Landslide Displacement Prediction Based on Variational Mode Decomposition and MIC-GWO-LSTM Model

Taorui Zeng (zengtaorui@163.com)
China University of Geosciences https://orcid.org/0000-0002-1241-3238
Hongwei Jiang
China University of Geosciences
Qingli Liu
China Institute of Geological Environmental Monitoring: China Institute of Geo-environment Monitoring
Kunlong Yin
China University of Geosciences https://orcid.org/0000-0001-8082-1500

Research Article

Keywords: Landslide displacement prediction, Variational mode decomposition, Maximum information coefficient, Grey wolf optimizer, Long short-term memory neural networks, Three Gorges Reservoir area

Posted Date: May 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-507356/v1

Landslide displacement prediction based on Variational mode decomposition and MIC-GWO-LSTM model

Zeng Taorui[1], Jiang Hongwei[2], Liu Qingli[3], Yin Kunlong[2*]

1. Institute of Geological Survey, China University of Geosciences, Wuhan 430074, China; zengtaorui@163.com

2. Faculty of Engineering, China University of Geosciences, Wuhan, China; jianghongwei@cug.edu.cn

3. Geo-environmental monitoring office, Wanzhou District, Chongqing, China; 1291592798@qq.com

*. Correspondence: yinkl@cug.edu.cn

Abstract: Landslide displacement prediction is essential to establish the early warning system (EWS). According to the dynamic characteristics of landslide evolution and the shortcomings of the traditional static prediction model, a dynamic prediction model of landslide displacement based on long short-term memory (LSTM) neural networks was proposed. Meanwhile, the Variational modal decomposition (VMD) theory was used to decompose the cumulative displacement and triggering factors, which not only give clear physical meaning to each displacement subsequence, but also closely connect the rock and soil conditions with the influence of external factors. Besides, the maximum information coefficient (MIC) was used to sort the redundant features. The LSTM is a dynamic model that can remember historical information and apply it to the current output. The hyperparameters of the LSTM model was optimized by the Grey wolf optimizer (GWO), and the dynamic one-step prediction was carried out for each displacement. All the predicted values were superimposed to complete the displacement prediction based on the time series model. The Tangjiao landslide in the Three Gorges Reservoir area (TGRA), China, was taken as a case study. The displacement data of monitoring sites GPS06 had step-like characteristics. Measured data from March 2007 to December 2016 were selected for analysis. The results indicate that the root mean square error (RMSE) of the test set and validation set are 23.240 mm and 64.714 mm, respectively, and the coefficient of determination (R^2) are 0.997 and 0.971, respectively. This model provides a new idea and exploration for the displacement prediction of step-like characteristics landslide in the Three Gorges Reservoir area.

Key words: Landslide displacement prediction; Variational mode decomposition; Maximum information coefficient; Grey wolf optimizer; Long short-term memory neural networks; Three Gorges Reservoir area

1. Introduction

Landslide, as a global geological disaster, caused at least 4300 deaths and 19.5 billion US dollars of losses every year (Froude and Petley, 2018). In recent years, China has suffered human casualties and economic losses due to landslides. For example, on July 23, 2019, a landslide occurred in Shuicheng County, Liupanshui City, Guizhou Province, causing 43 deaths, 11 injurers, and a direct economic loss of 190 million yuan (2.9 million dollars). There were 6181 geological disasters in 2019, 4220 of which were landslides. With climate change and human engineering activities, landslide-caused losses may increase (Gariano and Guzzetti, 2016). The United Nations International Strategy for Disaster Reduction (UNISDR, 2006) used the early warning system as a powerful tool to reduce the risk of landslides. One of the important components is prediction (Intrieri et al., 2013).

Landslides are affected by complex and diverse factors such as tectonics (Bennett et al., 2016), climate (Moreiras, 2005), earthquake (Sassa et al., 2010), and human activities (Petley et al., 2007). Since Satio (Saito, 1969) used creep theory to predict the failure time of landslide in 1969. The instability prediction methods of basic kinematics can be divided into three categories (Intrieri et al., 2019): (1) Empirical methods: based on the observation

that the displacement velocity critically increases before rupture and usually extrapolate the time of failure through geometrical arguments; (2) Semi-empirical methods: with analyzing displacement rate and acceleration; (3) Numerical methods: with using time series theory to predict future displacement values based on past observations. In the TGRA, landslides are affected by seasonal reservoir fluctuation and rainfall, which lead to the "step-like" characteristic in the displacement curves of these landslides. Lian (Lian et al., 2015) used artificial neural networks to predict "step-like" displacement, and achieved good prediction results. Cao (Cao et al., 2016) proposed an extreme learning machine model to establish the landslide displacement prediction model in relation to triggering factors. Numerical methods were widely used in displacement prediction of step-like landslide (Huang et al., 2016; Yang et al., 2019).

The decomposition of landslide displacement and the establishment of the prediction model were the key factors affecting the final prediction results (Li et al., 2018). Many researchers considered the landslide displacement as a time series, and decomposed the displacement into subsequences with different characteristics (Huang et al., 2017; Yang et al., 2018). At present, it is widely considered that there are three subsequences: trend displacement, periodic displacement, and random displacement. The traditional decomposition algorithm used the moving average method to extract the trend displacement and the periodic displacement from the original displacement. This method was simple and efficient, but it cannot extract the random displacement (Liao et al., 2020; Zhang et al., 2020). To extract more displacement subsequences, the wavelet decomposition (Zhou et al., 2016; Zhou et al., 2018; Zhu et al., 2018), the empirical mode decomposition (Xu and Niu, 2018), the complementary ensemble empirical mode (Zhang et al., 2020) were applied to decompose displacement. However, these algorithms decompose displacement sequences into more than three sequences with giving, which have no clear physical meaning to the redundant sequences. The Variational Mode Decomposition (VMD) has been widely used in signal noise reduction, power grid prediction, and fault diagnosis since 2014 (Dragomiretskiy and Zosso, 2014). The VMD theory overcomes the phenomenon of modal aliasing, and can specify decompose three intrinsic mode functions (IMFs), which represent the trend, periodic, and random displacement respectively (Guo et al., 2020; Li et al., 2018). Therefore, VMD was adopted to decompose the displacement series.

At present, there are many excellent machine learning models used to predict landslide displacement, such: PSO-ELM (Zhou et al., 2018), GWO-ELM (Liao et al., 2020; Zhang et al., 2020), GWO-BP (Guo et al., 2020), GWO-SVR (Li et al., 2018), PSO-SVR(Zhang et al., 2020), and so on. But there are still two problems in practice application: (1) Different triggering factors had different contributions to the prediction model. The redundant features will affect the prediction accuracy. (2) These models regarded the features and displacements at different time points as independent vectors. They considered only the mapping relationship between input features and output, which ignored the correlation between time series. Therefore, this paper adopted the maximum information coefficient (MIC) algorithm to sort the redundant features, and the long short-term memory (LSTM) neural network to reflect the long-term dependence between the time step and sequence data.

A landslide displacement prediction model based on the VMD theory and the MIC-GWO-LSTM model was proposed in this paper. The cumulative displacement and triggering factors were decomposed into three subsequences by the VMD theory. The MIC algorithm was used to sort the redundant features. A GWO-LSTM model was proposed to update and predict each sequence. The prediction displacement and measured displacement were compared for the model validation.

2. Methodology

2.1 Time series analysis

Landslide displacement is caused by internal geological conditions and external environmental factors (Zhang et al., 2015). The main factors were as follows：(1) Trend factors (e.g., topography, geotechnical properties). The trend displacement is an approximately monotonic increasing function (Zhang et al., 2015); (2) Periodic factors (rainfall, reservoir water level, etc.). The periodic displacement is an approximately periodic function (Guo et al., 2020); (3) Random factors (snow load、non-seasonal rainfall, etc). The random displacement presents a white noise distribution (Li et al., 2018). Therefore, the accumulated displacement can be decomposed as follow:

$$y_t = s_t + c_t + r_t \tag{1}$$

where y_t it the total displacement at time t； s_t is the trend displacement, c_t is the periodic displacement, and r_t is the random displacement.

2.2 Variational mode decomposition

The Variational Mode Decomposition (VMD) theory is an adaptive signal processing method proposed by Dragomiretskiy and Zosso. The theory searched the optimal solution of the variational mode, and obtains modal functions with a certain bandwidth (Dragomiretskiy and Zosso, 2014). It redefines intrinsic mode function (IMF) as an amplitude-modulated signal $u_k(t)$, the expression of which is as follows:

$$u_k(t) = A_k(t)\cos(\phi_k(t)) \tag{2}$$

where $A_k(t)$ is the instantaneous amplitude and $\varphi_k(t)$ is the phase; the instantaneous frequency is $\omega_k(t)= \phi_k'(t)$. The VMD theory established a set of modal functions u_k , that can accurately reconstruct a given signal $f(t)$. Each signal is limited to an estimated center frequency ω_k. The calculation formula of this model is as follows:

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}$$

$$s.t. \sum_{k=1}^{K} u_k = f(t) \tag{3}$$

where $\{ u_k \}$ is the component obtained after decomposition (k=1,2,…,K);$\{\omega_k \}$ is the center frequency of each IMF; $[\delta(t)+j/(\pi t)]u_k(t)$ is the analytical signal of each IMF component; $e^{-j\omega kt}$ is the center frequency of each analytical signal; K is the total number of modes; $f(t)$ is the initial signal. Then, Eq. (3) should be converted to an unconstrained model to obtain the optimal solution:

$$L(\{u_k\},\{\omega_k\},\lambda) = \alpha \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \otimes u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_{k=1}^{K} u_k \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t) \right\rangle \tag{4}$$

where α is the penalty parameter; λ is the Lagrange multiplier; $\otimes$ is the convolution operator. The unconstrained optimization problem can be solved by the alternating direction method of multipliers (ADMM):

i. u_k 、$f(t)$、 $u_i(t)$ 、λ(t) is transformed into $\hat{u}_k(\omega)$ 、 $\hat{f}(\omega)$ 、 $\hat{u}(\omega)$ 、 $\hat{\lambda}(\omega)$ by Fourier transform.

ii. When the number of iterations n=n+1(n=0, 1 , … , n_{max}), update $\hat{u}_k$ and ω_k of different modes k.

$$\hat{u}_k^{n+1}(\omega) \leftarrow \frac{\hat{f}(\omega) - \sum_{i<k} \hat{u}_i^n(\omega) - \sum_{i>k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \tag{5}$$

$$\omega_k^{n+1} \leftarrow \frac{\int_0^\infty \omega \left|\hat{u}_k^{n+1}(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{u}_k^{n+1}(\omega)\right|^2 d\omega} \tag{6}$$

iii. The Lagrange multiplier is updated, and τ is the ascending step.

$$\hat{\lambda}^{n+1}(\omega) \leftarrow \hat{\lambda}^n(\omega) + \tau\left(\hat{f}(\omega) - \sum_{k=1}^{K} \hat{u}_k^{n+1}(\omega)\right) \tag{7}$$

iv. When the result is less than the error limit ε or reaches the maximum number of iterations $n_{\max}$, the solution is stopped:

$$\sum_{k=1}^{K} \left\|\hat{u}_k^{n+1} - \hat{u}_k^n\right\|_2^2 / \left\|\hat{u}_k^n\right\|_2^2 < \varepsilon \tag{8}$$

2.3 Long Short-Term Memory Neural Networks

The long short-term memory (LSTM) neural network is a kind of time cycle neural network, which is suitable for predicting the data with long intervals and delay in time series (Hochreiter and Schmidhuber, 1997). The neurons of the LSTM model ensure data discovery and long-term memory through the forgetting gate, input gate, and output gate. These three gate functions provide a good nonlinear control mechanism for the input and deletion of the control information (Figure 1). The signal forward propagation and the error back propagation are used to train the network for the LSTM model. The hidden vector ($h^{(t)}$) in the LSTM neural network can be obtained as follows:

i. Update forgotten gate output:

$$f^{(t)} = \sigma\left(W_f h^{(t-1)} + U_f x^{(t)} + b_f\right) \tag{9}$$

ii Update input gate output：

$$i^{(t)} = \sigma\left(W_i h^{(t-1)} + U_i x^{(t)} + b_i\right) \tag{10}$$

$$c^{'(t)} = \tanh\left(W_c h^{(t-1)} + U_c x^{(t)} + b_c\right) \tag{11}$$

iii. Update cell state：

$$C^{(t)} = C^{(t-1)} \text{ e } f^{(t)} + i^{(t)} \text{ e } c^{'(t)} \tag{12}$$

iv. Update output gate output：

$$o^{(t)} = \sigma\left(W_o h^{(t-1)} + U_o x^{(t)} + b_o\right) \tag{13}$$

$$h^{(t)} = o^{(t)} \text{ e } \tanh\left(C^{(t)}\right) \tag{14}$$

v. Update current time forecast output:

$$y^{(t)} = \sigma\left(W_y h^{(t)} + b_y\right) \tag{15}$$

where $f^{(t)}$, $i^{(t)}$, $o^{(t)}$, and $c'^{(t)}$ are the values of the forget gate, the input gate, the forget gate, the output gate, and the memory cell in the memory block; $W_{(f,i,c,o)}$ is the weight of $h^{(t-1)}$;$U_{(f,i,c,o)}$ is the weight of input data $x^{(t)}$; b_f、 b_c、 b_i and b_o are their corresponding bias values; σ is the sigmoid function; $C^{(t)}$ is the updated value of the cell state; $y^{(t)}$ is the output value at time t. The back propagation through time (BPTT) algorithm is used to transfer the accumulated error back from the last time, and to calculate the gradient of error corresponding parameters after the forward propagation. Finally, the weights and thresholds are updated by the stochastic gradient descent algorithm.

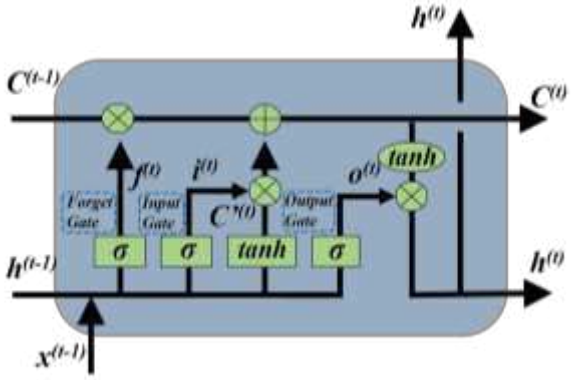

Figure 1. Long short-term memory (LSTM) module contains four interacting layers.

