Adaptive Hybrid Machine Learning Model For Forecasting The Step-Like Displacements of Reservoir Colluvial Landslides: A Case Study in The Three Gorges Reservoir Area, China

Li Linwei
China University of Geosciences
Yiping Wu (ypwu@cug.edu.cn)
China University of Geosciences
Miao Fasheng
China University of Geosciences
Xue Yang
China University of Geosciences
Huang Yepiao
PowerChina: Power Construction Corporation of China

Research Article

Keywords: Reservoir colluvial landslide, Step-like displacement, Landslide displacement prediction, Time series decomposition, Machine learning, Three Gorges Reservoir area

Posted Date: February 24th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-217782/v1

Adaptive hybrid machine learning model for forecasting the step-like displacements of reservoir colluvial landslides: A case study in the Three Gorges Reservoir area, China

Authors: Li Linwei[1], Wu Yiping[1], Miao Fasheng[1], Xue Yang[1], Huang Yepiao[2]

Author affiliations:

1. Faculty of Engineering, China University of Geosciences, Wuhan 430074, China

2. Guiyang Engineering Corporation Limited of Power China, Guiyang 550081, China

Correspondence information:

Name: Wu Yiping**.**

Affiliation: Faculty of Engineering, China University of Geosciences**.**

Address: No. 388 Lumo Road, Wuhan, PR China**.**

Email: ypwu@cug.edu.cn**.**

Acknowledgments：

The study is supported by the National Natural Science Foundation of China (No. 41977244) and the National Key R&D Program of China (2017YFC1501301).

Adaptive hybrid machine learning model for forecasting the step-like displacements of reservoir colluvial landslides: A case study in the Three Gorges Reservoir area, China

Abstract: Constructing an accurate and stable displacement prediction model is essential to build a capable early warning system for landslide disasters. To overcome the drawbacks of previous displacement prediction models for step-like landslides, such as the incomplete or excessive decompositions of cumulative displacements and input factors and the redundancy or lack of input factors, we propose an adaptive hybrid machine learning model. This model is composed of three parts. First, candidate factors are proposed based on the macroscopic deformation response of landslides. Then, the landslide displacement and its candidate factors are adaptively decomposed into different displacement and factor components by applying optimized variational mode decomposition (OVMD). Second, in the gray wolf optimizer-based kernel extreme learning machine (GWO-KELM) model, the global sensitivity analysis (GSA) of the prediction results of different displacement components to each decomposed factor is analyzed based on the PAWN method. Then, the decomposed factors are reduced according to the GSA results. Third, based on the reduced factors, the optimal GWO-KELM models of the different displacement components are established to predict the displacement. Taking the Baishuihe landslide as an example, we used the raw data of three representative monitoring sites from June 2006 to December 2016 to verify the validity, accuracy, and stability of the model. The results indicate that the proposed hybrid model can effectively determine the displacement decomposition parameters. In addition, this model performed well over a three-year forecast with low model complexity.

Keywords: Reservoir colluvial landslide; Step-like displacement; Landslide displacement prediction; Time series decomposition; Machine learning; Three Gorges Reservoir area

1. Introduction

Landslides, as one of the most widely distributed and common geological disasters of the natural world, not only endanger the safety of residents but also severely harm the living

environment, resources, and property of humankind (Petley 2012; Zhang et al., 2015; Haque et al., 2016). Especially since the impoundment of the Three Gorges Reservoir (TGR) of China in 2003, many ancient landslides in the reservoir area have been reactivated, and some recent landslides have also been induced (Zhang et al., 2015). Therefore, it is imperative to develop accurate early warning systems (EWSs) for landslide disasters in this district (Yin et al., 2010; Tang et al., 2019). Displacement prediction models, as a critical element of EWSs (Casagli et al., 2010; Intrieri et al., 2013; Liu et al., 2013), are receiving increasing attention with the increasing understanding of the need for disaster prevention and mitigation (Kirschbaum et al., 2010). Therefore, establishing accurate, stable, and dependable prediction models of landslide displacements has become one of the pressing issues in current geological disaster prediction research.

Among the current prediction methods of landslide displacements, the decomposition of landslide displacements and the establishment of prediction models affect the final prediction performance of models. From the perspective of displacement decomposition, the time series analysis (TSA) technique is taken as the theoretical basis for most current models. These models can be divided into two categories: partial decomposition (PD) models and complete decomposition (CD) models. In PD models, the TSA technique and the moving average method are combined to separate the trend and periodic displacements from the original displacement time series (Du et al., 2013; Zhou et al., 2016; Yang et al., 2019). It is very efficient and straightforward to decompose the displacement by using this model. The physical meaning of the decomposition components is clear. However, it is usually unable to obtain the random displacement proposed in the TSA due to the defects of the decomposition method, which leads to the incomplete decomposition problem.

In CD models, some novel signal decomposition technologies, including wavelet transform (WT), empirical mode decomposition (EMD), ensemble empirical mode decomposition (EEMD), and variational mode decomposition (VMD), are used to decompose the original displacement time series into several displacement components (Lian et al., 2013; Huang et al., 2016; Shihabudheen and Peethambaran 2017; Li et al., 2018). It offers a practical solution to the incomplete decomposition problem of random displacement. Nevertheless, considering the complexity of the method, two new problems have appeared during its practical application. The first problem is the overdecomposition of displacement time series, and the

second problem is the uncertainties in selecting the decomposition parameters. These two problems significantly limit the suitability, effectiveness, and generalizability of this model Therefore, we combine the TSA technique with the VMD method, which has achieved excellent performance in landslide displacement decomposition (Li et al., 2018; Guo et al., 2020), to propose a new adaptive decomposition method, named the optimized VMD (OVMD) method.

From the perspective of prediction model establishment, the development of landslide prediction models has experienced a leap from empirical models, semiempirical models, and mathematical-statistical models to nonlinear neural network models in recent decades (Saito, 1965; Fukuzono, 1985; Li et al., 2012; Du et al., 2013; Intrieri et al., 2019; Liu et al., 2020). Especially with the rise of artificial intelligence technologies, various neural network models have been introduced into landslide displacement predictions. As a nonlinear neural network prediction model with strong generalization, the kernel extreme learning machine (KELM) has been widely utilized (Zhou et al., 2018a; Zhou et al., 2018b; Liao et al., 2020). However, similar to other nonlinear models, such as the backpropagation neural network (BPNN), support vector regression (SVR), and extreme learning machine (ELM) models, there remains a crucial problem in the application of the KELM model except for the model parameter optimization, i.e., the selection of model input factors.

In general, qualitative analysis is the most commonly used method for input factor selection. That is, the input factors are directly proposed based on the experience judgment of engineering geologists. However, while using this method, the sensitivity of models to the interactions between some low-correlation factors is not fully considered. In some other cases, a few simple univariate correlation analyses, such as the Pearson correlation or gray correlation analysis (GRA), are used to carry out an a posteriori test on the correlation between the selected input factors and the displacement. However, usually, no input factors are reduced during this process. It is evident that using these qualitative or semiquantitative methods of factor selection is not conducive to the generalization of relevant prediction models.

A more reasonable and quantitative method for input factor selection is to propose candidate factors for different displacement components based on the experience judgment of engineering geologists at first. Subsequently, based on the sensitivity analysis results of the well-trained prediction model, these candidate factors are optimized by eliminating factors that have no contribution or low contribution to the prediction. Thus, in this study, we used this improved method

to select the input factors of KELM models. In this improved method, a global sensitivity analysis (GSA) method, called the PAWN method (Pianosi and Wagener, 2015), is introduced to analyze the sensitivity of the results obtained by prediction models to the input factors. Compared to the widely used factor selection methods and other variance-based sensitivity analysis methods, this method is easily implemented to characterize the input factors' significance in cases of small samples (Wang et al., 2020).

Overall, based on the variation characteristics of landslide displacements and the global response characteristics of input factors, we propose a adaptive hybrid prediction model, the OVMD and gray wolf optimizer (GWO)-based KELM (OVMD-GWO-KELM) model. In this model, the TSA model and the OVMD method are first combined to decompose the landslide displacements into displacement components. Second, based on the macroscopic deformation response of the landslide to various triggering factors, a suitable set of candidate factors is constructed for each displacement component. Some of these factors are also decomposed by the OVMD method according to their changing characteristics. Next, the Latin hypercube sampling (LHS) algorithm and the PAWN method are used to optimize these candidate factors. Finally, based on the optimization results of candidate factors, the GWO-KELM model is applied to conduct the multistep-ahead prediction of each displacement component. The prediction result is superposed to realize the prediction of the cumulative displacements of landslides. To verify the validity, superiority, and stability of the proposed model, we select the Baishuihe landslide as an example to predict the displacement of its three most representative monitoring sites. In total, fifty predictions for each monitoring site are implemented by using this model. The root mean square error (RMSE), mean absolute percentage error (MAPE), R-squared (R^2), and Akaike information criterion (AIC) are used to compare the average of these 50 forecasting results with the results of four different GWO-KELM models.

2. Methodology

2.1 TSA model of landslide displacements

The displacement time series of landslides are nonstationary. The internal and external factors that cause their variation can be mainly divided into three categories (Yang, 1992): trend, periodic, and random triggering factors.

149 The trend triggering factors mainly refer to geographical and geological conditions of the slope
150 body that constantly evolve under constant external factors (e.g., gravity). In some exceptional cases,
151 they also involve external triggering factors with an evident tendency to change over a long period,
152 e.g., the large-scale and long-period scheduling of the reservoir water level. These factors are the
153 most fundamental factors leading to the formation of recent landslides or the reactivation of ancient
154 landslides. Under the separate control of such trend factors, the displacement of landslides tends to
155 increase monotonically over time.

156 Periodic triggering factors are often described as external factors that have apparent periodic
157 variational regularity and have an important influence on the deformation and stability of landslides
158 during the formation of recent landslides or the reactivation of ancient landslides. They often involve
159 seasonal precipitation, seasonal flash floods, periodic scheduling of reservoir water levels.
160 Influenced by this kind of factor, the displacement time series often shows a periodical variation,
161 such as a step-like increase.

162 The random triggering factors are the external factors that have significant uncertainty and
163 impact the deformation and stability of the landslide during the formation of recent landslides or the
164 gradual reactivation of ancient landslides. These factors are generally composed of nonseasonal
165 precipitation, temporary small-scale scheduling of reservoir water levels, and accidental small-scale
166 human activities. Under the influence of these factors, the landslide displacement time series can be
167 characterized by approximate white noise sequences.

168 Based on the analysis mentioned above, to better clarify the physical significance of the
169 components of landslide displacement time series and display the variations in each component, the
170 displacement time series of landslides can be characterized as follows:

171 $$X(t)=\begin{cases}\alpha(t)+\beta(t)+\gamma(t) & \textit{if periodic inducing factors exist}\\ \alpha(t)+\gamma(t) & \textit{if periodic inducing factors do not exist}\end{cases} \quad (1)$$

172 where $X(t)$ represents the observed value of landslide displacement, $\alpha(t)$ represents the trend
173 displacement caused by trend triggering factors, $\beta(t)$ represents the periodic displacement caused
174 by periodic triggering factors, and $\gamma(t)$ represents the random displacement caused by random
175 triggering factors.

176 Unfortunately, due to the current monitoring instruments' limitations, these displacement

components cannot be directly obtained. Therefore, effectively obtaining the real displacement components within the original displacement time series has become a necessary foundation to establishing the nonlinear mapping relationship between triggering factors and the landslide displacement. In this study, the various displacement components with different physical significances are obtained by using the proposed OVMD technique which is described in detail in section 2.2.

