

## Article

# A Novel Model for Landslide Displacement Prediction Based on EDR Selection and Multi-Swarm Intelligence Optimization Algorithm

Junrong Zhang <sup>1</sup>, Huiming Tang <sup>1,2,3,\*</sup>, Dwayne D. Tannant <sup>4</sup>, Chengyuan Lin <sup>1</sup>, Ding Xia <sup>1</sup>, Yankun Wang <sup>5</sup> and Qianyun Wang <sup>2</sup>

<sup>1</sup> Faculty of Engineering, China University of Geosciences, Wuhan 430074, China; zjr@cug.edu.cn (J.Z.); chengyuanlin@cug.edu.cn (C.L.); cug\_xia@cug.edu.cn (D.X.)

<sup>2</sup> Three Gorges Research Center for Geohazards of Ministry of Education, China University of Geosciences, Wuhan 430074, China; wangqianyun@cug.edu.cn

<sup>3</sup> Badong National Observation and Research Station of Geohazards, China University of Geosciences, Wuhan 430074, China

<sup>4</sup> School of Engineering, University of British Columbia, Kelowna, BC V1V 1V7, Canada; dwayne.tannant@ubc.ca

<sup>5</sup> School of Geosciences, Yangtze University, Wuhan 430100, China; ykwang@yangtzeu.edu.cn

\* Correspondence: tanghm@cug.edu.cn; Tel.: +86-027-6788-3127

**Abstract:** With the widespread application of machine learning methods, the continuous improvement of forecast accuracy has become an important task, which is especially crucial for landslide displacement predictions. This study aimed to propose a novel prediction model to improve accuracy in landslide prediction, based on the combination of multiple new algorithms. The proposed new method includes three parts: data preparation, multi-swarm intelligence (MSI) optimization, and displacement prediction. In the data preparation, the complete ensemble empirical mode decomposition (CEEMD) is adopted to separate the trend and periodic displacements from the observed cumulative landslide displacement. The frequency component and residual component of reconstructed inducing factors that related to landslide movements are also extracted by the CEEMD and *t*-test, and then picked out with edit distance on real sequence (EDR) as input variables for the support vector regression (SVR) model. MSI optimization algorithms are used to optimize the SVR model in the MSI optimization; thus, six predictions models can be obtained that can be used in the displacement prediction part. Finally, the trend and periodic displacements are predicted by six optimized SVR models, respectively. The trend displacement and periodic displacement with the highest prediction accuracy are added and regarded as the final prediction result. The case study of the Shiliushubao landslide shows that the prediction results match the observed data well with an improvement in the aspect of average relative error, which indicates that the proposed model can predict landslide displacements with high precision, even when the displacements are characterized by stepped curves that under the influence of multiple time-varying factors.

**Keywords:** landslide displacement prediction; complete ensemble empirical mode decomposition (CEEMD); edit distance for real sequence (EDR); multi-swarm intelligence (MSI); support vector regression (SVR)



**Citation:** Zhang, J.; Tang, H.; Tannant, D.D.; Lin, C.; Xia, D.; Wang, Y.; Wang, Q. A Novel Model for Landslide Displacement Prediction Based on EDR Selection and Multi-Swarm Intelligence Optimization Algorithm. *Sensors* **2021**, *21*, 8352. <https://doi.org/10.3390/s21248352>

Academic Editor: Francesca Cigna

Received: 24 November 2021

Accepted: 12 December 2021

Published: 14 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Landslides reactivated by the impoundment of a reservoir or rainfall can cause catastrophic losses such as casualties, road burying, and house damages, which seriously threaten the property and life safety of human society [1,2]. In 2019, there were approximately 6181 geological hazard events in China, causing economic losses of 2.77 billion yuan. Among these, 4220 were landslides, accounting for 68% of the geological hazards [3]. The development of more accurate and effective landslide displacement prediction methods is

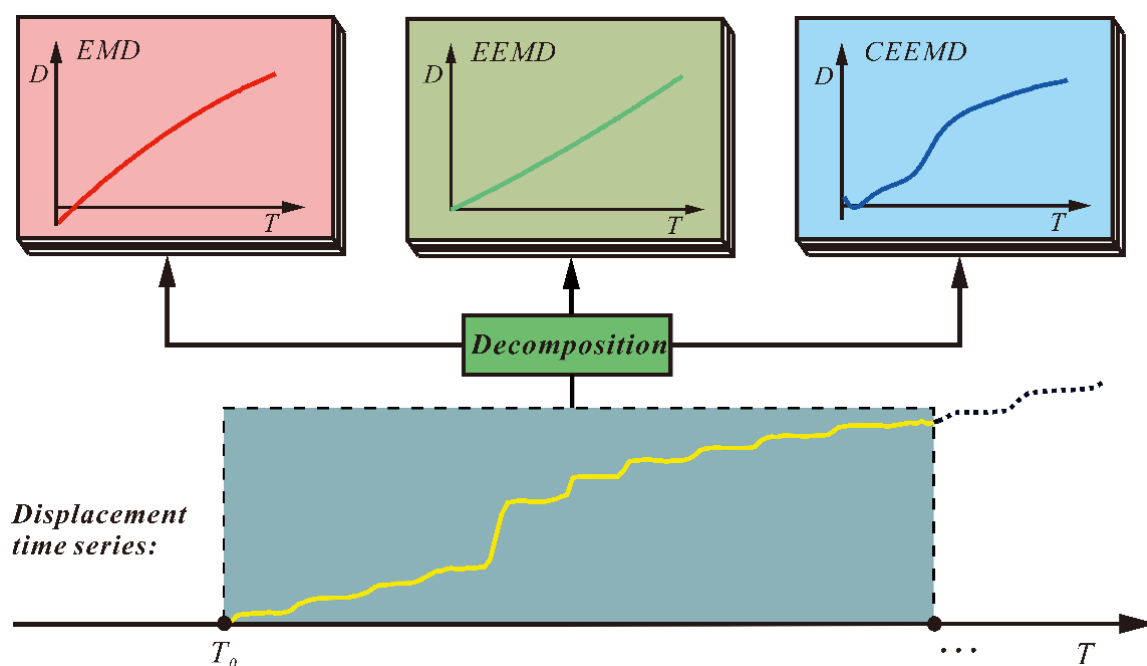
of great significance for the early warning of catastrophic landslide movements and is an active research area [4–7]. Through the information obtained from the prediction approaches, the landslide status can be evaluated, and the corresponding mitigation measures can be taken in advance to reduce the destructive effects of landslides.

Landslide prediction models can generally be divided into physical–mechanical and phenomenological models [8,9]. The physical–mechanical models are generally recognized as originating from the empirical formula proposed by Saito in 1965 [10], and a series of models have been developed based on creep theory in the following decades [11,12]. Owing to the complexity, strict application conditions, and time-consuming shortcomings of the physical–mechanical models, research on phenomenological models is becoming more and more popular nowadays [13]. By means of mathematical statistics and machine learning, measured landslide displacements are analyzed and modeled while considering the related factors, such as rainfall, the reservoir water level, groundwater level, etc., allowing for the prediction of landslide displacements [12,14].

The support vector machine (SVM) is a frequently used method among all phenomenological models. Nevertheless, when solving regression problems, the performance of the SVM model, also known as support vector regression (SVR), is highly influenced by the determination of penalty parameter  $C$  and kernel parameters  $g$  [15]. Therefore, research has focused on improving the predictive ability of SVR models for landslide displacements through optimization algorithms. In addition to some classical optimization algorithms such as the genetic algorithm (GA) [16], particle swarm optimization (PSO) [17–19], artificial bee colony (ABC) [20], and ant colony optimization (ACO) [21], recently, studies have advanced with the times, and some newly developed optimization algorithms start to be used [22,23]. Moreover, the continuous enhancement process of the optimization algorithm, as well as the evaluation of the prediction effect after using different frameworks, are also carried out at the same time. Miao et al. [24] adopted a variety of algorithms to optimize the SVR model and achieved a good application effect in the prediction of Baishuihe landslide displacement. Zhang et al. [25] made comparisons of the predictive capability of the SVR model optimized by ACO and GA and found the advantage of ACO-SVR with the consideration of the inducing factors' frequency component. At present, the application of optimization algorithms on SVR-based landslide prediction model parameter optimization is limited. It is still necessary to apply new optimization algorithms to these SVR-based models and compare their performance in landslide prediction.

Although based on the no free lunch (NFL) theorem, any optimization algorithms are equivalent when their performance is averaged across all possible problems; the swarm intelligence optimization algorithms (SIs) still show competitive results in solving optimization problems [26]. Similar to evolutionary algorithms (EA) [27] and artificial neural network algorithms (ANN) [28], the SIs also belong to the nature-inspired metaheuristics method [29]. With its high robustness, the SIs have been applied in many fields, including data clustering, network traffic forecast, data classification, UAV control, etc. Liu et al. [30] proposed a model of a global artificial fish swarm algorithm optimized support vector regression (GAFSA-SVR) for the network traffic forecast; the simulation shows an improvement of forecast precision and is superior to GA and chaos particle swarm optimization (CPSO)-optimized SVR model. Ali et al. [31] adopted the ant lion optimization algorithm (ALOA) in optimal allocation and sizing of renewable distributed generation sources in various distribution networks and results confirmed the effectiveness of the proposed algorithm. Jiang et al. [32] proposed an opposition-based seagull optimization algorithm (OSO) to overcome the shortage of classification models such as slow computation, instability, and sensitivity to noise. In this paper, six new SIs proposed after 2010, including the bat algorithm (BA) [33], grey wolf optimization (GWO) [34], dragonfly optimization algorithm (DA) [35], whale optimization algorithm (WOA) [36], grasshopper optimization algorithm (GOA) [37], and sparrow search algorithm (SSA) [38], have been tested and compared in the proposed model, and the most suitable optimization algorithm has been identified.

Decomposition of landslide displacement is also a vital step in a prediction model and will directly affect the prediction effect. At present, decomposition method based on signal processing technology, for instance, Fourier transform (FT), discrete wavelet transforms (DWT), wavelet transform (WT), empirical mode decomposition (EMD), variational mode decomposition (VMD), and ensemble empirical mode decomposition (EEMD), are massively used in this field [39–42]. With these methods, the landslide displacement can be decomposed into a trend term and a periodic term, and then these components of the displacement can be predicted by different models. However, when using the CEEMD (complete ensemble empirical mode decomposition), the residual term shows a trend of first decreasing and then increasing, which is difficult to predict as a trend term compared with the residual terms of EMD and EEMD (Figure 1). Hence, a novel prediction model needs to be designed when the CEEMD is adopted in the decomposition of landslide displacement.



**Figure 1.** Residual terms of Baishuihe landslide displacements obtained through EMD, EEMD, and CEEMD.

The screening of input parameters for an SVR model from related factors is an important part of prediction model optimization. Grey relational analysis (GRA) is a usual approach for this and has achieved convincing results [43]. Meanwhile, many other statistical methods such as maximal information coefficient (MIC) [23] and mean influence value (MIV) [44] have also been tried for this purpose. Zhang et al. [25] found that, as a similarity measuring method of time series, dynamic time warping (DTW) can be employed and works well in optimal input parameters selection of the SVR model. However, the DTW has the limitation of insensitive to the noise of the time series. To overcome this, the edit distance on real sequence (EDR) has been chosen and utilized in this study [45]. The EDR method is a classic trajectory similarity measurement that calculates the minimal number of editing operations needed for altering one sequence to another. With the advantages of robustness and accuracy, it has been utilized in traffic trajectory classification, physical movement similarity, and fiber segmentation, etc. [46]. It can also be applied in the related components selection for the prediction of landslides. Through calculating the similarity between restructured related factors sequence and periodic displacements sequence after normalization, two restructured related factors with minimum EDR value are the input variables of the SVR model.

This paper aims to improve the accuracy of landslide displacement prediction by constructing a novel model combined with the EDR method and multi-swarm intelligence (MSI). The new method can provide useful predictions of landslide displacements, allowing for the landslide status to be evaluated and the corresponding landslide mitigation measures to be taken before destructive movements occur.

In this paper, the next content is arranged as follows. In Section 2, the CEEMD, EDR, and MSI algorithms are briefly introduced. Section 3 considers the geological conditions and deformation features of the study case, the Shiliushubao landslide. The data preparation and statistical analysis of related factors are shown in Section 4. The predicted results and analysis are shown in Section 5. Section 6 discusses the proposed method, and conclusions are given in Section 7.

## 2. Methodology

### 2.1. Data Preprocessing with CEEMD

The CEEMD method is an effective improvement of the EMD method and EEMD method. By adding the white noise in the way of positive and negative pairs to the initial sequence of data, the residual auxiliary noise in the reconstruction signal can be better eliminated. Furthermore, the number of noise sets added can be very low, resulting in higher calculation efficiency. In CEEMD, based on local characteristics, the sequence can be converted to a limited number of intrinsic mode functions (IMF) and a residue. The operation of CEEMD includes three steps [47]:

Step 1: Add white noise consisting of positive and negative pairs to the original sequence data.

$$\begin{bmatrix} P \\ T \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \eta(t) \\ N \end{bmatrix} \quad (1)$$

where the original sequence is  $\eta(t)$ ,  $N$  is the added white noise, and  $P$  and  $T$  are two reverse white noise. The number of the decomposed sequences is  $2n$ , with  $j$  as the  $j$ th sample.

Step 2: Obtain a series of IMFs by decomposing  $P$  and  $N$  with the EMD method to generate two sets of IMFs.

$$\begin{cases} P = \sum_{i=1}^m IMF_{ji}^+ \\ T = \sum_{i=1}^m IMF_{ji}^- \end{cases} \quad (2)$$

where  $IMF_{ji}^+$  is the  $i$ th IMF after adding the positive white noise,  $IMF_{ji}^-$  is the  $i$ th IMF after adding the negative white noise, and  $m$  is the number of IMFs.

Step 3: Repeat step 1 and step 2 to get the corresponding IMF terms, and calculate the average of all the IMFs:

$$IMF_j = \frac{\sum_{i=1}^n (IMF_{ij}^+ + IMF_{ij}^-)}{2n} \quad (3)$$

Through this method, the original sequence can be expressed as the sum of some IMFs and a residue  $r_n(t)$ .

Zhang et al. pointed out that the CEEMD method combined with a  $t$ -test can obtain the high-frequency and low-frequency components from related factors such as rainfall and the reservoir water level through a fine-to-coarse reconstruction [25]. Moreover, according to the time series theory, the landslide displacement can be separated into a trend term and a periodic term by methods presented in the Introduction section. In this paper, the CEEMD is adopted as the decomposition method, and the obtained residual term is considered as the trend term. The result after the trend term is subtracted by the cumulative displacement of the landslide is regarded as the period term. Due to the special shape of the trend displacement time series after CEEMD decomposition, the displacement trend term and the period term will be predicted by the SVR model, respectively, later.



