \vec{X}_\beta(t) - \vec{A}_2 \cdot \vec{D}_\beta \\ \vec{X}_3(t+1) = \vec{X}_\delta(t) - \vec{A}_3 \cdot \vec{D}_\delta \end{cases} \tag{3}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1(t+1) + \vec{X}_2(t+1) + \vec{X}_3(t+1)}{3} \tag{4}$$

where t is the current iterations, the initial generation t=1; $\vec{X}$ is a vector of individual wolf positions relative to prey; $\vec{C}$ is a random vector between 0 and 2; $\vec{A}$ is a dynamic random vector between -2 and 2 contracting progressively towards 0, when $|\vec{A}| < 1$, the wolves attack the prey, otherwise the wolves move away from the prey and look for more suitable prey.

Then, use equation (5) for differential evolution to generate mutant individual.

$$P_i^{t+1} = X_{k1}^t + Z(X_{k2}^t - X_{k3}^t) \tag{5}$$

where Z is a scaling factor between 0.2 and 0.8; X_{k1}^t, X_{k2}^t, X_{k3}^t are three different random individuals in the t generation population.

Next, use crossover manipulation to generate test individuals U_{ij}^{t+1}.

$$U_{ij}^{t+1} = \begin{cases} P_{ij}^{t+1}, & rand\ (0, 1) \le CR \text{ or } j = randn(i) \\ X_{ij}^t, & rand\ (0, 1) > CR\ and\ j \ne randn(i) \end{cases} \tag{6}$$

where j is the jth variable of the test individual; randn(i) $\in$ [1, 2, …, D] represents a randomly selected index of dimensional variables, D is the dimension of the trial individual; X_{ij}^t is the target individual; P_{ij}^{t+1} is the variant individual and CR is the crossover probability, $CR \in [0, 1]$.

The generated test individuals are then competed by selection through equation (7) to determine whether the individual is passed on to the next one.

$$X_i^{t+1} = \begin{cases} U_i^{t+1}, f(U_i^{t+1}) < f(X_i^t) \\ X_i^t, \quad f(U_i^{t+1}) \ge f(X_i^t) \end{cases} \tag{7}$$

where f is the adaptation function of the optimization objective.

Finally update wolves X_α, X_β and X_δ of the wolf pack according to the order of adaptation. Proceed with the next iteration of the cycle in sequence, i.e., t= t+1, to determine the location of the prey.

$$X^{t+1} = \frac{X_1^{t+1} + X_2^{t+1} + X_3^{t+1}}{3} \tag{8}$$

Repeatedly carry out differential, crossover, and inter-individual competition to keep up with the new wolf pack rank. Until the maximum cycle algebra is reached, the X_α is returned. Finally, the selected alpha wolf is the optimal solution space of the problem.

B. THE PRINCIPLE OF SUPPORT VECTOR MACHINE REGRESSION

Support Vector Machine (SVM) is an approach based on the statistical theory of the VC dimension and the principle of structural risk minimization proposed by Vapnik [24]. It addresses the problems such as small samples, nonlinearity, high dimensionality, and local minima. The aim is to find the maximum spaced hyperplane that minimizes the distance between all sample points and the hyperplane. The input variables are mapped to a high-dimensional feature space by appropriate nonlinear function transformations, and a linear regression optimal hyperplane is sought in this space. Ultimately, the regression is attributed to the convex planning problem to find the global optimal solution. Its mathematical description is given in Eq. (9), as shown at the bottom of the next page, where α_i, , α_i^*, α_j, α_j^* are Lagrange multipliers, y_i is observed value, c is error penalty parameter, ε is the

upper error limit, and $K(x_i,x_j)$ is the nuclear function. There are three types of nuclear functions in common use: polynomial nuclear functions $k(x_i,x_j) =(x_ix_j+1)d$, $d=1,2, \ldots$; radial basis nuclear function (RBF) $k(x_i, x_j) =\exp(-\gamma x_i\text{-}x_j^2)$; Sigmoid nuclear function $k(x_i, x_j) =\tanh[\beta(x_ix_j)+c]$.

The RBF kernel function is widely used in low-dimensional, high-dimensional, small-sample and large-sample datasets with wide convergence domain. Therefore, it is chosen as the SVR kernel function in this article.

C. SUPPORT VECTOR MACHINE PREDICTION MODEL OPTIMIZED BY HYBRID GRAY WOLF ALGORITHM

The choice of parameters affects the prediction performance of the SVR. The values of the penalty factor C and the RBF kernel parameter g directly determine the prediction accuracy and the speed of the SVR. Therefore, the rational combination of C and g is the key to perform model building [25], [26].

Intelligent algorithms widely used for SVR parameter optimization include ant colony algorithm, gravity algorithm, and simulated annealing algorithm [27]. However, these algorithms have such drawbacks as precocious convergence, low solution accuracy, and slow convergence speed. Therefore, they cannot meet the convergence speed and accuracy of high-dimensional problems.

In contrast, the hybrid gray wolf algorithm uses an evolutionary, displacement model to achieve the entire solution space for optimal search. The parallel global search strategy based on the wolf pack is easy to implement and does not require complicated parameter adjustment. Therefore, it converges quickly and has a strong advantage in high-dimensional problems [28], [29].

To obtain the best combination of penalty factor C and kernel parameter g, the HGWO algorithm is employed to perform a fast global optimization search. It can reduce the blindness of trial calculations and improve the accuracy of model prediction. The implementation steps of the HGWO-SVM model are shown in Fig.1.

Step 1: Enter the time series of factors related to the target to form a multidimensional time series set. Set up a training set and a prediction set for the SVR. Initialize the SVR penalty factor C and the range of the kernel parameter g. Set up the HGWO-related parameters, including population size NP, scaling factor upper and lower bounds $\beta_{\max}$ and $\beta_{\min}$, cross probability CR, and evolutionary algebra N_iteration.