2.2 OVMD

The VMD method is a novel complicated-signal decomposition method based on the EMD method (Dragomiretskiy and Zosso, 2013). Three critical parameters—the number of modes to be recovered K, the balance parameter of the data-fidelity constraint α, and the dual ascent timestep τ—must be set before using this method. Many studies (Dragomiretskiy and Zosso, 2013; Tang and Wang, 2015; Tang and Wang, 2016) have shown that except for τ, how to set the parameters K and α reasonably is very complicated. Although Li et al. (2018) and Guo et al. (2020) have used the VMD algorithm to decompose the landslide displacement time series, the settings of relevant parameters still depend on experience. Thus, considering that the complexity and variability in the actual displacement time series, the parameters selected entirely by the empirical method are not conducive to the promotion of the relevant models in actual landslide disaster prediction.

To solve this problem, we propose an adaptive signal decomposition method, named the OVMD method. As shown in Fig. 1, this model can adaptively determine the optimal model parameters to achieve the necessary decomposition of the object automatically. In this study, the OVMD method is used to decompose both the cumulative displacement time series and the candidate triggering factors. The specific calculation process of OVMD method consists of three parts as follows.

[Insert Fig. 1]

(1) ***Optimization of*** K**.** The input time series is preliminarily decomposed according to the initial parameter setting (i.e., $K=2$, and $\alpha=1$). Then, the optimal value $K_{optimal}$ is found

according to the following two criteria.

1) Criterion 1 is whether there is a medium or high correlation between different components after the decomposition (i.e., whether the distance correlation coefficient (DCC) between different components is greater than 0.4). When this criterion is not satisfied, the different displacement components or the different factor components obtained after decomposition should theoretically be independent or poorly correlated, ensuring that the decomposition components obtained by using the OVMD method are consistent with reality and have real physical meaning.

2) Criterion 2 is whether there is a mode aliasing between decomposed components after the decomposition (i.e., whether the difference in the center frequencies (DCF) between the decomposed components is higher than half their sum of 3 dB bandwidths (HSB)). As shown in Fig. 2, when the center frequencies f_c of the components are close to each other, the risk of aliasing between components increases gradually. Therefore, in addition to ensuring that the decomposed components have different physical meanings, we should pay attention to the mode aliasing problem.

[Insert Fig. 2]

(2) *Optimization of* α. The comprehensive determination indexes (CDIs) corresponding to different α values and $K_{optimal}$ are calculated based on the signal-to-noise ratio (SNR), approximate entropy, center frequency, and 3 dB bandwidth of the decomposed components. The definition of the CDI is as follows:

$$CDI = \begin{cases} MAPEN, & if\ t = 0 \\ NaN, & if\ t = 1 \end{cases} \tag{2}$$

where CDI is the comprehensive determination index; $MAPEN$ is the mean value of approximate entropies of all decomposed components; NaN is a null variable; and t is the logical variable of the determination matrix of modal aliasing (DMMA), which is calculated based on the DCF and HSB between the decomposed components. If there is mode aliasing between decomposed components, $t = false$; otherwise, $t = true$.

The smaller the CDI index is, the more orderly the variation in each component is, the smaller

the uncertainty is, and the more meaningful the information it carries. Therefore, when the CDI of the decomposed components reaches its minimum value, $\alpha_{optimal}$ is found. It should be noted that when there are multiple identical minimum values of the CDI, $\alpha_{optimal}$ should be determined according to the maximum value of the SNR of the decomposed components, which can minimize the information loss of the reconstructed signals. Finally, the input time series is decomposed according to the optimized parameter settings, $K_{optimal}$ and $\alpha_{optimal}$.

In the OVMD method, the signal decomposition is still performed based on the traditional VMD method proposed by Dragomiretskiy and Zosso (2013). That is, based on the MATLAB platform, we only write the code of the parameter optimization part according to the above process as shown in Figure 1. Besides, in the calculation process, the optimization space of K and α are set to $[2, 6]$ and $[1, 1\times10^4]$, respectively.

2.3 GWO-KELM

The KELM model is an improved model proposed by Huang et al. (2006) based on the ELM model and kernel function technology. The regression function based on the KELM model can be expressed as follows:

$$f(x) = \begin{bmatrix} K(x,x_1) \\ \vdots \\ K(x,x_N) \end{bmatrix}^{\mathrm{T}} (I/C + \Omega_{ELM})^{-1} T \tag{3}$$

where I is the diagonal matrix, C is the penalty parameter, I/C is the bias constant of the diagonal elements in the symmetric matrix Ω_{ELM}, Ω_{ELM} is the kernel function matrix of ELM, and T is the output target vector-matrix ($T = [t_1, t_2, \cdots, t_i]^{\mathrm{T}}$). The function matrix Ω_{ELM} is defined as follows:

$$\begin{aligned} \Omega_{ELM} = HH^{\mathrm{T}} : \Omega_{ELM(i,D)} &= h(x_i)h(x_j) \\ &= K(x_i,x_j) \end{aligned} \tag{4}$$

where H is the output matrix of the hidden layer ($H = [h(x_1), h(x_2), \cdots, h(x_i)]^{\mathrm{T}}$), $h(x_i)$ is the output vector of the hidden layer corresponding to the input vector x_i, and $K(x_i, x_j)$ is the kernel function. Since the radial basis function (RBF) is the most widely used kernel function, it is used as the kernel function of the KELM model in this paper. It is defined as follows:

$$K(x_i,x_j) = \exp(-\lambda \| x_i - x_j \|^2) \tag{5}$$

where λ is the kernel function parameter ($\lambda > 0$), and $\|x_i - x_j\|^2$ is the square of the Euclidean distance between the input vector x_i and x_j.

The KELM model has strong applicability in landslide displacement prediction (Zhou et al., 2018b; Liao et al., 2020). Thus, in this paper, the KELM model is used as one of the most critical parts of the proposed adaptive model. It is well known that the model parameter optimization of the KELM model has significant effects on its performance. Since the KELM model is the core of the adaptive prediction model proposed in this paper, its parameters have also become the core parameters that affect the prediction performance of the hybrid model. To obtain more accurate and stable prediction results, we use the GWO algorithm proposed by Mirjalili et al. (2014) to optimize C and λ. In this paper, the optimization space of C and λ are both set to $[1\times10^{-13}, 1\times10^{13}]$. The numbers of search agents, iterations, and convergence errors used in the GWO algorithm are 40, 200, and 1×10^{-7}, respectively. For the factor selection of the KELM model, the GSA method, called the PAWN method, is used and introduced in detail in section 2.4.

2.4 PAWN method

As a novel distribution-based GSA methodology, the PAWN method was proposed by Pianosi and Wagener (2015) to calculate the sensitivity index S_i of the *i*-th input factor x_i. More concretely, the sensitivity y of the model output to x_i can be calculated according to the maximum distance between the unconditional output distribution $F_y(y)$ and the conditional output distribution $F_{y|x_i}(y|x_i)$, where $F_y(y)$ can be obtained by varying all input factors x simultaneously, and $F_{y|x_i}(y|x_i)$ is generated by changing all input factors x except x_i. The calculation of the maximum distance between $F_y(y)$ and $F_{y|x_i}(y|x_i)$ is implemented by the Kolmogorov–Smirnov (KS) test in this paper. Thus, for x_i, S_i is defined as follows:

$$S_i = stat\,\mathrm{KS}(x_i) \text{ where } \mathrm{KS}(x_i) = \max\left|F_y(y) - F_{y|x_i}(y|x_i)\right| \tag{6}$$

where $stat$ denotes a statistic determined by the user, such as the maximum, median, or mean.

In this study, we set $S_i = mean(\mathrm{KS}(x_i))$, in which the value range of S_i is $[0, 1]$. The smaller the value of S_i is, the smaller the influence of x_i on y.

In general, S_i is computed numerically. Unlike the PAWN index approximation approach based on the tailored sampling strategy (Pianosi and Wagener, 2015), the fundamental idea of the generic approach (Pianosi and Wagener, 2018) to estimate S_i is to divide the changing range of x_i into w equidistant intervals of $\mathcal{I}_k$ and determine the conditional samples YC_{ik} accordingly. The unconditional sample Y_U can be replaced with the entire sample Y or with a subsample of Y. Therefore, the approximative PAWN indexes calculated according to the new strategy can be formulated as follows:

$$\hat{S}_i = stat_{k=1,\ldots,w}\mathrm{KS}(\mathcal{I}_k) \text{ where } \mathrm{KS}(\mathcal{I}_k) = \max_y \left|\hat{F}_y(y) - \hat{F}_{y|x_i}(y|x_i \in \mathcal{I}_k)\right| \tag{7}$$

where $\hat{F}_y(y)$ is the empirical distribution of Y_U, and $\hat{F}_{y|x_i}(y|x_i \in \mathcal{I}_k)$ is the empirical distribution of YC_{ik}.

When the number of samples is relatively small and the above approximate formula is applied to compute the PAWN indexes, the robustness of the chosen samples' sensitivity values should be evaluated accurately and carefully by using the bootstrap method (Efron and Tibshirani, 1993). Additionally, the KS statistic $\hat{S}_{dummy}$ of the dummy factor is introduced to identify the inputs whose measured sensitivities are too low to distinguish them from approximation errors (Zadeh et al., 2017).

Due to the computational efficiency and robustness of the PAWN algorithm, we use this algorithm to select the input factors of the KELM models. The total number of calculations of the model N and the number of conditioning intervals w must be set up before starting the GSA with the PAWN method. Ordinarily, N should be set as large as possible to provide more available computing resources to the PAWN method (Wang et al., 2020). Nevertheless, the higher the value of N, the higher the computational cost. Therefore, we set $N = 10000$. In

terms of w, the default value ($w=10$) recommended by Pianosi and Wagener (2019) is used in this study. The SAFE toolbox (Pianosi et al. 2015 and 2018) is applied to implement the PAWN algorithm based on the new generic approach. The specific process of factor selection based on the PAWN method consists of four parts, as shown in section 2.5.2.

2.5 Prediction process of the OVMD-GWO-KELM model

Fig. 3 shows the schematic of the OVMD-GWO-KELM model for forecasting the step-like displacement of colluvial landslides. This model mainly consists of the following three parts:

[Insert Fig. 3]

2.5.1 Data processing and adaptive decomposition

(1) ***Data cleaning.*** This step involves filling in the missing values of raw data, reducing noise in the raw data, identifying outliers in the raw data, and correcting the inconsistencies in the raw data.

(2) ***Isochronous processing.*** The raw data of the landslide displacements, reservoir water levels, precipitation, and so on are preprocessed to unify the data sampling frequency at a fixed value (i.e., monthly).

(3) ***Normalization processing.*** All raw data are normalized to [-1, 1] by different dimensions of the triggering factors.

(4) ***Selection of candidate input factors.*** According to the landslide deformation evolution law and its deformation response relationship with external factors, a candidate set of factors is constructed.

(5) ***Decomposition of the input and output variables based on the OVMD method.*** The OVMD method is applied to adaptively decompose the cumulative displacements and their corresponding triggering factors into multiple components with distinctive characteristics. The detailed calculation flow of the OVMD method is described in detail in section 2.2.

(6) ***Dataset construction.*** According to the definitions of trend displacement, periodic displacement, and random displacement, the training datasets of the different displacement components are formed.

2.5.2 Factor adaptive selection

(1) *Response surface modeling*. Based on the constructed datasets, the GWO-KELM model for different displacement components is separately established and well trained as the response surface models to represent the natural response of the landslide displacement component to external induced factors. To make full use of every sample in the datasets, we used the leave-one-out method ten times during the parameter optimization for each GWO-KELM model.

(2) ***Data resampling and response fitting***. LHS was used to generate N samples from the triggering factors of each displacement component. In this process, all the factors were assumed to follow uniform distributions. The generated samples are then taken in the response surface model of each displacement component to obtain their corresponding predicted displacement values.

(3) *Calculation of the KS statistic based on the PAWN method*. The PAWN method is applied to calculate the mean KS statistics of all triggering factors for the prediction model of each displacement component. In this process, the KS statistic of a dummy factor is obtained to judge whether the triggering factors play a role in the GWO-KELM models. Then, according to the mean KS statistic, the input factors of the GWO-KELM model for each displacement component are sorted in descending order. Finally, factors whose mean KS statistics are less than the mean KS statistics of their corresponding dummy factors are deleted. The initial valid factors are determined according to the remaining factors.