## 2.2. Selection of Optimal Related Factors via EDR

The EDR, which is based on Levenshtein distance, is a traditional and well-established similarity measurement method proposed by Chen et al. [45] and has been used for judging trajectory similarity since [48,49]. The EDR calculates the number of insertions, deletions, or replacement operations required to change the sequence  $R$  to  $T$  when the threshold is  $\varepsilon$ . It reduces the effect of noise by quantifying the distance into 0 and 1, and the Levenshtein distance method itself improves the local time-shifting situation (especially when the local time-shifting is not very large). Based on this, the displacement trend term sequence and residue of restructured related factors sequence were set as a reference sample sequence  $R = \{r_1, r_2, \dots, r_n\}$  and a test sample sequence  $T = \{t_1, t_2, \dots, t_m\}$  after normalization. Then, the EDR( $R, T$ ) can be calculated as follows:

$$\text{match}(r_i, t_j) = \text{true}; \text{ if } |r_{ix} - t_{jx}| \leq \varepsilon \text{ and } |r_{iy} - t_{jy}| \leq \varepsilon \quad (4)$$

$$D_{EDR}(R, T) = \begin{cases} n; \text{ if } m = 0 \\ m; \text{ if } n = 0 \\ \text{Min} \begin{cases} D_{EDR}(\text{Rest}(R), \text{Rest}(T)) + \text{subcost}, \\ D_{EDR}(\text{Rest}(R), T) + 1, \\ D_{EDR}(R, \text{Rest}(T)) + 1 \end{cases} & ; \text{ otherwise} \end{cases} \quad (5)$$

$$\text{subcost} = \begin{cases} 0, \text{ match}(r_1, t_1) = \text{true} \\ 1, \text{ otherwise} \end{cases} \quad (6)$$

where the real number  $0 < \varepsilon < 1$  is the matching threshold. The cost for a replace, insert, or delete operation is set to 1. Therefore, through calculating the edit distance between two sequences, the smaller the EDR is, the greater the similarity will be. After calculating the EDR between the displacement trend term sequence and residues of original related factors sequence and restructured related factors sequence, three residues with the highest similarity were chosen as the input variable of the SVR model for predicting the displacement trend term.

Similarly, three optimal input variables for predicting the displacement periodic term with an SVR model can be obtained by calculating the EDR between displacement periodic term sequence and original related factors, restructured related factors and related factors frequency sequence.

## 2.3. Support Vector Regression (SVR)

The support vector regression (SVR) algorithm is a classic landslide displacement prediction model developed from statistical learning theory. With a powerful generalization ability and robust performance, the SVR model can easily solve quadratic programming problems with constraints. The main steps of an SVR model are summarized as follows [50].

Suppose that a nonlinear sample set in low dimensional space is:  $\{x_i, y_i\}$ , where  $x_i = \{x_{i1}, x_{i2}, \dots, x_{ip}\}$  is the input vector,  $y_i$  is the corresponding output vector,  $i$  is the number of samples and  $j$  is the number of input vectors. Then, the regression estimation function is:

$$f(x) = w^T \varphi(x) + b \quad (7)$$

where  $w$  is the weight vector,  $\varphi(x)$  is the nonlinear mapping function and  $b$  is the offset. Through minimizing the following equation, the value of  $w$  and  $b$  can be obtained:

$$\min J = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i^+ + \xi_i^-) \quad (8)$$

$$\text{s.t.} \begin{cases} y_i - w^T \varphi(x_i) - b \leq \varepsilon + \xi_i^+ \\ w^T \varphi(x_i) + b - y_i \leq \varepsilon + \xi_i^- \\ \xi_i^+, \xi_i^- \geq 0, i = 1, 2, \dots, n \end{cases} \quad (9)$$

where  $C$  and  $\varepsilon$  are the penalty parameter and the size of the insensitive loss function, respectively.  $\zeta_i^+$  and  $\zeta_i^-$  are the relaxation factors. By solving the quadratic optimization problem, the weight vector  $w$  can be expressed as:

$$w = \sum_{i=1}^n (\beta_i^* - \beta_i) \varphi(x_i) \quad (10)$$

where  $\beta_i^*$  and  $\beta_i$  are Lagrange multipliers. Therefore, the SVR model can be denoted as follows:

$$f(x) = \sum_{i=1}^n (\beta_i^* - \beta_i) K(x_i, x_p) + b \quad (11)$$

where  $K(x_i, x_p)$  is the kernel function. The SVR kernel function has various forms; in this study, the Gaussian radial basis function (RBF function) is chosen and adopted. Since algorithms for the determination of the penalty factor and the kernel function parameter ( $C, g$ ) vary, the approach for selecting  $C$  and  $g$  must be further studied. Different forms of MSI algorithms were explored for the parameter optimization of the SVR model and all of them are briefly described next.

#### 2.4. Multiple Swarm Intelligence

##### 2.4.1. Bat Algorithm (BA)

The bat algorithm (BA), proposed in 2010 by Yang et al., is a novel swarm intelligence optimization technique that simulates the echolocation behavior of microbats [33]. Based on iteration, this algorithm describes the echolocation of microbats and uses it to minimize any objective function and solve optimization problems. In BA, after initializing a group of random solutions, the optimal solution is searched by iteration, and a new local solution is generated by a random flight around the optimal solution, which strengthens the local search. BA is an accurate and effective method of finding the optimal parameter values for an SVR model with few parameters to adjust.

##### 2.4.2. Grey Wolf Optimization (GWO)

The grey wolf optimization (GWO) algorithm is a new swarm intelligent optimization algorithm proposed by Mirjalili et al. [23,51]. Based on the predatory behavior and strict social dominant hierarchy of grey wolves, this algorithm first randomly generates a group of gray wolves in the search space. Then, the wolves are divided into four social hierarchies according to the fitness from high to low, each marked with alpha, beta, delta, and omega. The location and distance between the grey wolves and the prey, which is the possible solution of the optimized SVR model, is obtained through iterative calculation. Finally, through the evolution of the wolf group itself, the distance between them is gradually reduced to realize the optimal hunting of prey. The algorithm has the advantages of strong convergence, few parameters, and easy implementation.

##### 2.4.3. Dragonfly Algorithm (DA)

The dragonfly optimization algorithm (DA) is a swarm intelligent optimization algorithm proposed by Mirjalili et al. [35,52]. The algorithm is based on the dynamic and static swarm behavior of dragonflies in nature, which includes separation behavior, alignment behavior, cohesion behavior, foraging behavior, and distraction from enemy behavior. By establishing a mathematical model of all these behaviors, the dragonfly's latest position vector, which is a possible solution of the objective function, is calculated. This algorithm has the advantages of simple calculation, low complexity, few control parameters, and fast convergence speed.

#### 2.4.4. Whale Optimization Algorithm (WOA)

The WOA algorithm is a new heuristic optimization algorithm. The key idea is to simulate the behavior of humpback whales [36]. The humpback whales hunt in a special way using bubble nets, which can be described as two mechanisms: upward spirals and double loops. The WOA optimization algorithm has three steps: searching and encircling prey, the bubble-net argument attacking method (exploitation phase), and search for prey (exploration phase). Through this, the position vector of humpback whales with the best fitness value can be obtained by satisfying a termination criterion, and the final position vector is chosen as the best solution of the optimized SVR model parameters. The algorithm has the advantages of simple operation, few parameters to adjust, and a strong ability to jump out of a local optimum.

#### 2.4.5. Grasshopper Optimization Algorithm (GOA)

The grasshopper optimization algorithm (GOA), proposed by Saremi et al., in 2017, is a metaheuristic bionic optimization algorithm that mimics the swarming behavior of grasshoppers during population migration (exploration) and foraging behavior (exploitation) [37]. The grasshoppers' position vector is equal to the value of an objective function [53]. When the grasshoppers reach a food source, the parameters reach the optimal variable, and the optimal value of the SVR model parameters is obtained. The algorithm provides a balanced condition between local and global search operators to achieve the final target. Two forces in grasshoppers, attraction and repulsion, provide global search and local search, respectively. To obtain effective solutions, the influence of the grasshopper's current position, its relative position to other grasshoppers, and the position of the target point are regarded as the effective agents to determine the search vector. It has higher search efficiency and faster convergence speed, and its special adaptive mechanism can balance the global and local search processes with better optimization accuracy.

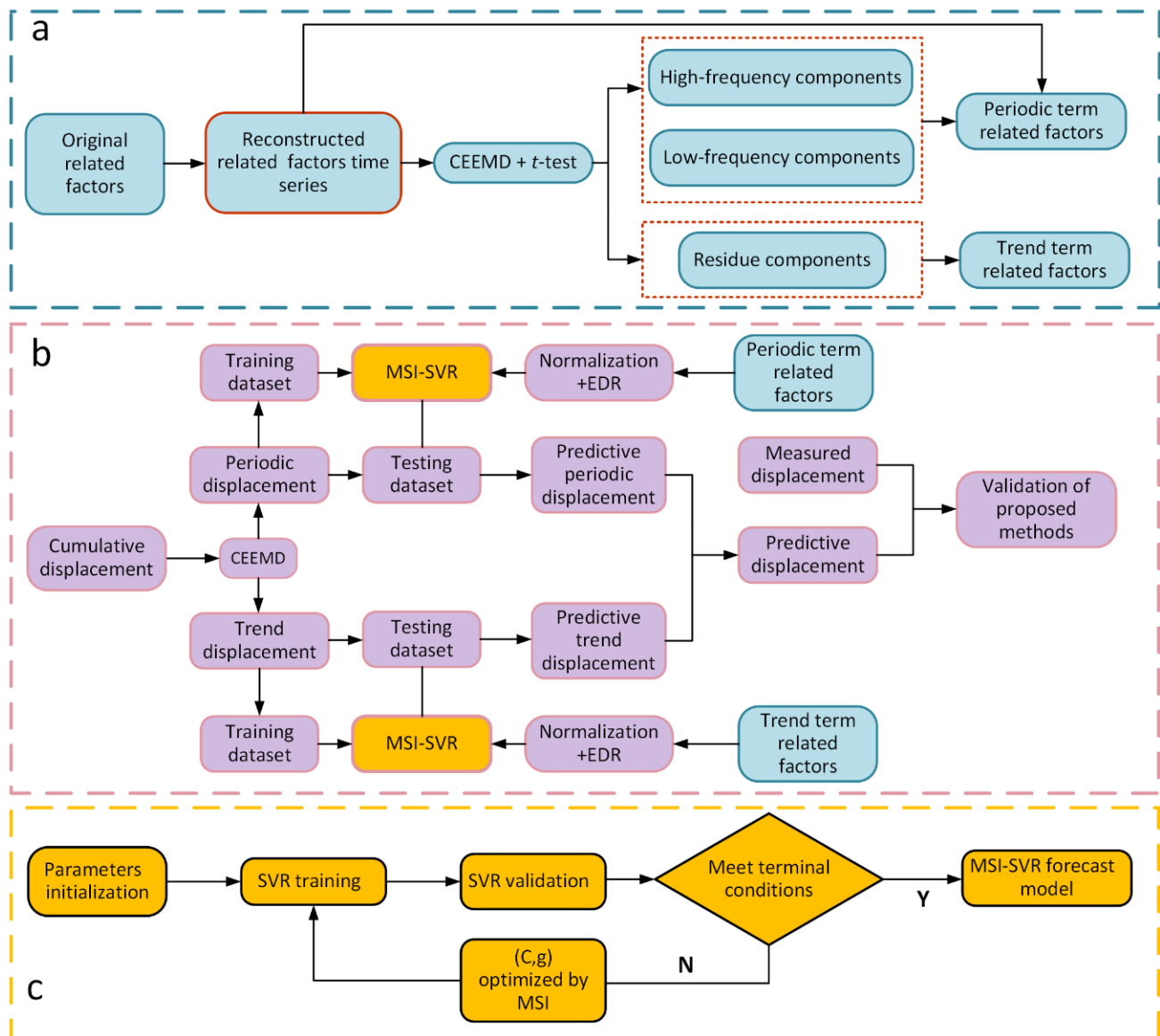
#### 2.4.6. Sparrow Search Algorithm (SSA)

The sparrow search algorithm (SSA), as proposed by Xue et al. [38], was mainly inspired by the foraging behavior and anti-predation behavior of sparrows. Some sparrows are in charge of seeking food and providing locations for the entire population, while the remaining sparrows use the locations to obtain food. Meanwhile, when a sparrow is aware of the danger and alarms, the entire population will immediately take anti-predation behavior. Although idealized, these behaviors are formulated with corresponding rules, and the algorithm classifying the sparrows into producers and scroungers. Their positions are updated according to their own rules, separately. In SSA, the position of each sparrow is equal to a possible solution of the objective function, and the best solution can be obtained when meeting iteration conditions. The algorithm is novel and has the advantages of a strong optimization ability, fast convergence speed, fewer adjustment parameters, and simple calculation.

### 2.5. Procedure of the Proposed Hybrid Algorithm

The framework of the proposed ensemble prediction model is shown in Figure 2. The entire forecasting process is divided into three steps: data preparation, multi-swarm intelligence (MSI) optimization, and displacement prediction. In the data preparation step, the time-sequences of factors related to the landslide movements, such as rainfall and reservoir water level, are restructured. The frequency component and residual component of all original and restructured sequences are then obtained through the combined application of CEEMD and *t*-test. In the MSI optimization step, MSI optimization algorithms are used to select the optimal *C* and *g* for the SVR model. In the displacement prediction step, the trend and periodic displacements are extracted from the observed cumulative landslide displacement through CEEMD. Then EDR is used to select the input variables of the periodic displacement prediction SVR model by calculating the EDR value between the periodic displacement and original related factors, restructured related factors, and

frequency related factors after normalization. Similarly, the input variables of the trend displacement prediction SVR model are obtained by calculating the EDR value between the trend term displacement and all residue terms after normalization. Finally, the predictions of the trend and the periodic displacements are performed separately, and the total predicted displacement is obtained by adding them together.



**Figure 2.** Framework of the proposed ensemble prediction model, (a) data preparation step, (b) displacement prediction step, and (c) MSI optimization step.

### 2.6. Performance Evaluation Formula

The most commonly used indicators to evaluate the performance of prediction models are coefficient of determination ( $R^2$ ), root mean square error (RMSE), mean absolute error (MAE), and mean average percentage error (MAPE). These indicators were used in this study and are defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_t - \hat{y}_t)^2}{\sum_{i=1}^N (y_t - \bar{y}_t)^2} \quad (12)$$



$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_t - \hat{y}_t)^2} \quad (13)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_t - \hat{y}_t| \quad (14)$$

$$\text{MAPE} = \frac{1}{N} \left( \sum_{i=1}^N \left| \frac{\hat{y}_t - y_t}{y_t} \right| \right) \times 100\% \quad (15)$$

where  $y_t$  is the  $t^{\text{th}}$  measured value,  $\bar{y}_t$  is the mean of the measured value,  $\hat{y}_t$  is the  $t^{\text{th}}$  predicted value, and  $\bar{\hat{y}}_t$  is the mean value of the prediction.