2.4 Grey wolf optimizer

The Grey wolf optimizer (GWO) algorithm was proposed by S. mirjalili et al. (Mirjalili et al., 2014) in 2014, and is a new meta-heuristic algorithm, which has the advantages of strong global search ability and fast operation speed. The GWO algorithm imitated the leadership and hunting mechanism of gray wolves in nature. It used three head wolves (the α wolf, the β wolf, and the δ Wolf) to determine the fitness position. While the ω wolf calculate the distance between them and their prey according to the determined position of their prey. After n iterations, the hunting of the prey can be realized. The main definitions of this algorithm are as follows:

i. The optimal solution is regarded as the α wolf, the second, and third optimal solutions are the β wolf, and the δ wolf respectively. The other candidate solutions are the ω wolf.

ii. The distance $\vec{D}$ between the prey and the grey wolves is determined before preying：

$$\vec{D} = \left|\vec{C} \cdot \vec{X}_P(t) - \vec{X}(t)\right| \tag{16}$$

where $\vec{C}$ is the swing factor, $\vec{C}=2\vec{r}$; $\vec{r}$ is the random vector, $\vec{r} = random[0,1]$; $\vec{X}_P(t)$ is the position vector of the prey of the t th generation of the grey wolves; $\vec{X}(t)$ is the position vector of the t th generation of the grey wolves.

iii. The distance between a gray wolf and its prey is shorten by iterative updating：

$$\vec{X}(t+1) = \vec{X}_P(t) - \vec{A} \cdot \vec{D} \tag{17}$$

$$\vec{A}=2\vec{a} \cdot \vec{r} - \vec{a} \tag{18}$$

where $\vec{A}$ is the convergence factor; $\vec{a}$ is the weight factor, and its initial value set as 2, which decreases to 0 as the number of iterations increases.

iv. In the abstract search space, the ω wolf moves closer to the prey according to the position of the three wolves

(Figure 2). The final position is determined by the random position in the circle defined by the α, β and δ wolves in the search space. The distance between the ω wolves and the three wolves in the t th generation can be obtained according to the following formula:

$$\vec{D}_\alpha = \left|\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}\right|, \vec{D}_\beta = \left|\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}\right|, \vec{D}_\delta = \left|\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}\right| \tag{19}$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{D}_a), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot (\vec{D}_\beta), \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot (\vec{D}_\delta) \tag{20}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{21}$$

The number of hidden layer neurons, iterations, and learning rate have a significant impact on the fitting and prediction ability of the LSTM model. Therefore, The GWO algorithm is introduced to optimize the hyperparameter of the LSTM model in landslide displacement prediction.

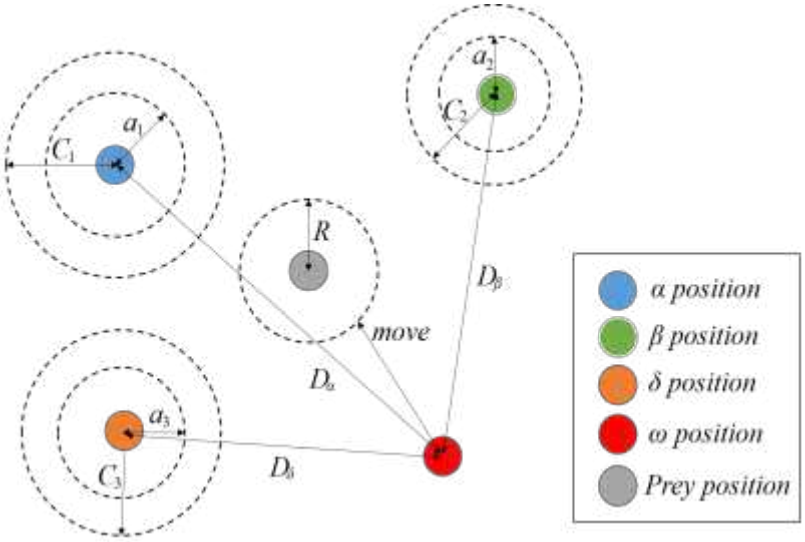

Figure 2. Position updating in GWO

2.5 Maximal Information Coefficient

The Maximum information coefficient (MIC) algorithm is a new variable correlation analysis method proposed by D. n. Reshef et al.(Reshef et al., 2011), which based on mutual information (MI) theory. This algorithm can find the linear functional and nonlinear functional relationship between variables (exponential, periodic, etc). The main steps of the algorithm are as follows:

i. There is a connection between two variables $V_1=\{v_1(i)\}$, i=1,2,…,n and $V_2==\{v_2(i)\}$, i=1,2,…,n. D is the set of ordered pairs $\{v_1(i),v_2(i)\}$, i=1,2,…,n. The grid G_1 ($x_1 \times y_1$) is used to divide V_1 sample points into x_1 and V_2 sample points into y_1. Some cells are allowed to be empty sets.

ii. The characteristic matrix $M(D)_{x_1,y_1}$ can be obtained from the maximum mutual information value $\max I(D|_{G_1})$ as follows:

183 $$M(D)_{x_1,y_1}=\frac{\max I(D|_{G_1})}{\ln\min(x_1,y_1)}=\max\left(\sum_{i=1}^{x_1}\sum_{j=1}^{y_1}\frac{n_{ij}}{N}\ln\frac{n_{ij}}{N}-\sum_{i=1}^{x_1}\frac{\sum_{j=1}^{y_1}n_{ij}}{N}\ln\frac{\sum_{j=1}^{y_1}n_{ij}}{N}-\sum_{i=1}^{y_1}\frac{\sum_{j=1}^{x_1}n_{ij}}{N}\ln\frac{\sum_{j=1}^{x_1}n_{ij}}{N}\right)\Big/\ln\min(x_1,y_1) \quad (22)$$

184 where $D|_{G_1}$ is the probability mass distribution function of all cells in grid G_1; n_{ij} is the sample point in column i
185 of row j in grid G_1; N is the total number of samples.
186 The global optimal grid G_0 ($x_0 \times y_0$) is determined by exhaustive search of characteristic matrix. The maximum
187 information coefficients of variables V_1 and V_2 are as follows:

188 $$MIC(D)=\max_{xy<B(N)}\left\{M(D)_{x,y}\right\}=\max\frac{\max I(D|_G)}{\ln\min(x,y)}=M(D)_{x_0,y_0} \quad (23)$$

189 where $B(N)$ is the maximum grid area for searching. In general, The MIC algorithm is a normalized maximum mutual
190 information, and its value range is [0, 1]. The stronger the correlation between the two variables, the greater the MIC
191 will be. Therefore, to express the importance of different triggering factors in the process of landslide displacement
192 prediction and get a more accurate regression estimation, the MIC algorithm is introduced to calculate the weight
193 between the triggering factors and the displacement.

194 **2.4 Displacement prediction flow**

195 Based on the time series theory, the total displacement can be obtained by adding the three predicted
196 subsequences. The whole prediction process shown in Figure 3, and divided into three parts: the model establishment,
197 the model prediction, and the model verification.
198 To verify the accuracy of prediction, the root mean square error (RMSE) and the coefficient of determination
199 (R^2) was used to evaluate the prediction accuracy. The calculation formula is as follows:

200 $$RMSE=\sqrt{\frac{\sum_{i=1}^{N}\left(x_i-\hat{x}_i\right)^2}{N}} \quad (24)$$

201 $$R^2=1-\frac{\sum_{i=1}^{N}\left(x_i-\hat{x}_i\right)^2}{\sqrt{\sum_{i=1}^{N}\left(x_i-\overline{x}\right)^2}} \quad (25)$$

202 where x_i and $\hat{x}_i$ are the true and predicted values respectively; $\overline{x}$ is the average values of the true values; N is
203 the number of samples.
204

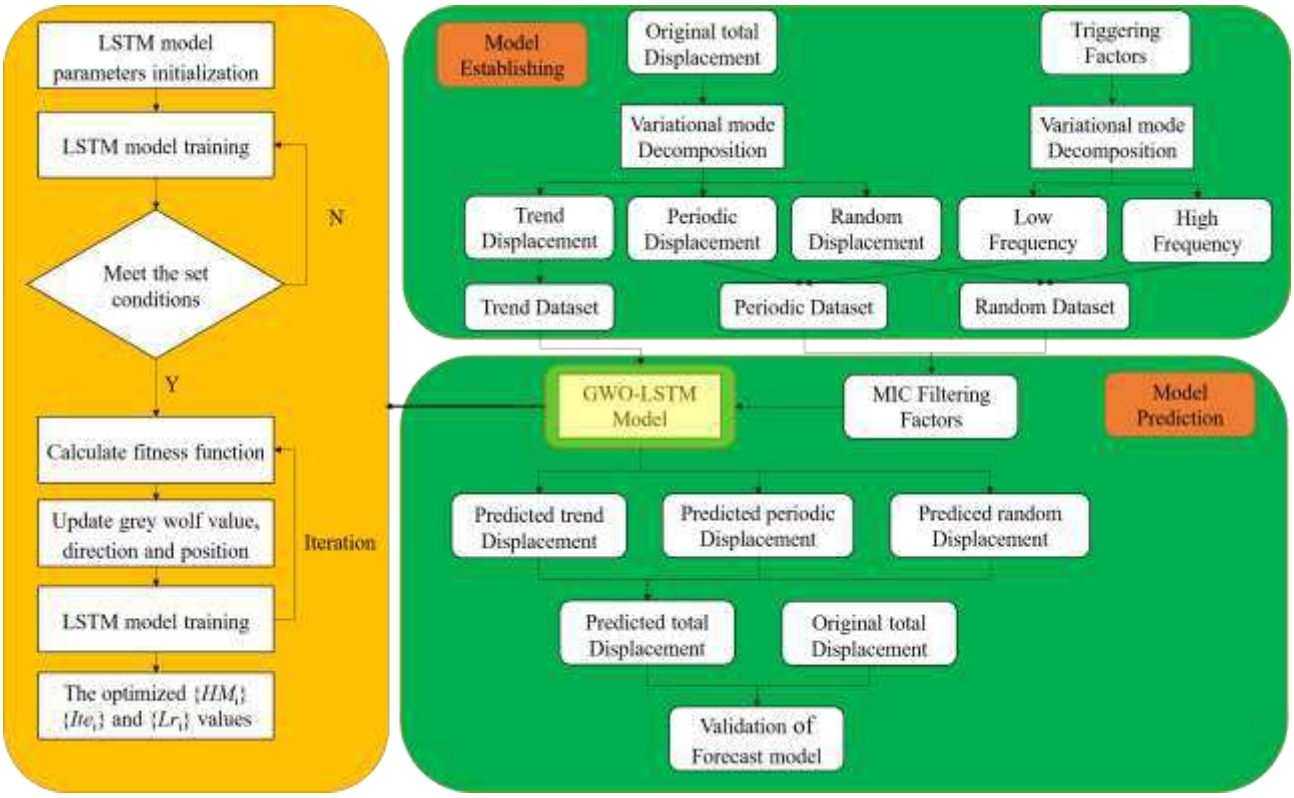

Figure 3. Flowchart of the proposed forecast model.

3. Case study: Tangjiao landslide

3.1 Geological conditions

The Tangjiao landslide is located in the Wanzhou District, Chongqing (108°27′49.26″E, 30°49′32.60″N), which is on the south bank of the Yangtze River (Figure 4). The landslide has an armchair shape in plain view with an elevation between 188 and 266 m. It has an estimated volume of 2900 $\times 10^4 m^3$, and covers an area of 115 $\times$ $10^4 m^2$, with a length of 916 m and a width of 1270 m. The sliding body depth varies from 20 to 30 m. The average depth is 25 m.

The engineering geological profile of the landslide is stepped form with a main sliding direction in 358° (Figure 5). Three-level platforms are developed in the landslide area, which is located at the front elevation of 188 m~248 m, the middle elevation of 248 m~266 m, and the rear elevation of 266 m~330 m respectively. According to the terrace division of the Yangtze River Valley in Wanzhou District, the platforms of the landslide belong to the III, IV, and V terrace of the Yangtze River respectively.

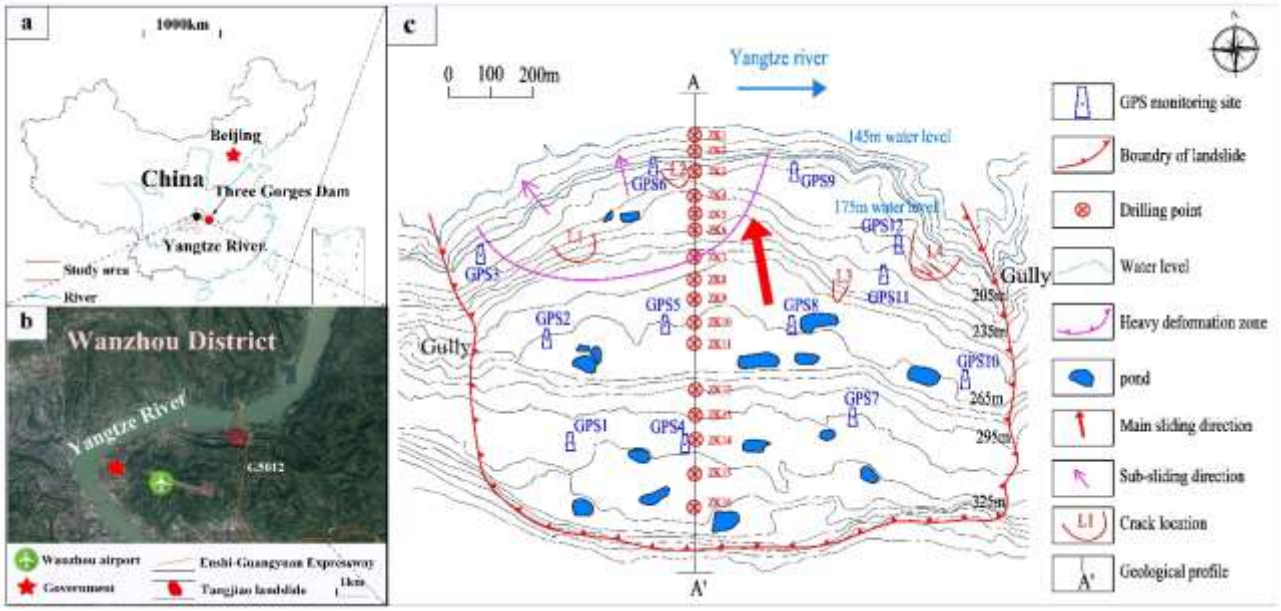

Figure 4. (a) Location of the study area. (b) Location of Tangjiao landslide. (c) Topographical map of the Tangjiao landslide, with the location of the monitoring network (unit: m)