(4) *Deletion of redundant factors*. First, the sorted mean KS statistics of the initial optimized factors are fit with an appropriate exponential fitting function as follows:

$$y_{ks} = Ae^{Bx_{order}} + Ce^{Dx_{order}} \tag{8}$$

where y_{ks} is the fitted value of the mean KS statistics of all triggering factors. x_{order} is the order number of the triggering factors sorted in descending order according to the mean KS statistics. A, B, C, and D are the function parameters to be fitted by the nonlinear least-squares method.

Second, the second-order differences of y_{ks} (i.e., $\triangle\left(\triangle y_{ks}\right)$) are calculated. Third, all factors whose $\triangle\left(\triangle y_{ks}\right)$ values are less than 0.001 are eliminated. Finally, the retained factors

are used as the final optimized factors to reconstruct the datasets of the different displacement components.

2.5.3 Model training, prediction, and validation

(1) ***Model training and prediction.*** Based on the reconstructed optimal datasets, the GWO-KELM prediction models of the different displacement components are rebuilt. During this process, the training method and optimization parameters of the KELM model are consistent with those of the above method. The optimal input factors are introduced into the well-trained models to predict the different displacement components. The predicted cumulative displacements of the landslide are obtained by summing all the prediction results.

(2) ***Performance validation of the model.*** In the prediction performance evaluation, the R^2 and MAPE values are not easily affected by extreme values and can better evaluate the mean prediction performance of the model as a whole than the other metrics. The RMSE is more sensitive to extreme value prediction results due to the squared amplification of the prediction errors, thus highlighting the larger error values with significant influences. Furthermore, the AIC is usually used to measure the computational complexity of the models. Therefore, in this study, the above four indexes are comprehensively used to validate the performance of the proposed model. The definitions of the AIC are as follows:

$$AIC = 2k - 2\ln(MSE) \tag{9}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \tag{10}$$

where $\hat{y}_i$ denotes the predicted value, y_i denotes the measured value, n denotes the number of predicted values, $\overline{y}$ denotes the mean of the measured values, and k denotes the number of model input factors.

3. Materials

3.1 Brief introduction of Baishuihe landslide

The Baishuihe landslide is located on the south bank of the Yangtze River (Fig. 4 (a)). It is approximately 56 km from the site of the Three Gorges Dam. As shown in Fig. 4 (b), the toe of the

389 landslide is directly below the reservoir water level of 135 m, and its rear is located at the
390 geotechnical boundary with an elevation of 410 m. A bedrock ridge bounds the east and west sides
391 of the landslide. The main sliding direction, north-south length, and east-west width of the landslide
392 are approximately 20°NE, 600 m, and 700 m, respectively. The front and back parts of the landslide
393 are steeper than its middle part, which has an overall slope of approximately 30°. The average
394 thickness of the sliding mass is approximately 30 m, and its estimated volume is approximately 1.26
395 $\times 10^7$ m^3.

396 The monitoring data of inclinometer QZK1 (located under monitoring site ZG118) show two
397 sliding zones with different depths (Li et al., 2013; Miao et al., 2018; Yang et al., 2019). As shown
398 in Fig. 5, the shallow sliding zone is the contact zone between Quaternary loose deposits (i.e.,
399 gravelly soils with silty clay and fragmented rubble) and cataclastic rocks, with a buried depth of
400 12 m to 25 m. The deep sliding zone is the contact zone between the bottom of the cataclastic rock
401 and the underlying bedrock (i.e., the silt mudstone of the Jurassic Xiangxi Formation), with a buried
402 depth of 18.9 m to 34.1 m. The deep sliding zone is the original sliding zone of the Baishuihe
403 landslide, while the shallow sliding zone is the new sliding zone formed when the Baishuihe
404 landslide was reactivated by water-level fluctuations in the TGR in 2003.

405

406 **[Insert Fig. 4]**

407

408 **[Insert Fig. 5]**

409

410 Since 2003, the Baishuihe landslide has undergone five large deformations, among which the
411 most severe occurred in July 2004. According to the characteristics and scope of the deformation
412 that appeared in July 2004, the early warning area of the landslide (i.e., the main deformation area)
413 is determined (Fig. 4 (c)). The front edge of the early warning area is below the water level of 145
414 m, and the back edge is approximately 295 m in elevation. Huangtubao and Shanyanggou bound
415 the east and west sides of the area, respectively. It is approximately 430 m from north to south and
416 500 m from east to west, in which the plane space is approximately 5.5×10^6 m^2.

417 At present, the deformation of the landslide is mainly concentrated in the early warning area at
418 the front and middle parts of the whole landslide, and the displacement of the monitoring point at

the back of the landslide is still stable, which shows that the whole landslide mainly presents retrogressive deformation characteristics. As shown in Fig. 4 (b), 11 GPS displacement monitoring sites are on the Baishuihe landslide. Compared with the other eight monitoring sites, monitoring sites ZG93, ZG118, and XD01 have the following two advantages: (1) the monitoring period is the longest, and the data are the most complete; (2) they are all located in the middle of the landslide, and their monitoring data are representative, which can better reflect the whole landslide movement process. Therefore, we select the data of these three monitoring sites for the research in this paper.

3.2 Initial selection of the candidate factors based on engineering geological analysis

It is generally known that for machine learning models, choosing the appropriate input factors is essential to guarantee the prediction accuracy and generalization performance of models. Numerous studies have shown that precipitations, fluctuations in reservoir water level, and earthquakes are the main triggering factors of slope deformation and instability (Crosta 2004; Lee et al., 2008; Tang et al., 2019). No massive seismic disaster has been reported in the TGR area recently, and abundant precipitation occurs year-round in this area (Yao et al., 2015). Additionally, the research object (i.e., the Baishuihe landslide) is a typical riverside landslide in this region. Thus, it is quite evident that precipitation and fluctuations in the reservoir water level should be the major triggering factors of the Baishuihe landslide.

[Insert Fig. 6]

Fig. 6 shows that the increasing extents of cumulative displacements generally increase with the increase in precipitation intensity during the rainy season and decrease with the decline in precipitation intensity during the nonrainy season. This finding illustrates that precipitation is one of the triggering factors affecting the deformation of the Baishuihe landslide, which has a positive effect on the increase in the landslide displacements. Considering that precipitation infiltration is usually a continuous and relatively slow process and that the time intervals between two continuous heavy precipitation processes on the Baishuihe landslide are relatively long (the time intervals range from 12.38 days to 24.6 days), it is not appropriate to consider

447 the cumulative effect of precipitation infiltration over a long period. Because the sharp annual
448 increases in displacements and the heavy precipitation processes are both mainly concentrated
449 in June, July, and August every year, it is quite clear that there is a significant correlation
450 between the occurrence of antecedent heavy precipitation processes before 1 or 2 months and
451 the sharp increase in the monthly landslide displacement. Thus, it is appropriate to consider the
452 cumulative effect of precipitation infiltration within two months. Some earlier research (Du et
453 al., 2013; Roering et al., 2015; Krkač et al., 2017; Bogaard and Greco, 2018; Li et al., 2020)
454 was consistent with this analysis. Therefore, in this paper, the monthly cumulative antecedent
455 precipitation, the bimonthly cumulative antecedent precipitation, and the monthly maximum
456 continuous effective precipitation with four different reduction factors (i.e.,
457 $r = 1, 0.8, 0.6, and\ 0.4$) are selected as the candidate factors. These six precipitation factors are
458 defined as CF_1, $CF_2, \cdots, CF_6$.

459 Fig. 6 also shows that the cumulative displacements increase sharply approximately one
460 or two months later when the reservoir level declines. However, when the reservoir water level
461 rises, the landslide displacements remain the same. This finding illustrates that fluctuations in
462 the reservoir water level are also among the triggering factors that cause the deformation of the
463 Baishuihe landslide, which has both positive and negative effects on the increase in the
464 landslide displacement because the pore water pressure in the sliding mass is difficult to
465 dissipate (Tang et al., 2019). Hence, considering the hysteresis effect of groundwater level that
466 is caused by the fluctuation of the reservoir water level (Miao et al., 2018; Xue et al., 2020),
467 the influence of both the current and past water levels on landslide displacements should be
468 considered in the prediction of landslide displacements. Therefore, in this study, we select the
469 monthly and bimonthly average elevations of reservoir levels, the monthly and bimonthly
470 increases in reservoir levels, the monthly and bimonthly decreases in reservoir levels, the
471 monthly and bimonthly relative variations in reservoir levels, and the monthly and bimonthly
472 absolute variations in reservoir levels as candidate factors. These ten water-level factors are
473 defined as CF_7, $CF_8, \cdots, CF_{16}$.

474 Moreover, due to the continuous evolution of geological conditions, there is a remarkable
475 difference in the deformation responses of landslides between their different evolution stages

(Glade et al., 2005). For landslides with progressive failure characteristics, even extremely heavy precipitation may result in only minor deformation when they are in a steady evolution stage (i.e., the primary or secondary creeping stage). In contrast, when they are under an unstable evolution stage (i.e., the tertiary creeping stage), slight precipitation may easily break the limit equilibrium state of the original system and cause extremely intense deformation (Zhou et al., 2018b; Yang et al., 2019). Thus, it is necessary to consider the influence of historical landslide deformation on current landslide evolution (Cao et al., 2016). The cumulative displacements and their increments over the preceding 1-3 months are adopted as the supplement of the rainfall and reservoir level factors to characterize the influence of the evolution of geological conditions on landslide displacement. These six displacement factors are defined as CF_{17}, $CF_{18}, \cdots, CF_{22}$.

4. Results

In total, 408 data points collected from each monitoring site of ZG93, ZG118, and XD01 between June 2006 and December 2016 were used in this simulation. The data points from June 2006 to December 2013 and January 2014 to December 2016 were used as the training and testing sets, respectively. The data preprocessing step was conducted before splitting the entire dataset. The multistep-ahead prediction method was used in the prediction process of the proposed model, which means that all 36 samples in the testing set were predicted at once after building the prediction model based on all 100 samples in the training set. Due to the limitation of this paper's length, we mainly took the data of monitoring site ZG93 as an example to introduce the details of the prediction process and its results obtained by using the OVMD-GWO-KELM model.

4.1 Decomposition of landslide displacements and triggering factors based on the OVMD method

4.1.1 Decomposition parameters and decomposition results of landslide displacements

1) *The optimal number of modes.*

Fig. 7 shows the variations in DCC between the decomposed components with increasing K. When $K=4$, two groups of medium-correlation or high-correlation decomposed

504 components appeared. With the increase in K from 4 to 6, the number of medium-correlation
505 and high-correlation components increased from 2 to 3. Clearly, $K=3$ is the threshold at
506 which high- or medium-correlation components appeared in the displacement decomposition
507 result of monitoring site ZG93.
508
509 **[Insert Fig. 7]**
510
511 Fig. 8 shows the variations in the DCF and HSB between different components with
512 increasing K. When $K=4$, the HSB between different components is larger than the DCF
513 between different components for the first time. Accompanying the increase in K from 4 to 6,
514 the number of points with HSBs higher than the DCFs also increased from 3 to 6. It is clear that
515 $K=3$ is also the threshold at which the aliasing components appeared in the displacement
516 decomposition result of monitoring site ZG93.
517
518 **[Insert Fig. 8]**
519
520 According to the definition of the convergence conditions for the optimization of K
521 described in section 2.5.1, $K=3$ is the threshold of the underdecomposition and
522 overdecomposition of landslide displacements for monitoring site ZG93, which means that
523 $K_{optimal}=3$. Monitoring sites ZG118 and XD01 have the same optimal mode number as
524 monitoring site ZG93 in their displacement time series.
525 2) *The optimal balancing parameter.*
526 After determining the optimal value of K, the optimal value of α was searched
527 immediately. Fig. 9 shows that the MAPEN gradually decreases with increasing α. This
528 finding means that the greater the value of α is, the better the regularity of the displacement
529 components. However, it is not true that the greater the value of α is, the better the

decomposition effect of the results, because when α is more significant than a specific value, mode aliasing appears in the decomposed components. Therefore, according to the definition of the CDI and its corresponding optimization principle, which are both described in section 2.2, the optimal values α for monitoring sites ZG93, ZG118 and XD01 were determined, as shown in Fig. 9 (d).