### 3. Cases Study

#### 3.1. Geological Conditions

The Shiliushubao landslide is part of the famous Huanglashi landslide group, one of the large-scale landslides in the Three Gorges Reservoir Area (TGRA). It is located on the north bank of the Yangtze River, 1.5 km east of Badong county, 66 km away from the Three Gorges Dam (TGD) (Figure 3). The landslide's geographical coordinates are  $110^{\circ}26'$  east longitude and  $31^{\circ}02'$  north latitude. The Shiliushubao landslide is bordered by the Lijiawan valley on the east and the Gan valley on the west, with a tongue-like shape. It is bigger than the well-known Baishuihe landslide with an estimated volume of  $11.8 \times 10^6 \text{ m}^3$  and covers an area of  $0.34 \text{ km}^2$ . The top of the landslide is at an elevation of 340 to 358 m with a width of 140 m, and the toe of the landslide is at an elevation of 68 to 80 m with a width of 570 m.

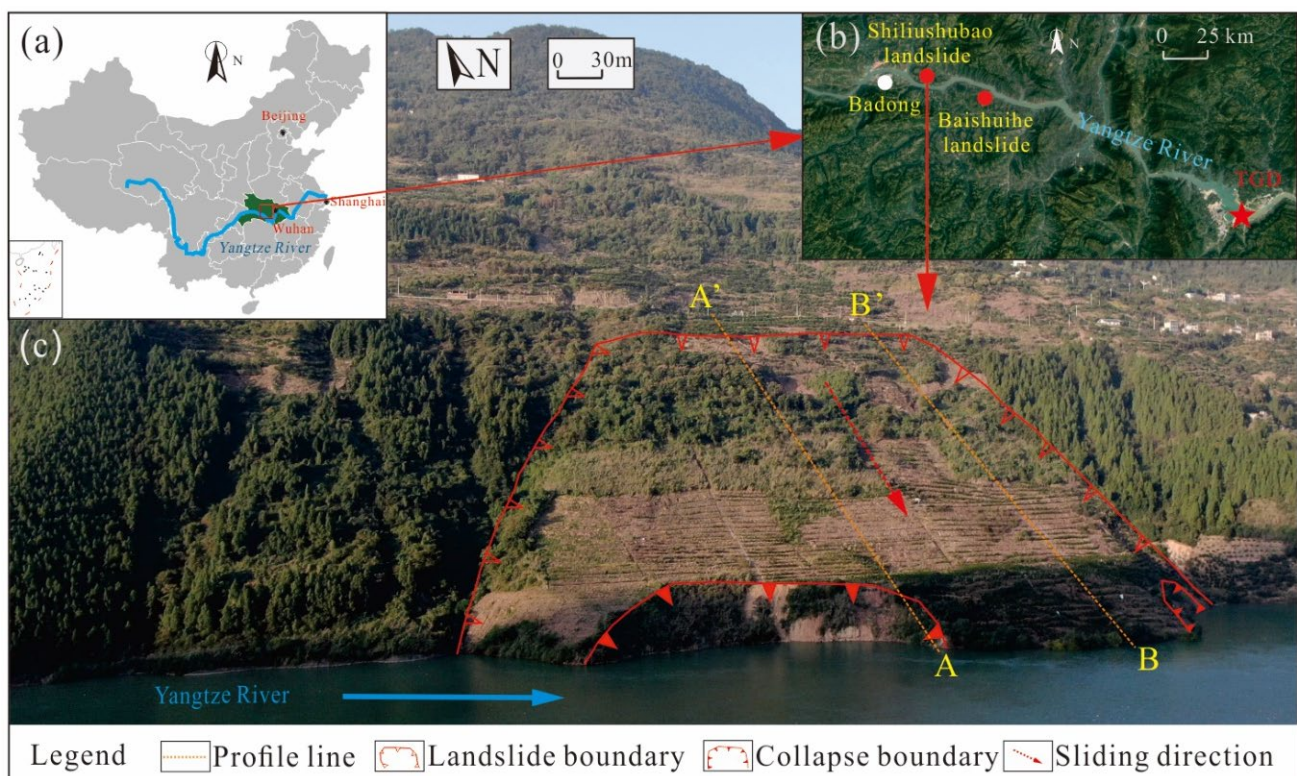
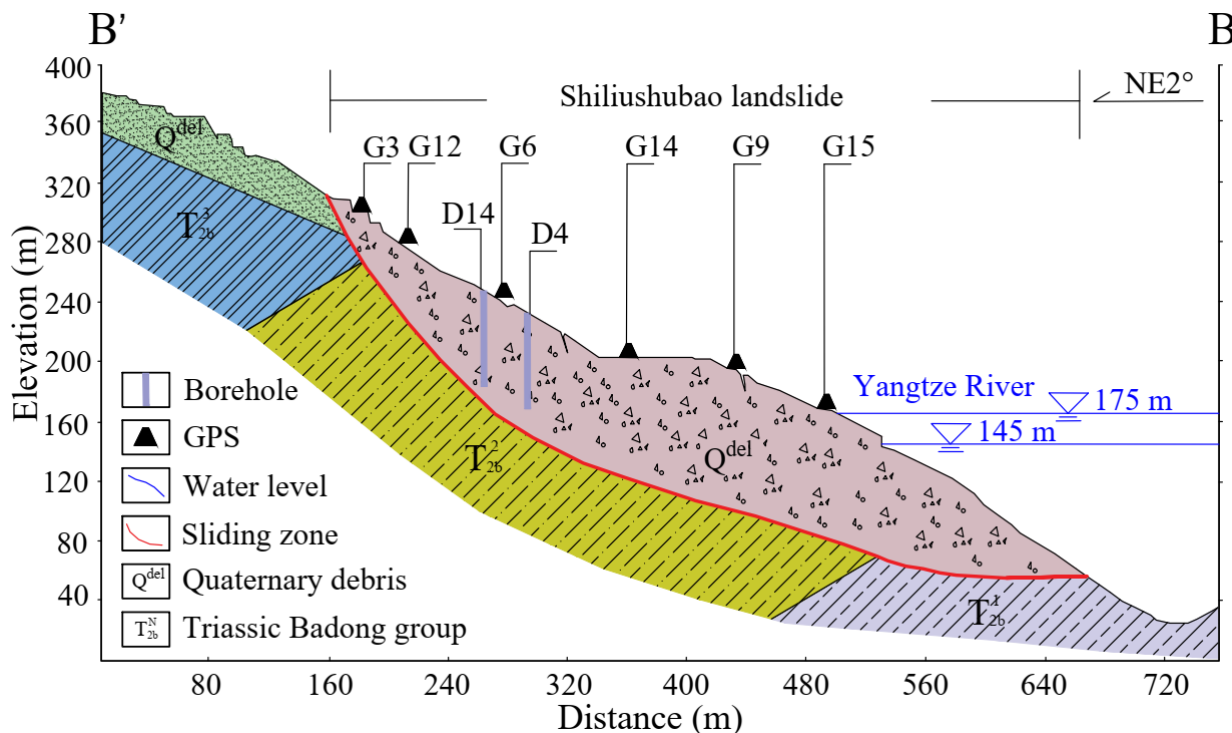


Figure 3. Location (a,b) and an oblique view (c) of Shiliushubao landslide captured by UAV, October 2020.



The cross-section of the ground surface is shown in Figure 4 by the profile B-B'. The average slope angle is  $26^\circ$  along the sliding direction. However, the slope contains a gently sloping bench at an elevation near 200 m, and the slope is much steeper than  $26^\circ$  above and below the bench. The slope angle is up to  $40^\circ$  at elevations below the reservoir level.



**Figure 4.** Geological section of the Shiliushubao landslide (B-B').

The geological profile B-B' in Figure 4. shows that the Shiliushubao landslide occurs in the Triassic Badong Group consisting of red mudstone, siltstone, gray-green marl, and limestone. These rocks are characterized by high clay mineral content (about 68%). Exposure of the rock to water allows the rock to soften and weaken. The sliding mass also includes near-surface Quaternary soils. The rear edge of the landslide is mainly a loose accumulation of gravel and clay. This soil is weak and is prone to collapses or sliding along the bedrock surface. The sliding zone consists of clay or silty clay with some gravel. The thickness of the sliding zone varies from 1.0 to 4.9 m, with an average thickness of 2.0 m.

The topography of the lower part of the Shiliushubao landslide was mostly altered by the newly formed Hengping landslide (Figure 5), and some landslide materials under 100 m elevation have been removed by erosion. There are some small gullies near the landslide's front edge caused by surface water runoff, which are the main channels for gathering and draining surface water.

### 3.2. Rainfall and Reservoir Levels

The Shiliushubao landslide is located in a subtropical zone, in which rainfall is continuous and concentrated in the summer. The rainy season generally occurs from May to September, which accounts for 70% of the yearly rainfall. Rainfall is one factor that increases the movement of the Shiliushubao landslide. Fluctuation in the reservoir level in the TGRA is another factor influencing the landslide movements, especially the sudden reservoir drops before the flood season.

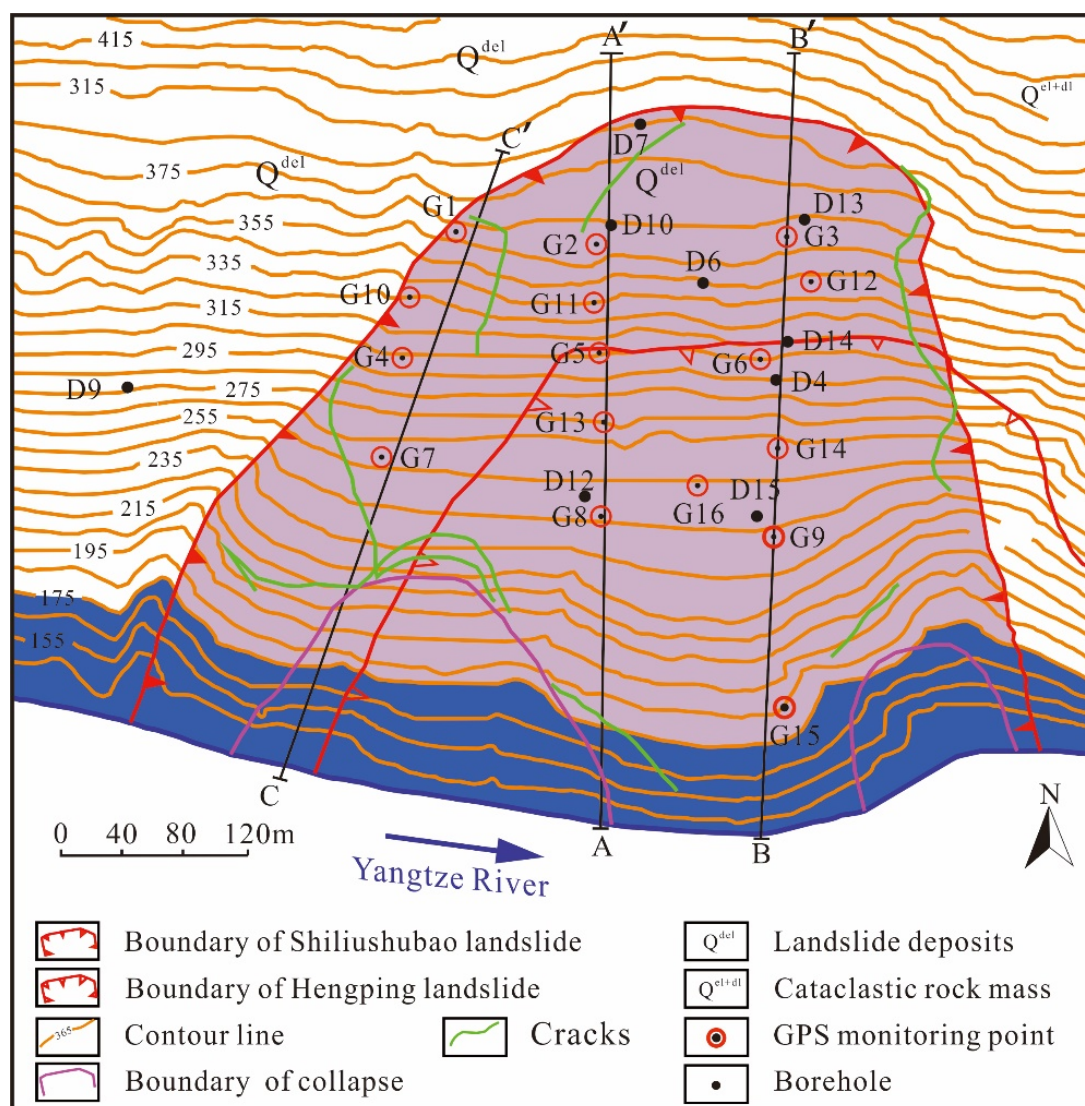


Figure 5. Geological map and monitoring points for the Shiliushubao landslide.

### 3.3. Deformation Characteristics

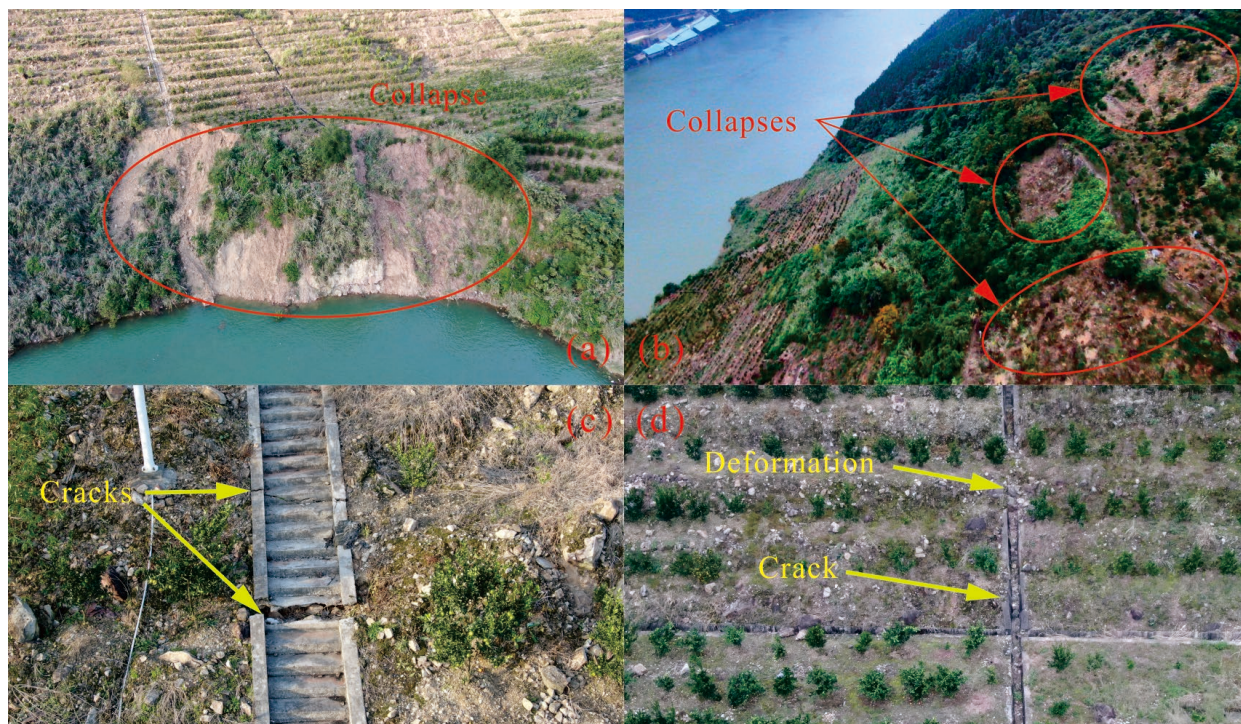
Since the reservoir was first impounded in June 2003, the toe area of the slope has experienced repeated small collapses (Figure 6). From 4 to 14 June 2004, four sliding events occurred during a period of rainfall, involving an estimated volume of 6000 m<sup>3</sup>. The toe area is very unstable, and slope movements at the toe affect the rest of the slope. At present, the slope's deformation processes are causing small collapses under the influence of rainfall or reservoir level fluctuations.