Step 2: Generate the initial parent wolf pack according to the initial C and g. The training set is learned and the adaptation degree of each wolf is calculated. The wolves are divided into three levels: α, β, and δ according to the adaptation values. According to equations (2)-(4), each individual position in the new parent wolf pack is followed. After that, perform variation, crossover, and selection operations of differential evolution according to Eqs. (5)-(7) to produce a variant wolf pack and a daughter wolf pack. The adaptation of each individual gray wolf to the new location was calculated and compared with the best adaptation Fg of the previous iteration. If the adaptation degree is better than Fg, the individual gray wolf replaces the group's best adaptation degree, otherwise the original best adaptation degree is kept. If the number of iterations exceeds the set evolutionary algebra N_iteration, the training ends with the output of the global optimal position, i.e., the optimal combination of C and g parameters of the SVR model, otherwise the parameter optimization continues.

Step 3, choose the best combination of (C, g) parameters to build the prediction model, and perform prediction and result analysis on the validation set. If the error does not meet the requirement, skip to step 1.

III. ENGINEERING APPLICATION

A. ENGINEERING BACKGROUND

Relying on the right line of a two-lane highway tunnel project in Guizhou province, this article carried out a specific investigation and analysis. According to the hydrological conditions, geological structure and the degree of karst development in the tunnel site area, the right tunnel is divided into three sections. They are YK20+480 (tunnel entrance)∼YK21+100, YK21+100∼YK21+960 and YK21+960∼YK24+300 (tunnel exit). The risk of water intrusion is low in the first and last sections of the tunnel. YK20+480 (tunnel entrance) to YK21+100 (tunnel exit) is the middle and upper Silurian (S_{2+3}), which mainly consists of purple-red and gray-green medium-thickness siltstone, siltstone shale, shale and limestone lenses, and yellow-gray thin to medium-thickness quartz sandstone, siltstone and siltstone shale at the upper part, which is a water-insulating rock layer. The YK21+960∼YK24+300 (tunnel exit) section at the end of the tunnel includes the Upper Intermediate Ordovician (O_{2+3}) and the Lower Shillouian (S_1). The upper and middle Ordovician mainly consists of gray-green shale, gray-purple thin to medium-thickness limestone, mudstone, marl, and calcareous shale, with black carbonaceous shale at the top, and the formation is a water-insulating layer. The Lower

$$\begin{cases} \left\{ \max_{a_i a^*} -\left[\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\left(a_i-a_i^*\right)\left(a_j-a_j^*\right)K\left(x_i,x_j\right) - \sum_{i=1}^{l}\left(a_i+a_i^*\right)\varepsilon + \sum_{i=1}^{l}\left(a_i-a_i^*\right)y_i \right] \right. \\ \text{s.t.} \begin{cases} \sum_{i=1}^{l}\left(a_i-a_i^*\right)=0 \\ 0\le a_i\le C \\ 0\le a_i^*\le C \end{cases} \end{cases} \tag{9}$$

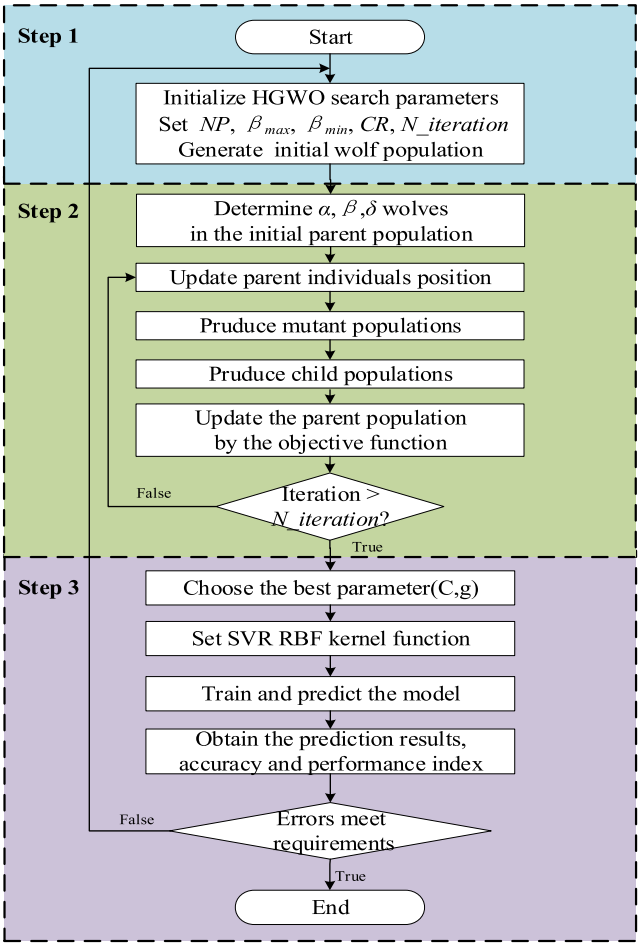

FIGURE 1. Flowchart of the HGWO-SVR algorithm.

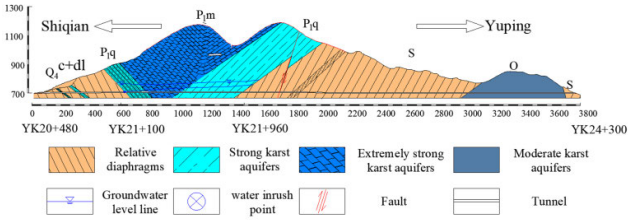

FIGURE 2. Hydrogeological map of the study site.

FIGURE 3. Water inrush from the upper left of the palm surface at YK21+122.8.

Silurian (S_1) mainly consists of gray siltstone, carbonaceous shale with calcareous siltstone and sandy limestone, marl or limestone, and the whole stratum is a water-insulating rock layer, in which the limestone interlayer contains a small amount of fissure water. Among them, the first and last sections are the Shillou system insulation layers. They prevent the movement of groundwater, thus the possibility of water inrush is extremely low. The middle section of the tunnel YK21+100~YK21+960 crosses the water-bearing karst layer of the Permian system, Sumida-Maukou group (P_{1q+m}) in the oblique core. The strong soluble rocks in the region have a high $CaCO_3$ content, so there is a high risk of water inrush disaster. The geological and hydrological conditions of the study area and its sections are shown in Fig.2.