[Insert Fig. 9]

3) *Displacement decomposition results.*

The decomposition of the actual cumulative displacements was carried out based on $K_{optimal}$ and $\alpha_{optimal}$ described above. The actual cumulative displacements and their corresponding decomposed displacement components and the residuals of monitoring sites ZG93, ZG118, and XD01 are shown in Fig. 10.

[Insert Fig. 10]

4.1.2 Decomposition parameters and the decomposition results of the triggering factors

OVMD was also used to decompose the candidate factors proposed in section 3.2. The decomposition parameters of these factors are shown in Table 1. Not all candidate factors must be decomposed. In addition, for the same type of factors, the decomposition parameters are nearly the same.

[Insert Tab. 1]

The representative decomposition results of the different types of factors for monitoring site ZG93, including CF_3, CF_9, CF_{11}, CF_{14} and CF_{17}, are shown in Fig. 11. For the candidate factors to be decomposed, their components obtained from the decomposition have different variation characteristics, which describes the changes in the original signal from different perspectives.

[Insert Fig. 11]

4.2 Optimization of the triggering factors based on the PAWN method

4.2.1 GSA of triggering factors for prediction models

As described in section 2.5.2, after the prediction models for different displacement components were established based on the decomposed training data, the validity of each model's input factors needed to be analyzed based on the PAWN method. Fig. 12 shows the mean KS statistics of the triggering factors for the prediction models of the different displacement components. Most of the mean KS statistics of triggering factors are less than or remarkably close to those of the dummy factors. This finding means that except for only a few valid factors, most triggering factors have little or no influence on the prediction model results in general. Additionally, for the prediction models of the same type of displacement components, there is a certain similarity between their valid input factors.

[Insert Fig. 12]

4.2.2 Selection of valid input factors

Taking monitoring site ZG93 as an example, Fig. 13 shows the valid and optimized factors for the different displacement component predictions obtained using the reduction method described in section 2.5.2. The red dotted line represents the initial dividing line of the optimized factors based on the KS statistic of the dummy factor. The green dotted line represents the final dividing line of the optimized factors determined according to the second-order difference results in the KS statistical fitting values of the initial valid factors. By using this approach, the number of input factors is substantially reduced. For instance, the number of input factors is reduced from 52 to 7 within the trend displacement prediction process of ZG93. Moreover, their factors after reduction are quite different for the prediction models of different displacement components.

[Insert Fig. 13]

Fig. 14 shows the DCC matrixes between the triggering factors before and after the reduction in factors, in which the diagonal elements of the matrixes of the different displacement components represent the DCC between the triggering factors and the predicted targets. For the prediction models of the different displacement components, the correlations among the triggering factors and between the triggering factors and the prediction objects are quite different. The trend displacement prediction model has the highest average correlations among the triggering factors and between the triggering factors and the predicted targets, followed by the random displacement prediction model. The periodic displacement prediction model has the lowest average correlations among the triggering factors and between the triggering factors and the predicted targets.

[Insert Fig. 14]

4.2.3 Computational stability of factor selection

To evaluate the stability of the PAWN method, the different displacement component prediction model factors were reduced 50 times for each prediction model. Fig. 15 shows the statistical results of the reduced factors. Regarding the prediction model of any displacement component for any monitoring site, the reduced result of the triggering factors is very stable, and the remaining factors are mainly concentrated on several of the same factors. In addition, factor reductions still have some similarities in choosing several core factors for the prediction model of the same displacement component of different displacement sites.

[Insert Fig. 15]

4.3 Cumulative displacement prediction based on the proposed modified model

To verify the accuracy and repeatability of the proposed OVMD-GWO-KELM model, we established fifty such models to predict the cumulative displacements for each monitoring site. The final prediction results of each monitoring site were the mean values of the prediction results provided by these fifty corresponding models. To compare the model performance, four different VMD-GWO-KELM models, which were separately established based on the initial

candidate factors, the candidate factors optimized by the GRA method, the adaptive decomposed factors, and the adaptive decomposed factors optimized by the GRA method, were taken as the comparison models. For the specific process of optimizing the input factors of the prediction models by using the GRA method, please refer to the research published by Miao et al. (2017) and Yang et al. (2019).

Fig. 16 and Table 2 show that the R2 and MAPE values of the proposed modified model are slightly higher or very close to those of the other comparable models. For the RMSE, the improved models have relatively low RMSEs in the training process and have the lowest RMSEs in the prediction process. For the AIC, all proposed models have the minimum AIC values, whether in prediction or training. Additionally, the standard deviations of the prediction results based on the proposed model changed only slightly with increasing prediction period.

Table 2 also shows the variations in each performance index caused by the change in the number of factors. In most cases, the optimization of the number of input factors is very conducive to improving each index of the model, whether in prediction or training. Only in very few cases does the optimization of factors lead to the decline in two performance indexes (i.e., the R^2 and MAPE) of the OVMD-GWO-KELM model in the process of prediction. However, compared with the number of factors to be increased to improve these two indexes in the VMD-GWO-KELM model, this decrease in R^2 and MAPE in the proposed model performance is negligible.

[Insert Fig. 16]

[Insert Tab. 2]

5. Discussion

5.1 Component extraction of landslide displacements

As shown in Fig. 10, the decomposed trend displacements accurately retain the changing trend in the original cumulative displacements. For the decomposed periodic displacements, there is a good correspondence between their periodic variations and the step-like change in the

original cumulative displacements. For the decomposed random displacements, their variations have a particular uncertainty and randomness. These results are consistent with the definition of the landslide displacement components in section 2.1, which supports the finding that the OVMD method can decompose the landslide displacements into different displacement components with real physical meanings and different variation characteristics.

Moreover, the residuals between the reconstructed cumulative displacements (i.e., the sum of trend displacement, periodic displacement, and random displacement) and the original cumulative displacements account only for approximately one-thousandth of the original cumulative displacements and do not have very obvious regularity. Obviously, using the reconstructed cumulative displacements obtained by the OVMD method for the prediction would not yield misjudgment of landslides' actual displacements. This further illustrates the feasibility of using the OVMD method to decompose the displacement.

5.2 Component extraction of triggering factors

The decomposition of input factors has been proposed only in recent years. Its purpose is to extract the components in each factor that are more related to their corresponding displacement components to build prediction models with better performance. In this kind of research (Deng et al., 2017; Li et al., 2018; Guo et al., 2020), it is usually assumed that there are two components with different frequencies in the time series of factors, and then the correlations between the two components of factors and the prediction objects are analyzed. According to the analysis results, the prediction model based on the high-correlation components was built to achieve a more accurate prediction than the traditional model based on the original undecomposed factors.

However, whether there are two components with different physical meanings in factor time series data is often ignored. Through our study, as shown in Table 1, it is found that not all factors have the above two components with different frequencies. Most factors do not have two components with different physical meanings, but a few factors have multiple components with different physical meanings. This fact reminds us that we cannot directly use this decomposition factor method to improve the prediction performance of the models. We should pay more attention to whether there are components with different physical meanings in the factors to avoid performance improvements based on fake factor components, thus increasing

677 the risk of incorrect predictions.

678 **5.3 Influence of considering factor coupling on the prediction results**

679 Considering the influence of increasing the number of input factors on the complexity,
680 generalization ability, and calculation efficiency of prediction models, how to extract valid
681 factors from numerous factors has become the key to improving the general performance of
682 prediction models. In past research, GRA has been widely used as a method to determine the
683 input factors of the model with the advantage of strong generality (Zhou et al., 2018b; Yang et
684 al., 2018; Liao et al., 2020). It is very suitable for selecting input factors to remove a large
685 number of irrelevant and low-correlation input factors quickly without the need for training
686 models.

687 As shown in Fig. 14, overall, only some of the optimization factors obtained by the
688 proposed PAWN-based method have high correlations with the prediction target, which is
689 different from the factor optimization results obtained by the GRA approach. Locally, the
690 factors selected by these two methods have a certain degree of similarity. For example, in the
691 trend displacement prediction factor of monitoring site ZG93, the top three factors selected by
692 these two methods are the same: F_{20}, $F_{21,}$ and F_{22}. The reason for these situations is that the
693 coupling influence between low-correlation factors on the prediction results has not been
694 considered when the GRA method is used. Thus, although there is a certain similarity between
695 the factors highly related to the prediction target and the high-sensitivity factors that have a
696 greater impact on the prediction results, these two kinds of factors are not entirely equivalent.
697 In addition, when using the proposed PAWN-based method to select factors, the correlations
698 between the selected factors are relatively low, which can effectively suppress the influence of
699 multicollinearity on the prediction results to a certain extent. This multicollinearity problem is
700 usually ignored, thus leading to the failure of prediction models.

701 The results in Table 2 show that compared with no optimization of the models' input factors,
702 using the GRA method can improve the comprehensive prediction performance of the VMD-
703 GWO-KELM model to a certain extent. However, the GRA method's optimization effect is still
704 lower than that of the proposed PAWN-based method. Obviously, considering the impact of
705 factor coupling on the prediction results of models is beneficial to improve the comprehensive
706 prediction performance of the model. Thus, in constructing the prediction model, we should

choose the high-sensitivity factors as the input factors of the model instead of the traditional high-correlation factors. This finding indicates the superiority of the proposed PAWN-based method by reducing the number of input factors for prediction models.

5.4 Performance of the OVMD-GWO-KELM model and its application prospects

As described in section 2.5.3, the different evaluation indexes used in this study can quantify the performance of the model from different aspects. By comparing and analyzing the prediction results of the three models, it is easy to determine that the average prediction performance of the proposed OVMD-GWO-KELM model is slightly higher than or remarkably close to that of the other four VMD-GWO-KELM models. However, from the view of extreme value prediction, especially when the displacements of landslides under the mutation state are predicted, the proposed model has a better generalization ability than the comparison models. Because the input factors used in the prediction process are obviously optimized, the complexity and computational efficiency of the proposed model are significantly higher than those of the comparison models. Moreover, the computational stability and reliability of the proposed model results are also confirmed according to the variation characteristics of the results of the fifty multistep-ahead predictions. Obviously, adopting the proposed model is very useful for improving the comprehensive performance of the EWS of landslides.

As described in section 2.5, two automatic parameter determination processes are applied in the proposed model. The first is the automatic determination process of the displacement time series decomposition parameters and input factor time series. It enables the proposed model to adaptively decompose the displacement time series or factor time series regardless of whether there are different components with clear physical meaning and different change characteristics in these time series. The second is the automatic determination process of the optimal factor set of displacement components to be predicted. This process considers the overall prediction performance and the extreme value prediction accuracy of the model and ensures that the constructed model has low complexity and high computational efficiency. Therefore, in theory, the proposed model also has a certain application and popularization value in predicting other landslides with different displacement variation characteristics. It should be noted that when using this model to predict the displacement of other types of landslides, it is vital to put forward more targeted candidate factors according to the response analysis of macro

deformations to establish a more accurate prediction model. In addition, this model cannot be updated effectively with the growth of prediction time at present, and its prediction accuracy is bound to be reduced. Therefore, in future research or applications of this model, its adaptive updating function should be subjected to more targeted optimization.