Slope movements have created ground fissures that have gradually intensified. Areas of subsidence have also occurred. While the existing main cracks continued to expand, a series of new cracks gradually formed at the landslide's rear edge. These cracks have connected and coalesced inside the sliding mass. The maximum crack length obtained by field monitoring is 345 m with opening widths up to 0.5 m and depths over 1 m. Many cracks have occurred in a concrete-lined drainage ditch at the front edge of the landslide. Moreover, some feathery cracks are also scattered along both sides of the landslide.



### 3.4. Landslide Monitoring

From February 2004 to December 2009, field monitoring was conducted to study the Shiliushubao landslide movements, based on which, the deformation evolution characteristics and development trend of the Shiliushubao landslide can be mastered. A total of sixteen GPS monitoring points and 15 boreholes were arranged on the surface of the sliding mass (Figure 5). Some monitoring points were destroyed due to rainfall, landslide movement, and other reasons. Thus, only monitoring data from February 2004 to December 2009 have been recorded and preserved. The cumulative displacement data from GPS points G1, G2, G4, G8, plus the rainfall and reservoir water levels were selected and shown in Figure 7.

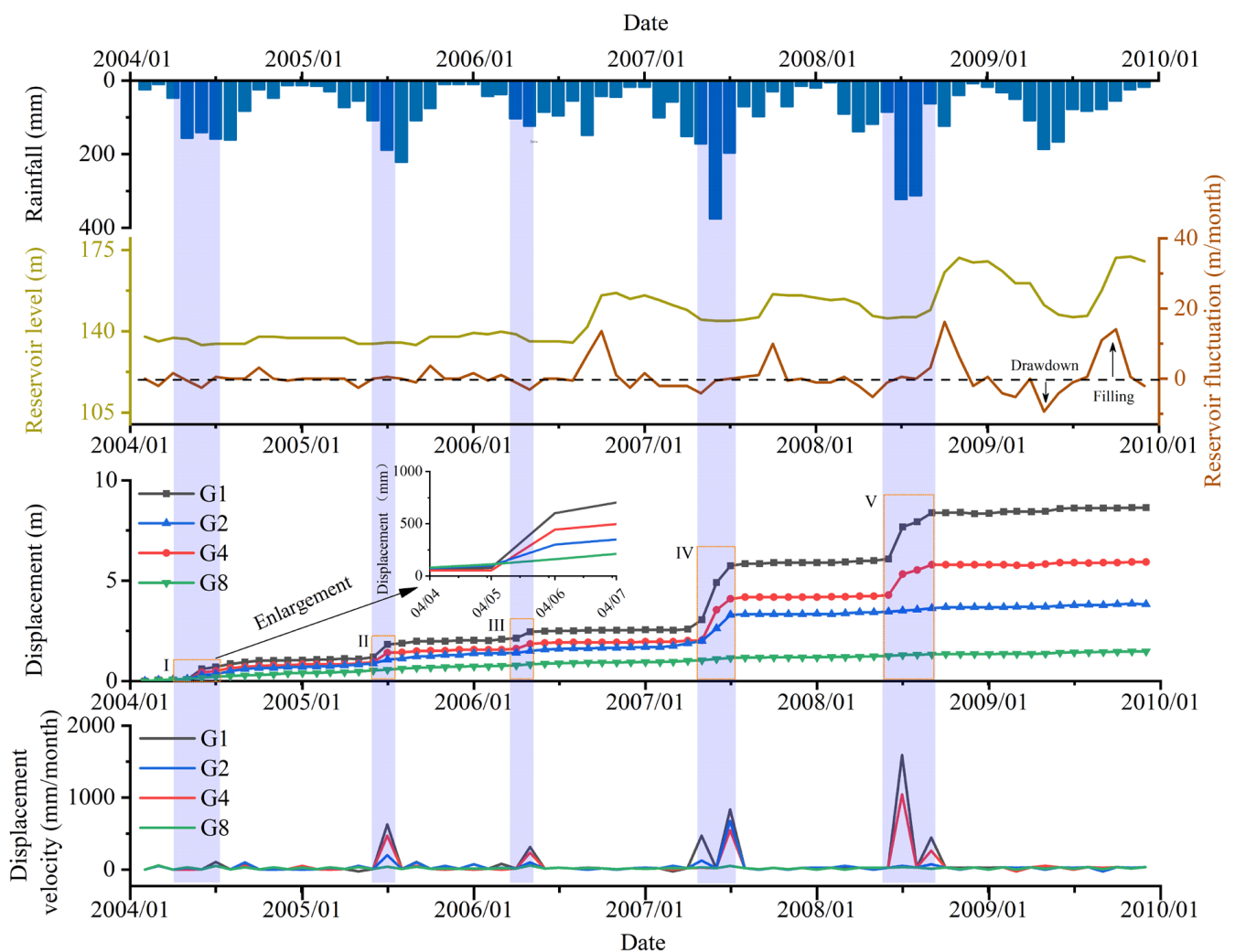


**Figure 6.** The ground collapses (a,b) and cracking (c,d) in the toe area captured by UAV, October 2020.

### 3.5. Analysis of Monitoring Data

Monitoring data show that the displacement of the sliding mass increases with time in an obvious stepped shape. From February 2004 to December 2009, due to rainfall, the five displacement jumps occurred in the rainy season (May to September). After the rainy season, the landslide resumes movement at a slow, roughly constant speed.

The fluctuation of the reservoir water level is another factor affecting the deformation of the sliding mass. When the reservoir level drops sharply, the movement of the sliding mass accelerates. For example, from January to May 2007, the water level dropped from 155.4 m to 144.7 m, and the landslide displacement rate reached 16.2 mm/month in March when the water level dropped by 5 m. In May, when the water level dropped by 10.7 m, the landslide displacement rate was 44.4 mm/month. Similarly, when the reservoir water level fell in other periods, such as January to July 2009, the landslide displacement rate increased from 1.5 to 47.1 mm/month.



**Figure 7.** Displacement data from GPS points G1, G2, G4, G8, and rainfall and reservoir level data.

For a better understanding of this seasonal deformation acceleration's related factors, a correlation analysis between displacement velocity at G1 (located at the northeast edge of the landslide) and rainfall, rate of reservoir level change, and reservoir level are shown in Figure 8. The size of the bubbles represents the deformation velocity. The larger bubbles tend to plot where the rainfall is higher. Meanwhile, the large bubbles are mainly concentrated where the reservoir level is between 140 and 150 m and are located where the reservoir level fluctuates slowly (between  $-4.4$  and  $9.0$  m/month). This indicates that reservoir level fluctuations mainly trigger accelerated landslide movements when the reservoir level is low. The maximum size bubble appears where the rainfall is about  $325$  mm/month, and the water level rises between  $4.5$  and  $9.0$  m/month. The combined effect of heavy rainfall and rising reservoir level on landslide deformation is more significant than low rainfall combined with reservoir level drawdown.

Inclinometer D7 indicates that the main sliding zone is located at a depth of 22 to 26 m (Figure 9). The data show that before June 2003, the shear deformation in the slip zone was slow. Then, with the operation of the TGRA, the displacement in the shear zone increases. Therefore, it can be judged that the Shiliushubao landslide is in the stage of accumulative creep deformation, and the deformation tends to be intensified under the influence of reservoir water.

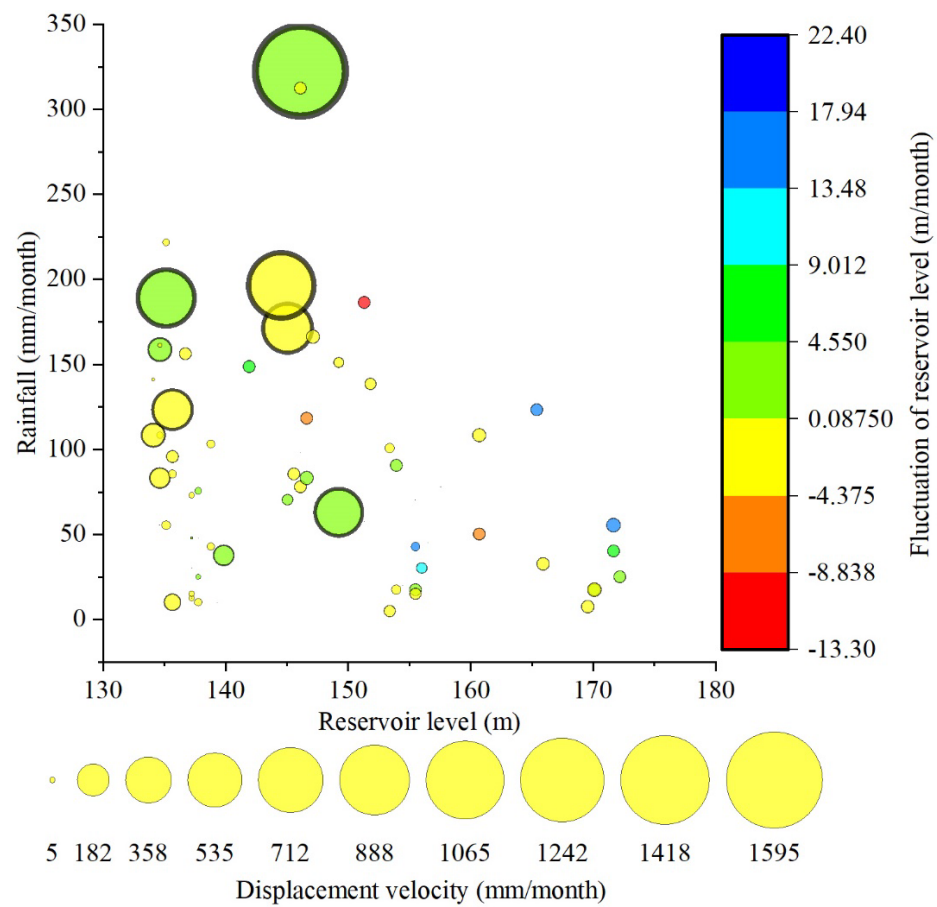


Figure 8. Correlation of displacement velocity at G1 versus the reservoir level, rainfall, and fluctuation of reservoir level.

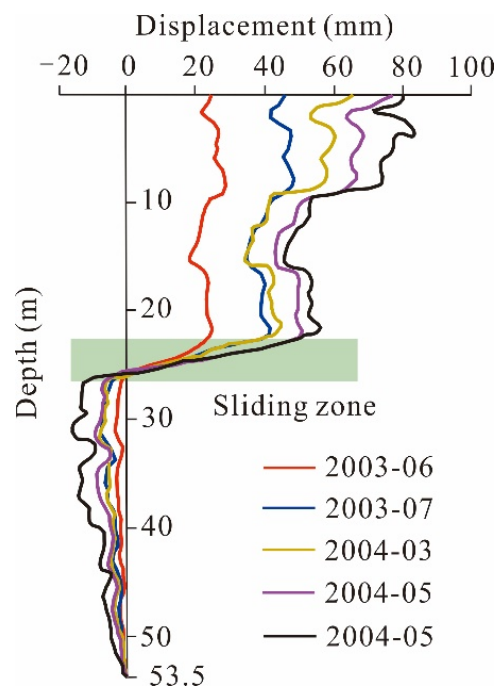


Figure 9. Lateral deformation versus depth in inclinometer D7.



In conclusion, the formation of the Shiliushubao landslide is the result of a series of related factors including internal inducing factors and external inducing factors. Weak rock formations are the inherent cause of deformation. In the Badong Formation, soft rocks characterized by high clay mineral content account for about 68%. The hydrophilicity of the rock determines that the rock has the characteristics of easy softening, muddy and weathering, and lays the material foundation for the deformation and failure of the slope. Water is the external cause of deformation. The impact of concentrated high intensity rainfall and periodic water storage activities in the TGRA, especially the sudden drop before the flood season, are main external inducing factors for the reactive of the Shiliushubao landslide.

#### 4. Data Processing and Statistical Analysis

##### 4.1. CEEMD Decomposition of Landslide Displacement Versus Time Data

Since all displacements at the Shiliushubao landslide show a similar step-like deformation curve, only the displacement data at site G1 is chosen for model validation in this study. The CEEMD method can be used to extract the trend displacements and the periodic displacements. The following parameters were used [2]:

- ensemble member = 200
- standard deviation of added white noise in each ensemble member = 0.2
- threshold variance = 0.2
- threshold for first iteration = 4

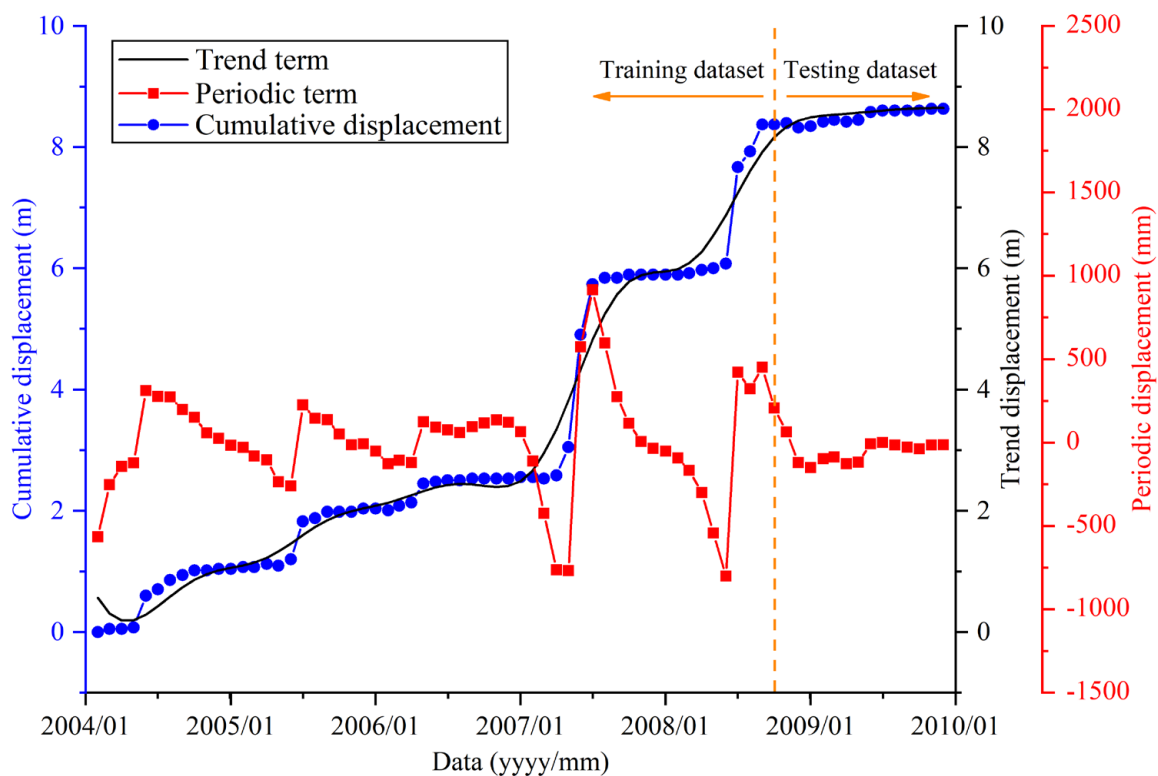
The landslide displacement sequence was decomposed into a few IMFs and a residue through CEEMD. The residue is considered to be the trend displacement of the landslide, and the periodic displacement was obtained by adding all the IMFs together.