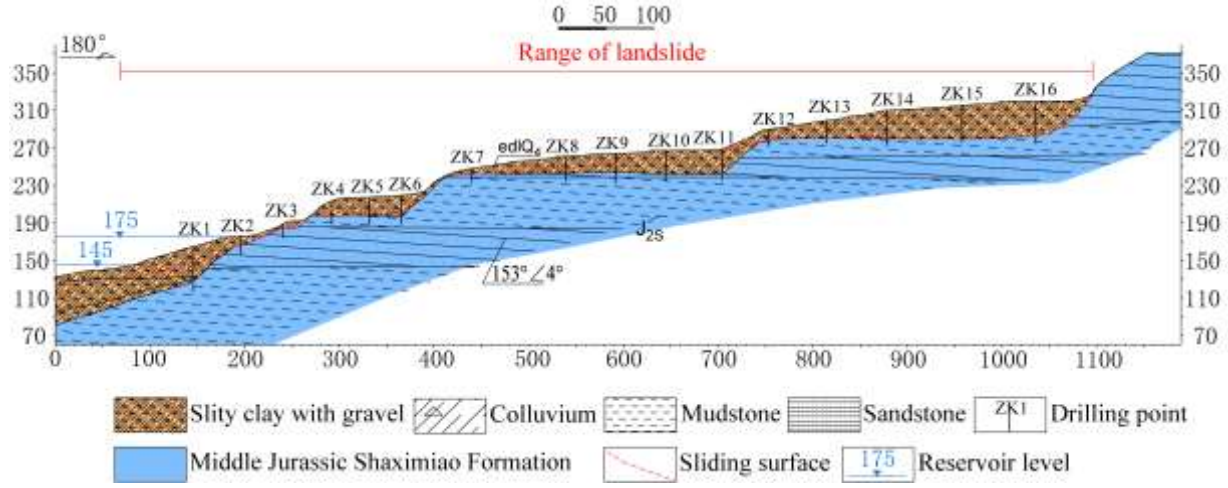

Figure 5. Engineering geological profile A-A' of the Tangjiao landslide

3.2 Deformation characteristic analysis

The initial monitoring network of the Tangjiao landslide was established in March 2007, which includes 12 GPS monitoring points (Figure 4). The Tangjiao landslide began to deform in the 1990s. In May 2007, after the Three Gorges Reservoir Area (TGRA) was impounded to 156m, the deformation of the landslide was intensified due to the fluctuation of the reservoir water level and seasonal rainfall. There were a lot of through cracks and secondary collapses in the middle-front area of the landslide (Figure 6). Based on the surface displacement data collected from March 2007 to December 2016 (Figure 7), the deformation behavior of the landslide can be analyzed:

i. There are spatial differences in the cumulative displacement of landslides. Based on the perspective of GPS cumulative displacement, the deformation of the front area of the landslide is the largest, followed by the middle area, and the smallest deformation is at the rear area, which presents the deformation characteristics of traction landslide. The main cause of this phenomenon is that there is no mechanical connection between the platforms but only a hydraulic connection. The deformation of the middle and rear area of the landslide controlled by rainfall is independent of the front deformation. The GPS06 in the heavy deformation zone has the largest cumulative

displacement, and also has a step-like displacement characteristic.

ii. Take the GPS06 as an example to analyze the relationship between the displacement and the fluctuation of the reservoir water level. From March 2007 to June 2007, the landslide displacement increased significantly during the discharge period after TGRA's first test impoundment. The maximum monthly displacement reached 1899.9mm (May 2007). From June 2007 to September 2007, the reservoir water level fluctuated slightly. However, the cumulative displacement was still increasing, which indicated that the landslide displacement is also affected by rainfall.

iii. From June 2008 to December 2016, the reservoir water level fluctuated periodically from 145m to 175m, and the cumulative displacement curve of GPS06 showed a "step-like" characteristic. From April to June every year, the reservoir water level dropped from 162m to 145m at a rate of about 0.6m/d (Rapid drawdown period of reservoir water level). During this period, the displacements all showed "step-like" characteristics. Especially when the reservoir water level dropped from 175m to 145m for the first time, the cumulative displacement increased to 14852.43mm, which indicated that the monthly displacement increased by 7291.6 mm in June 2009. Since then, the cumulative displacement in this period of each year was more than 200 mm. In general, the first impoundment of the reservoir had the greatest impact on the landslide deformation. The landslide had reached a new stress balance with the long-term effect of the reservoir water level. The displacement growth also tended to be stable.

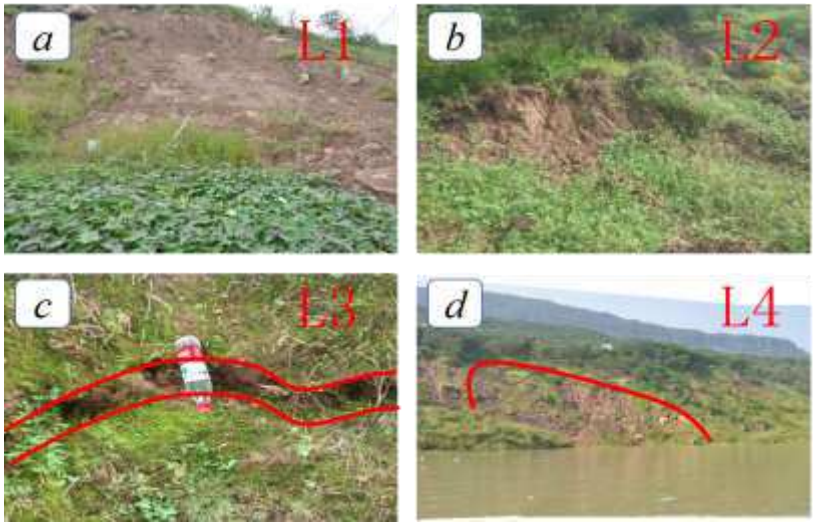

Figure 6. Deformation indications on Tangjiao landslide. (a, b) small scale collapse in heavy deformation zone;(c) transverse tension cracks in the middle-front of landslide;(d) bank collapse in front of landslide.

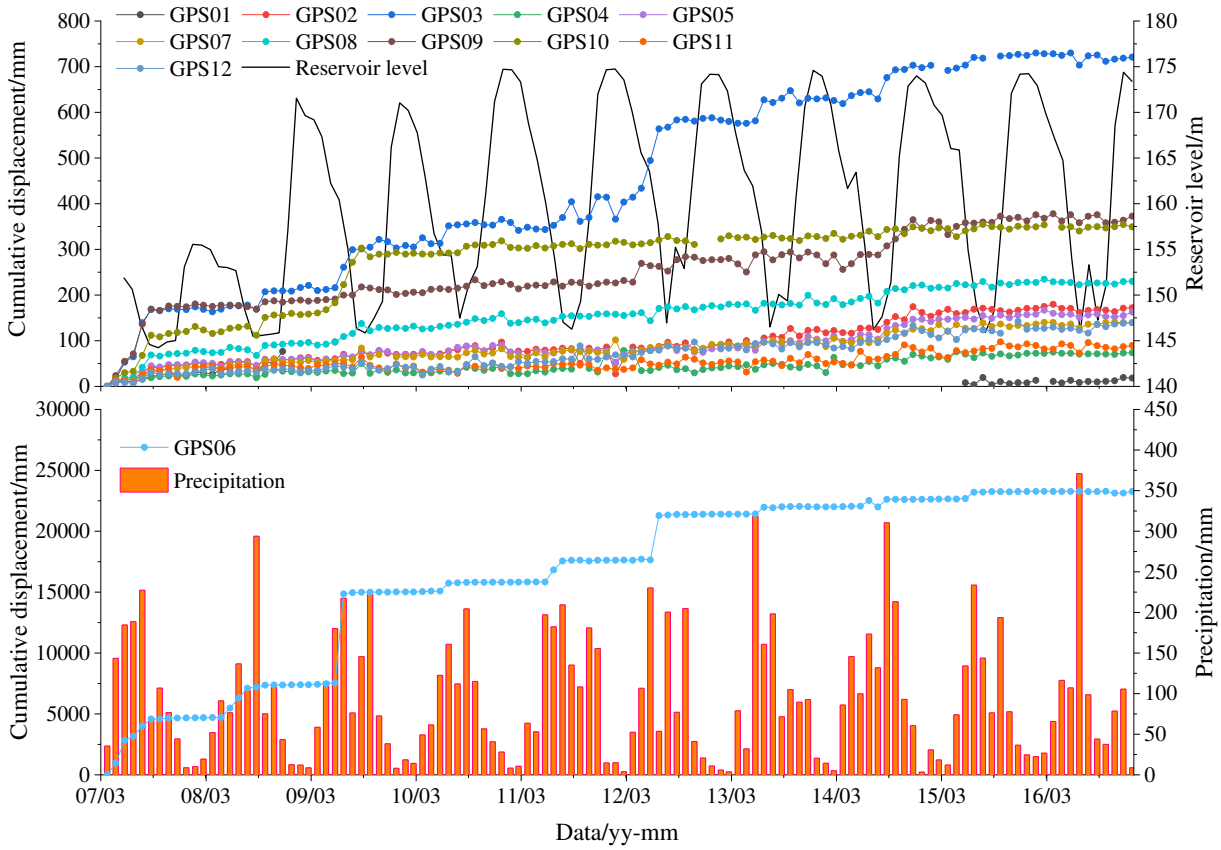

Figure 7. Monitoring data of rainfall, reservoir water level, and cumulative displacement of Tangjiao landslide

4. Prediction Process

4.1 Cumulative displacement decomposition

The monitoring data of GPS06 is the continuous, and the longest one. The surface deformation characteristics near GPS06 are the most obvious. This point was taken as an example for displacement prediction in this paper. Eighty-four measured values from January 2008 to December 2014 were adopted to establish the predicted models. The data from January 2008 to December 2012 were selected as the training samples. The subsequent 12 groups of data (Jan 2014~Dec 2014) were used as the test samples to evaluate the accuracy of the model and determine the optimal prediction model. The data from January 2014 to December 2014 were used as a validation set to verify the application feasibility of the optimal prediction model. According to Eq. (2) to Eq. (8), a VMD program was written in a MATLAB environment to decompose the time series samples. The VMD parameters were set as follows：

i. The model number K was set to 3 according to the analysis result of the landslide displacement-time curve (Eq. (1)). The decomposed subsequences also have actual physical meanings.

ii. The fidelity of the VMD theory is mainly determined by the penalty parameter α and the update rate τ. When the α was set to 22.1, the τ was set to 3.9, the RMSE was the minimum value $2.4656e^{-6}$. After decomposition, three different subsequences corresponded to different displacements were obtained, which indicated the trend displacement, the periodic displacement, and the random displacement (Figure 8).

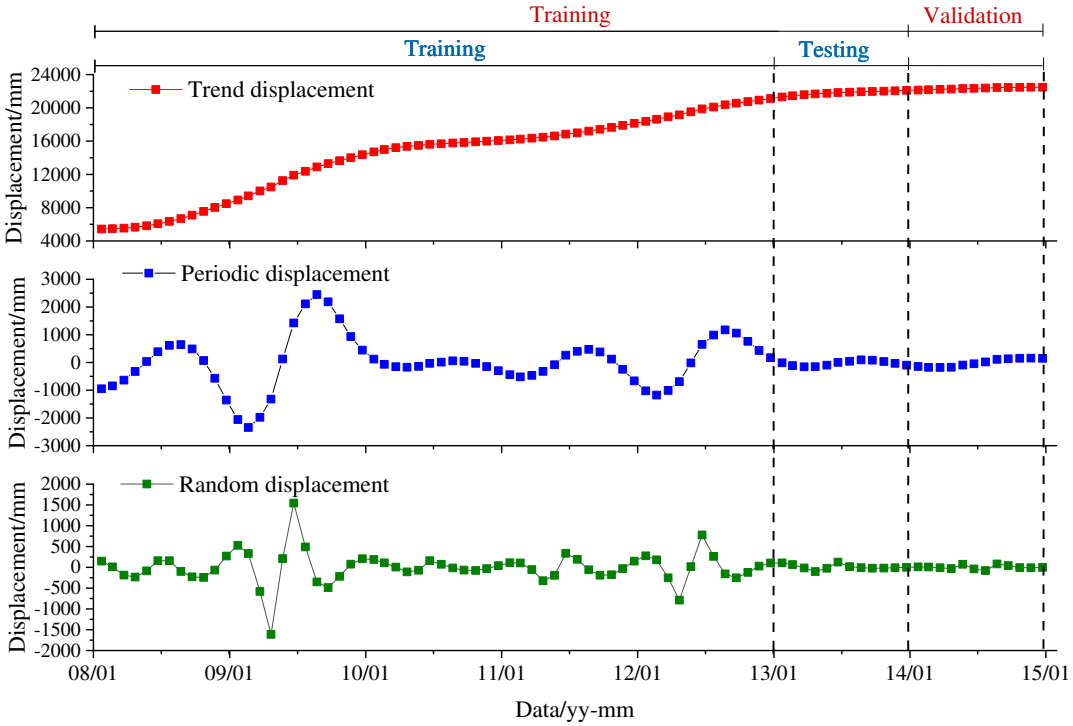

Figure 8. Time series decomposition results of the cumulative displacement

4.2 Decomposition of triggering factors

The selection of triggering factors is crucial to the effectiveness of model training. The rainfall and reservoir water level fluctuations are the dominant factors leading to this displacement feature (Roering et al., 2015). Besides, the deformation state of the landslide also has an important influence on the accurate prediction of the displacement (Jiang et al., 2020; Zhou et al., 2018). On this basis, the following factors were selected to build the initial database (Figure 9).

i. The displacement monitoring data is collected once a month. The displacement of n th month is D_n. Therefore, the candidate rainfall factors are selected as the accumulated rainfall in the n th month (F_1), and the maximum rainfall in the n th month (F_2).

ii. The significant deformation of the Tangjiao landslide occurred during the rapid decline of the reservoir water level. So the selected reservoir water level factors are the average elevation of the reservoir level in the n th month (F_3), the change of the reservoir level during the n th month (F_4), and the change of the reservoir level during the past two months (F_5).

iii. The displacement factors are selected as the change of the displacement that during the last month (F_6), the change of the displacement that during the last two months (F_7), and the change of the displacement that during the last three month (F_8).