The YK21+100~YK21+960 interval in the middle of the tunnel is the Lower Permian Qixia-Maukou Group (P_{1q+m}). The lower part is mainly grayish-black thin-layered carbonaceous tumbrous limestone and limestone with carbonaceous shale, containing a large number of flint nodules, while the middle part is dark gray to gray medium-thick to thick laminated flint-bearing mass limestone, and the upper part is grayish-white massive limestone. The stratum thickness of the Qixia Group (P_{1q}) is about 150-160m, which is a strong karst aquifer. The stratum thickness of Maukou Group (P_{1m}) is about 130~221m, which is an extremely strong karst aquifer. The saturation compressive strength of the surrounding rocks in this section is low, only 12.3Mpa, which is about 1/5 of the strength of the surrounding rocks in other sections of this section. The surrounding rock body is relatively broken, so it is presumed that the tunnel arch and side walls are less self-stable. Then, the tunnel is prone to accidents such as falling blocks and collapsing during excavation. Therefore, the risk of water inrush is higher. According to the preliminary geological investigation, there are a number of karst channels in the rock body that are approximately parallel to each other and rich in groundwater. The water-bearing rock body is easy to be uncovered and destroyed, resulting in water inrush accident. In the actual excavation, the first occurrence of water inrush in the tunnel right line YK21+122.8 is above the left side of the palm surface, as shown in Fig.3.

As the tunnel palm face advances, the carbonaceous perimeter rock behind the initial support is lost with water, resulting in concrete stripping. The left arch foot of YK21+128 is the main gushing point, as shown in Fig.4.

With the excavation to YK21+139, a vertically developed cave about 2.2m×2.2m was found on the left side of the palm surface. Through the field observation, the height of the cave from the upper step upward development is not yet identified, but there is no filler and water inrush phenomenon. After part of the cave slag fell into the cave during the blasting, the depth of the cave is about 7m from the upper step

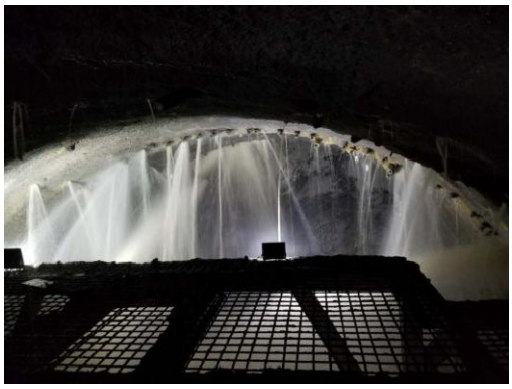

FIGURE 4. Concrete stripping due to water inrush.

FIGURE 5. Cave development at YK21+139, left is up view and right is down.

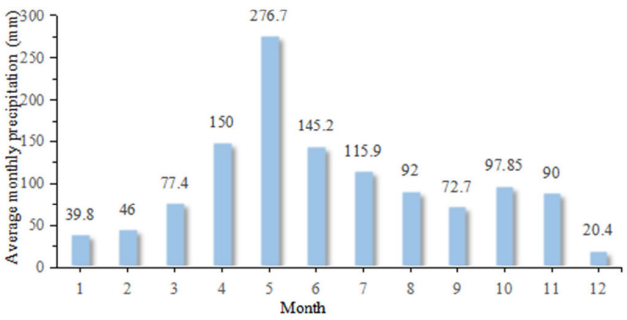

FIGURE 6. Multi-year monthly average precipitation in the county.

FIGURE 7. The on-site rainfall automatic rain gauge layout diagram.

downward through the probe rope. The cave site conditions are shown in Fig.5.

The severe water inrush brought the tunnel to shutdown for nearly 10 months from late September 2018 to mid-July 2019. It resulted in delays and a sharp increase in personnel and equipment costs. As a result, a more accurate prediction of water inrush during the downtime is a key to ensure safe working environment and can guide workers to resume work in a scientific manner.

B. DATA COLLECTION

For the time series prediction of water inrush in karst tunnel during shutdown, rainfall is the main influencing factor [30], [31]. Following the rainfall in the tunnel area, surface water eventually becomes tunnel inrush water, after the absorption and infiltration of the upper overburden and the hydraulic transport of the lower aquifer. Thus, there is an obvious time lag between surface rainfall and groundwater inrush. Both rainfall size and duration have an impact on the water inrush [32].

The historical mean values of rainfall in this county of Guizhou Province are shown in Fig.6.

The histogram distribution shows that the county has a high rainfall period from April to July, with an average monthly rainfall of 100 mm or more. Therefore, a focused analysis is required.

In order to carry out long-term dynamic hydrological monitoring of a tunnel in Guizhou province, the experimental team set up a monitoring device on site to monitor the actual rainfall in the tunnel area. An automatic rain gauge with a time step of 10min and a precision of 0.1mm was used to collect rainfall data directly. The on-site equipment layout is shown in Figure 7.

The time series of rainfall from February to July 2019 was obtained through field measurements. The value distribution is shown in Fig.8.

From Fig.8, it can be seen that April-July are high rainfall months during the monitoring period (February-July). And the maximum daily rainfall is more than 30mm in four months. Also, the daily rainfall changes more drastically, with the maximum value reaching 88.9mm.

On-site monitoring of water inrush is carried out using the thin-walled rectangular weir flow method. The arrangement of the weir is shown in Fig.9.

A thin-walled rectangular weir was constructed at the tunnel entrance at the drainage culvert. Water level probes were installed inside the weir to monitor the water level at the mouth of the weir. The thin-walled rectangular weir flow formula (10) is used to calculate the daily flow values.

$$Q = 1000mb\sqrt{2g}h^{\frac{3}{2}} \tag{10}$$

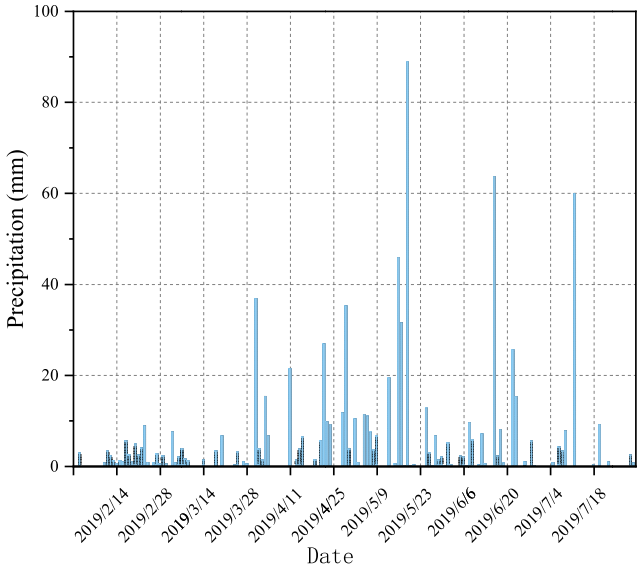

FIGURE 8. Time-series observation of rainfall in a tunnel in Guizhou province from February to July.