The results show that the trend displacement component of G1 has local fluctuations and an increasing trend over time, which is consistent with a long-term trend of cumulative displacements. The periodic displacement component shows a cyclical variation in displacements ranging from  $-800$  to  $917$  mm. The maximum variation range of periodic displacement occurred in the 2007 rainy season when the TGR was first impounded. As the periodic displacement and trend displacement are important components of the cumulative displacement, they will be separately modeled and predicted. Once the best prediction for each component is obtained, the best prediction for cumulative displacement is obtained.

The displacement data are divided into training and testing data sets to establish the SVR prediction model of periodic and trend displacements (Figure 10). The SVR model is organized with the training dataset to establish the regression relationships between displacement and selected variables. The trained SVR model can then be used to predict the current month periodic displacement and compared with the testing dataset to verify the model's accuracy. In this study, the displacement data from February 2004 to September 2008 were selected as the training dataset, and the rest were used as the testing dataset.

##### 4.2. CEEMD Decomposition of Related Factors

Before selecting the input various parameters, the factors related to the landslide deformation are usually restructured first [24]. Original related factors such as the rainfall, reservoir level, and date of displacement were restructured. The current monthly rainfall sequence (L1) was restructured as the accrued precipitation of the previous two months (L2), as were the accrued precipitation of the previous month and the current month (L3), and the accrued precipitation of the previous two and the current month (L4). The current monthly reservoir level data (X1) were restructured as the reservoir level monthly change (X2) and the change of reservoir level between two months (X3). The displacement data (D) were restructured as the previous month displacement (D1) and the accrued displacement of the previous month and the current month (D2).



**Figure 10.** Periodic and trend displacement at site G1 obtained through CEEMD.

Keeping the CEEMD parameters fixed, L1–L4, X1–X3, and D1–D2, can be decomposed into a few IMFs sorted by frequency from highest to lowest and a residue. The mean of  $IMF_1$  was compared to the other IMFs by a paired  $t$ -test with a significance set at 0.05 (two-tailed) for each decomposed and restructured factor. If the significance values of  $IMF_i$  are greater than 0.05, the difference between  $IMF_1$  and  $IMF_i$  is not significant. Therefore, the superposition of IMFs from  $IMF_1$  to  $IMF_i$  is the high-frequency component, and the superposition of the remaining IMFs is the low-frequency component. The IMFs of each restructured factors are shown in Figure 11, and the results of the paired  $t$ -test are shown in Table 1.

The results reveal that the IMFs obtained from the decomposition of all factors show a certain periodicity. Their frequency varies, and  $IMF_1$  usually has the highest frequency and fluctuation amplitude. Since there is only one IMF after the CEEMD decomposition of D2, it is considered that there are only high-frequency components in D2. The paired  $t$ -test results indicate that only  $IMF_3$  in X1 and  $IMF_4$  in X3 has a significance value that is less than 0.05, which denotes that the low-frequency components only exist in X1 and X3. Taking these two as the low-frequency components of X1 and X3, the high-frequency components of the other factors will be the sum of the remaining IMFs. Therefore, in addition to the variables mentioned above, new variables can also be chosen as input to an SVR model of the periodic displacements after reconstruction: high-frequency current monthly rainfall sequence ( $L1^H$ ), high-frequency accrued precipitation of the previous two months ( $L2^H$ ), high-frequency accrued precipitation of the previous month and the current month ( $L3^H$ ), high-frequency accrued precipitation of the previous two and the current months ( $L4^H$ ), high-frequency current monthly reservoir level data ( $X1^H$ ), low-frequency current monthly reservoir level data ( $X1^L$ ), high-frequency reservoir level monthly change ( $X2^H$ ), high-frequency change of reservoir level between two months ( $X3^H$ ), low-frequency change of reservoir level between two months ( $X3^L$ ), high-frequency previous month displacement ( $D1^H$ ), and high-frequency accrued displacement of the previous month and the current month ( $D2^H$ ).

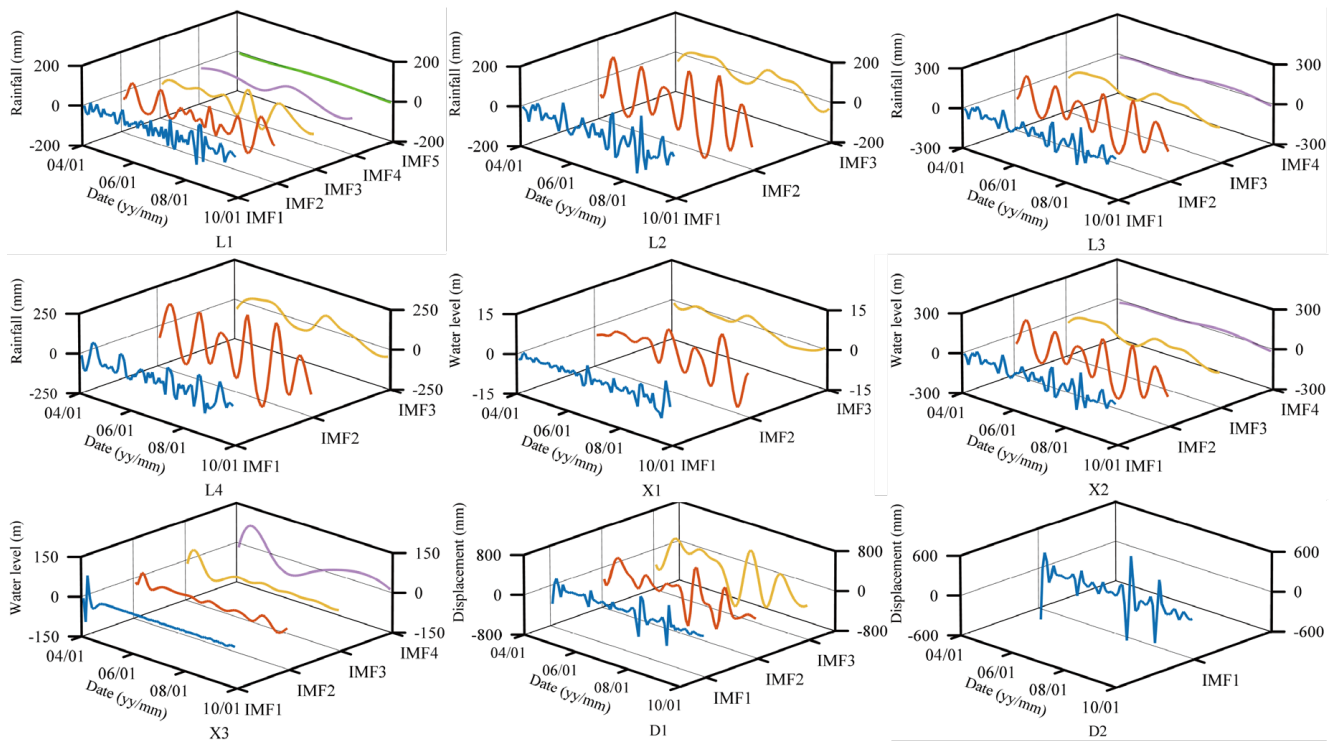
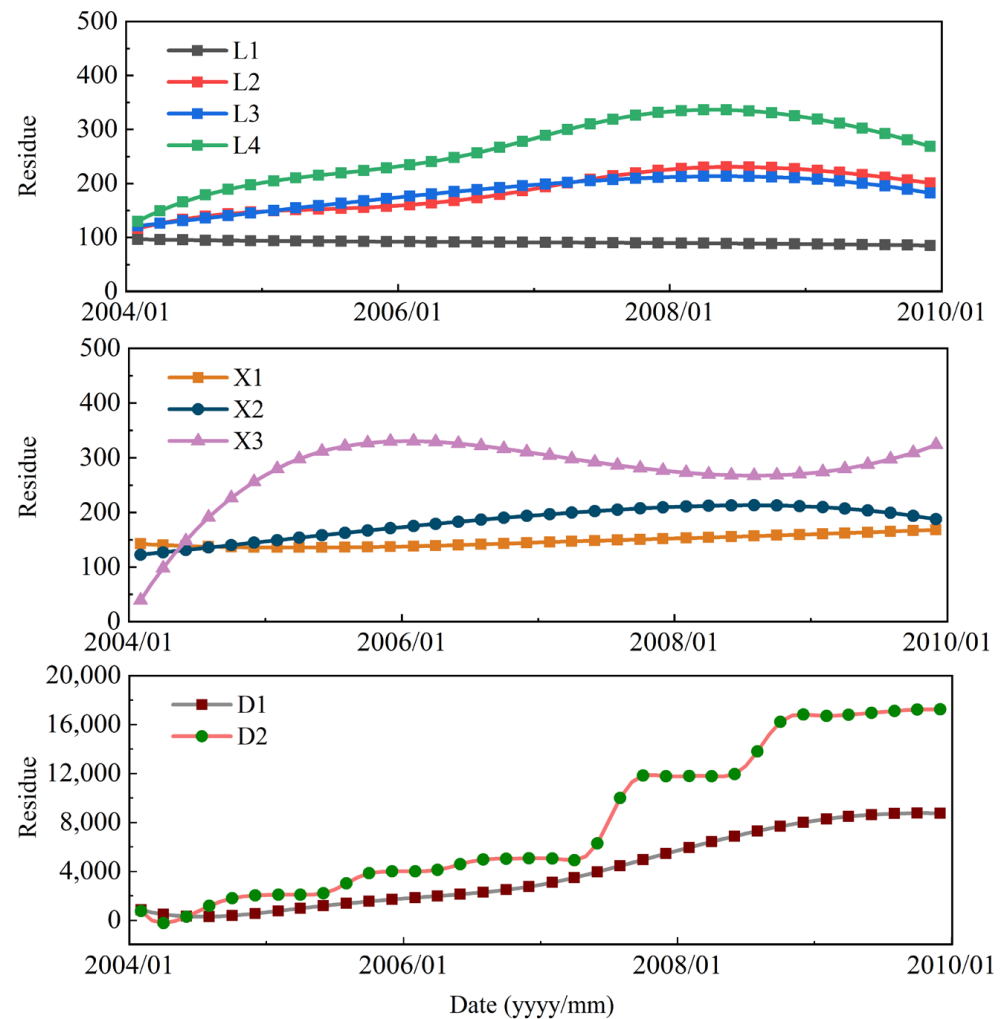


Figure 11. IMFs of restructured factors derived through CEEMD.

Table 1. Paired *t*-test results of all decomposed IMF.

Groups	Restructured Factor	Component	<i>t</i>	Sig.	Mean (mm)	Std. Deviation (mm)
Rainfall	L1	IMF <sub>2</sub>	0.22	0.83	1.50	56.92
		IMF <sub>3</sub>	−0.20	0.84	−1.23	50.75
		IMF <sub>4</sub>	0.60	0.55	3.02	42.24
		IMF <sub>5</sub>	0.10	0.92	0.41	36.49
	L2	IMF <sub>2</sub>	0.47	0.64	6.16	110.1
		IMF <sub>3</sub>	−1.34	0.18	−10.09	63.28
	L3	IMF <sub>2</sub>	0.23	0.82	3.05	111.1
		IMF <sub>3</sub>	−1.70	0.09	−12.42	61.56
L4	IMF <sub>2</sub>	0.38	0.70	5.43	120.1	
	IMF <sub>3</sub>	−0.61	0.54	−5.02	69.11	
Reservoir water level	X1	IMF <sub>2</sub>	0.47	0.64	0.26	4.73
		IMF <sub>3</sub>	2.07	0.04	0.66	2.70
	X2	IMF <sub>2</sub>	0.22	0.83	2.91	111.6
		IMF <sub>3</sub>	−1.58	0.12	−11.65	62.23
	X3	IMF <sub>4</sub>	−0.98	0.33	−6.13	52.64
		IMF <sub>2</sub>	−0.17	0.86	−0.37	18.16
	X3	IMF <sub>3</sub>	−0.52	0.61	−1.49	24.30
		IMF <sub>4</sub>	−2.19	0.03	−10.94	42.05
Displacement	D1	IMF <sub>2</sub>	−0.04	0.97	−1.16	229.4
		IMF <sub>3</sub>	−1.07	0.29	−40.82	320.5

The residue terms of restructured factors derived through CEEMD are shown in Figure 12. The results demonstrate that, except for L1, all the residue terms show a roughly increasing trend that is similar to the trend displacement term. This suggests that the residual terms roughly reflect the trend of the cumulative displacement, allowing the residue terms to be used as input parameters for the SVR to predict the displacement trend term.