6. Conclusion

To solve the inherent defects in the existing landslide displacement prediction models and construct an accurate and useful EWS for landslide disasters, based on the OVMD approach, PAWN method, and GWO-KELM model, we proposed an adaptive hybrid machine learning model, the OVMD-GWO-KELM model. This model contained three phases: data processing and adaptive decomposition, adaptive factor selection, and model training, prediction, and validation. Taking the Baishuihe landslide as an example, the process of applying this model for landslide cumulative displacement prediction with multistep-ahead prediction was discussed in detail, and the following conclusions were drawn from this study:

(1) The adaptive decomposition results obtained by using OVMD show three displacement components with clear physical meaning and different variation characteristics in the cumulative displacement time series of the Baishuihe landslide—namely, trend, periodic, and random displacements. This finding is consistent with the conclusion of the traditional TSA model of landslide displacement for step-like landslides.

(2) It is worth noting that not all candidate factors can be decomposed, because some of them do not contain many components with clear physical meanings and different variation characteristics. Therefore, a high level of care is necessary when using these decomposition factors to improve prediction model performance because this kind of improvement is more likely to be subjective and artificial rather than objective and real.

(3) Compared to traditional input factor selection methods, such as the GRA method, the PAWN method can analyze the interaction between various low-correlation factors on the results of prediction models. Thus, by considering the redundancy and lack of information in raw data, both the low-sensitivity factors and combinations of low-sensitivity factors could be correctly reduced for prediction models with different displacement components by applying this method.

(4) Compared with the traditional VMD-GWO-KELM model, the generalization performance of the proposed adaptive hybrid model can be significantly improved without significantly reducing its training effect so that the overfitting and underfitting phenomena can be effectively suppressed. In addition, the complexity of the model is typically low, and the stability of its calculation results is still high with varying time. Thus, considering its adaptive prediction characteristics, this proposed adaptive hybrid model also has a certain applicability for displacement prediction of all kinds of landslides in addition to step-like landslides.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (41977244) and the National Key R&D Program of China (2017YFC1501301). The authors thank the colleagues in our laboratory for their constructive comments and assistance. In addition, the authors would like to express their special thanks to Dr. Yao Wenmin for their sincere assistance in improving this study.

References：

Bogaard, T., Greco, R., 2018. Invited perspectives: hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: proposing hydrometeorological thresholds. Nat Hazard Earth Sys, 18(1), 31–39.

Cao, Y., Yin, K., Alexander, D. E., Zhou, C., 2016. Using an extreme learning machine to predict the displacement of step-like landslides in relation to controlling factors. Landslides, 13(4), 725-736.

Casagli, N., Catani, F., Del Ventisette, C., Luzi, G., 2010. Monitoring, prediction, and early warning using ground-based radar interferometry. Landslides, 7(3), 291-301.

Crosta, G. B., 2004. Introduction to the special issue on rainfall-triggered landslides and debris flows. Engineering Geology, 3(73), 191-192.

Deng, D., Liang, Y., Wang, L., Wang, C., Sun, Z., Wang, C., Dong, M., 2017. Displacement prediction method based on ensemble empirical mode decomposition and support vector machine regression- a case of landslides in Three Gorges Reservoir area. Chin J Rock Mech Eng 38(12), 3660–3669

Dragomiretskiy, K., Zosso, D., 2013. Variational mode decomposition. IEEE transactions on

signal processing, 62(3), 531-544.

Du, J., Yin, K., Lacasse, S., 2013. Displacement prediction in colluvial landslides, three Gorges reservoir, China. Landslides, 10(2), 203-218.

Efron, B., Tibshirani, R. J., 1994. An introduction to the bootstrap. CRC press.

Fukuzono, T., 1985. A new method for predicting the failure time of a slope. In Proceedings of 4th International Conference and Field Workshop on Landslide., 1985 (pp. 145-150).

Glade, T., Crozier, M. J., 2005. Landslide hazard and risk: concluding comment and perspectives. Landslide hazard and risk. Wiley, Chichester, 767-774.

Guo, Z., Chen, L., Gui, L., Du, J., Yin, K., Do, H. M., 2019. Landslide displacement prediction based on variational mode decomposition and WA-GWO-BP model. Landslides, 1-17.

Haque, U., Blum, P., da Silva, P. F., Andersen, P., Pilz, J., Chalov, S. R., Malet, J. P.,Auflič, M. J., Andres, N, Poyiadji, E., Lamas, P. C., Zhang, W., Peshevski, I., Pétursson, H. G., Kurt, T., Dobrev, N., García-Davalillo, J. C., Halkia, M., Ferri, S., Gaprindashvili, G., Engström, J., Keellings, D., 2016. Fatal landslides in Europe. Landslides 13(6):1545–1554

Huang, F., Yin, K., Zhang, G., Gui, L., Yang, B., Liu, L., 2016. Landslide displacement prediction using discrete wavelet transform and extreme learning machine based on chaos theory. Environmental Earth Sciences, 75(20), 1376.

Huang, G. B., Zhu, Q. Y., Siew, C. K., 2006. Extreme learning machine: theory and applications. Neurocomputing, 70(1-3), 489-501.

Intrieri, E., Carlà, T., Gigli, G., 2019. Forecasting the time of failure of landslides at slope-scale: A literature review. Earth-science reviews.

Intrieri, E., Gigli, G., Casagli, N., Nadim, F., 2013. Brief communication: landslide early warning system: toolbox and general concepts. Nat Hazards Earth Syst Sci 13:85–90

Jiang, P., Chen, J., 2016. Displacement prediction of landslide based on generalized regression neural networks with K-fold cross-validation. Neurocomputing, 198, 40-47.

Kirschbaum, D.B., Adler, R., Yang, H., Hill, S., Lerner-Lam, A., 2010. A global landslide catalog for hazard applications: method, results, and limitations. Nat Hazards 52(3):561–575

Krkač, M., Špoljarić, D., Bernat, S., Arbanas, S, M., 2017. Method for prediction of landslide movements based on random forests. Landslides,14(3), 947–960.

Lee, C. T., Huang, C. C., Lee, J. F., Pan, K. L., Lin, M. L., Dong, J. J., 2008. Statistical approach

825 to earthquake-induced landslide susceptibility. Engineering Geology, 100(1-2), 43-58.
826 Li, D., Yin. K., Leo, C., 2010. Analysis of Baishuihe landslide influenced by the effects of
827 reservoir water and rainfall. Environ Earth Sci, 60(4), 677–687.
828 Li, H., Xu, Q., He, Y., Deng, J., 2018. Prediction of landslide displacement with an ensemble-
829 based extreme learning machine and copula models. Landslides, 15(10), 2047-2059.
830 Li, H., Xu, Q., He, Y., Fan, X., Li, S., 2020. Modeling and predicting reservoir landslide
831 displacement with deep belief network and EWMA control charts: a case study in Three Gorges
832 Reservoir. Landslides, 17(3), 693-707.
833 Li, L., Wu, Y., Miao, F. et al. 2020. A hybrid interval displacement forecasting model for
834 reservoir colluvial landslides with step-like deformation characteristics considering dynamic
835 switching of deformation states. Stochastic Environmental Research and Risk Assessment.
836 https://doi.org/10.1007/s00477-020-01914-w
837 Li, L., Wu, Y., Miao, F., Liao, K., Zhang, L., 2018. Displacement prediction of landslides based
838 on variational mode decomposition and GWO-MIC-SVR model. Chin J Rock Mech Eng, 37(6),
839 1395-1406.
840 Li, S. H., Wu, L. Z., Chen, J. J., Huang, R. Q., 2020. Multiple data-driven approach for
841 predicting landslide deformation. Landslides, 17(3), 709-718.
842 Li, X., Kong, J., Wang, Z., 2012. Landslide displacement prediction based on combining
843 method with optimal weight. Natural hazards, 61(2), 635-646.
844 Lian, C., Zeng, Z., Yao, W., Tang, H., 2013. Displacement prediction model of landslide based
845 on a modified ensemble empirical mode decomposition and extreme learning machine. Natural
846 hazards, 66(2), 759-771.
847 Lian, C., Zeng, Z., Yao, W., Tang, H., 2014a. Ensemble of extreme learning machine for
848 landslide displacement prediction based on time series analysis. Neural Computing and
849 Applications, 24(1), 99-107.
850 Lian, C., Zeng, Z., Yao, W., Tang, H., 2014b. Extreme learning machine for the displacement
851 prediction of landslide under rainfall and reservoir level. Stochastic environmental research and
852 risk assessment, 28(8), 1957-1972.
853 Liao, K., Wu, Y., Miao, F., Li, L., Xue, Y., 2020. Using a kernel extreme learning machine with
854 grey wolf optimization to predict the displacement of step-like landslide. Bulletin of

855 Engineering Geology and the Environment, 79(2), 673-685.
856 Liu, Y., Xu, C., Huang, B., Ren, X., Liu, C., Hu, B., Chen, Z., 2020. Landslide displacement
857 prediction based on multi-source data fusion and sensitivity states. Engineering Geology,
858 105608.
859 Liu, Z., Shao, J., Xu, W., Chen, H., Shi, C., 2014. Comparison on landslide nonlinear
860 displacement analysis and prediction with computational intelligence approaches. Landslides,
861 11(5), 889-896.
862 Miao, F., Wu, Y., Xie, Y., Li, Y., 2018. Prediction of landslide displacement with step-like
863 behavior based on multialgorithm optimization and a support vector regression model.
864 Landslides, 15(3), 475-488.
865 Miao, F., Wu, Y., Xie, Y., Yu, F., Peng, L., 2017. Research on progressive failure process of
866 Baishuihe landslide based on Monte Carlo model. Stochastic Environmental Research and Risk
867 Assessment, 31(7), 1683-1696.
868 Mirjalili, S., Mirjalili, S. M., Lewis, A., 2014. Grey wolf optimizer. Advances in engineering
869 software, 69, 46-61.
870 Petley, D., 2012. Global patterns of loss of life from landslides. Geology, 40(10), 927-930.
871 Pianosi, F., Wagener, T., 2015. A simple and efficient method for global sensitivity analysis
872 based on cumulative distribution functions. Environmental Modelling and Software, 67, 1-11.
873 Pianosi, F., Wagener, T., 2018. Distribution-based sensitivity analysis from a generic input-
874 output sample. Environmental Modelling and Software, 108, 197-207.
875 Ren, F., Wu, X., Zhang, K., Niu, R., 2015. Application of wavelet analysis and a particle swarm-
876 optimized support vector machine to predict the displacement of the Shuping landslide in the
877 Three Gorges, China. Environmental earth sciences, 73(8), 4791-4804.
878 Saito, M., 1965. Forecasting the time of occurrence of a slope failure. In: Proceedings of the
879 6th International Mechanics and Foundation Engineering, Montr al, Que. Pergamon Press,
880 Oxford 537-541
881 Shihabudheen, K. V., Peethambaran, B., 2017. Landslide displacement prediction technique
882 using improved neuro-fuzzy system. Arabian Journal of Geosciences, 10(22), 502.
883 Tang, G. J., Wang, X. L. 2016. Variational mode decomposition method and its application on
884 incipient fault diagnosis of rolling bearing. Journal of Vibration Engineering, 29(4), 638-648.