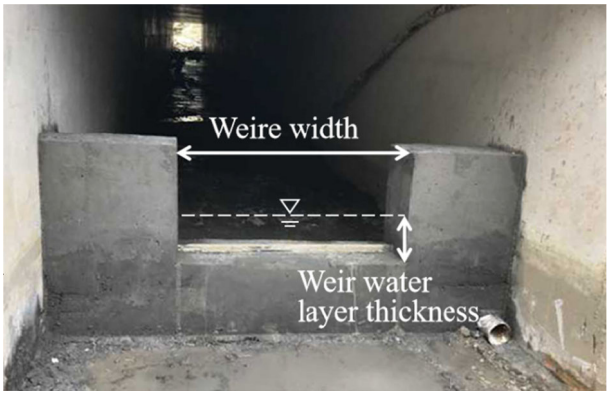

FIGURE 9. The thin wall rectangular weir flow arrangement.

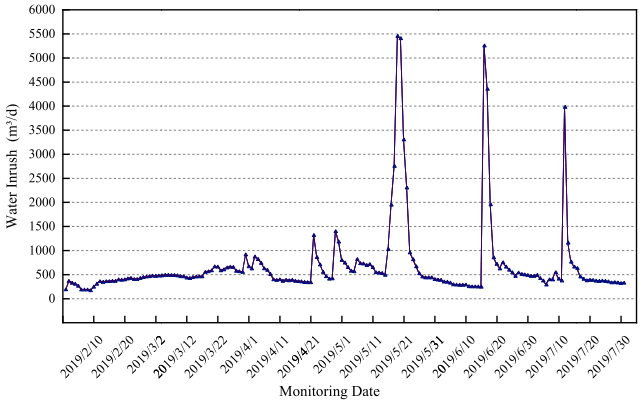

FIGURE 10. Time series of water inrush in a tunnel in Guizhou province from February to September.

where Q is the water flow rate (l/s); m is the flow coefficient; b is the width of the weir; h is the thickness of the water layer at the mouth of the weir; and g is the gravity acceleration.

The time series of water inrush from February to July 2019 was obtained from field measurements. The distribution of their values is shown in Fig.10.

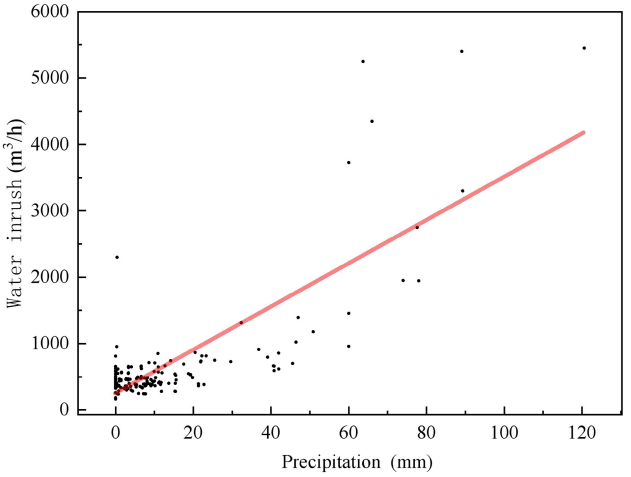

FIGURE 11. Fit of the linear relationship between rainfall and water inrush.

As can be seen in Fig.10, the February-July water inrush time series is roughly about 500m^3/h. As a result of increased rainfall from May to July, the value from the middle to the end of the month are showing a cumulative increase to the peak and then cumulative decline. In terms of days, there is still an overall time dependence in the amount of water inrush. The amount of the current day's inrush is influenced by the magnitude of the previous. The linear fit analysis was plotted as shown in Fig.11, with the sum of the current day, yesterday and the previous day's rainfall as the horizontal coordinate and the same-day inrush as the vertical coordinate.

The linear fit analysis consists of 181 sample points from February to July 2019, and the fitted R^2 is 0.63616. A clear linear relationship between precipitation and the amount of water inrush in the tunnel area can also be seen in the figure.

C. DATA PROCESSING AND ANALYSIS

1) DATA PREPROCESSING

The noisy data collected are cleaned and abnormal data are processed. The rainfall is mainly monitored by the automatic rain gauge. When there is a malfunction of the gauge or the data deviates greatly due to other reasons, the weather station is used to monitor the rainfall data for correction. The daily gush of water is monitored by a thin-walled rectangular weir, and the first record is made at 6:00 a.m. and the second record is made every four hours. When there is a significant deviation in the data, the data at that time is discarded and the mean value is filled in.

2) CORRELATION ANALYSIS

In order to analyze whether there is a correlation between the tunnel rainfall and the amount of water inrush, the correlation coefficient is used to quantitatively characterize.

$$\gamma_{xy} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{(x-\bar{x})^2(y-\bar{y})^2}} \tag{11}$$

where the correlation coefficient γ_{xy} ranges from [−1,1].

TABLE 1. Correlation coefficients between precipitation and water inrush during the downtime of a karst tunnel in Guizhou province from February to July.

Month	February	March	April	May	June	July
γ_{xy}	0.42	0.65	0.67	0.47	0.69	0.91

TABLE 2. Time lagged correlation coefficients between precipitation and water inrush data for April-July.

Lag days	0	1	2	3
γ_{xy}	0.5484	0.5126	0.0131	0.3059

The water inrush during shutdown is influenced by atmospheric evaporation, environmental conditions and other factors, so the water inrush and rainfall does not exist an absolute correlation. Therefore, we set that the two are not correlated when $|\gamma_{xy}| < 0.1$, weakly correlated when $0.1 \leq |\gamma_{xy}| < 0.3$, strongly correlated when $0.3 \leq |\gamma_{xy}| < 0.5$, and highly correlated when $|\gamma_{xy}| \geq 0.5$ in this article. According to equation (10), the correlation coefficient value of the tunnel's February-July rainfall and water inrush was calculated as shown in Table 1.