**Figure 12.** Residue term of restructured factors derived through CEEMD.

#### 4.3. Factors Affecting Landslide Displacement Selected by EDR

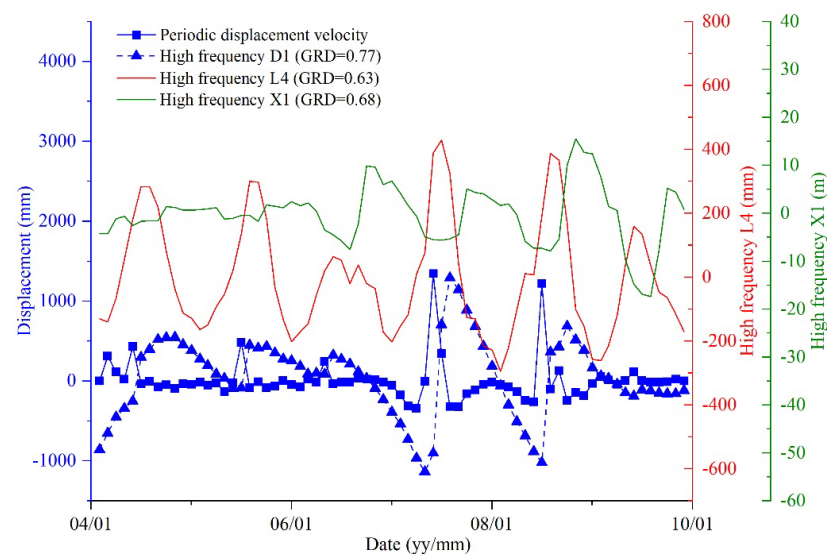
Previous analyzes demonstrated a strong association between the landslide displacement and the aforementioned factors. Thus, it is vital to determine which factors that have the greatest influence on landslide displacement. The EDR distance was determined between each factor and the displacements to determine the specific factors most closely related to the landslide's periodic displacement and trend displacement, respectively. This helps to identify the best factors to use the SVR model. The original restructured factors and their frequency components were chosen to compute the EDR distance with the periodic displacement. Simultaneously, the residue term for each factor and restructured factors were utilized to compute the EDR distance with the trend displacement. Normalization can be used to eliminate the influence of the numerical magnitude on analysis results due to the dimension difference between the displacement time series and the related factors. The calculated EDR distances are shown in Table 2.

**Table 2.** EDR distance between periodic displacements and related factors.

Groups	Component	Periodic Displacement			Trend Displacement
		Origin	High	Low	
Rainfall	L1	61	60	/	68
	L2	56	53	/	24
	L3	56	53	/	42
	L4	54	49	/	42
Reservoir level	X1	53	33	44	22
	X2	56	53	/	41
	X3	69	41	58	60
Displacement	D1	67	21	/	3
	D2	66	32	/	2

After dividing all the related factors into rainfall, reservoir water level, and displacement groups, the factors with a smaller EDR distance can be regarded as more interrelated with the landslide displacement component in each group. The results show that, for periodic displacement, the high-frequency accrued precipitation of the previous two and current months ( $L4^H$ ), the high-frequency current monthly reservoir level data ( $X1^H$ ), and the high-frequency previous month displacement ( $D1^H$ ) are the most relevant factors in each group. Thus, when predicting periodic displacement,  $L4^H$ ,  $X1^H$ , and  $D1^H$  are the input variables for the periodic displacement SVR model. Similarly, related factors for predicting trend displacement are the residual terms of L2, X1, and D2 according to the EDR results in each group, and these were chosen as the input parameters for the trend displacement SVR model.

To verify the effectiveness of the EDR method, grey relational analysis (GRA), a common method for selecting input variables in landslide displacement prediction, was used to calculate the grey relational degree (GRD) between the selected factors and the displacement component. The periodic displacement component is chosen as the research object, and the factor's GRD and periodic displacement velocity are shown and compared in Figure 13. The factors with a GRD value higher than 0.6 are regarded as closely interrelated with the periodic displacement. Therefore, the high-frequency accrued precipitation of the previous two and current months ( $L4^H$ ), the high-frequency current monthly reservoir level data ( $X1^H$ ), and the high-frequency previous month displacement ( $D1^H$ ) are the most relevant related factors in each group, which is consistent with the results selected by EDR.

**Figure 13.** Landslide periodic displacement compared with selected factors affecting landslide movement.



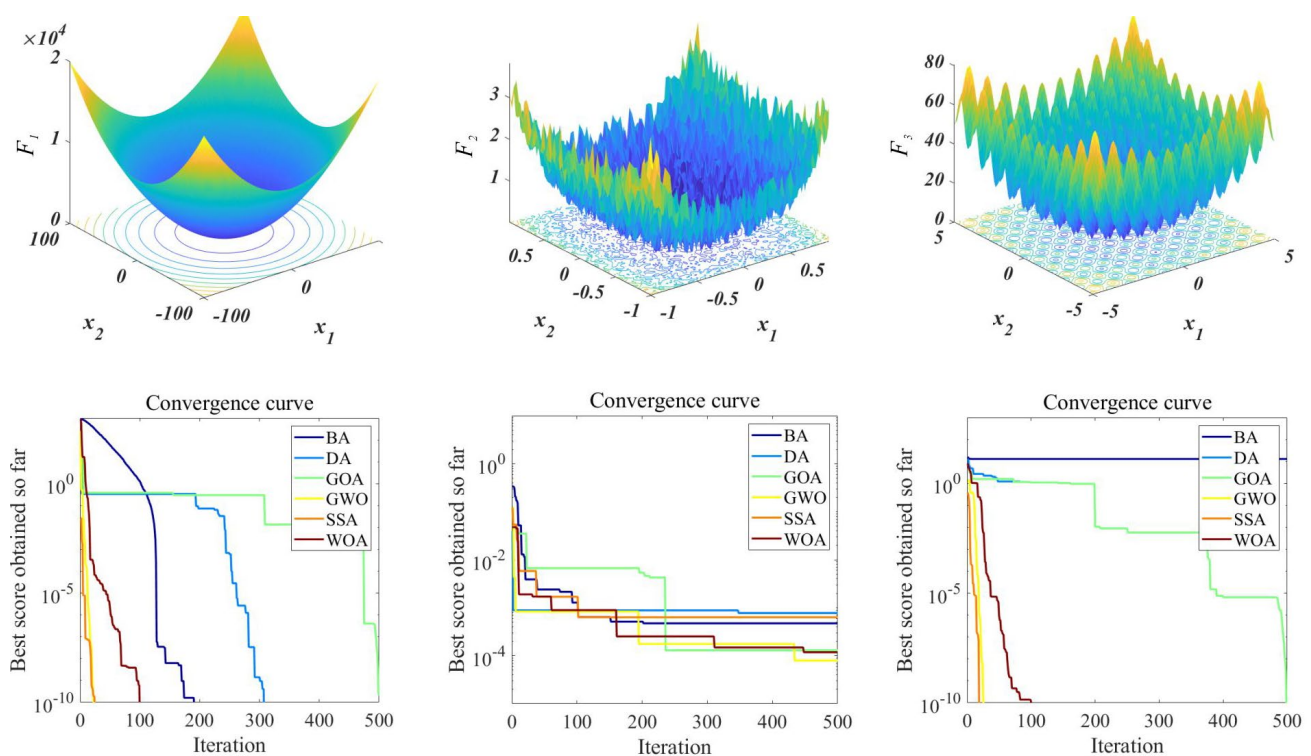
## 5. Prediction Results and Comparison

### 5.1. Parameter Optimization

For quantitatively measuring the optimization performance of the six SIs adopted in this study, three selected benchmark functions (Table 3) with different features are employed as test functions and results are shown in Figure 14. Different from  $F_2(x)$  and  $F_3(x)$ , the  $F_1(x)$  is smoother and has a unique extreme point in the solution space of  $x_1$  and  $x_2$ . The calculation results and process show that the slopes of the convergence curves of SSA and GWO are close, indicating that the convergence performance of the two is close and is the best among the six algorithms. The solutions obtained by each SI in  $F_1(x)$  and  $F_3(x)$  are relatively scattered, and some algorithms (such as BA) will fall into a local optimum.

**Table 3.** Three benchmark functions.

Function	Range	Theoretical Minimum Value
$F_1(x) = \sum_{i=1}^n x_i^2$	$x_i \in [-100, 100], i = 1, 2$	0
$F_2(x) = \sum_{i=1}^n ix_i^4 + \text{random}(0, 1)$	$x_i \in [-1.28, 1.28], i = 1, 2$	0
$F_3(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	$x_i \in [-5.12, 5.12], i = 1, 2$	0



**Figure 14.** Iterative curves of three benchmark functions solved using multiple-SI.

Determining the optimal value of the penalty factor  $C$  and the kernel function parameter  $g$  of the SVR model is a vital procedure dominating the accuracy of a displacement prediction. The parameters  $C$  and  $g$  in this study are optimized with MSI algorithms and are conducted independently for periodic and trend terms. For each MSI algorithm, the parameters  $C$  and  $g$  make a two-dimensional searching space. A population of simple agents communicate locally with each other and with their environment and move in specific patterns to search for the best result. The parameter settings and initial conditions in the MSI algorithm jointly affect the result. The parameter settings are iteratively adjusted and recalculated according to the optimal prediction effect. The results of the optimization

are shown in Table 4. The optimized C and g are later used in the SVR-based model to predict the periodic and trend displacements.

**Table 4.** Parameter and results of each optimization algorithm.

Algorithm	Parameters			Periodic		Trend	
				C	g	C	g
BA-SVR	Sizepop = 20	Max_iter. = 200	A = 0.2	220.67	0.00109	657.16	0.00106
	Lb = $1 \times 10^{-2}$	Ub = $1 \times 10^2$	r = 0.5				
	Freq_min = 0.1	Freq_min = 0.2	Alpha = 0.2				
DA-SVR	Sizepop = 30	Max_iter. = 200	e = f = 0.1	66506	0.00001	83702	0.00001
	lb = $1 \times 10^{-5}$	ub = $1 \times 10^5$	c = 0.7				
	w = 0.5	s = 0.1	a = 0.1				
GOA-SVR	Sizepop = 30	Max_iter. = 200	l = 1.5	16.13	0.00100	29.68	0.01000
	lb = $1 \times 10^{-3}$	ub = $1 \times 10^3$	f = 0.5				
GWO-SVR	Sizepop = 30	Max_iter. = 200	dim = 2	474.94	0.00100	706.29	0.00100
	lb = $1 \times 10^{-3}$	ub = $1 \times 10^3$	/				
SSA-SVR	Sizepop = 30	Max_iter. = 200	pNum = 20%	16.17	0.00100	9677.9	0.00014
	lb = $1 \times 10^{-4}$	ub = $1 \times 10^4$	sNum = 20%				
OA-SVR	Sizepop = 20	Max_iter. = 200	dim = 2	1.74	0.01000	48277.4	0.00001
	lb = $1 \times 10^{-5}$	ub = $1 \times 10^5$	b = 1				

### 5.2. Prediction of Periodic and Trend Displacements

An MSI-based SVR prediction model was developed with the optimized input factors to predict the periodic displacements and the trend displacements separately, as shown in Figure 15. The prediction accuracy and error of each model are shown and compared in Figure 16. For the periodic displacements, the prediction accuracy with the largest  $R^2$  and smallest MAPE, RMSE, and MAE was obtained using the DA algorithm among all of the given models. The corresponding result of MAPE, RMSE, MAE, and  $R^2$  is 3.654173, 63.0435, 119.2786, 0.824217, respectively. Meanwhile, the GWO-based SVR model gave the best prediction for the trend displacements compared to the other optimization algorithms, with the result of MAPE, RMSE, MAE, and  $R^2$  being 0.010273, 95.9178, 184.4194, and 0.99473, respectively. Overall, the prediction results provided by the SVR model optimized by MSI matched well with the observation results.

### 5.3. Prediction of Cumulative Displacements

The predicted cumulative displacements of the Shiliushubao landslide can be obtained by adding the predicted periodic and trend displacements. The predicted cumulative displacements are shown in Figure 17, and these are in good agreement with the observed displacements. The maximum relative error of monthly displacement is generally less than 3% and the average relative error of less than 1%. The results show the usefulness of the proposed model. The most appropriate optimization algorithm and the most relevant landslide related factors were selected and applied.

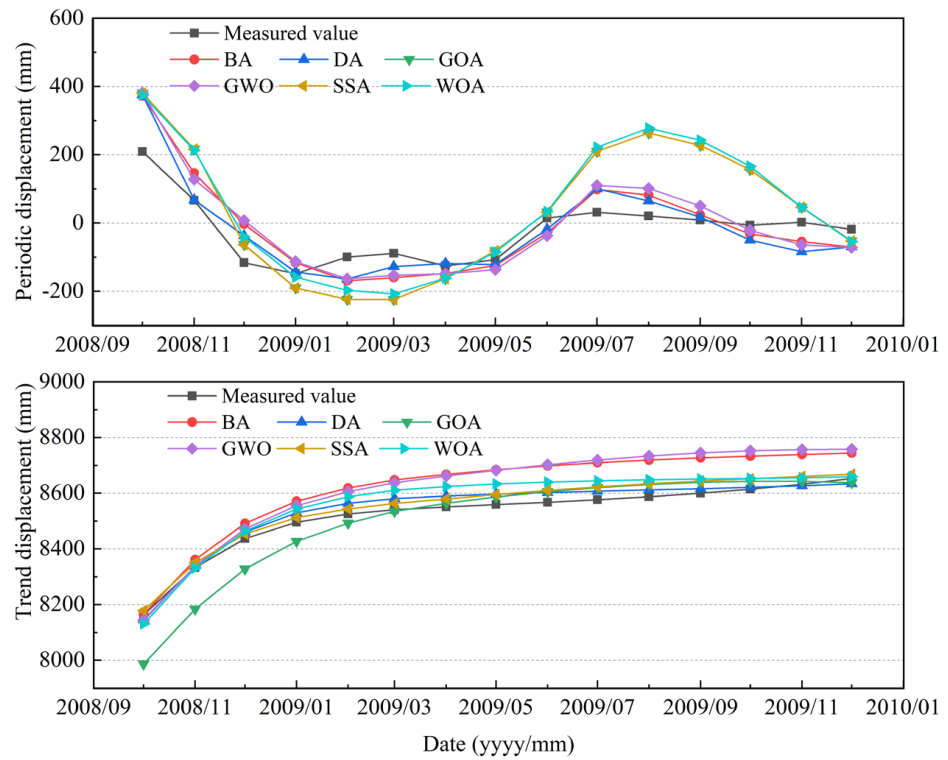


Figure 15. Comparison of prediction results by MSI-SVR model with monitoring data.

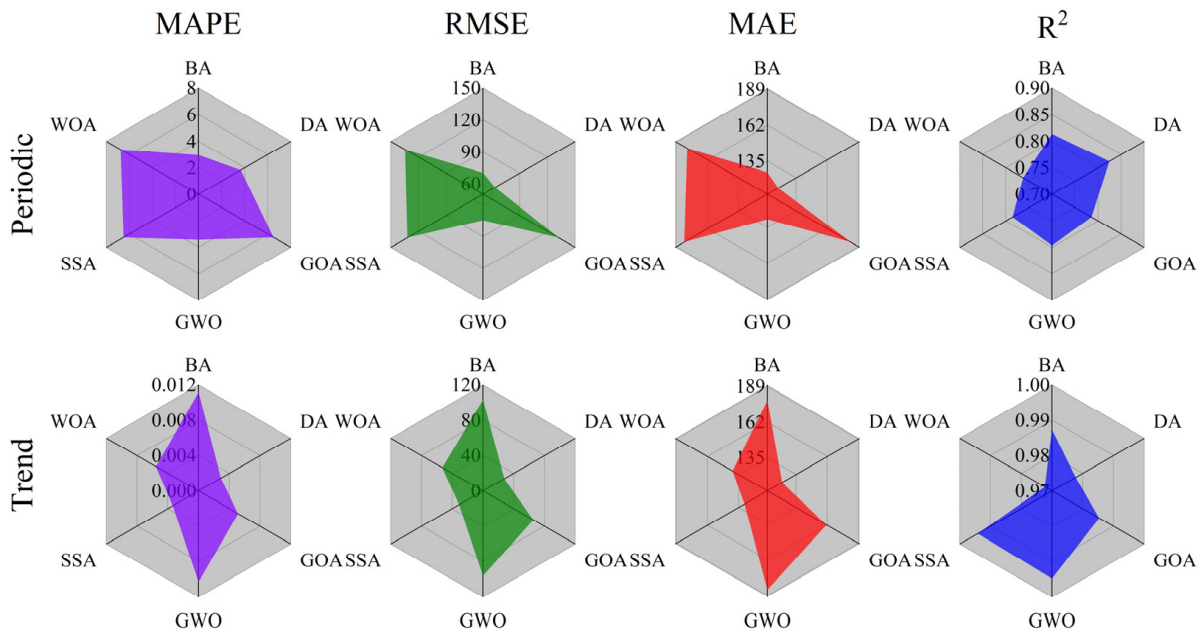


Figure 16. Rose diagram for each model's performance.