885 Tang, G. J., Wang, X. L., 2015. Parameter optimized variational mode decomposition method
886 with application to incipient fault diagnosis of rolling bearing. Journal of xi'an jiaotong
887 University, 49(5), 73-81.
888 Tang, H., Wasowski, J., Juang, C. H., 2019. Geohazards in the three Gorges Reservoir Area,
889 China–Lessons learned from decades of research. Engineering Geology, 105267.
890 Wang, Y., Huang, J., Tang, H., 2020. Global sensitivity analysis of the hydraulic parameters of
891 the reservoir colluvial landslides in the Three Gorges Reservoir area, China. Landslides, 17(2),
892 483-494.
893 Xue, Y., Wu, Y., Miao, F., Li, L., Liao, K., Ou, G., 2020. Effect of spatially variable saturated
894 hydraulic conductivity with nonstationary characteristics on the stability of reservoir landslides.
895 Stochastic Environmental Research and Risk Assessment, 34, 311–329.
896 Yang, B., Yin, K., Lacasse, S., Liu, Z., 2019. Time series analysis and long short-term memory
897 neural network to predict landslide displacement. Landslides, 16(4), 677-694.
898 Yang, S., 1992. Engineering application of time series analysis. Huazhong University of
899 Science and Technology Press, Wuhan
900 Yao, W., Zeng, Z., Lian, C., Tang, H., 2015. Training enhanced reservoir computing predictor
901 for landslide displacement. Engineering Geology, 188, 101-109.
902 Yin, Y., Wang, H., Gao, Y., Li, X., 2010. Erratum to: Real-time monitoring and early warning
903 of landslides at relocated Wushan Town, the Three Gorges Reservoir, China. Landslides, 7(3),
904 389.
905 Zadeh, F. K., Nossent, J., Sarrazin, F., Pianosi, F., van Griensven, A., Wagener, T., Bauwens,
906 W., 2017. Comparison of variance-based and moment-independent global sensitivity analysis
907 approaches by application to the SWAT model. Environmental Modelling & Software, 91, 210-
908 222.
909 Zhang L., Liao M., Balz T., Shi X., Jiang Y., 2015. Monitoring Landslide Activities in the Three
910 Gorges Area with Multi-frequency Satellite SAR Data Sets. In: Scaioni M. (eds) Modern
911 Technologies for Landslide Monitoring and Prediction. Springer Natural Hazards. Springer,
912 Berlin, Heidelberg
913 Zhang, J., Gurung, D. R., Liu, R., Murthy, M. S. R., Su, F., 2015. Abe Barek landslide and
914 landslide susceptibility assessment in Badakhshan Province, Afghanistan. Landslides, 12(3),

597-609.

Zhou, C., Yin, K., Cao, Y., Ahmed, B., 2016. Application of time series analysis and PSO–SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir, China. Engineering geology, 204, 108-120.

Zhou, C., Yin, K., Cao, Y., Ahmed, B., Fu, X., 2018a. A novel method for landslide displacement prediction by integrating advanced computational intelligence algorithms. Scientific reports, 8(1), 1-12.

Zhou, C., Yin, K., Cao, Y., Intrieri, E., Ahmed, B., Catani, F., 2018b. Displacement prediction of step-like landslide by applying a novel kernel extreme learning machine method. Landslides, 15(11), 2211-2225.

Zhu, X., Xu, Q., Tang, M., Nie, W., Ma, S., Xu, Z., 2017. Comparison of two optimized machine learning models for predicting displacement of rainfall-induced landslide: A case study in Sichuan Province, China. Engineering geology, 218, 213-222.