Based on the multi-year average monthly rainfall distribution in the region as shown in Fig.8, and the correlation coefficients mentioned above, specific analyses were performed. Compared to the rainy season (April-July), the dry season (November-March) has relatively little rainfall. The rainfall is scattered and only wets the surface soil after a single rainfall. Even if there is an occasional heavy rainfall, the amount of rainfall infiltrated by the infiltration and recharge of the surface air pocket is negligible, which will not have a significant impact on the change of water inrush in the tunnel. Therefore, when examining the correlation between rainfall and water inrush in the tunnel, only specific analyses of the rainfall and water inrush during the rainy season (April-July) are available in this article.

According to equation (11), the time-lagged correlation between the precipitation and the water inrush in the tunnel from April to July is calculated, and the results are shown in Table 2. As can be seen from the table, the number of correlations between the amount of water inrush and precipitation on the day and the day before is significantly greater than the number of other days with a lag.

Therefore, the long-term downtime tunnel inrush prediction model is constructed in this article by selecting three factors: the day's rainfall, yesterday's rainfall, and yesterday's water inrush. Also, from the parameter sensitivity analysis, the order of the importance for the results was as follows: the day's rainfall, yesterday's water inrush and yesterday's rainfall.

3) SELECTION OF TRAINING SETS

Based on the above model construction, each factor was quantitatively generalized. The sample from April-June was selected as the training set and the sample from July was used as the test set. The training sample set is shown in Fig.12.

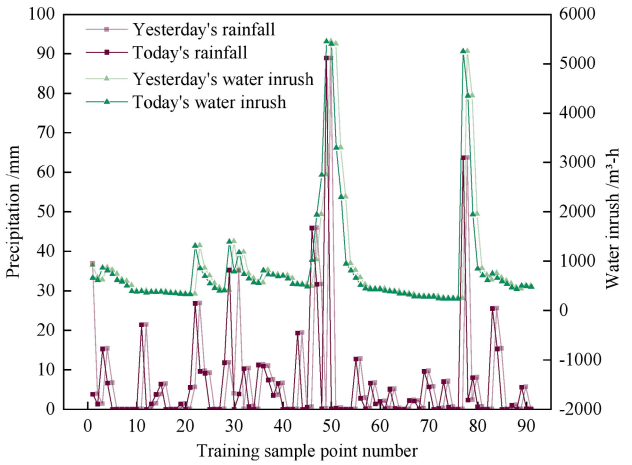

FIGURE 12. Training samples set.

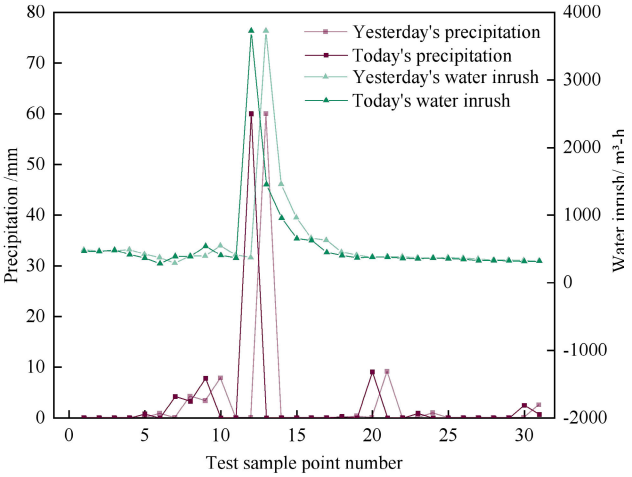

FIGURE 13. Test samples set.

Where yesterday's rainfall, today's rainfall and yesterday's water inrush are used as model inputs and today's water inrush is used as model output.

The set of test samples is constructed according to the same logic to form a consistent input or output correspondence. The specific distribution of the test sample set is shown in Fig.13.

4) PARAMETER SELECTION AND OPTIMIZATION

The HGWO-SVR program was written using a programming software. The parameters of the algorithm were set as: population size nPop=40, independent variable dimension nVar=2 (penalty factor C and nuclear parameter g respectively), search factor $\beta_{\min}$ =0.2, $\beta_{\max}$ =0.8, cross probability Pcr=0.2, and evolutionary algebra N_iteration=2000. The specific optimization steps were carried out step by step according to the algorithm flow Fig. 1. The model runs for 80.051150 seconds. After the running, the search yields the optimal parameter of the support vector machine as C=92.2261 and the kernel parameter g=42.5804. The adaptation optimization process is shown in Fig.14.

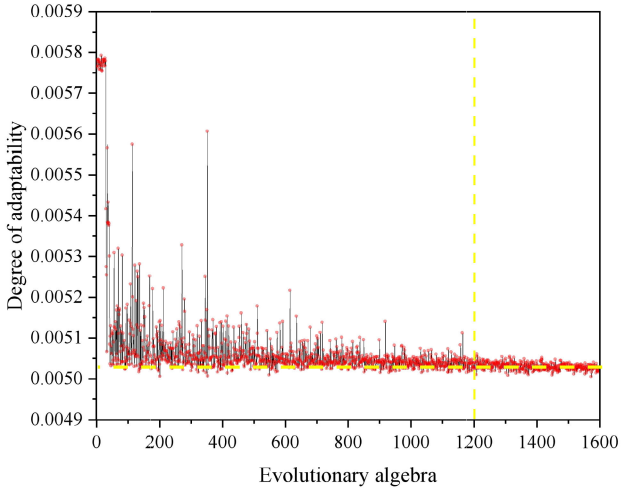

FIGURE 14. HGWO-SVR parameter optimization process.

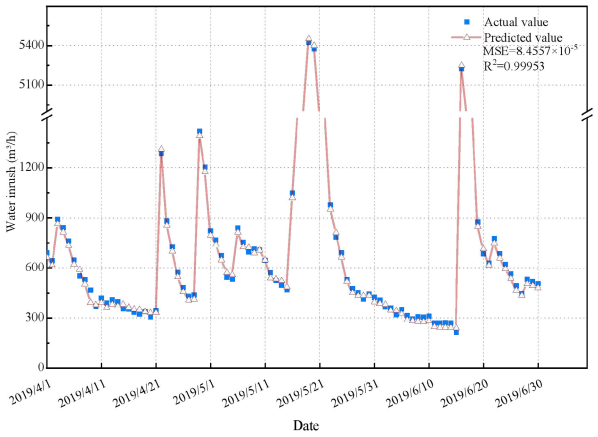

FIGURE 15. Fitting result of the training set based on the HGWO-SVR model.