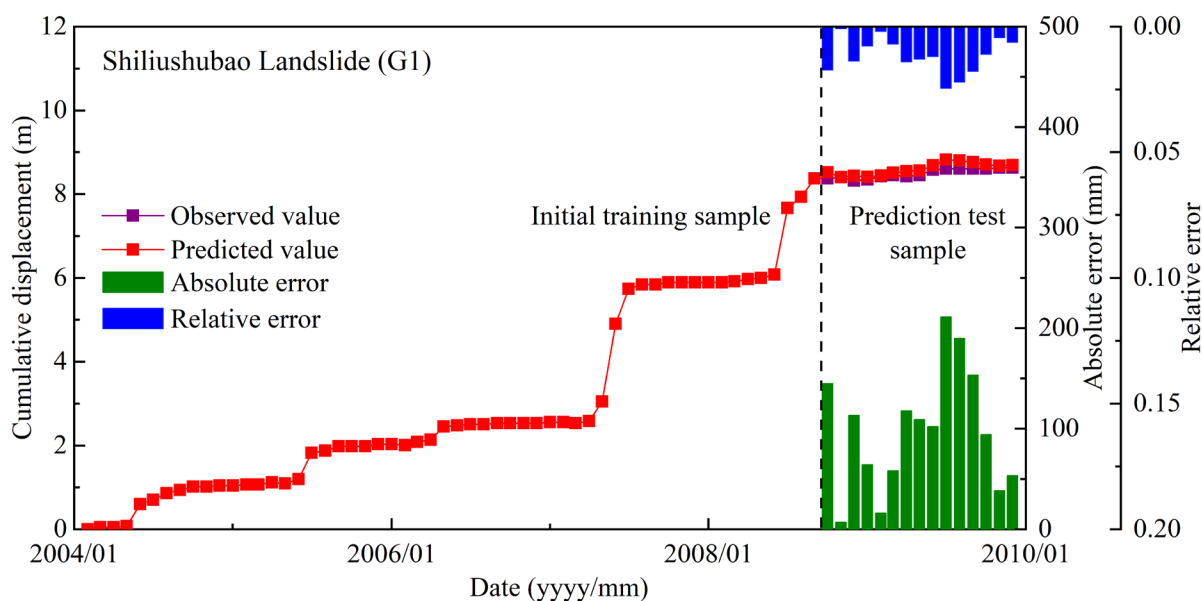


Figure 17. Comparison of predicted displacement and observed displacement.

To further verify the effectiveness of the proposed prediction model, the displacement at ZG93 of the well-known Baishuihe landslide is chosen as another case and predicted. The prediction accuracy of each SI and cumulative displacement prediction result are shown and compared in Table 5 and Figure 18.

Table 5. Prediction accuracy of each SI in Baishuihe landslide.

Optimization Algorithm	Periodic Displacement				Trend Displacement			
	MAPE	RMSE	MAE	R <sup>2</sup>	MAPE	RMSE	MAE	R <sup>2</sup>
BA	0.688	13.691	30.118	0.757	0.395	1065.132	926.683	0.8621
DA	0.788	13.652	30.367	0.761	0.008	20.448	66.336	0.9997
GOA	0.692	13.663	30.110	0.758	0.008	19.649	66.214	0.9997
GWO	0.680	13.592	29.558	0.751	0.008	20.448	66.336	0.9997
SSA	0.786	13.589	30.307	0.762	0.009	22.766	64.733	0.9998
WOA	0.788	13.629	30.329	0.761	0.008	20.448	66.336	0.9997

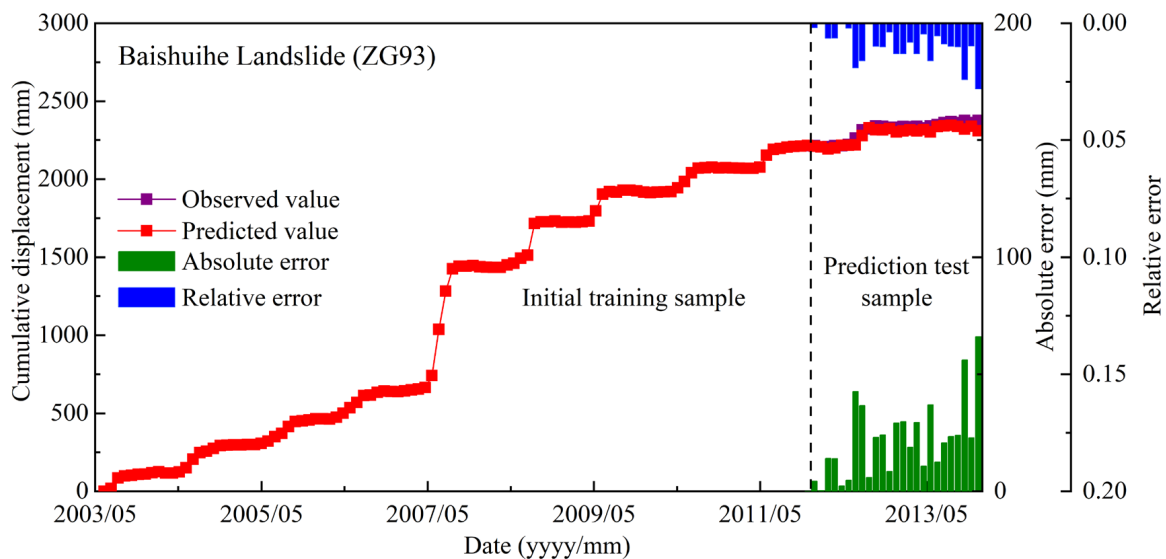


Figure 18. Comparison of predicted displacement and measured displacement.

The SSA method has achieved the best results in predicting both periodic displacement and trend displacement, with the largest value of  $R^2$ , which is 0.762 and 0.9998, respectively. The cumulative displacement prediction results are in good agreement with the measured displacement, with an absolute error of monthly displacement that is generally less than 67mm and the maximum relative error of monthly displacement that less than 3%. The average relative error of the proposed prediction model is 0.898%, which is slightly smaller than the result obtained by the prediction model of Deng et al. [54]. The comparative study shows the effective improvement of the proposed model in terms of prediction performance and the universality of it to predict the displacement of slow-moving landslides all around the world.

## 6. Discussion

This paper aims to improve the accuracy of landslide displacement prediction by constructing a novel prediction model combined with the CEEMD method, EDR method, and multi-swarm-intelligence (MSI) algorithm. The new prediction model can forecast landslide movements so that the landslide status can be evaluated, and appropriate stabilization measures can be implemented in advance to reduce the destructive effects of landslide movements. The CEEMD method was first employed for the landslide displacement decomposition, and a new prediction based on this was proposed to overcome its defect by optimizing the model's framework. The trend displacement obtained from CEEMD decomposition can reflect the long-term trend of landslide deformation. The periodic displacement obtained from CEEMD decomposition shows a cyclical variation in displacements consistent with periodic changes of related factors such as rainfall and reservoir levels. The frequency components of related factors that change periodically can be decomposed by the CEEMD. Combining with the  $t$ -test, the high-frequency and low-frequency components of related factors can be separated. With the EDR method, the most relevant factors related to the landslide displacements among the original related factors, reconstruction related factors, and frequency related factors can be selected by calculating the distance between all related factors and the extracted displacement component. The relevant factors that were identified are consistent with the results obtained by GRA.

The factors related to landslide displacement prediction can be separated into three groups: rainfall, reservoir level, and previous displacement. The most relevant factors for the Shiliushubao landslide's periodic displacement are  $L4^H$ ,  $X1^H$ , and  $D1^H$ , and the most relevant factors for the trend displacements are the residual terms of  $L2$ ,  $X1$ , and  $D2$ . MSI (BA, DA, GOA, GWO, SSA, and WOA) was used to optimize the proposed prediction model. For the Shiliushubao landslide, the DA-based SVR model performs best to predict periodic displacements, and the GWO-based SVR model works best for predicting trend displacements. The prediction of cumulative displacements is in good agreement with the measured displacements with a maximum relative error of monthly displacement of less than 3%. The trail of the proposed model on the Baishuihe landslide, another landslide in the reservoir area, is also satisfied with the average relative error of 0.898%, which performs slightly better than that from the previous study.

While the proposed methodology yielded satisfactory results, there are also some limitations. First, the CEEMD method has limits in the decomposition of measured displacements and related factors when the time series does not have enough extreme points, which limits the applicability of this method. When there is only one IMF sequence after CEEMD decomposition, a  $t$ -test cannot be carried out, and the IMF itself is high frequency. Second, the trend displacement and residue term for related factors after CEEMD decomposition may still have local fluctuations. It might contain some periodic fluctuation information, which can lead to prediction error, which needs to be further studied in the future. Third, the values of  $thrsh$ ,  $sthresh$ ,  $N$ , and  $\alpha$  used in CEEMD will have an indirect impact on the prediction results. The appropriate range of these parameters and their influence on the results are still unclear. Fourth, when using MSI to optimize the parameters of the SVR model, the search for the  $g$  value is usually close to the lowest value of the search interval,



and the result does not gradually increase as the search boundary continues to widen. Different SI optimization algorithms may perform differently for different landslides, so for new landslide data, the trial of different optimization algorithms for the best results is needed.

The deformation and failure of landslides are usually closely bonded with the groundwater effect [55]. The evaluation of the landslide stability with groundwater nowadays has developed into several hotspot branches, which includes analytical methods, such as the Limit Equilibrium Analysis with the Reliability Analysis and the Intelligent Algorithms on sliding zone searching, and numerical methods, such as the Fast Lagrange Analysis, the Finite Element Method, and the Discrete Element Method coupled with hydraulic calculations. These advanced evaluation methods have their status in the practical industry on the slope stability and deformation assessment, based on the current state and data gathered in the field and laboratory; however, these mechanism-based methods took insufficient account of the history state and data of the slope. The novel prediction model proposed in this paper can consider the historical influence of rainfall and reservoir fluctuation that precisely related to the displacement periodic component and displacement trend component with the help of the CEEMD method and EDR, thus improving the accuracy of landslide displacement prediction. It is a profitable attempt and a good way to improve the accuracy of landslide movement prediction. Although some in-depth research in consideration of historical factors of inducing factors has been carried out in this study, the predictive capability of the proposed model is still flawed in the sense that they cannot say anything about changes that are caused by external factors not captured by the available data series. Therefore, it is very important to develop multi-field (displacement field, seepage field, stress field, etc.) monitoring technology for the landslides, and the innovative prediction models based on this can more reflect the evolution process of the sliding mass.

Landslides in complex water environments could develop different deformation patterns, both categorized by history data and potential failure mechanism [4]. The pattern is highly related to the interaction between soil and water in a certain engineering geological condition. In the proposed novel displacement prediction model, the interaction mechanism is still not included, which limits the adaptability and comparability among different landslide cases. A better insight into the landslide development patterns is to be developed, combining the failure-mechanism-based evaluation method, in the future model for displacement prediction. Other than from the pure displacement prediction based on displacement, rainfall, and water level data sequence, an evaluation of the critical rainfall intensity and critical water level fluctuation rate is needed to be conducted under certain landslide development patterns in the further study.

In addition, landslide displacement is a noisy and non-stationary process that varies with time, which is highly affected by internal factors such as formation lithology and geological structure and external factors such as the rainfall, reservoir water level, and snow melting. Due to the complex nonlinear relationship between all these various inducing factors and landslide displacement, the landslide displacement prediction is subject to considerable uncertainties [56]. The limitations of the machine learning model, parameter selection, and data noise will increase the uncertainty of prediction [57]. The prediction model proposed in this paper is a deterministic point prediction model which cannot estimate the variability and uncertainty related to a given landslide displacement prediction, which limits its reliability under uncertain conditions. This should be addressed in the future study.

## 7. Conclusions

A reservoir landslide's movement is closely associated with the related factors including reservoir level fluctuations, rainfall intensity, and previous deformations. The complex nonlinear relationship between all these various inducing factors and landslide displacement increased the challenge of forecasting in the form of considerable uncertainties. In this study, a novel prediction model for landslide displacement prediction was proposed

to improve the accuracy by the combination of multiple algorithms. The EDR method can identify the most relevant factors influencing a landslide's movements to use as input variables for an SVR model. The CEEMD method is suitable for the decomposition of various time series and can be used to extract the trend displacement of slow-moving landslide displacement. The CEEMD method can also highlight local fluctuations in the time series of related factors, and the frequency components of these time series can be extracted by combining the *t*-test method. With the help of MSI optimization algorithms, the optimal value of the penalty factor *C* and the kernel function parameter *g* for an SVR model can be obtained. This paper proposes an SVR model based on the CEEMD method, EDR selection, and MSI optimization algorithm that can capture the deformation characteristics of the landslide before failure.

Measurements of landslide displacements for the Shiliushubao landslide in the TGRA were used to demonstrate the novel displacement prediction model. The predicted displacements, including season fluctuations and the long-term trend, were found to be consistent with the observed data, which indicates that the proposed model has good predictive performance, even when the displacement characteristics are cyclic and complex. The DA- and GWO-based SVR model provided the best prediction of periodic displacement and trend displacement, respectively. The prediction model proposed in this paper has wider applicability. It can enhance the prediction of landslide displacements characterized by slow-moving, step-like displacements that are influenced by multiple related factors with frequency conversion characteristics.

**Author Contributions:** Conceptualization, J.Z.; methodology, J.Z.; supervision, H.T.; writing—original draft preparation, J.Z. and D.D.T.; funding acquisition, H.T.; resources, C.L.; software, C.L.; investigation, Y.W. and D.X.; visualization, D.X.; writing—review and editing, Q.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Major Program of the National Natural Science Foundation of China (No. 42090055), the National Key Research and Development Program of China (No. 2017YFC1501305), the National Major Scientific Instruments and Equipment Development Projects of China (No. 41827808), the National Natural Sciences Foundation of China under Grant Nos. 42177147, 41807263 and China Postdoctoral Science Foundation (Grant No. 2021M703002).