Figures

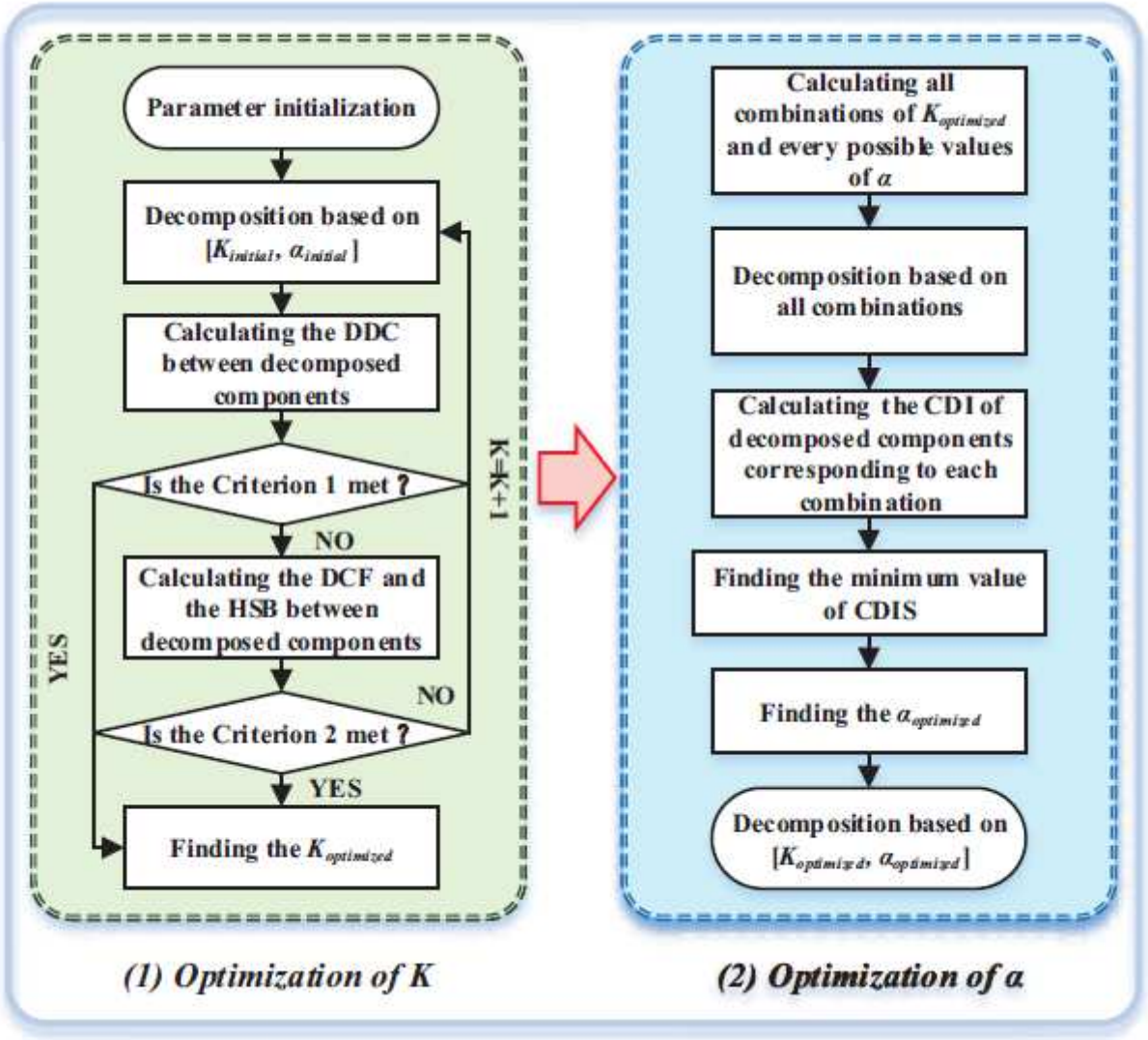

Figure 1

Calculation flowchart of the OVMD method

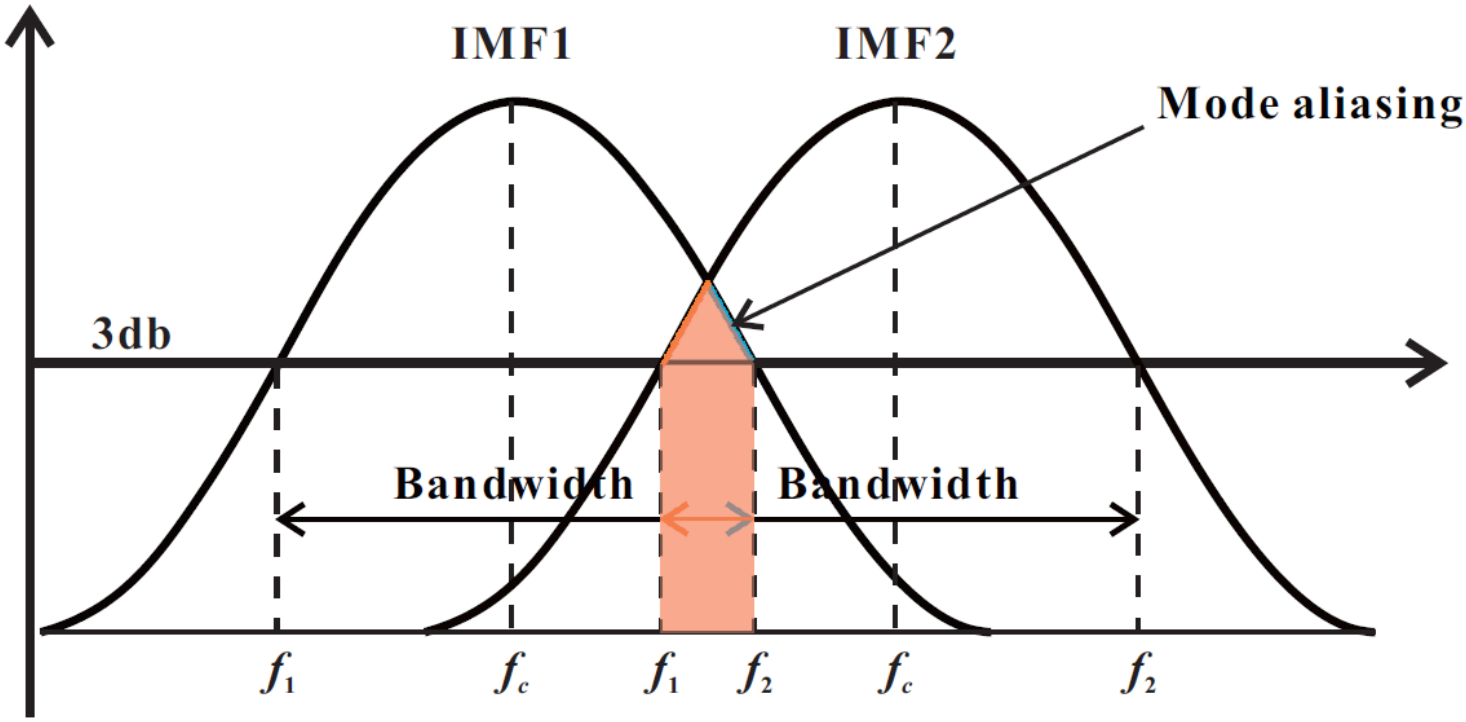

Figure 2

Schematic diagram of the mode aliasing between different decomposed components

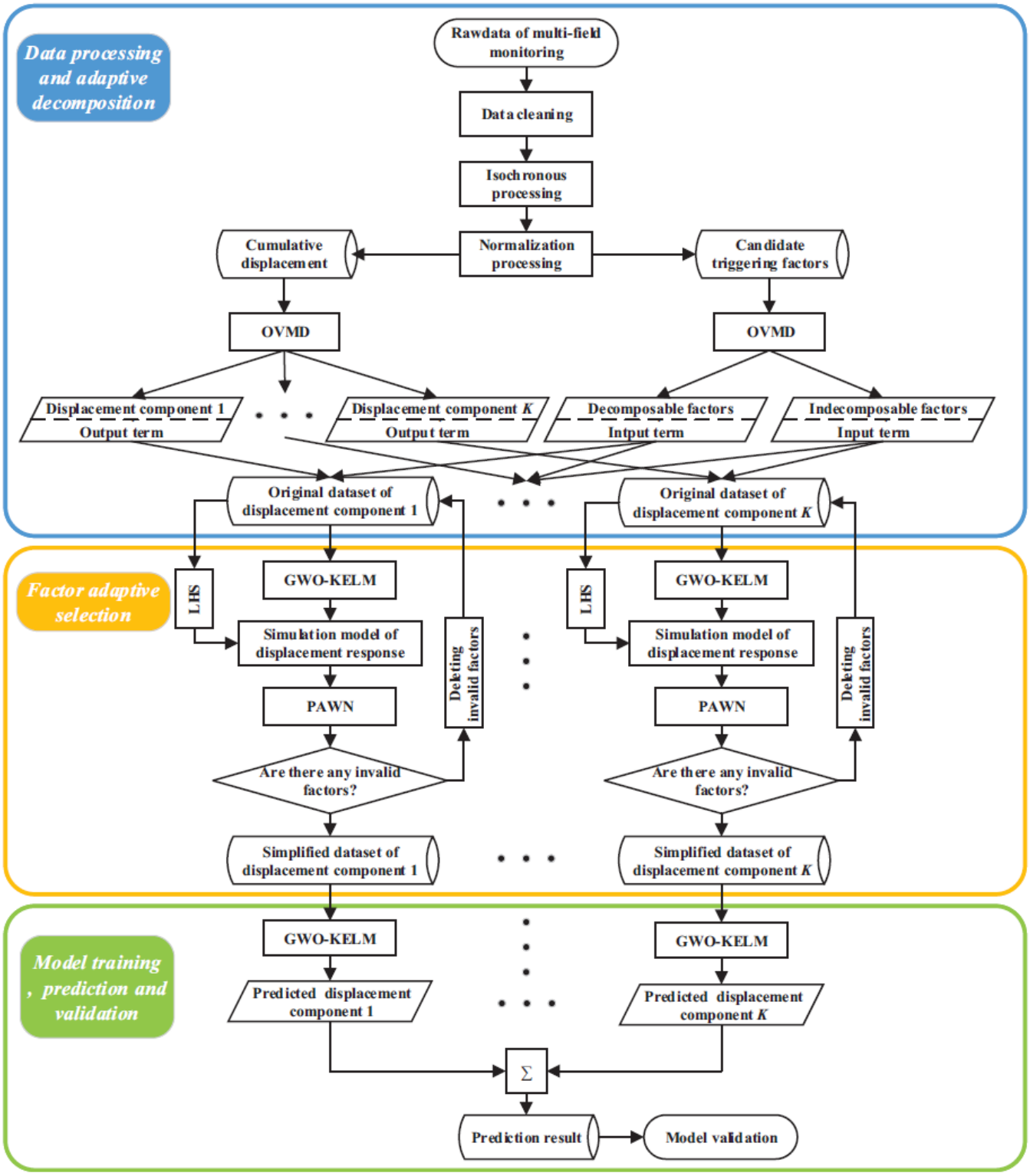

Figure 3

Schematic of the steps of the OVMD-GEO-KELM model

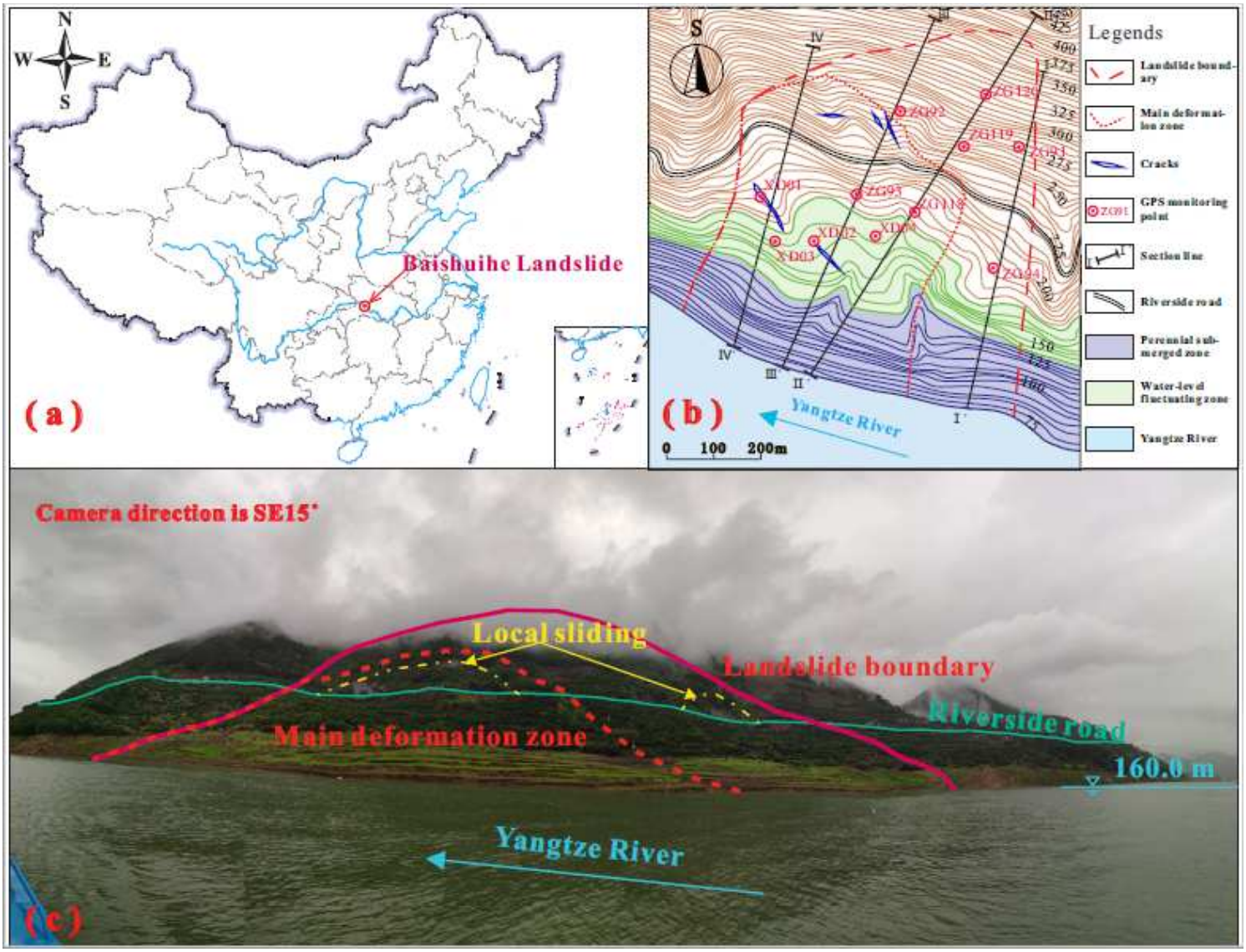

Figure 4

(a) Geographic location of the Baishuihe landslide in the TGR area, Hubei Province, China; (b) topographical map of the Baishuihe landslide; (c) panoramic image of the Baishuihe landslide 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.

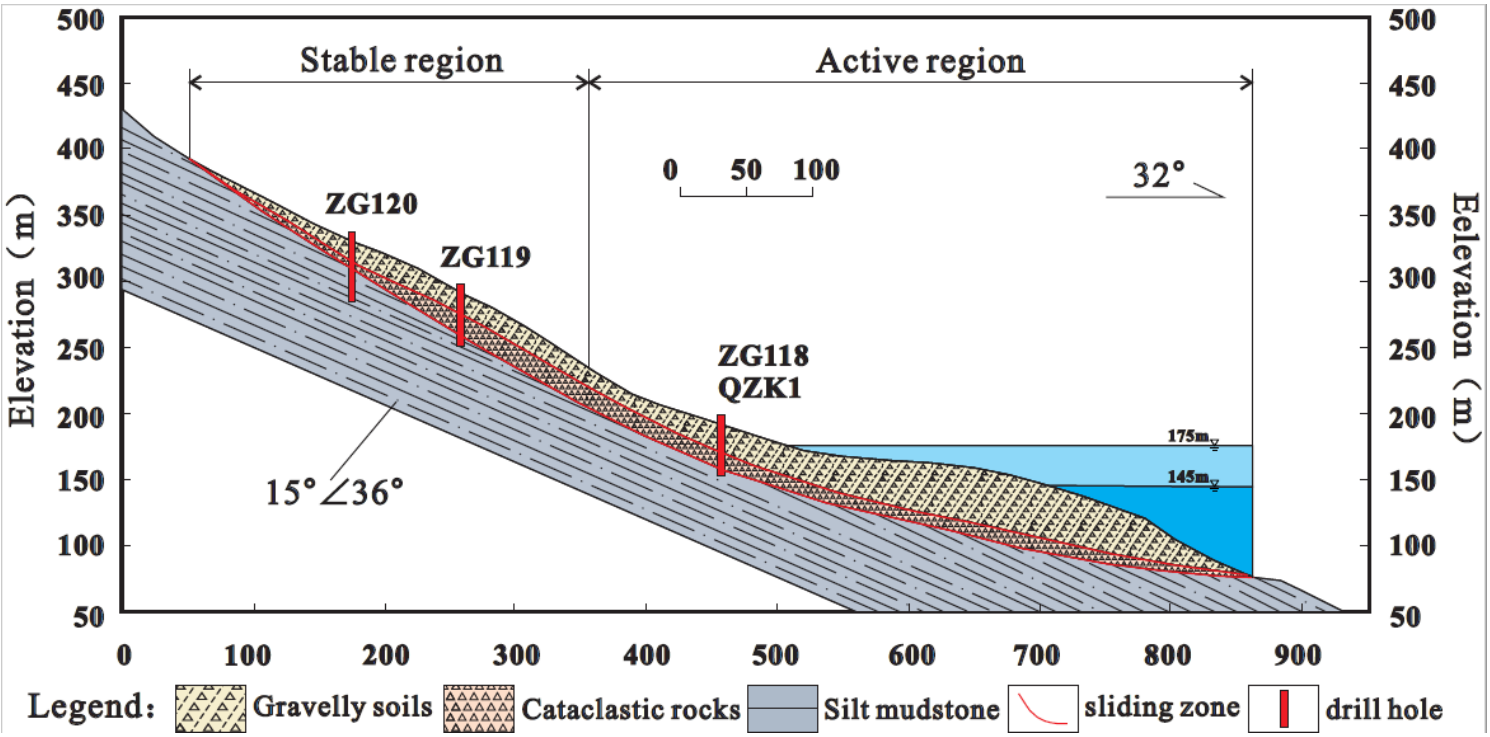

Figure 5

Monthly cumulative displacements, daily precipitation, and daily reservoir water level

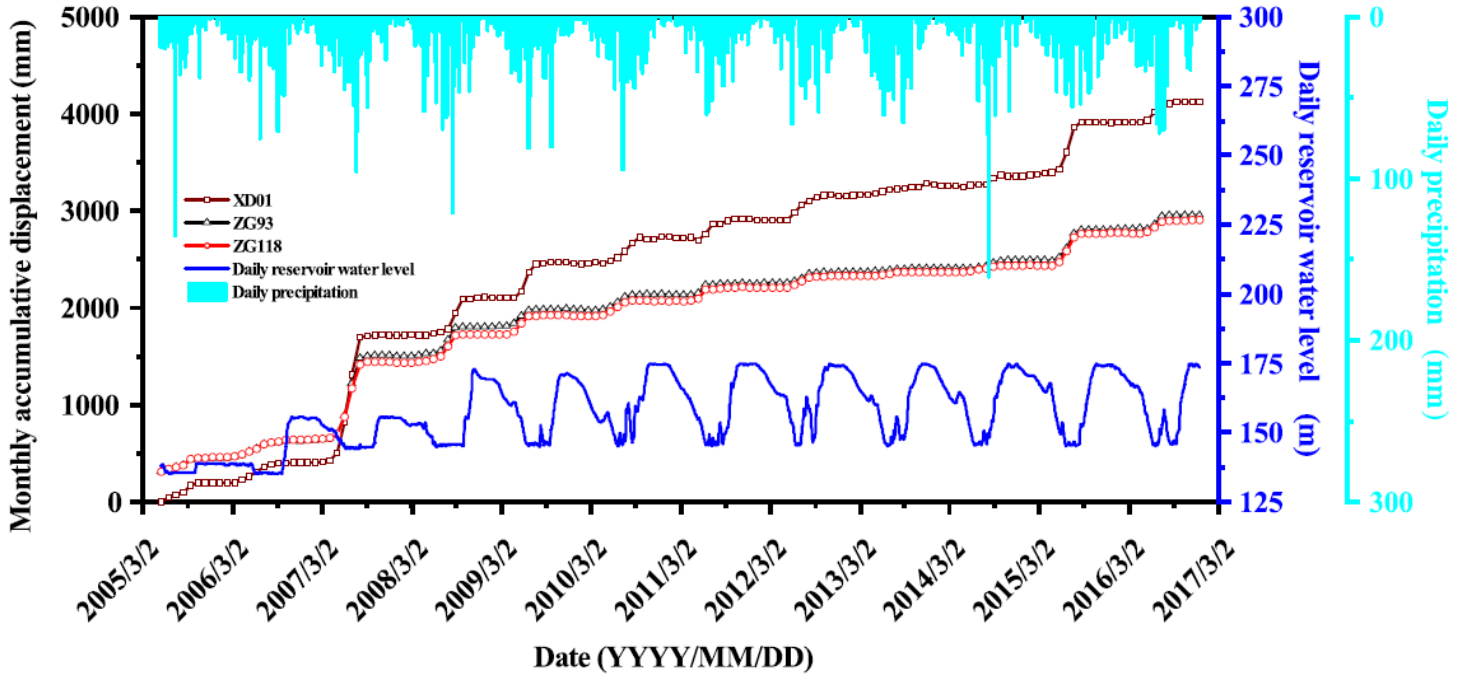

Figure 6

Monthly cumulative displacements, daily precipitation, and daily reservoir water level