It can be seen from Fig.14 that the evolutionary algebra is selected at 1600 reasonable steps, the population adaptation tends to stabilize at about the 1200th step, and the target function fg adaptation value eventually stabilizes at about 0.00504. In addition, as the number of iterations increases and the parameter search is closer to the target value, the estimated value becomes more stable, the coefficients of the selected model become more reliable, and the fitted prediction model becomes more accurate.

D. ANALISIS OF RESULTS

The best parameters $c = 92.2261$ and $\sigma = 42.5804$ optimized by HGWO were substituted into the SVR prediction model for the inrush prediction, and the sample training is shown in Fig.15.

The running results show that the training set has a mean square error (MSE) of 8.4557×10^{-5} and a squared correlation coefficient (R^2) of 0.99953.

The test set mean square error (MSE) of the HGWO-SVR model is 1.945×10^{-3}, and the test set coefficient of determination (R^2) is 0.97794. The smaller the MSE is, and

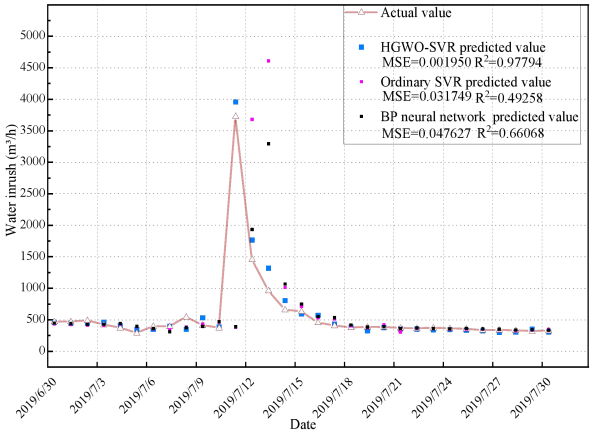

FIGURE 16. Predicted result of the test set by three methods.

the closer the R^2 is to 1, the better the prediction obtained. Thus, the HGWO-SVR model results are in good agreement with the measured data for the inrush prediction of water-rich karst long-term downtime tunnels. It shows the strong generalization ability and high accuracy in nonlinear prediction.

To show the advantages of the model, this article also uses the traditional grid search method to find the best SVM and an artificial neural network to compare the prediction effect. From the prediction curves, the SVM of the ordinary grid search method for parameter optimization and BP neural network also show some ability to make predictions for real problems. However, the accuracy of the prediction results is limited. In particular, the BP neural network exhibits large fluctuations and hysteresis near the inflection point, and both the prediction accuracy and the fitting effect are lacking compared to the HGWO-SVR model.

In order to further specifically compare the effects of the three prediction methods, this article introduces three statistical indicators of prediction, mean absolute percentage error (MAPE), Theil inequality coefficient (TIC), and prediction direction accuracy (DA) to comprehensively evaluate the prediction results [33].

The smaller the value of the calculated mean absolute percent error, the higher the accuracy of the prediction results, via Eq. (12).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i} \times 100\right| \tag{12}$$

The Hill's inequality coefficient is between 0 and 1. The smaller the value, the smaller the fit error, via Eq. (13).

$$TIC = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i)^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i)^2}} \tag{13}$$

The directional accuracy DA is calculated by Eqs. (14-15), and the closer its value to 1, the better the overall prediction

TABLE 3. The comparison of predicted effects.

Predictive methods	Statistical evaluation indicators		
	MAPE	TIC	DA
HGWO-SVM	10.7917	0.0909	0.5667
Ordinary SVM	41.1939	0.6360	0.4667
BP	43.1847	0.691034854	0.5

accuracy.

$$DA = \frac{1}{n}\sum_{i=1}^{n} A_i \tag{14}$$

$$A_i = \begin{cases} 1 & (\hat{y_{i+1}} - \hat{y_i})(y_{i+1} - y_i) > 0 \\ 0 & (\hat{y_{i+1}} - \hat{y_i})(y_{i+1} - y_i) \leq 0 \end{cases} \tag{15}$$

where $\hat{y_i}$ is the predicted value and y_i is the actual value.

In order to verify the stability of the model results, this article calculates the mean values of the evaluation indicators through a ten-fold cross-validation for inter-algorithm comparison. The mean values of the above three statistical indicators are used to comprehensively evaluate the prediction results, see Table 3.

As shown in Fig.16 and Table 3, the prediction curves of the three methods agree well with the real value curves. However, the HGWO-SVM prediction curve is closest to the true value. The evaluation index reflects the closeness of the prediction data from different aspects. The accuracy and directional accuracy of HGWO-SVM prediction are higher than the other, and the overall evaluation degree reaches the ideal prediction state. Therefore, the HGWO-SVM prediction is significantly better than the ordinary SVM and BP models. It not only has strong parametrization and approximation ability, but also has simple algorithm and good robustness, which is suitable for practical engineering applications.

IV. CONCLUSION

In this study, an intelligent water inrush prediction algorithm based on HGWO-SVR is proposed for long-lasting shutdown tunnel. The method embeds HGWO in support vector regression, searches for the optimal combination of parameters (C, g), and performs an error analysis on the results of the HGWO-SVR model. The generalization ability and prediction accuracy of the model are guaranteed. Then, according to the data collected from field measurements, the improved HGWO-SVR intelligent prediction algorithm was used to predict the water consumption in a long-lasting shutdown tunnel. The applicability of this method was verified through a long downtime tunnel project in Guizhou Province. The results of this study are as follows.

(1) The results of precipitation monitoring data and water inrush accident analysis show that there is a significant linear relationship between cumulative rainfall and tunnel water inrush time series. Through the correlation coefficient, it is concluded that the current day's water inrush is affected by yesterday's and today's rainfall and yesterday's water inrush, which provides a reference index for the time series prediction.