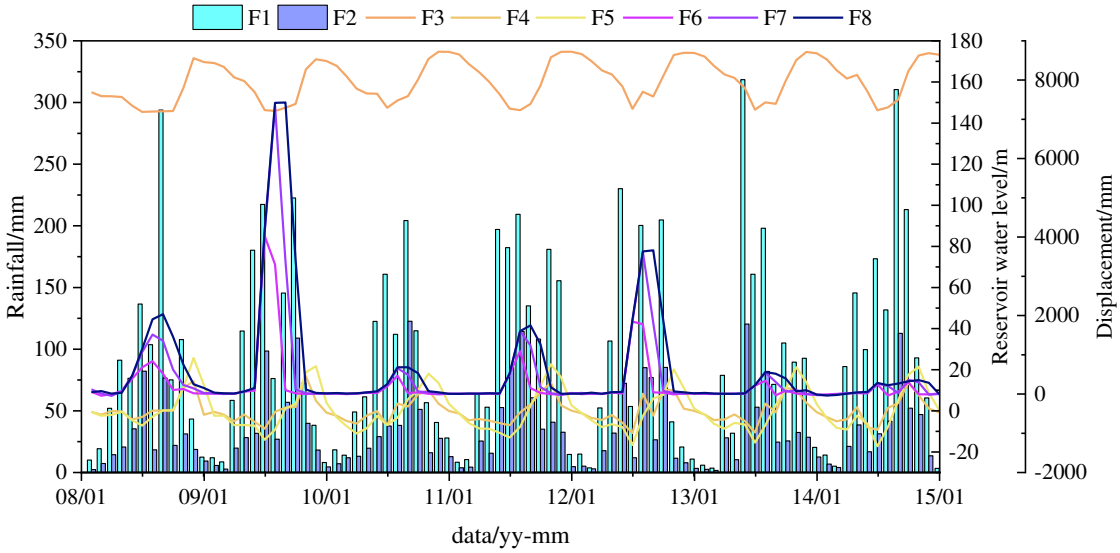

Figure 9. The Candidate Triggering factors of the displacement prediction

According to Li (Li et al., 2018), The VMD theory was used to decompose the sequence of each factor to obtain high-frequency and low-frequency sequences. The total value of triggering factors was no longer used as an input variable. A Matlab program was written to calculate the decomposition effect of the low-frequency factor. The number of modes K was set to 2, the penalty parameter α was set to 0.05, and the update rate τ was set to 0.1. The decomposition result is shown in Figure 10. The curve of the low-frequency sequence (Figure 10) of each factor is relatively smooth, whereas the high-frequency sequence (Figure 11) curve more closely reflects a random displacement (Guo et al., 2020). Therefore, the low-frequency factor obtained by decomposition was used as the triggering factor of the periodic term displacement, and the high-frequency factor was used as the triggering factor of the random term displacement.

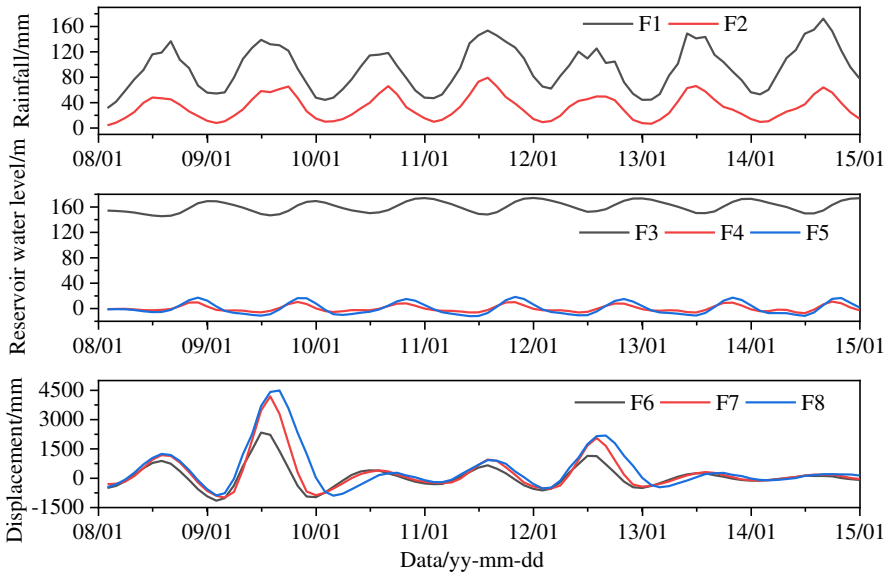

Figure 10. The Candidate low-frequency triggering factors

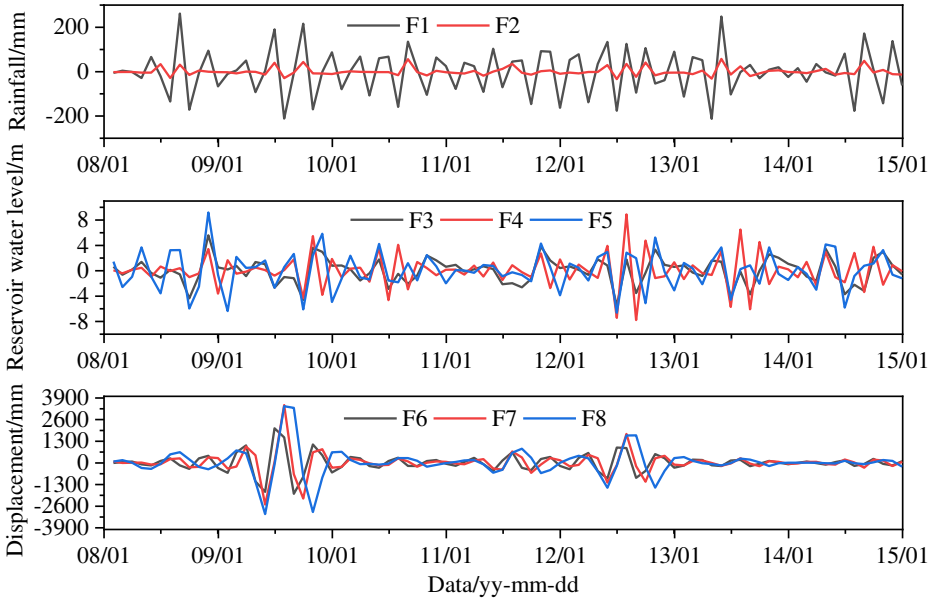

Figure 11. The Candidate high-frequency triggering factors.

4.3 Correlation analysis of triggering factors

Eight triggering factors were selected as the evaluation indicators of landslide displacement. Different triggering factors had different contributions to displacement prediction modeling. The MIC theory is used to calculate the maximum mutual information value of each triggering factor with the landslide displacement (Figure 12). The larger the MIC of triggering factors is, the stronger the correlation between the two variables. In the prediction model, different landslide triggering factors have different prediction abilities. Unimportant triggering factors may bring the noise. The contribution of these factors to the improvement of predictive modeling accuracy might be weaker than the error caused by them (Zhou, 2018). Therefore, the importance of triggering factors should be calculated before prediction modeling. The unimportant factors should be eliminated.

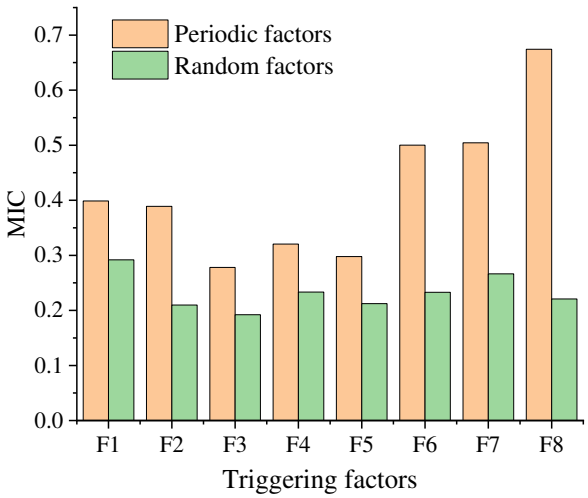

Figure 12. Maximum mutual information value of triggering factors

4.4 Trend displacement prediction

The trend displacement represents the main trend of landslide deformation, which was predicted by the GWO-

LSTM model. Based on the self-written program in Matlab R2019b, the LSTM model was used to learn and train the data in the training samples. When the network error converged to the expected value, the network was used to predict the data in the test samples to test the prediction effect. The parameters and calculation process of the GWO-LSTM model used in this paper are as follows:

i. Assuming that the validation set data is unknown, the optimal hyperparameters of the LSTM algorithm are obtained by using the training set and the testing set. Then the training set and the testing set are combined into the training set of the validation set. To prevent information leakage, the maximum and minimum values of input and output for normalization and inverse normalization are obtained respectively from the training data set. Each triggering factor is normalized to [-1,1] using linear normalization:

$$x^* = 2 \times \frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1 \tag{26}$$

where $x*$ is the normalized value, x is the original value, x_{max} is the maximum value of the samples, and x_{min} is the minimum value of the samples.

ii. An LSTM model with n_h (Number of hidden units) layer, n_f (Number of triggering factors) input nodes, and one output node was constructed. The triggering factors which have a greater influence on displacement were selected as an input variable. The displacement was selected as the output variable. When the trend displacement is predicted, the triggering factors were not considered. The number of triggering factors n_f was set as 1.

iii. The number of grey wolf groups and generations were set as 30, 50 respectively. The root mean square error (RMSE) between the neural network measured values and predicted values was set as the fitness function.

iv. The optimized hyperparameters were used to train the LSTM model, which selected the sigmoid function as the activation function. The PredictAndUpdateState function was used to train the recurrent neural network and update the network state.

v. A univariate GWO-LSTM model was adopted for the trend displacement prediction with the number of hidden units was set as 85, the Max Epochs was set as 166, and the initial learning rate was set as 0.004. The RMSE of the test set was 10.747, the R^2 was 0.998. The RMSE of the verification set was 16.184, the R^2 was 0.980 (Figure 13). The results show that the optimal hyper-parameters obtained by the GWO algorithm have a good prediction effect on the test set and the verification set, which can be used for long-term prediction of trend term displacement.

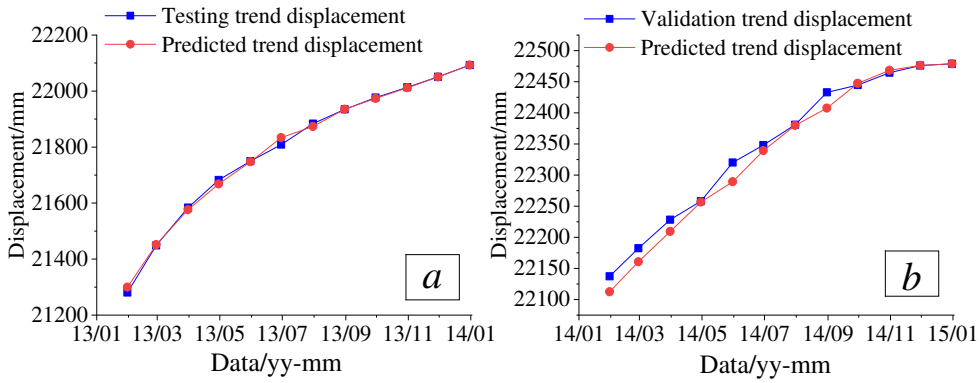

Figure 13. Prediction results of the trend displacement. (*a*) Testing trend displacement. (*b*) Validation trend displacement.

4.5 Periodic displacement prediction

The periodic displacement represents the deformation characteristics caused by periodic action, which is usually predicted by a multivariable model (Miao et al., 2018; Zhou et al., 2018). In these models, monthly rainfall, elevation changes of reservoir water level, and groundwater level were taken as an input variable. However, the

elected triggering factors have different influences on the landslide displacement. The unimportant factors will produce noise and affect the accuracy of prediction modeling. Xu (Xu et al., 2016) used the MIC algorithm to select variables. The unimportant triggering factors can be removed one by one. The prediction accuracy values after removing these factors can be tested by the GWO-LSTM model. A good prediction model should have high generalization ability in both the testing set and verification set. A G_{ave} values was adopted to represent the generalization ability of the model, the formula is as follows:

$$G_{ave} = \frac{R^2_{testing} + R^2_{validation}}{2} \tag{27}$$

where $R^2_{testing}$ is the determination coefficient of the testing set and $R^2_{validation}$ is the determination coefficient of the verification set. As shown in Table 1, when the unimportant factors were removed from the index system, the generalization ability of prediction modeling is improved. The generalization ability of the predicted model achieved to the highest level with the G_{ave} are 0.964, after removing F3, F5, F4, and F2. However, when the factors were removed continuously, the generalization ability decreased. It can be seen from model 9 that without consideration of triggering factors, the model also has good generalization ability, but its RMSE value is significantly higher than that of model 5. Therefore, the four factors were removed from the periodic index system, which improved the accuracy and the generalization ability of the predicted model. The final prediction effect is shown in Figure 14.

Table 1. Prediction modeling accuracy of eliminating unimportant periodic factors

Model	Eliminating unimportant factors	Number of hidden units	Max Epochs	Initial learn rate	G_{ave}	Testing RMSE	Validation RMSE
1	No elimination	2	84	0.238	0.321	41.650	145.423
2	F3	2	163	0.265	0.414	46.359	127.587
3	F3、F5	5	95	0.309	0.470	42.075	123.197
4	F3、F5、F4	132	80	0.037	0.795	17.053	83.153
5	F3、F5、F4、F2	11	193	0.036	0.964	9.337	22.645
6	F3、F5、F4、F2、F1	9	200	0.114	0.953	13.817	35.279
7	F3、F5、F4、F2、F1、F6	8	198	0.131	0.940	15.726	39.917
8	F3、F5、F4、F2、F1、F6、F7	143	103	0.012	0.810	15.442	80.479
9	Without consideration of triggering factors.	21	179	0.117	0.967	14.339	28.675

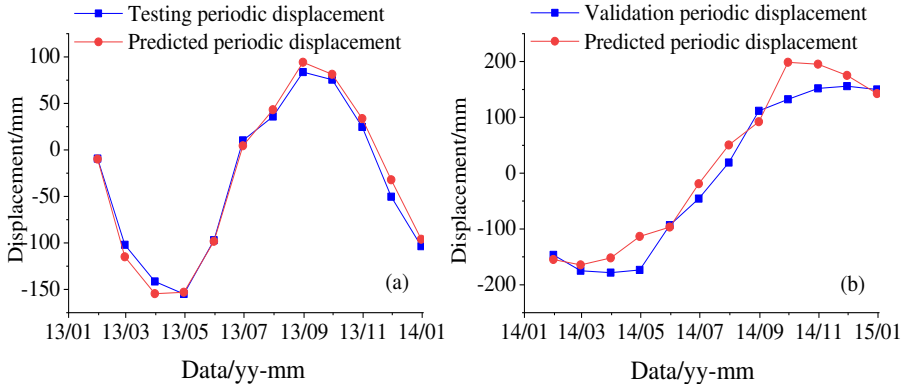

Figure 14. Prediction results of the periodic displacement. (a) Testing periodic displacement. (b) Validation periodic displacement.

4.6 Random displacement prediction

A MIC-GWO-LSTM model was used to predict the random displacement. The prediction results are shown in Table 2. Although the prediction accuracy is improved with the removal of redundant features, the prediction accuracy of various models is not high, which is mainly due to the lack of monitoring data of random factors. Moreover, due to the lack of regularity of random term displacement itself, the univariate LSTM model performs poorly in the prediction process. Model 4 was selected as the random term displacement prediction model. The prediction results are shown in Figure 15. Compared with the original value, the values obtained by model 4 can accurately reflect the characteristics of random displacement fluctuation.