**Acknowledgments:** The authors gratefully appreciate the editor and anonymous reviewers for their constructive criticism and comments on the earlier version of this paper and for offering valuable suggestions that have helped to improve this paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Tang, H.; Wasowski, J.; Juang, C.H. Geohazards in the three Gorges Reservoir Area, China—Lessons learned from decades of research. *Eng. Geol.* **2019**, *261*, 105267. [[CrossRef](#)]
2. Niu, M.; Wang, Y.; Sun, S.; Li, Y. A novel hybrid decomposition-and-ensemble model based on CEEMD and GWO for short-term PM2.5 concentration forecasting. *Atmos. Environ.* **2016**, *134*, 168–180. [[CrossRef](#)]
3. Wang, J.; Schweizer, D.; Liu, Q.; Su, A.; Hu, X.; Blum, P. Three-dimensional landslide evolution model at the Yangtze River. *Eng. Geol.* **2021**, *292*, 106275. [[CrossRef](#)]
4. Wang, Y.; Tang, H.; Wen, T.; Ma, J. A hybrid intelligent approach for constructing landslide displacement prediction intervals. *Appl. Soft Comput.* **2019**, *81*, 105506. [[CrossRef](#)]
5. Zhang, Y.G.; Tang, J.; He, Z.Y.; Tan, J.; Li, C. A novel displacement prediction method using gated recurrent unit model with time series analysis in the Erdaohe landslide. *Nat. Hazards* **2020**, *105*, 783–813. [[CrossRef](#)]
6. Zhang, J.; Tang, H.; Tannant, D.D.; Lin, C.; Xia, D.; Liu, X.; Zhang, Y.; Ma, J. Combined forecasting model with CEEMD-LCSS reconstruction and the ABC-SVR method for landslide displacement prediction. *J. Clean. Prod.* **2021**, *293*, 126205. [[CrossRef](#)]
7. Yin, Y.P.; Huang, B.L.; Wang, W.P.; Wei, Y.J.; Ma, X.H.; Ma, F.; Zhao, C.J. Reservoir-induced landslides and risk control in Three Gorges Project on Yangtze River, China. *J. Rock Mech. Geotech. Eng.* **2016**, *8*, 577–595. [[CrossRef](#)]
8. Zhou, C.; Yin, K.; Cao, Y.; Intrieri, E.; Ahmed, B.; Catani, F. Displacement prediction of step-like landslide by applying a novel kernel extreme learning machine method. *Landslides* **2018**, *15*, 2211–2225. [[CrossRef](#)]
9. Ma, J.; Niu, X.; Tang, H.; Wang, Y.; Wen, T.; Zhang, J. Displacement Prediction of a Complex Landslide in the Three Gorges Reservoir Area (China) Using a Hybrid Computational Intelligence Approach. *Complexity* **2020**, *2020*, 2624547. [[CrossRef](#)]

10. Saito, M. Forecasting the time of occurrence of a slope failure. In Proceedings of the 6th International Congress on Soil Mechanics and Foundation Engineering, Montreal, QC, Canada, 8–15 September 1965; pp. 537–541.
11. Crosta, G.B.; Agliardi, F. Failure forecast for large rock slides by surface displacement measurements. *Can. Geotech. J.* **2003**, *40*, 176–191. [[CrossRef](#)]
12. Ma, J.W.; Tang, H.M.; Liu, X.; Hu, X.L.; Sun, M.J.; Song, Y.J. Establishment of a deformation forecasting model for a step-like landslide based on decision tree C5.0 and two-step cluster algorithms: A case study in the Three Gorges Reservoir area, China. *Landslides* **2017**, *14*, 1275–1281. [[CrossRef](#)]
13. Zou, Z.; Yang, Y.; Fan, Z.; Tang, H.; Zou, M.; Hu, X.; Xiong, C.; Ma, J. Suitability of data preprocessing methods for landslide displacement forecasting. *Stoch. Environ. Res. Risk Assess.* **2020**, *34*, 1105–1119. [[CrossRef](#)]
14. Ma, J.; Liu, X.; Niu, X.; Wang, Y.; Wen, T.; Zhang, J.; Zou, Z. Forecasting of landslide displacement using a probability-scheme combination ensemble prediction technique. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4788. [[CrossRef](#)]
15. Tharwat, A.; Gabel, T. Parameters optimization of support vector machines for imbalanced data using social ski driver algorithm. *Neural. Comput. Appl.* **2019**, *32*, 6925–6938. [[CrossRef](#)]
16. Cai, Z.; Xu, W.; Meng, Y.; Shi, C.; Wang, R. Prediction of landslide displacement based on GA-LSSVM with multiple factors. *Bull. Eng. Geol. Environ.* **2015**, *75*, 637–646. [[CrossRef](#)]
17. Zhou, C.; Yin, K.; Cao, Y.; Ahmed, B. Application of time series analysis and PSO-SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir, China. *Eng. Geol.* **2016**, *204*, 108–120. [[CrossRef](#)]
18. Zhang, J.; Yin, K.; Wang, J.; Huang, F. Displacement prediction of Baishuihe landslide based on time series and PSO-SVR model. *Chin. J. Rock Mech. Eng.* **2015**, *34*, 382–391.
19. Peng, L.; Niu, R.; Wu, T. Time series analysis and support vector machine for landslide displacement prediction. *J. Zhejiang Univ. (Eng. Sci.)* **2013**, *47*, 1672–1679.
20. Zhou, C.; Yin, K.; Cao, Y.; Ahmed, B.; Fu, X. A novel method for landslide displacement prediction by integrating advanced computational intelligence algorithms. *Sci. Rep.* **2018**, *8*, 7287. [[CrossRef](#)]
21. Ding, L.; Lv, J.; Li, X.; Li, L. Support vector regression and ant colony optimization for HVAC cooling load prediction. In Proceedings of the 2010 International Symposium on Computer, Communication, Control and Automation (3CA), IEEE, Tainan, Taiwan, 5–7 May 2010; Volume 1, pp. 537–541.
22. Balogun, A.L.; Rezaie, F.; Pham, Q.B.; Gigović, L.; Drobnjak, S.; Aina, Y.A.; Panahi, M.; Yekeen, S.T.; Lee, S. Spatial prediction of landslide susceptibility in western Serbia using hybrid support vector regression (SVR) with GWO, BAT and COA algorithms. *Geosci. Front.* **2021**, *12*, 101104. [[CrossRef](#)]
23. Li, L.W.; Wu, Y.P.; Miao, F.S.; Liao, K.; Zhang, F.L. Displacement prediction of landslides based on variational mode decomposition and GWO-MIC-SVR model. *Chin. J. Rock Mech. Eng.* **2018**, *37*, 1395–1406. (In Chinese)
24. Miao, F.; Wu, Y.; Xie, Y.; Li, Y. Prediction of landslide displacement with step-like behavior based on multialgorithm optimization and a support vector regression model. *Landslides* **2018**, *15*, 475–488. [[CrossRef](#)]
25. Zhang, J.; Tang, H.; Wen, T.; Ma, J.; Tan, Q.; Xia, D.; Liu, X.; Zhang, Y. A Hybrid Landslide Displacement Prediction Method Based on CEEMD and DTW-ACO-SVR—Cases Studied in the Three Gorges Reservoir Area. *Sensors* **2020**, *20*, 4287. [[CrossRef](#)]
26. Zhang, Y.; Agarwal, P.; Bhatnagar, V.; Balochian, S.; Yan, J. Swarm Intelligence and Its Applications. *Sci. World J.* **2013**, *2013*, 528069. [[CrossRef](#)]
27. Chen, W.; Tsangaratos, P.; Ilia, I.; Duan, Z.; Chen, X. Groundwater spring potential mapping using population-based evolutionary algorithms and data mining methods. *Sci. Total Environ.* **2019**, *684*, 31–49. [[CrossRef](#)]
28. Kışi, Ö. Streamflow Forecasting Using Different Artificial Neural Network Algorithms. *J. Hydrol. Eng.* **2007**, *12*, 532–539. [[CrossRef](#)]
29. Beni, G. From Swarm Intelligence to Swarm Robotics. In *Swarm Robotics*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 1–9.
30. Liu, Y.; Wang, R. Study on network traffic forecast model of SVR optimized by GAFSA. *Chaos Solitons Fractals* **2016**, *89*, 153–159. [[CrossRef](#)]
31. Ali, E.S.; Abd Elazim, S.M.; Abdelaziz, A.Y. Ant Lion Optimization Algorithm for Renewable Distributed Generations. *Energy* **2016**, *116*, 445–458. [[CrossRef](#)]
32. Jiang, H.; Yang, Y.; Ping, W.; Dong, Y. A Novel Hybrid Classification Method Based on the Opposition-Based Seagull Optimization Algorithm. *IEEE Access* **2020**, *8*, 100778–100790. [[CrossRef](#)]
33. Yang, X.S.; He, X. Bat algorithm: Literature review and applications. *Int. J. Bio-Inspir. Comput.* **2013**, *5*, 141. [[CrossRef](#)]
34. Emary, E.; Zawbaa, H.M.; Hassanien, A.E. Binary grey wolf optimization approaches for feature selection. *Neurocomputing* **2016**, *172*, 371–381. [[CrossRef](#)]
35. Mirjalili, S. Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Comput. Appl.* **2015**, *27*, 1053–1073. [[CrossRef](#)]
36. Mirjalili, S.; Lewis, A. The Whale Optimization Algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [[CrossRef](#)]
37. Mirjalili, S.Z.; Mirjalili, S.; Saremi, S.; Faris, H.; Aljarah, I. Grasshopper optimization algorithm for multi-objective optimization problems. *Appl. Intell.* **2017**, *48*, 805–820. [[CrossRef](#)]
38. Xue, J.; Shen, B. A novel swarm intelligence optimization approach: Sparrow search algorithm. *Syst. Sci. Control* **2020**, *8*, 22–34. [[CrossRef](#)]

39. Du, H.; Song, D.; Chen, Z.; Shu, H.; Guo, Z. Prediction model oriented for landslide displacement with step-like curve by applying ensemble empirical mode decomposition and the PSO-ELM method. *J. Clean. Prod.* **2020**, *270*, 122248. [[CrossRef](#)]
40. Li, Y.; Sun, R.; Yin, K.; Xu, Y.; Chai, B.; Xiao, L. Forecasting of landslide displacements using a chaos theory based wavelet analysis-Volterra filter model. *Sci. Rep.* **2019**, *9*, 19853. [[CrossRef](#)]
41. Xu, S.; Niu, R. Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China. *Comput. Geosci.* **2018**, *111*, 87–96. [[CrossRef](#)]
42. Ren, F.; Wu, X.; Zhang, K.; Niu, R. Application of wavelet analysis and a particle swarm-optimized support vector machine to predict the displacement of the Shuping landslide in the Three Gorges, China. *Environ. Earth Sci.* **2014**, *73*, 4791–4804. [[CrossRef](#)]
43. Yang, B.; Yin, K.; Lacasse, S.; Liu, Z. Time series analysis and long short-term memory neural network to predict landslide displacement. *Landslides* **2019**, *16*, 677–694. [[CrossRef](#)]
44. Huang, H.F.; Wu, Y.I.; Yi-Liang, L. Study on variables selection using SVR-MIV method in displacement prediction of landslides. *Chin. J. Undergr. Space Eng.* **2016**, *12*, 213–219. (In Chinese)
45. Chen, L.; Özsu, M.T.; Oria, V. Robust and fast similarity search for moving object trajectories. In *SIGMOD '05, Proceedings of the 24th ACM International Conference on Management of Data, New York, NY, USA, 13–15 June 2005*; ACM Press: New York, NY, USA, 2005; pp. 491–502. [[CrossRef](#)]
46. Mai, S.T.; Goebel, S.; Plant, C. A Similarity Model and Segmentation Algorithm for White Matter Fiber Tracts. In *Proceedings of the 2012 IEEE 12th International Conference on Data Mining, IEEE, Washington, DC, USA, 10–13 December 2012*.
47. Xu, Y.; Zhang, M.; Zhu, Q.; He, Y. An improved multi-kernel RVM integrated with CEEMD for high-quality intervals prediction construction and its intelligent modeling application. *Chemometr. Intell. Lab. Syst.* **2017**, *171*, 151–160. [[CrossRef](#)]
48. Ranacher, P.; Tzavella, K. How to compare movement? A review of physical movement similarity measures in geographic information science and beyond. *Cartogr. Geogr. Inf. Sci.* **2014**, *41*, 286–307. [[CrossRef](#)]
49. Moayedi, A.; Abbaspour, R.A.; Chehregan, A. An evaluation of the efficiency of similarity functions in density-based clustering of spatial trajectories. *Ann. GIS* **2019**, *25*, 313–327. [[CrossRef](#)]
50. Liu, H.; Mi, X.; Li, Y.; Duan, Z.; Xu, Y. Smart wind speed deep learning based multi-step forecasting model using singular spectrum analysis, convolutional Gated Recurrent Unit network and Support Vector Regression. *Renew. Energy* **2019**, *143*, 842–854. [[CrossRef](#)]
51. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey Wolf Optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [[CrossRef](#)]
52. Li, Z.; Xie, Y.; Li, X.; Zhao, W. Prediction and application of porosity based on support vector regression model optimized by adaptive dragonfly algorithm. *Energy Sources Part A Recovery Util. Environ. Eff.* **2019**, *43*, 1073–1086. [[CrossRef](#)]
53. Barman, M.; Dev Choudhury, N.B. Hybrid GOA-SVR technique for short term load forecasting during periods with substantial weather changes in North-East India. *Procedia Comput. Sci.* **2018**, *143*, 124–132. [[CrossRef](#)]
54. Deng, D.; Liang, Y.; Wang, L.; Wang, C.-S.; Sun, Z.-H.; Wang, C.; Dong, M.-M. Displacement prediction method based on ensemble empirical mode decomposition and support vector machine regression—A case of landslides in Three Gorges Reservoir area. *Rock Soil Mech.* **2017**, *38*, 3660–3669.
55. Hongtao, N. Smart safety early warning model of landslide geological hazard based on BP neural network. *Saf. Sci.* **2020**, *123*, 104572. [[CrossRef](#)]
56. Adnan, M.S.G.; Rahman, M.S.; Ahmed, N.; Ahmed, B.; Rabbi, M.F.; Rahman, R.M. Improving Spatial Agreement in Machine Learning-Based Landslide Susceptibility Mapping. *Remote. Sens.* **2020**, *12*, 3347. [[CrossRef](#)]
57. Ahmad, H.; Ningsheng, C.; Rahman, M.; Islam, M.M.; Pourghasemi, H.R.; Hussain, S.F.; Habumugisha, J.M.; Liu, E.; Zheng, H.; Ni, H.; et al. Geohazards Susceptibility Assessment along the Upper Indus Basin Using Four Machine Learning and Statistical Models. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 315. [[CrossRef](#)]