(2) The HGWO-SVR model is used to predict the water inrush in a long-lasting shutdown karst tunnel. The results show that proposed model can improve the search performance of the SVR model, as well as having a good ability to predict the water inrush. Using the obtained hyper-parameter combination, the prediction results are small in error and high in accuracy. The prediction error is within the acceptable range from the engineering point of view.

(3) Moreover, the proposed HGWO-SVR prediction model is better than using any other prediction model(e.g. the ordinary SVM and BP models) in terms of reducing the number of field tests and improving economic efficiency. According to the promising results of the proposed approach (HGWO-SVR) in predicting the water inrush for a long-lasting shutdown karst tunnel, we will combine this model with real-time tunnel water inrush monitoring system in future, its real-time guidance advantages will be highlighted, and the construction workers will be able to judge the flood risk of the tunnel in a more timely and effective manner.

REFERENCES

[1] S. Li, R. Liu, Q. Zhang, and X. Zhang, "Protection against water or mud inrush in tunnels by grouting: A review," *J. Rock Mech. Geotech. Eng.*, vol. 8, no. 5, pp. 753–766, Oct. 2016, doi: 10.1016/j.jrmge.2016.05.002.

[2] X. H. Xu, M. G. Xiao, D. H. Xu, and J. Cao, "Monitoring and results analysis of tunnel engineering," *Appl. Mech. Mater.*, vol. 256, pp. 1406–1410, Dec. 2013, doi: 10.1109/TTHZ.2016.2544142.

[3] M. Jun and Z. Yingmei, "A new dynamic assessment for multi-parameters information of water inrush in coal Mine*," *Energy Procedia*, vol. 16, pp. 1586–1592, Jan. 2012, doi: 10.1016/j.egypro.2012.01.247.

[4] S. Shi, L. Bu, S. Li, Z. Xiong, X. Xie, L. Li, Z. Zhou, Z. Xu, and D. Ma, "Application of comprehensive prediction method of water inrush hazards induced by unfavourable geological body in high risk karst tunnel: A case study," *Geomatics, Natural Hazards Risk*, vol. 8, no. 2, pp. 1407–1423, Dec. 2017, doi: 10.1080/19475705.2017.1337656.

[5] R. Liu, Y. Liu, D. Xin, S. Li, Z. Zheng, C. Ma, and C. Zhang, "Prediction of water inflow in subsea tunnels under blasting vibration," *Water*, vol. 10, no. 10, p. 1336, Sep. 2018, doi: 10.3390/w10101336.

[6] X. Kang, S. Luo, M. Xu, Q. Zhang, and Y. Yang, "Dynamic estimating the karst tunnel water inflow based on monitoring data during excavation," *Acta Carsologica*, vol. 48, no. 1, pp. 117–127, Mar. 2019, doi: 10.3986/ac.v48i1.4654.

[7] Y. Yang, J. Yue, J. Li, and Z. Yang, "Mine water inrush sources online discrimination model using fluorescence spectrum and CNN," *IEEE Access*, vol. 6, pp. 47828–47835, Aug. 2018, doi: 10.1109/ACCESS.2018.2866506.

[8] K. Bian, M. Zhou, F. Hu, W. Lai, and M. Huang, "CEEMD: A new method to identify mine water inrush based on the signal processing and laser-induced fluorescence," *IEEE Access*, vol. 8, pp. 107076–107086, 2020, doi: 10.1109/ACCESS.2020.3000333.

[9] H.-N. Wu, S.-L. Shen, R.-P. Chen, and A. Zhou, "Three-dimensional numerical modelling on localised leakage in segmental lining of shield tunnels," *Comput. Geotechnics*, vol. 122, Jun. 2020, Art. no. 103549, doi: 10.1016/j.compgeo.2020.103549.

[10] X.-X. Liu, S.-L. Shen, Y.-S. Xu, and Z.-Y. Yin, "Analytical approach for time-dependent groundwater inflow into shield tunnel face in confined aquifer," *Int. J. for Numer. Anal. Methods Geomech.*, vol. 42, no. 4, pp. 655–673, Mar. 2018, doi: 10.1002/nag.2760.

[11] Y.-X. Wu, H.-M. Lyu, S.-L. Shen, and A. Zhou, "A three-dimensional fluid-solid coupled numerical modeling of the barrier leakage below the excavation surface due to dewatering," *Hydrogeology J.*, vol. 28, no. 4, pp. 1449–1463, Jun. 2020, doi: 10.1007/s10040-020-02142-w.