Table 2. Prediction modeling accuracy of eliminating unimportant random factors

Model	Eliminating unimportant factors	Number of hidden units	Max Epochs	Initial learn rate	G_{ave}	Testing RMSE	Validation RMSE
1	No elimination	91	169	0.001	-0.979	29.382	83.917
2	F3	140	22	0.001	0.010	24.993	58.464
3	F3、F2	116	138	0.001	-0.087	12.549	63.547
4	F3、F2、F5	18	77	0.001	0.550	19.495	38.764
5	F3、F2、F5、F8	185	24	0.001	0.459	25.005	41.312
6	F3、F2、F5、F8、F6	195	40	0.001	0.555	27.774	35.726
7	F3、F2、F5、F8、F6、F4	130	38	0.005	-0.114	30.558	61.027
8	F3、F2、F5、F8、F6、F4、F7	77	24	0.023	-0.408	17.540	72.721
9	Without consideration of triggering factors.	25	114	0.187	-3.130	18.498	124.414

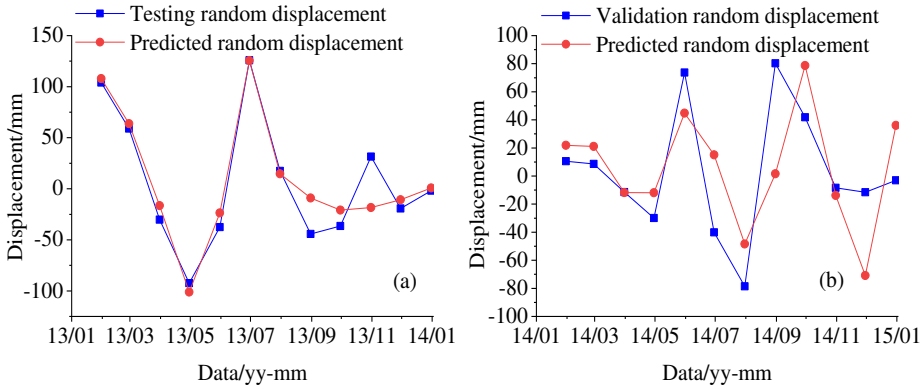

Figure 15. Prediction results of the random displacement. (a) Testing random displacement. (b) Validation random displacement.

4.7 Cumulative displacement prediction and accuracy assessment

According to formula (1), the cumulative displacement of landslide is obtained by adding the predicted values of trend displacement, periodic displacement, and random displacement (Figure 16). The results indicate that the predicted and measured values of landslide cumulative displacement is consistent, the RMSE and the R^2 of the test set are 23.240mm and 0.997 respectively, the RMSE and the R^2 of the verification set are 64.714mm and 0.971 respectively. The prediction results have high accuracy and show the Step-like characteristic of the landslide displacement. Also, the displacement data from May to September 2014 represent fluctuations due to the instruments

and measurement errors. But the predicted result can well represent the increasing trend. It shows that the model can avoid wrong representations through long-term learning.

It is necessary to compare the prediction effect of the MIC-GWO-LSTM model with the existing prediction models. At present, many novel displacement prediction models had been proposed, such as the GWO-BP model (Guo et al., 2020)、the GWO-SVR model (Li et al., 2018), etc. As shown in Table 3, compared with these models, the prediction model proposed in this paper has certain advantages in the perspective of the prediction accuracy and generalization ability.

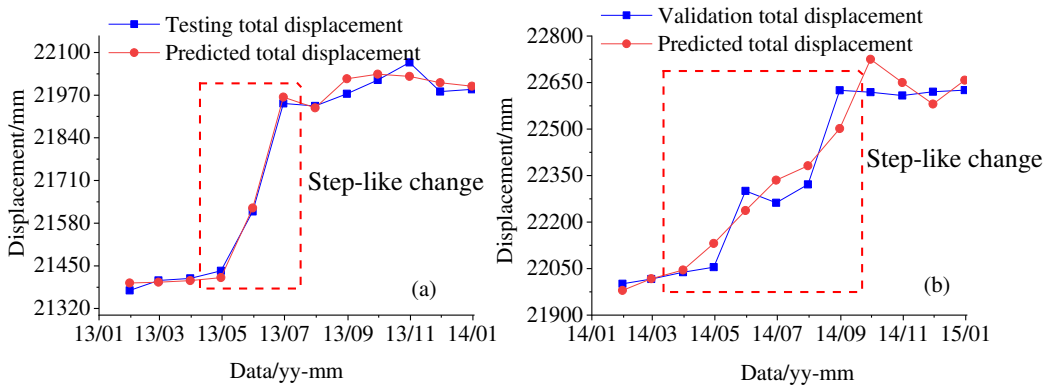

Figure 16. Prediction results of the total displacement. (a) Testing total displacement. (b) Validation total displacement.

a) Performance of various displacement prediction model

Table 3. Performance of various displacement prediction model

Model	Testing RMSE	Testing R^2	Validation RMSE	Validation R^2
MIC-GWO-LSTM	23.240	0.997	64.714	0.971
GWO-BP	48.681	0.969	65.707	0.935
GWO-SVR	68.203	0.940	68.242	0.929

5. Discussion

The high uncertainty in landslide displacement prediction makes it difficult to apply any single model or specific model as the most suitable model for all scenarios (Zhang et al., 2020). Although various methods have been proposed to predict landslide displacement, the accuracy of these methods is still controversial and uncertain (Ma et al., 2020). The stability of GPS monitoring points is a key control factor for the reliability of monitoring data. The evaluation of the stability of the GPS monitoring control point is an important content that cannot be ignored in the processing of displacement observation data (Jiang et al., 2020).

The accurate prediction of landslide displacement needs to know the main control factors (Crosta and Agliardi, 2003). For landslides with step-like characteristics, the rainfall and reservoir water level are the two main factors, which play a vital role in deformation analysis. Rainfall infiltration and reservoir water level fluctuation cause dynamic and hydrostatic pressure, reduce the shear strength of soil, and cause landslide deformation (Hu et al., 2017). The landslide rainfall data usually used regional rainfall, which has a certain error with on-site rainfall. The data of the reservoir water level used the upstream water level of the Three Gorges Reservoir Dam. But the certain reservoir water level has a lag. These uncertainties reduce the correlation between the triggering factors and the displacement sequence. Although the "black box" principle of machine learning can be used to describe the nonlinear relationship. The imprecise input parameters will reduce the reliability of the prediction results when predicting the future trend.

The VMD theory can adaptively decompose the time series according to the scale characteristics of each displacement component. When the model number k is set to 3, the trend displacement, periodic displacement, and random displacement can be directly separated. Each IMF has a clear physical meaning (Guo et al., 2020; Li et al.,

2018). The random factors were further decomposed from rainfall and reservoir water level. Despite this method established the prediction model of the three subsequences of displacement time series. It still ignored the influence of vehicle load, wind load, and other random factors. One should consider means to develop optimal models for predicting the random displacement in the future.

It is not that the more the input factors are, the better the prediction result is. On the contrary, if the input factors are independent of the displacement, the prediction accuracy will be reduced (Deng et al., 2017; Huang and Dun, 2008; Zhang et al., 2020). The MIC-based embedded method integrated the factor selection process with the learner training process, which effectively improved the prediction accuracy and generalization ability of the prediction model. In this paper, eight triggering factors were selected. Finding more factors to improve the generalization ability of the model is the focal point of the future research.

The LSTM model saves and uses the historical information, give full play to its advantage of extracting correlation information (Yang et al., 2019). The hyperparameters have a great influence on the training effect of the LSTM model. The GWO algorithm is used to further optimize the LSTM model to achieve the optimization of the LSTM model. Compared with the grid search (GS), the genetic algorithm (GA), and the particle swarm optimization (PSO) algorithms, the GWO algorithm avoids falling into local optima in high-dimensional, and has a lower convergence speed in the iterative process. The GWO-LSTM model has better generalization performance and improves the prediction accuracy of landslide displacement by automatically determining the hyperparameters of the LSTM model in this paper.

The practical use of machine learning methods in the landslide early warning systems may be very limited because it is not able to provide an indication of whether and when the landslide damage occurred (Intrieri et al., 2019). The practical purpose of these methods lies in the possibility of detecting when the actual displacement exceeds the forecasted values, implying the onset of an unprecedented acceleration (Zhou et al., 2016). In this study, an ensemble model of the VMD theory and the MIC-GWO-LSTM model was established for the landslide displacement prediction, and the prediction results performed well. Meanwhile, many machine learning algorithms have been developed for landslide susceptibility mapping and landslide risk analysis (Huang et al., 2017; Zhou et al., 2017). Thus, the novel method proposed in this paper is recommended for conducting landslide susceptibility mapping, landslide risk analysis, and other fields.

6. Conclusion

Based on the time series method and the VMD algorithm, the landslide displacement of the Tangjiao landslide in the TRGA was decomposed into three subsequences: the trend displacement, the periodic displacement, and the random displacement. A MIC-GWO-LSTM model for predicting the cumulative displacement of the landslide was established. The conclusions were summarized as below:

i.The VMD theory has high decomposition accuracy and high operating efficiency, which can highlight the local fluctuation characteristics of the time series. This theory is suitable for the extraction of landslide displacement. After decomposing the landslide displacement monitoring data with a clear physical meaning of each displacement component. The rock and soil conditions can be linked with the influence of external factors closely.

ii.The MIC algorithm evaluates the relationship between the factors and displacement effectively. The unimportant triggering factors can be removed one by one, which improved the generalization ability of prediction modeling.

iii. The results show that the prediction accuracy of the displacement prediction model based on the MIC-GWO-LSTM model is better than that of other prediction models.

Although this method has made some progress, it was only applied to the data set of landslide displacement triggering factors. The determination of the redundancy threshold between the evaluation factors of the maximum

information coefficient should still be considered according to the actual situation on site. As a deep learning method, the LSTM model needs more training parameters, which may lead to unstable output and long calculation time. Besides, the amount of data is also an important factor affecting the prediction ability of the model. Limited training samples may lead to insufficient training of the dynamic network models, and affect the prediction accuracy of the model.

Author Contributions: Writing—original draft preparation, T.Z.; writing—review and editing, T.Z., H.J., Q.L., and K.Y.; supervision, K.Y. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: We would like to thank the editor and anonymous reviewers providing valuable contributions and constructive comments.

Funding: This research was funded by the National Key R&D Program of China (No.292 2018YFC0809400). The National Natural Science Foundation of China (No. 4187752).

References:

Bennett, G.L., Miller, S.R. and Roering, J.J. et al., 2016. Landslides, threshold slopes, and the survival of relict terrain in the wake of the Mendocino Triple Junction. GEOLOGY, 44(5): 363-366.DOI:10.1130/G37530.1

Cao, Y., Yin, K. and Alexander, D.E. et al., 2016. Using an extreme learning machine to predict the displacement of step-like landslides in relation to controlling factors. Landslides, 13(4): 725-736.DOI:10.1007/s10346-015-0596-z

Crosta, G.B. and Agliardi, F., 2003. Failure forecast for large rock slides by surface displacement measurements. Canadian Geotechnical Journal, 40(1): 176-191.DOI:10.1139/t02-085

Deng, D., Liang, Y. and Wang, L. et al., 2017. Displacement prediction method based on ensemble empirical mode decomposition and support vector machine regression— a case of landslides in Three Gorges Reservoir area. Rock and soil Mechanics, 38(12): 3660-3669.DOI:10.16285/j.rsm.2017.12.034. (in Chinese)

Dragomiretskiy, K. and Zosso, D., 2014. Variational Mode Decomposition. IEEE Transactions on Signal Processing, 62(3): 531-544.DOI:10.1109/TSP.2013.2288675

Froude, M.J. and Petley, D.N., 2018. Global fatal landslide occurrence from 2004 to 2016. Nat. Hazards Earth Syst, 18: 2161–2181.DOI:https://doi.org/10.5194/nhess-18-2161-2018

Gariano, S.L. and Guzzetti, F., 2016. Landslides in a changing climate. EARTH-SCIENCE REVIEWS, 162: 227-252.DOI:10.1016/j.earscirev.2016.08.011

Guo, Z., Chen, L. and Gui, L. et al., 2020. Landslide displacement prediction based on variational mode decomposition and WA-GWO-BP model. Landslides, 17(3): 567-583.DOI:10.1007/s10346-019-01314-4

Hochreiter, S. and Schmidhuber, J., 1997. Long Short-Term Memory. Neural Computation, 8(9): 1735-1780

Hu, X., Tan, F. and Tang, H. et al., 2017. In-situ monitoring platform and preliminary analysis of monitoring data of Majiagou landslide with stabilizing piles. Engineering Geology, 228: 323-336.DOI:10.1016/j.enggeo.2017.09.001

Huang, C. and Dun, J., 2008. A distributed PSO–SVM hybrid system with feature selection and parameter optimization. Applied Soft Computing, 8(4): 1381-1391.DOI:10.1016/j.asoc.2007.10.007

Huang, F., Huang, J. and Jiang, S. et al., 2017. Landslide displacement prediction based on multivariate chaotic model and extreme learning machine. ENGINEERING GEOLOGY, 218: 173-186.DOI:10.1016/j.enggeo.2017.01.016

Huang, F., Yin, K. and Huang, J. et al., 2017. Landslide susceptibility mapping based on self-organizing-map network and extreme learning machine. Engineering Geology

Huang, F., Yin, K. and Zhang, G. et al., 2016. Landslide displacement prediction using discrete wavelet transform and extreme learning machine based on chaos theory. Environmental Earth Sciences, 75(20).DOI:10.1007/s12665-016-6133-0

Intrieri, E., Carla, T. and Gigli, G., 2019. Forecasting the time of failure of landslides at slope-scale: A literature review. EARTH-SCIENCE REVIEWS, 193: 333-349.DOI:10.1016/j.earscirev.2019.03.019