[12] S.-L. Shen, P. G. A. Njock, A. Zhou, and H.-M. Lyu, "Dynamic prediction of jet grouted column diameter in soft soil using bi-LSTM deep learning," *Acta Geotechnica*, pp. 1–13, Jul. 2020, doi: 10.1007/s11440-020-01005-8.
[13] Y. Jin, Z. Yin, W. Zhou, and J. Shao, "Bayesian model selection for sand with generalization ability evaluation," *Int. J. Numer. Anal. Methods Geomech.*, vol. 43, no. 14, pp. 2305–2327, Oct. 2019, doi: 10.1002/nag.2979.
[14] Y.-F. Jin and Z.-Y. Yin, "An intelligent multi-objective EPR technique with multi-step model selection for correlations of soil properties," *Acta Geotechnica*, vol. 15, no. 8, pp. 2053–2073, Aug. 2020, doi: 10.1007/s11440-020-00929-5.
[15] S. Mnasri, N. Nasri, M. Alrashidi, A. van den Bossche, and T. Val, "IoT networks 3D deployment using hybrid many-objective optimization algorithms," *J. Heuristics*, vol. 26, no. 5, pp. 663–709, Oct. 2020, doi: 10.1007/s10732-020-09445-x.
[16] S. Mnasri, N. Nasr, VDB. Adrien, and V. Thierry, "A comparative analysis with validation of NSGA-III and MOEA/D in resolving the 3D indoor redeployment problem in DL-IoT," in *Proc. IEEE Int. Conf. Internet Things, Embedded Syst. Commun. (IINTEC)*, Gafsa, Tunisia, Oct. 2017, doi: 10.1016/j.egypro.2012.01.247.
[17] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014, doi: 10.1016/j.advengsoft.2013.12.007.
[18] M. H. Sulaiman, Z. Mustaffa, M. R. Mohamed, and O. Aliman, "Using the gray wolf optimizer for solving optimal reactive power dispatch problem," *Appl. Soft Comput.*, vol. 32, pp. 286–292, Jul. 2015, doi: 10.1016/j.asoc.2015.03.041.
[19] A. A. M. El-Gaafary, Y. S. Mohamed, A. M. Hemeida, and A.-A.-A. Mohamed, "Grey wolf optimization for multi input multi output system," *Universal J. Commun. Netw.*, vol. 3, no. 1, pp. 1–6, Feb. 2015, doi: 10.13189/ujcn.2015.030101.
[20] R. Storn, "Differential evolution-a simple and efficient heuristic for global optimization over continuous space," *J. Global Optim.*, vol. 11, no. 1, pp. 341–359, Dec. 1997, doi: 10.1023/A:1008202821328.
[21] A. Zhu, C. Xu, Z. Li, J. Wu, and Z. Liu, "Hybridizing grey wolf optimization with differential evolution for global optimization and test scheduling for 3D stacked SoC," *J. Syst. Eng. Electron.*, vol. 26, no. 2, pp. 317–328, Apr. 2015, doi: 10.1109/JSEE.2015.00037.
[22] M. Niu, Y. Hu, S. Sun, and Y. Liu, "A novel hybrid decomposition-ensemble model based on VMD and HGWO for container throughput forecasting," *Appl. Math. Model.*, vol. 57, pp. 163–178, May 2018, doi: 10.1016/j.apm.2018.01.014.
[23] M. Niu, Y. Wang, S. Sun, and Y. Li, "A novel hybrid decomposition-and-ensemble model based on CEEMD and GWO for short-term PM2.5 concentration forecasting," *Atmos. Environ.*, vol. 134, pp. 168–180, Jun. 2016, doi: 10.1016/j.atmosenv.2016.03.056.
[24] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 988–999, Sep. 1999, doi: 10.1109/72.788640.
[25] S. M. Clarke, J. H. Griebsch, and T. W. Simpson, "Analysis of support vector regression for approximation of complex engineering analyses," *J. Mech. Des.*, vol. 127, no. 6, pp. 1077–1087, Nov. 2005, doi: 10.1115/1.1897403.
[26] Z. Chen and W. Liu, "An efficient parameter adaptive support vector regression using K-Means clustering and chaotic slime mould algorithm," *IEEE Access*, vol. 8, pp. 156851–156862, Aug. 2020, doi: 10.1109/ACCESS.2020.3018866.
[27] E. R. Hruschka, R. J. G. B. Campello, A. A. Freitas, and A. C. P. L. F. de Carvalho, "A survey of evolutionary algorithms for clustering," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 39, no. 2, pp. 133–155, Mar. 2009, doi: 10.1109/TSMCC.2008.2007252.
[28] X. Cheng, Z.-K. Feng, and W.-J. Niu, "Forecasting monthly runoff time series by single-layer feedforward artificial neural network and grey wolf optimizer," *IEEE Access*, vol. 8, pp. 157346–157355, 2020, doi: 10.1109/ACCESS.2020.3019574.
[29] J. Pan, B. Jing, X. Jiao, and S. Wang, "Analysis and application of grey wolf optimizer-long short-term memory," *IEEE Access*, vol. 8, pp. 121460–121468, 2020, doi: 10.1109/ACCESS.2020.3006499.
[30] G.-J. Wu, W.-Z. Chen, J.-Q. Yuan, D.-S. Yang, and H.-B. Bian, "Formation mechanisms of water inrush and mud burst in a migmatite tunnel: A case study in China," *J. Mountain Sci.*, vol. 14, no. 1, pp. 188–195, Jan. 2017, doi: 10.1007/s11629-016-4070-8.
[31] X. Zhang, S. N. Mohanty, A. K. Parida, S. K. Pani, B. Dong, and X. Cheng, "Annual and non-monsoon rainfall prediction modelling using SVR-MLP: An empirical study from odisha," *IEEE Access*, vol. 8, pp. 30223–30233, Feb. 2020, doi: 10.1109/ACCESS.2020.2972435.
[32] NM. Ramli, MA. Hussain, BM. Jan, and B. Abdullah, "Protection against water or mud inrush in tunnels by grouting: A review," *Neurocomputing*, vol. 131, no. 1, pp. 59–76, May 2014, doi: 10.1016/j.neucom.2013.10.039.
[33] P. Du, J. Wang, Z. Guo, and W. Yang, "Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting," *Energy Convers. Manage.*, vol. 150, pp. 90–107, Oct. 2017, doi: 10.1016/j.enconman.2017.07.065.

DUNWEN LIU received the Ph.D. degree in geotechnical engineering from Central South University, Changsha, China, in 2001. From 2003 to 2004, he was a Visiting Scholar with the KTH Royal Institute of Technology, Sweden. He is currently a Professor with the School of Resources and Safety Engineering, Central South University. His current research interests include the theory and technology of environment friendly rock breaking, prevention and control of tunnel, and underground engineering disasters. He has authored more than 80 articles published in academic journals. He is the Expert of Hunan Province Safety Production Expert Committee, China. He awards the Hunan Province New Century 121 Talent Project, China.

QIAN XU received the bachelor's degree in mining engineering from the Wuhan University of Science and Technology, Wuhan, China, in 2018. He is currently pursuing the master's degree with the School of Resources and Safety Engineering, Central South University, China. His current research interest includes civil safety.

YU TANG received the Ph.D. degree in civil engineering from Central South University, Changsha, China, in 2019. He is currently a Postdoctoral Researcher with the School of Resources and Safety Engineering, Central South University. He has authored 18 articles published in academic journals. He has presided more than one Scientific and Technological Innovation Project of Central South University and one Postdoctoral Research Project in China. His current research interests include the method and technology of rock fracturing by high-pressure gas expansion, the rock breaking method using thermal shock, and gas expansion coupling effect.

YINGHUA JIAN received the bachelor's degree in mining engineering and the master's degree in engineering mechanics from Central South University, Changsha, China, in 2013 and 2016, respectively, where he is currently pursuing the Ph.D. degree with the School of Resources and Safety Engineering. His current research interests include durability and anticorrosion of concrete.