Intrieri, E., Gigli, G. and Casagli, N. et al., 2013. Brief communication "Landslide Early Warning System: toolbox and

general concepts”. Natural Hazards AND EARTH SYSTEM SCIENCES, 13: 85-90.DOI:10.5194/nhess-13-85-2013
Jiang, H., Li, Y. and Zhou, C. et al., 2020. Landslide Displacement Prediction Combining LSTM and SVR Algorithms: A Case Study of Shengjibao Landslide from the Three Gorges Reservoir Area. Applied Sciences, 10(21): 7830.DOI:10.3390/app10217830
Li, L., Wu, Y. and Miao, F. et al., 2018. Displacement prediction of landslides based on variational mode decomposition and GWO-MIC-SVR model. Chinese Journal of Rock Mechanics and Engineering, 37(6): 1395-1406.DOI:10.13722/j.cnki.jrme.2017.1508 (in Chinese)
Lian, C., Zeng, Z. and Yao, W. et al., 2015. Multiple neural networks switched prediction for landslide displacement. Engineering Geology, 186: 91-99.DOI:10.1016/j.enggeo.2014.11.014
Liao, K., Wu, Y. and Miao, F. et al., 2020. Using a kernel extreme learning machine with grey wolf optimization to predict the displacement of step-like landslide. Bulletin of Engineering Geology and the Environment, 79(2): 673-685.DOI:10.1007/s10064-019-01598-9
Ma, J., Liu, X. and Niu, X. et al., 2020. Forecasting of Landslide Displacement Using a Probability-Scheme Combination Ensemble Prediction Technique. International Journal of Environmental Research and Public Health, 17(13): 4788.DOI:10.3390/ijerph17134788
Miao, F., Wu, Y. and Xie, Y. et al., 2018. Prediction of landslide displacement with step-like behavior based on multialgorithm optimization and a support vector regression model. Landslides, 15(3): 475-488.DOI:10.1007/s10346-017-0883-y
Mirjalili, S., Mirjalili, S.M. and Lewis, A., 2014. Grey Wolf Optimizer. Advances in Engineering Software, 69: 46-61.DOI:10.1016/j.advengsoft.2013.12.007
Moreiras, S.M., 2005. Climatic effect of ENSO associated with landslide occurrence in the Central Andes, Mendoza Province, Argentina. Landslides(2): 53–59.DOI:10.1007/s10346-005-0046-4
Petley, D.N., Hearn, G.J. and Hart, A. et al., 2007. Trends in landslide occurrence in Nepal. Nat Hazards, 42: 23–44.DOI:10.1007/s11069-006-9100-3
Reshef, D.N., Reshef, Y.A. and Finucane, H.K. et al., 2011. Detecting Novel Associations in Large Data Sets. SCIENCE, 334(6062): 1518-1524.DOI:10.1126/science.1205438
Saito, M., 1969. Forecasting Time of Slope failure by Tertiary Creep. Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering, 2: 677–683.DOI:10.1007/s11069-006-9100-3
Sassa, K., Nagai, O. and Solidum, R. et al., 2010. An integrated model simulating the initiation and motion of earthquake and rain induced rapid landslides and its application to the 2006 Leyte landslide. Landslides, 7: 219–236.DOI:10.1007/s10346-010-0230-z
UNISDR, 2006. The International Early Warning Programme-The four elements of effective early warning systems-Brochure, Platform for the Promotion of Early Warning. PPEW
Xu, S. and Niu, R., 2018. Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China. Computers & Geosciences, 111: 87-96.DOI:10.1016/j.cageo.2017.10.013
Xu, Z., Xuan, J. and Liu, J. et al., 2016. MICHAC: Defect Prediction via Feature Selection Based on Maximal Information Coefficient with Hierarchical Agglomerative Clustering. IEEE, pp. 370-381.10.1109/SANER.2016.34.
Yang, B., Yin, K. and Du, J., 2018. A model for predicting landslide displacement based on time series and long and short term memory neural network. Chinese Journal of Rock Mechanics and Engineering, 37(10): 2334-2343.DOI:10.13722/j.cnki.jrme.2018.0468 (in Chinese)
Yang, B., Yin, K. and Lacasse, S. et al., 2019. Time series analysis and long short-term memory neural network to predict landslide displacement. Landslides, 16(4): 677-694.DOI:10.1007/s10346-018-01127-x
Zhang, J., Tang, H. and Wen, T. et al., 2020. A Hybrid Landslide Displacement Prediction Method Based on CEEMD and DTW-ACO-SVR — Cases Studied in the Three Gorges Reservoir Area. Sensors, 20(15): 4287.DOI:10.3390/s20154287

Zhang, J., Yin, K. and Wang, J. et al., 2015. Displacement prediction of BaiShuihe Landslide based on time series and Pso-Svr model. Chinese Journal of Rock Mechanics and Engineering, 34(02): 382-391.DOI:10.13 722/j .cnki j rme.2015.02.017

Zhang, L., Chen, X. and Zhang, Y. et al., 2020. Application of GWO-ELM Model to Prediction of Caojiatuo Landslide Displacement in the Three Gorge Reservoir Area. Water, 12(7): 1860.DOI:10.3390/w12071860

Zhang, L., Shi, B. and Zhu, H. et al., 2020. PSO-SVM-based deep displacement prediction of Majiagou landslide considering the deformation hysteresis effect. Landslides.DOI:10.1007/s10346-020-01426-2

Zhou, C., 2018. Landslide identification and prediction with the application of time series InSAR, China University of Geosciences, China,Wuhan, 156 pp.

Zhou, C., Yin, K. and Cao, Y. et al., 2016. Application of time series analysis and PSO-SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir. Engineering Geology. Engineering Geology, 204: 108–120.DOI:http://dx.doi.org/10.1016/j.enggeo.2016.02.009

Zhou, C., Yin, K. and Cao, Y. et al., 2017. Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. Computers and Geosciences

Zhou, C., Yin, K. and Cao, Y. et al., 2018. Displacement prediction of step-like landslide by applying a novel kernel extreme learning machine method. Landslides, 15(11): 2211-2225.DOI:10.1007/s10346-018-1022-0

Zhu, X., Xu, Q. and Tang, M. et al., 2018. A hybrid machine learning and computing model for forecasting displacement of multifactor-induced landslides. Neural Computing and Applications, 30(12): 3825-3835.DOI:10.1007/s00521-017-2968-x

Figures

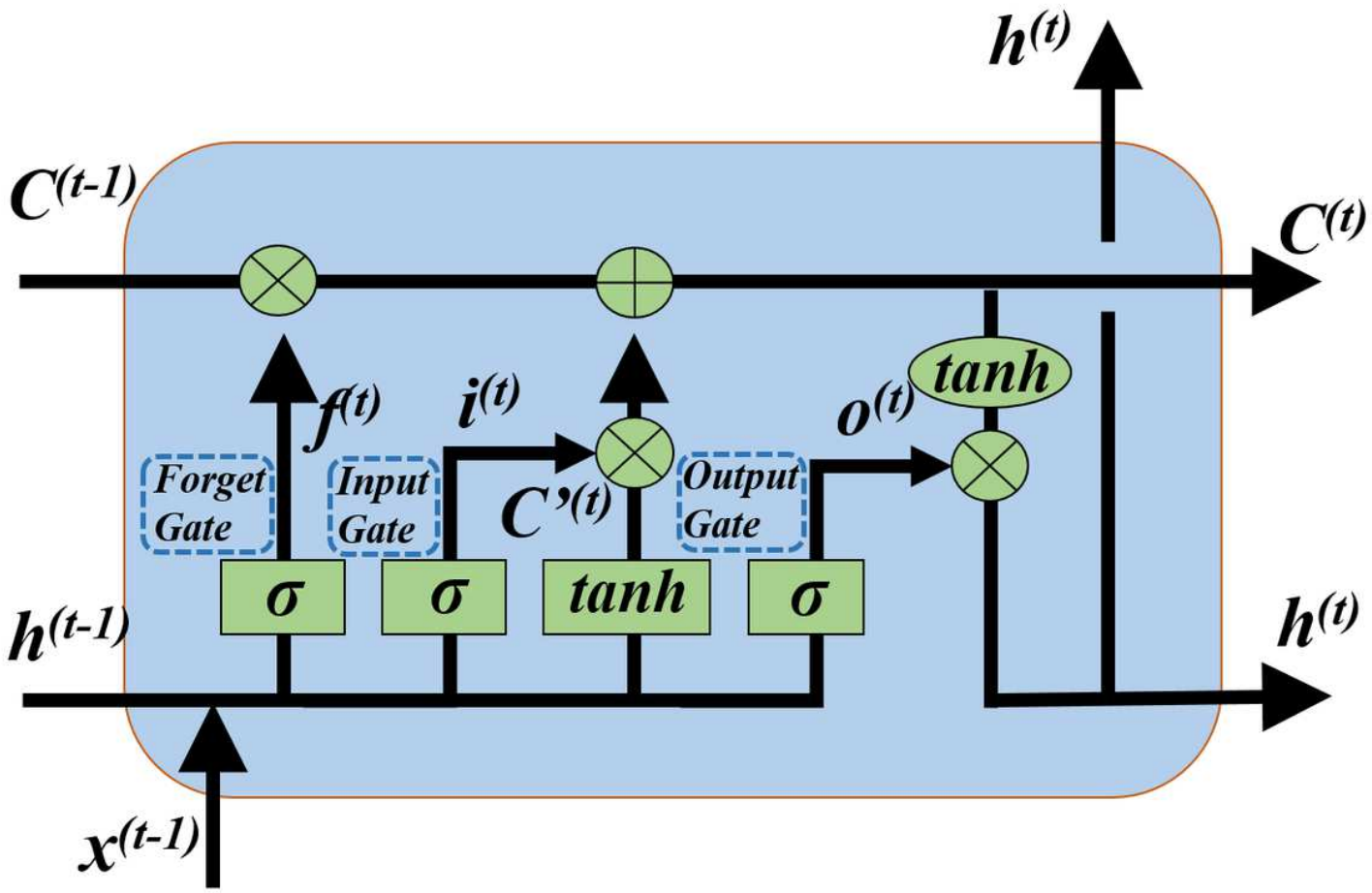

Figure 1

Long short-term memory (LSTM) module contains four interacting layers.

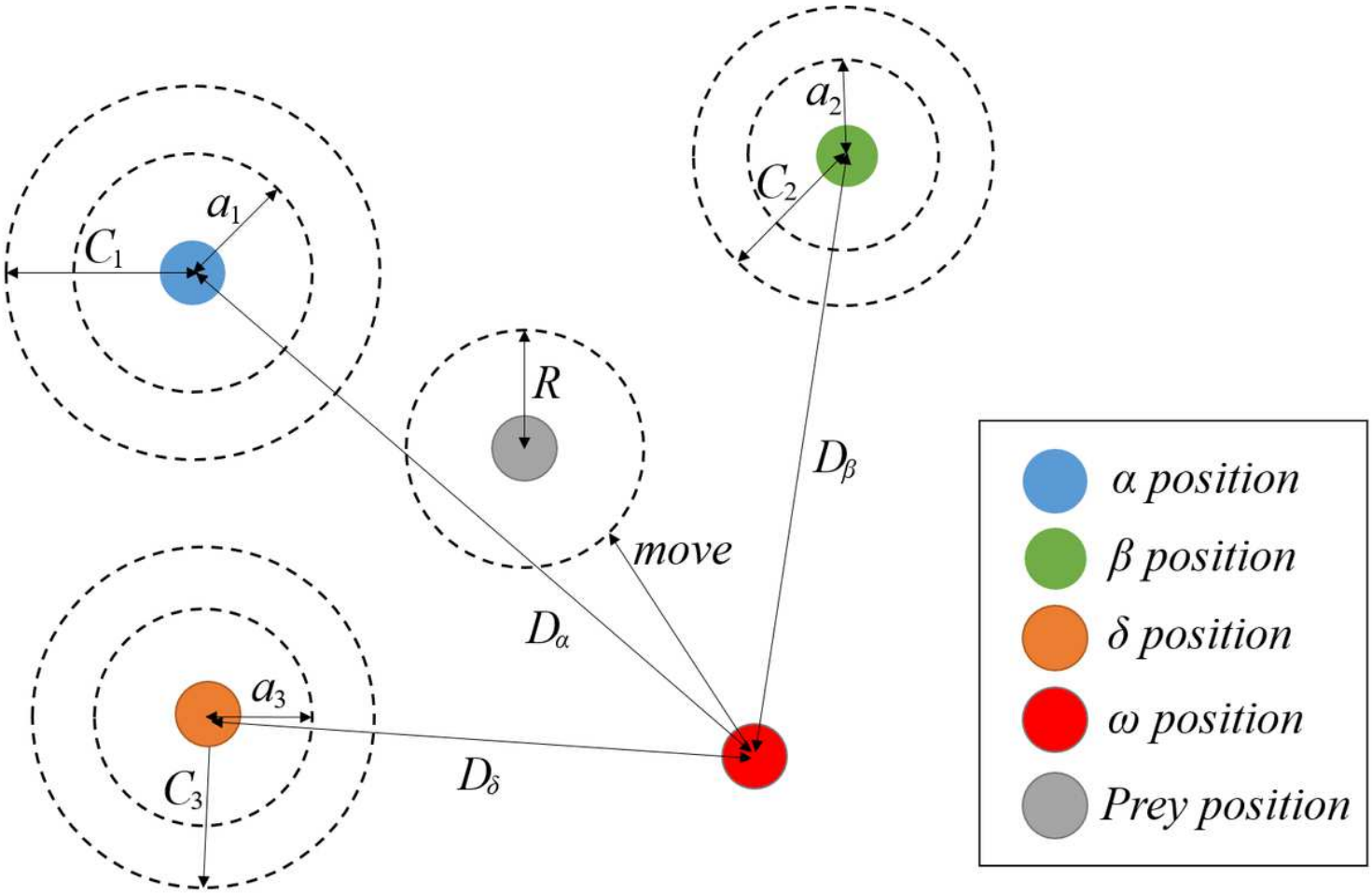

Figure 2

Position updating in GWO

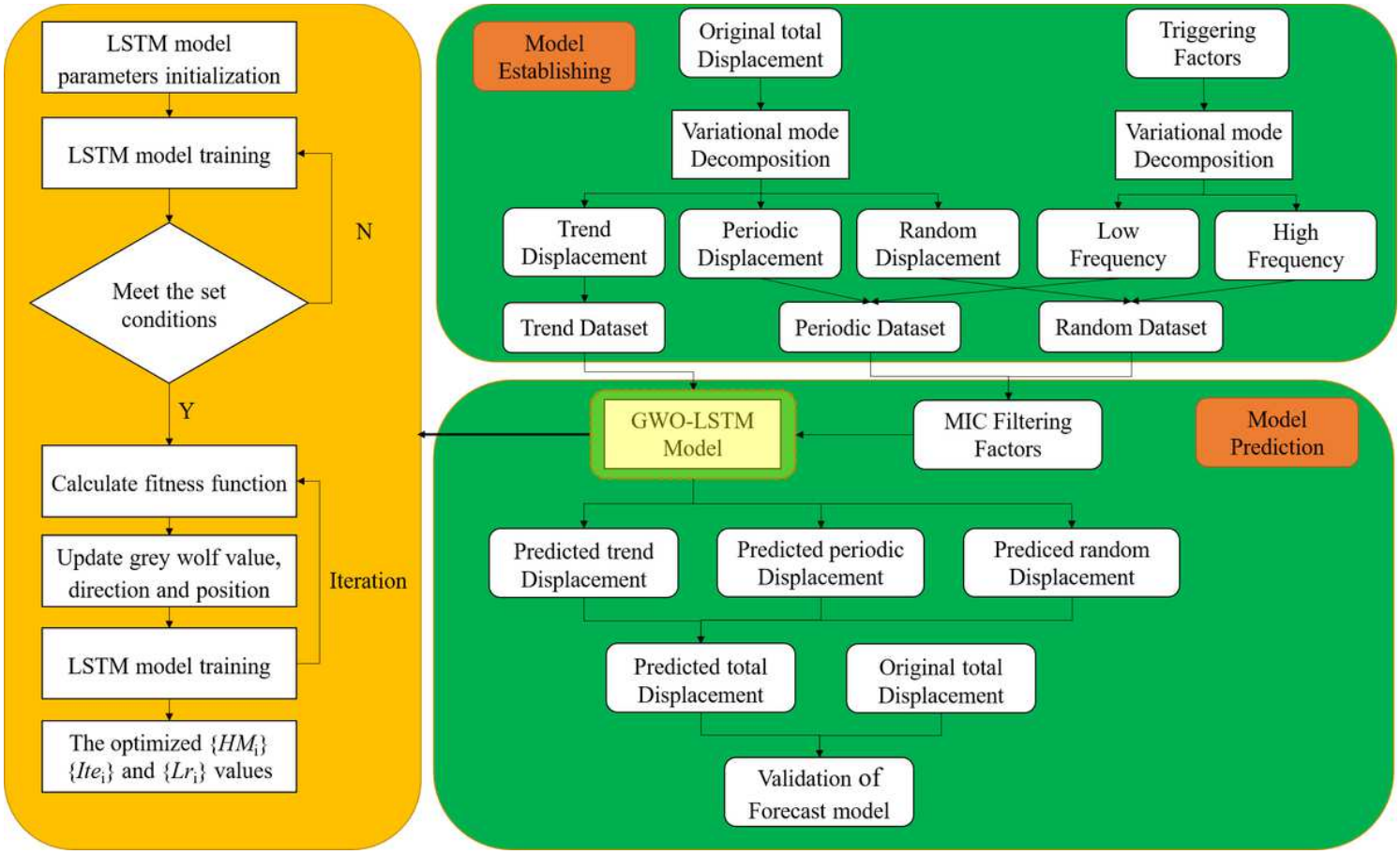

Figure 3

Flowchart of the proposed forecast model.

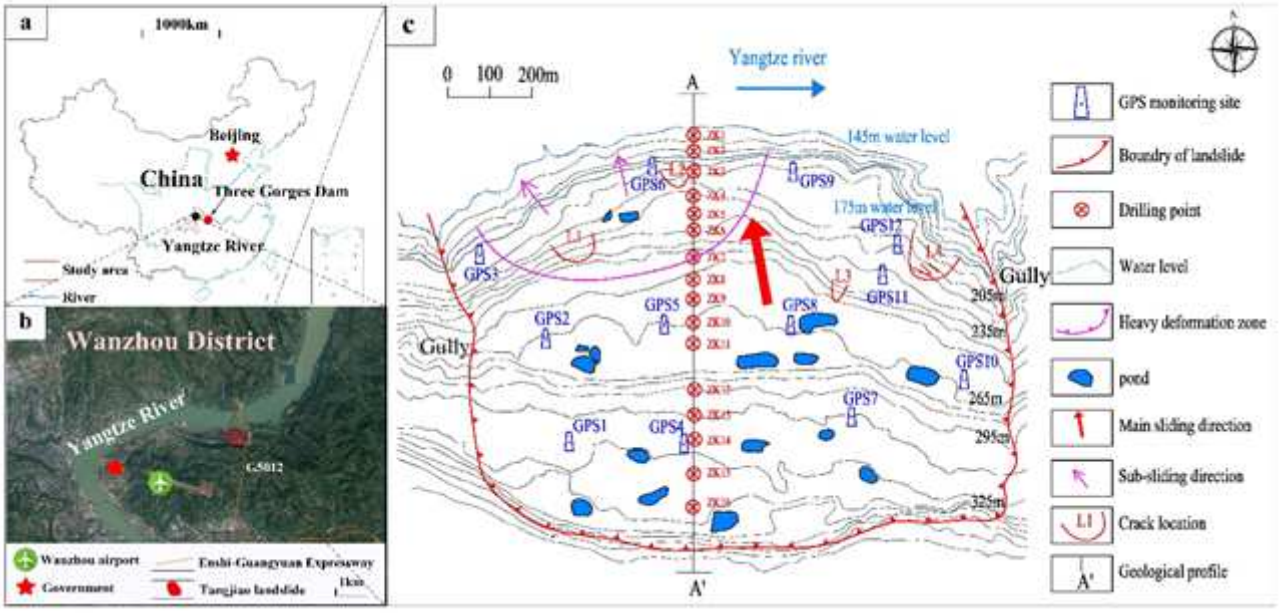

Figure 4

(a) Location of the study area. (b) Location of Tangjiao landslide. (c) Topographical map of the Tangjiao landslide, with the location of the monitoring network (unit: m) Note: The designations employed and the

presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

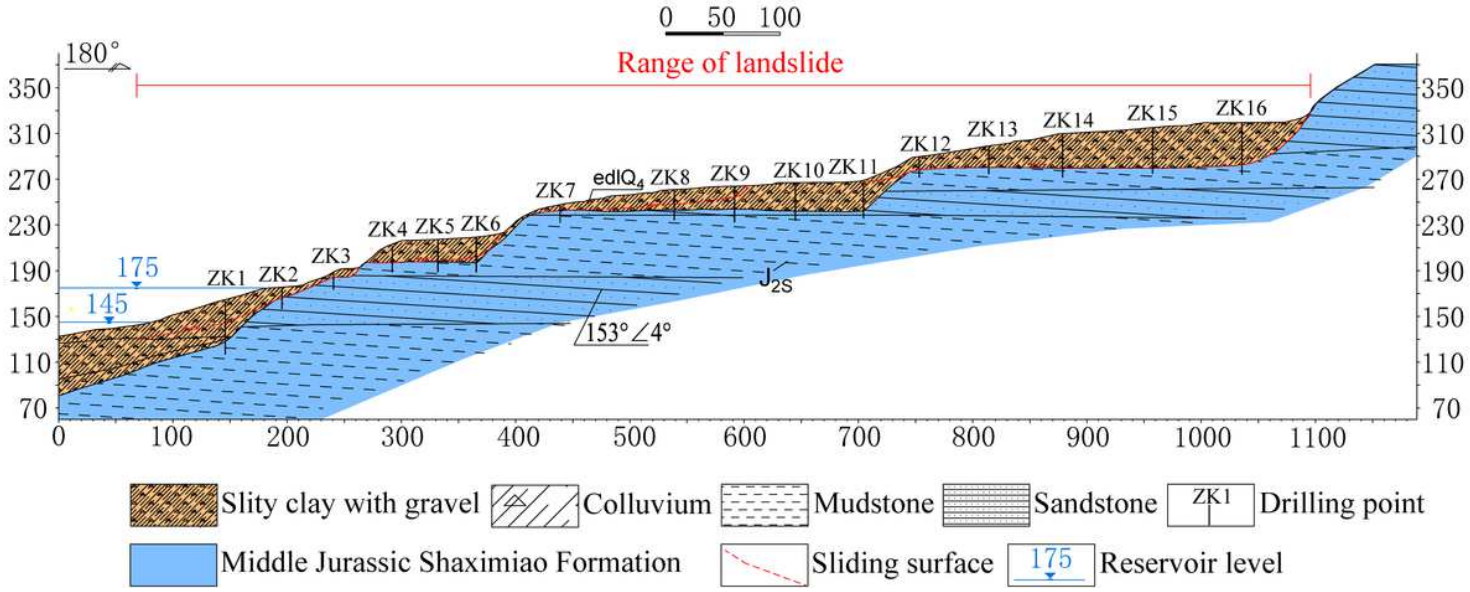

Figure 5

Engineering geological profile A-A’ of the Tangjiao landslide

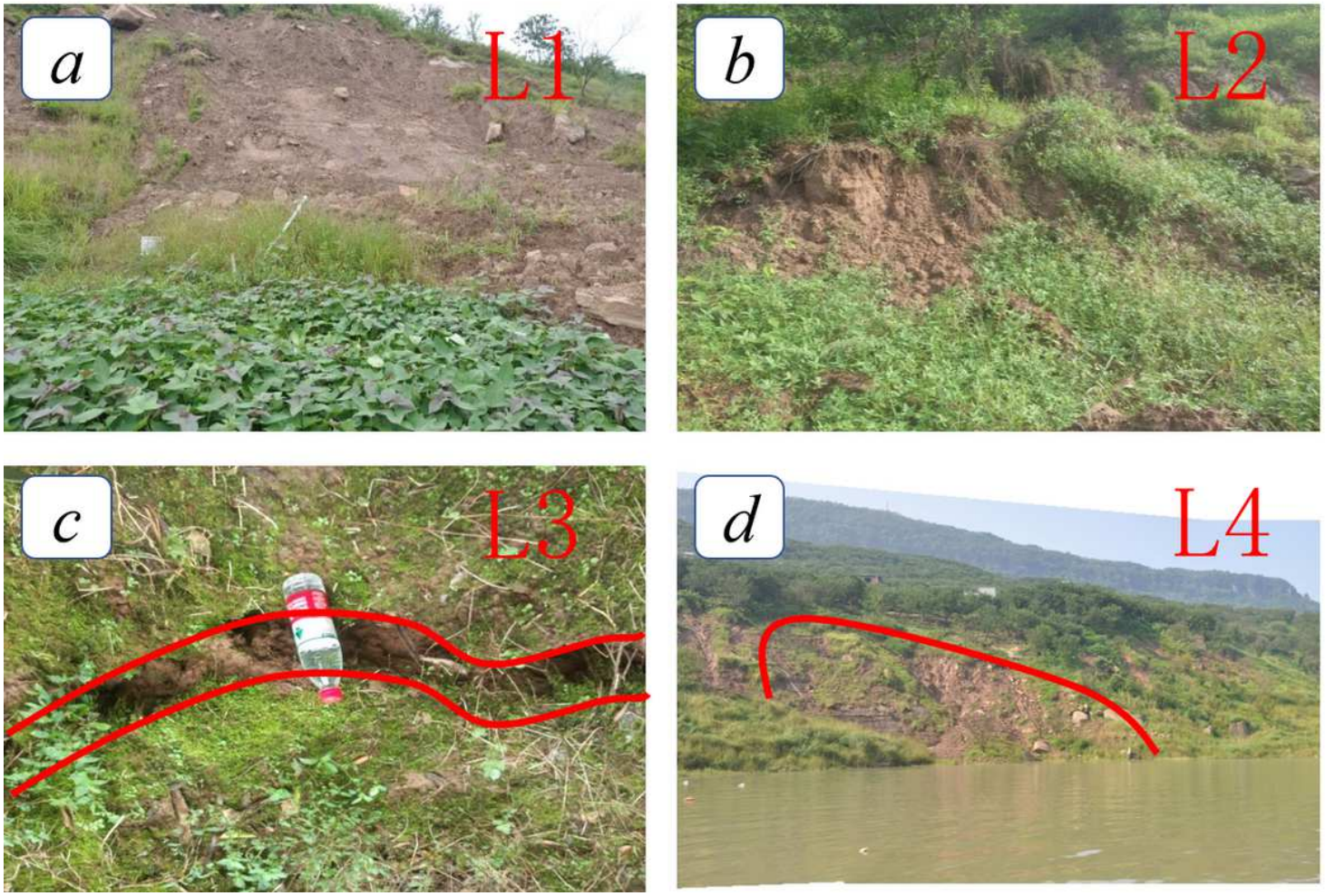

Figure 6

Deformation indications on Tangjiao landslide. (a, b) small scale collapse in heavy deformation zone;(c) transverse tension cracks in the middle-front of landslide;(d) bank collapse in front of landslide.

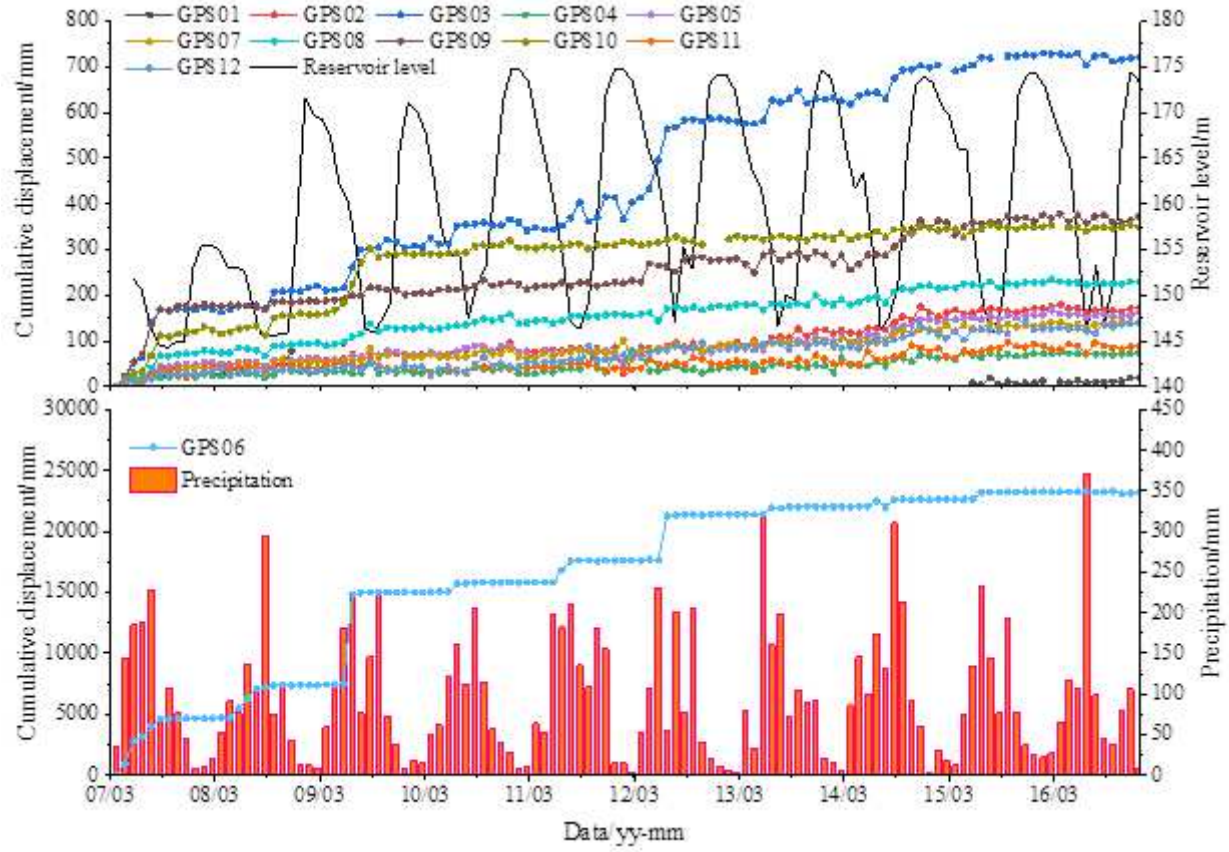

Figure 7

Monitoring data of rainfall, reservoir water level, and cumulative displacement of Tangjiao landslide

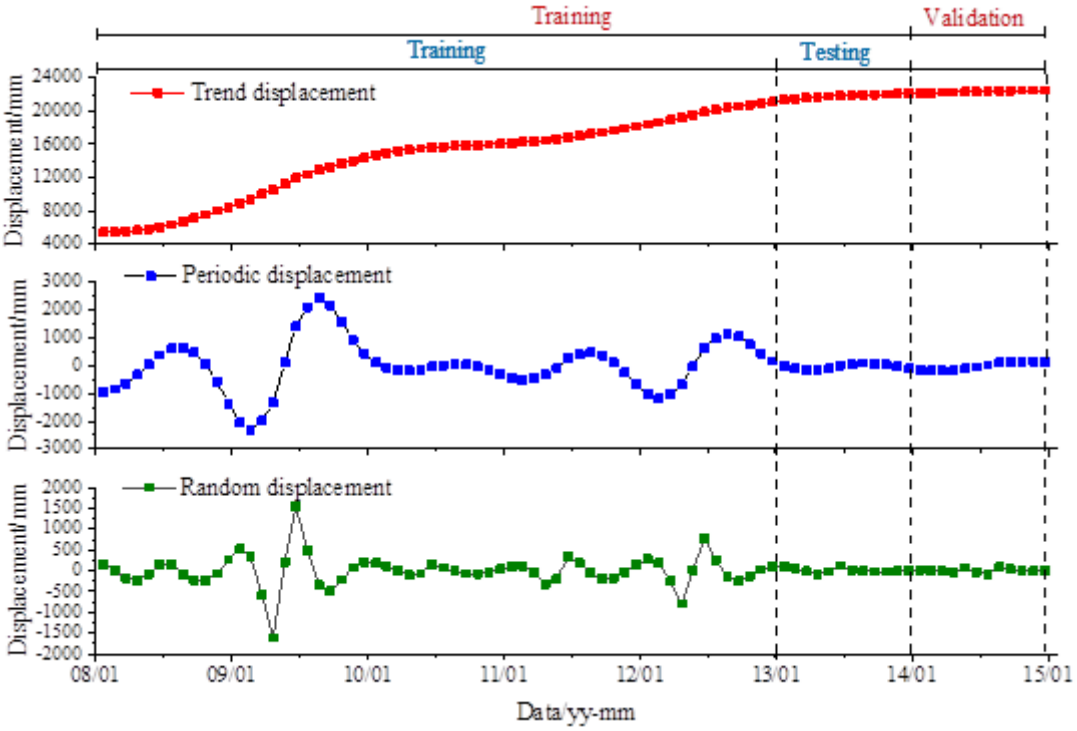

Figure 8

Time series decomposition results of the cumulative displacement

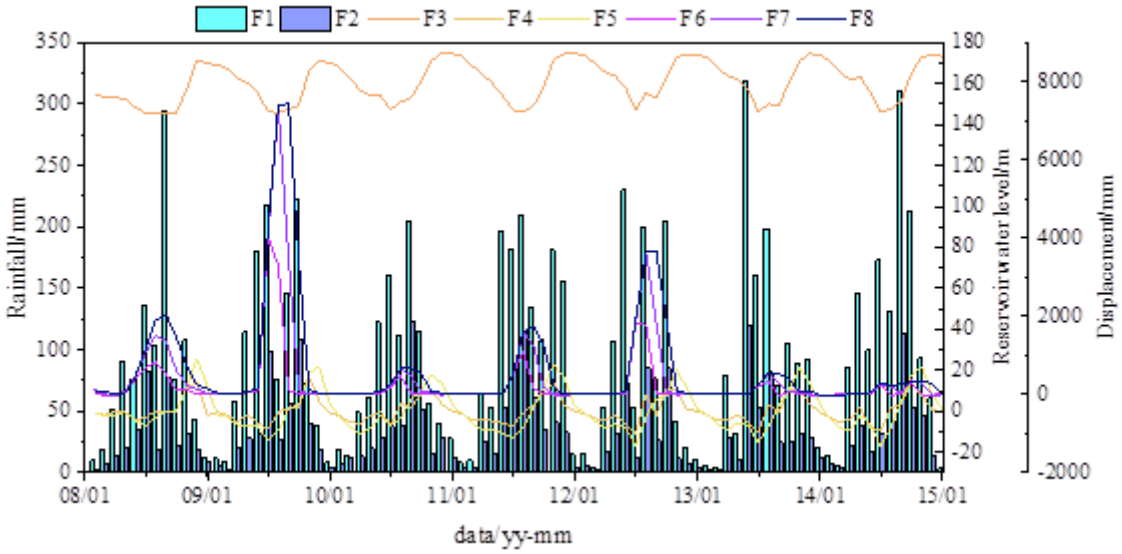

Figure 9

The Candidate Triggering factors of the displacement prediction

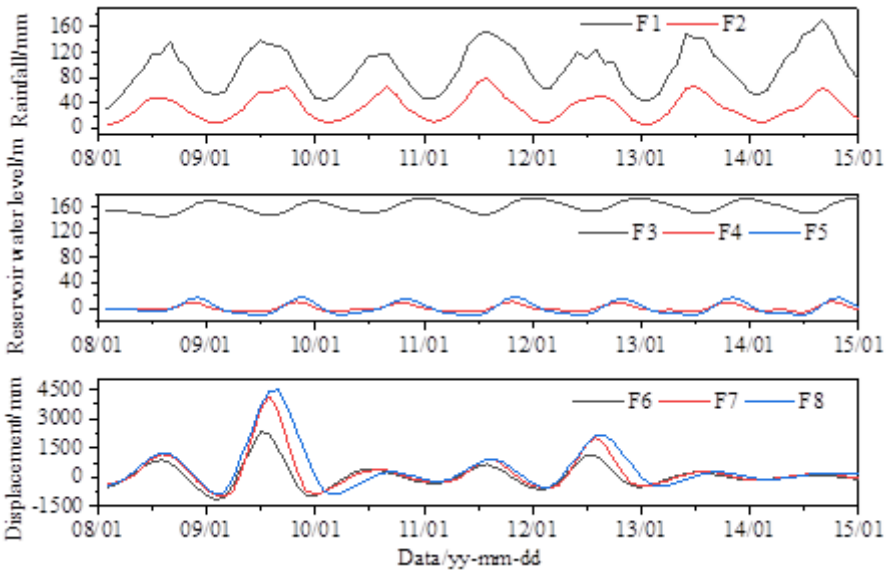

Figure 10

The Candidate low-frequency triggering factors

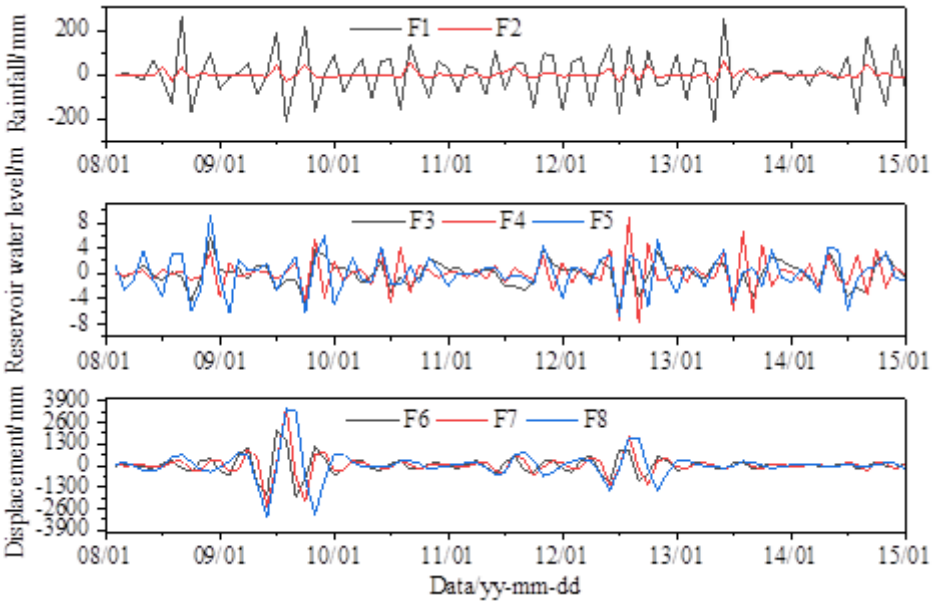

Figure 11

The Candidate high-frequency triggering factors.

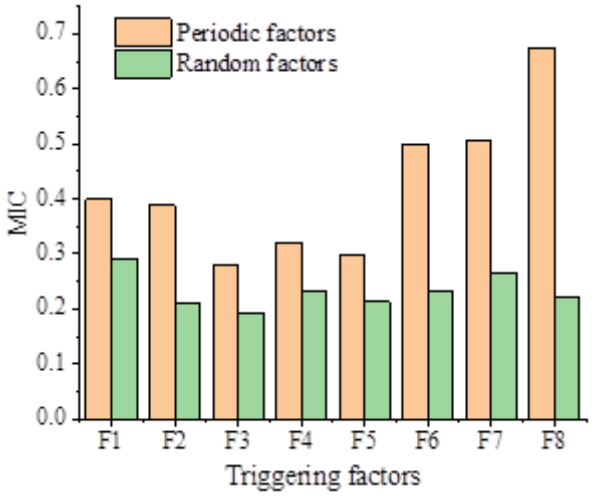

Figure 12

Maximum mutual information value of triggering factors

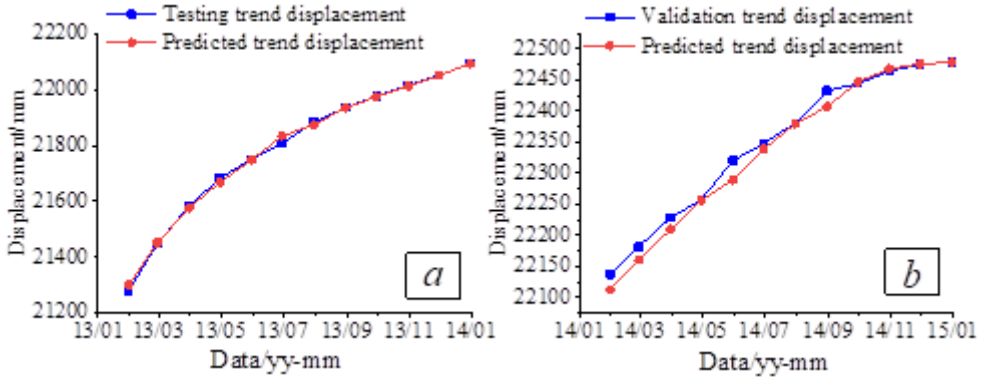

Figure 13

Prediction results of the trend displacement. (a) Testing trend displacement. (b) Validation trend displacement.

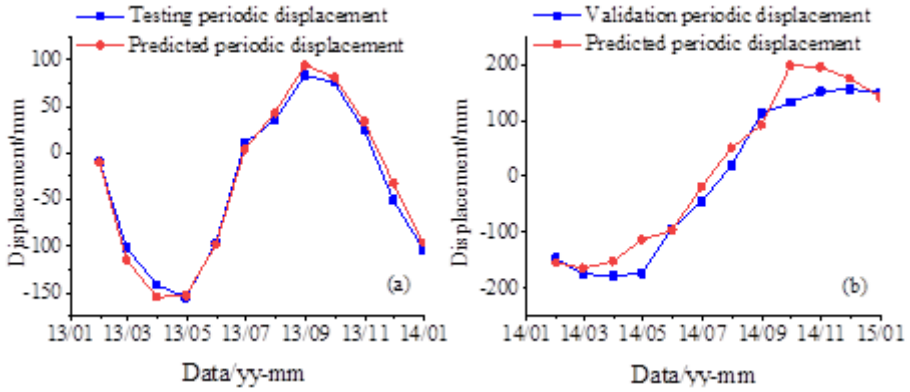

Figure 14

Prediction results of the periodic displacement. (a) Testing periodic displacement. (b) Validation periodic displacement.

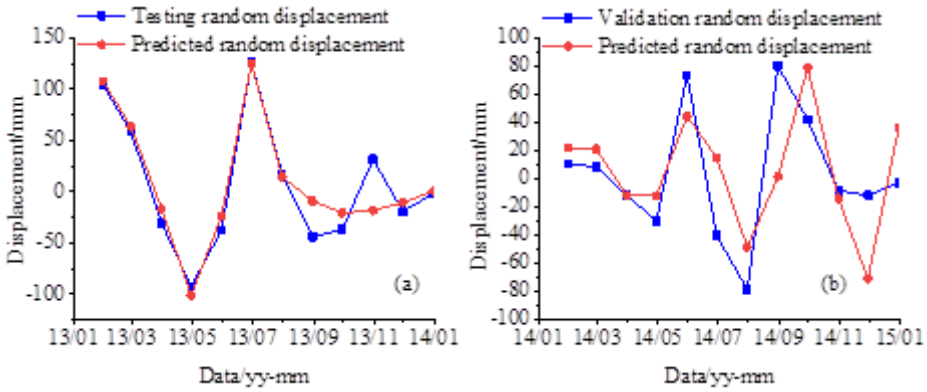

Figure 15

Prediction results of the random displacement. (a) Testing random displacement. (b) Validation random displacement.

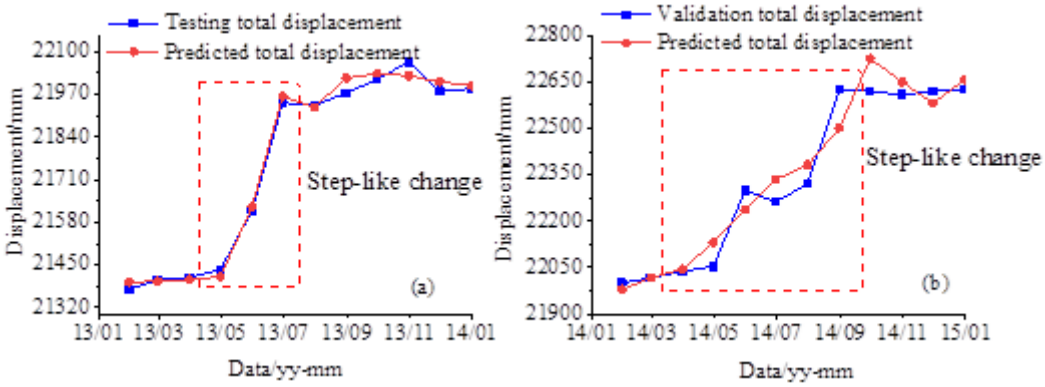

Figure 16

Prediction results of the total displacement. (a) Testing total displacement. (b) Validation total displacement. a) Performance of various displacement prediction model