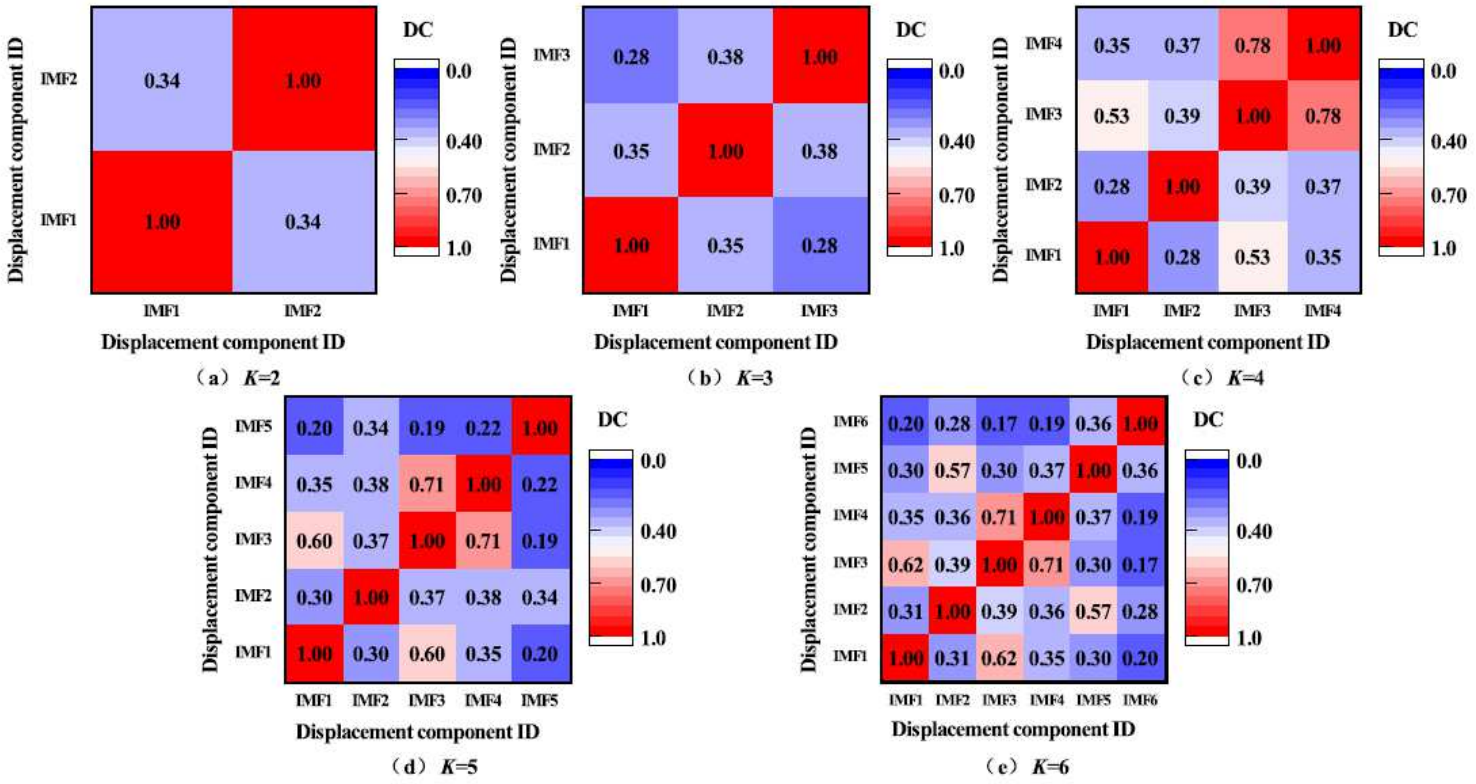

Figure 7

Distance correlation matrix between the different components obtained by using VMD based on the different settings of K(a=1)

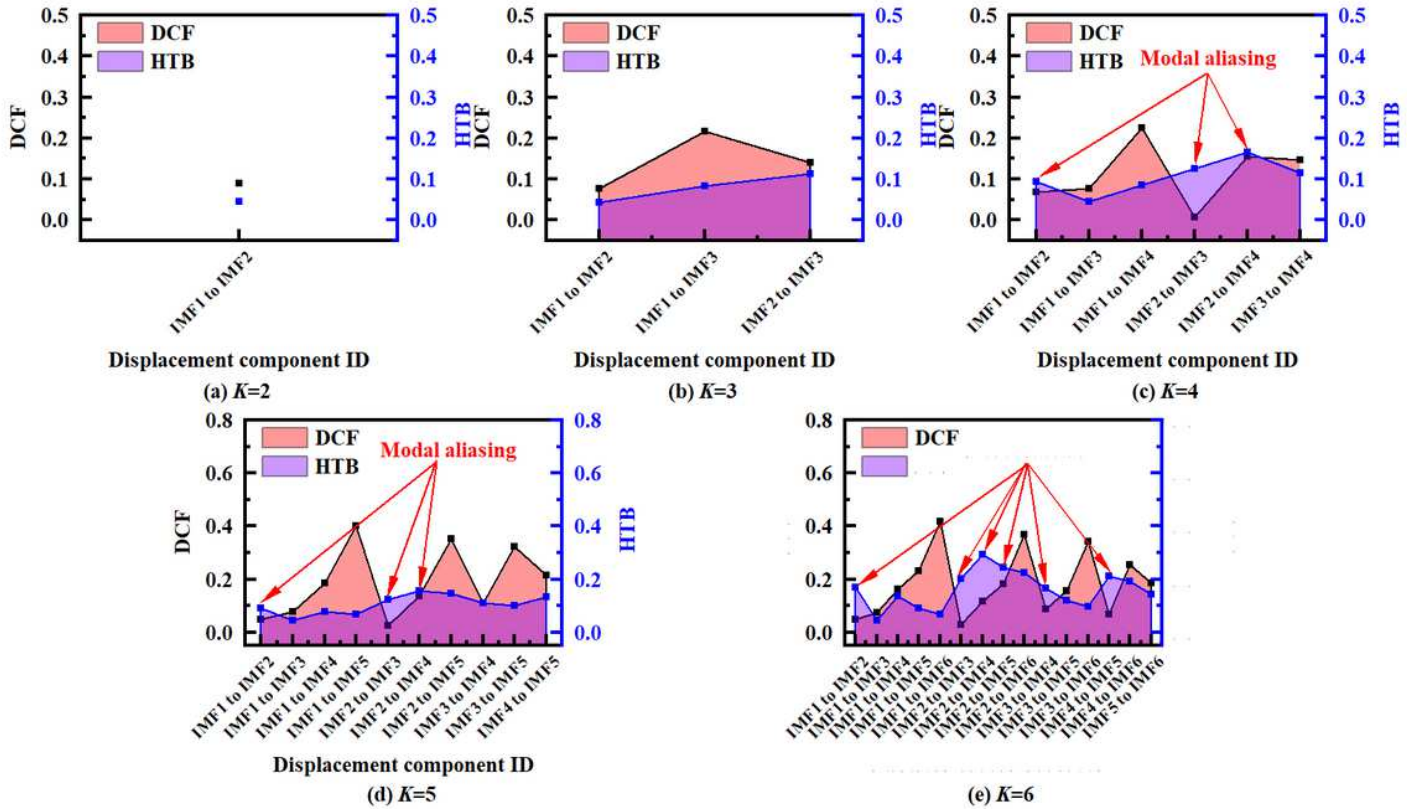

Figure 8

Normalized DCFs and their corresponding HSBs between the different components obtained by using VMD based on different settings of K(a=1)

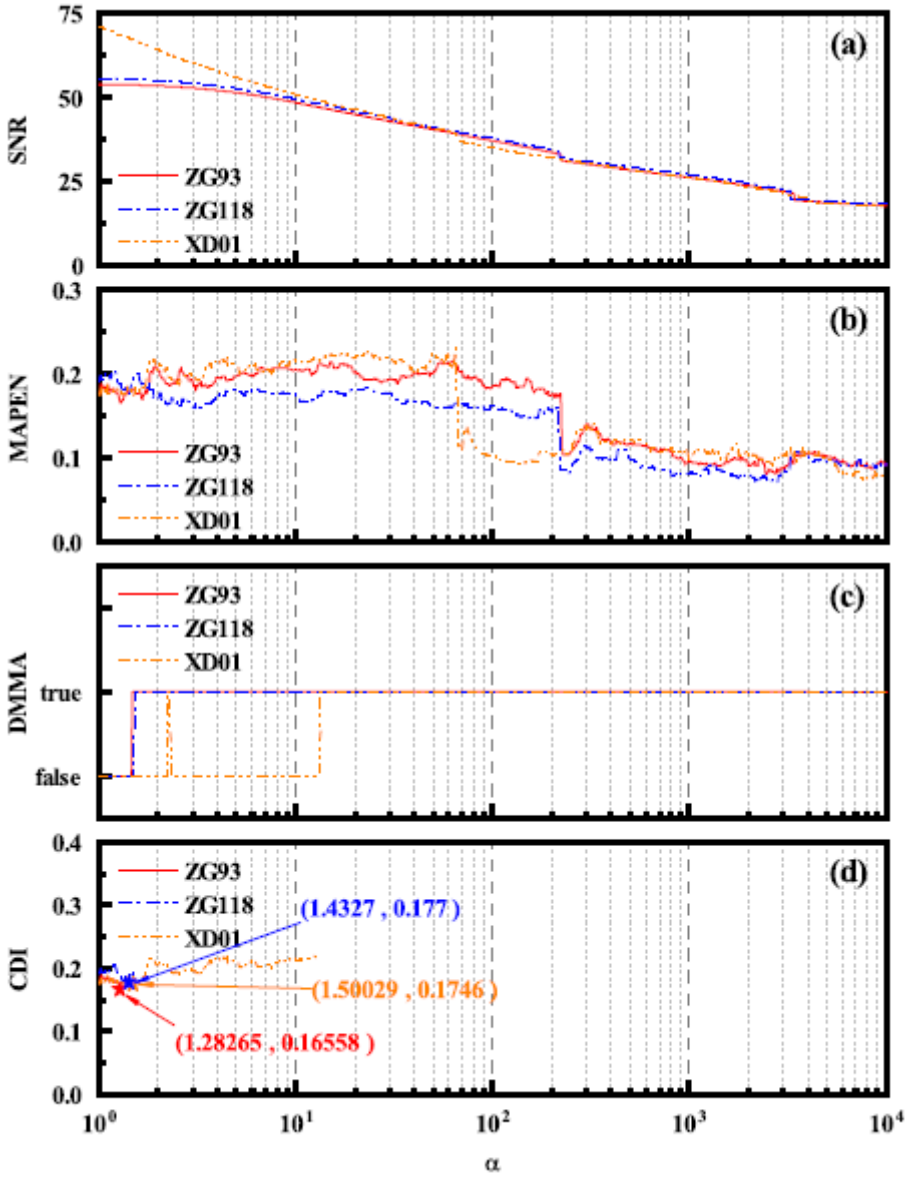

Figure 9

SNR, MAPEN, DMMA, and CDI of the decomposed components obtained by using VMD based on different settings of a (k=3)

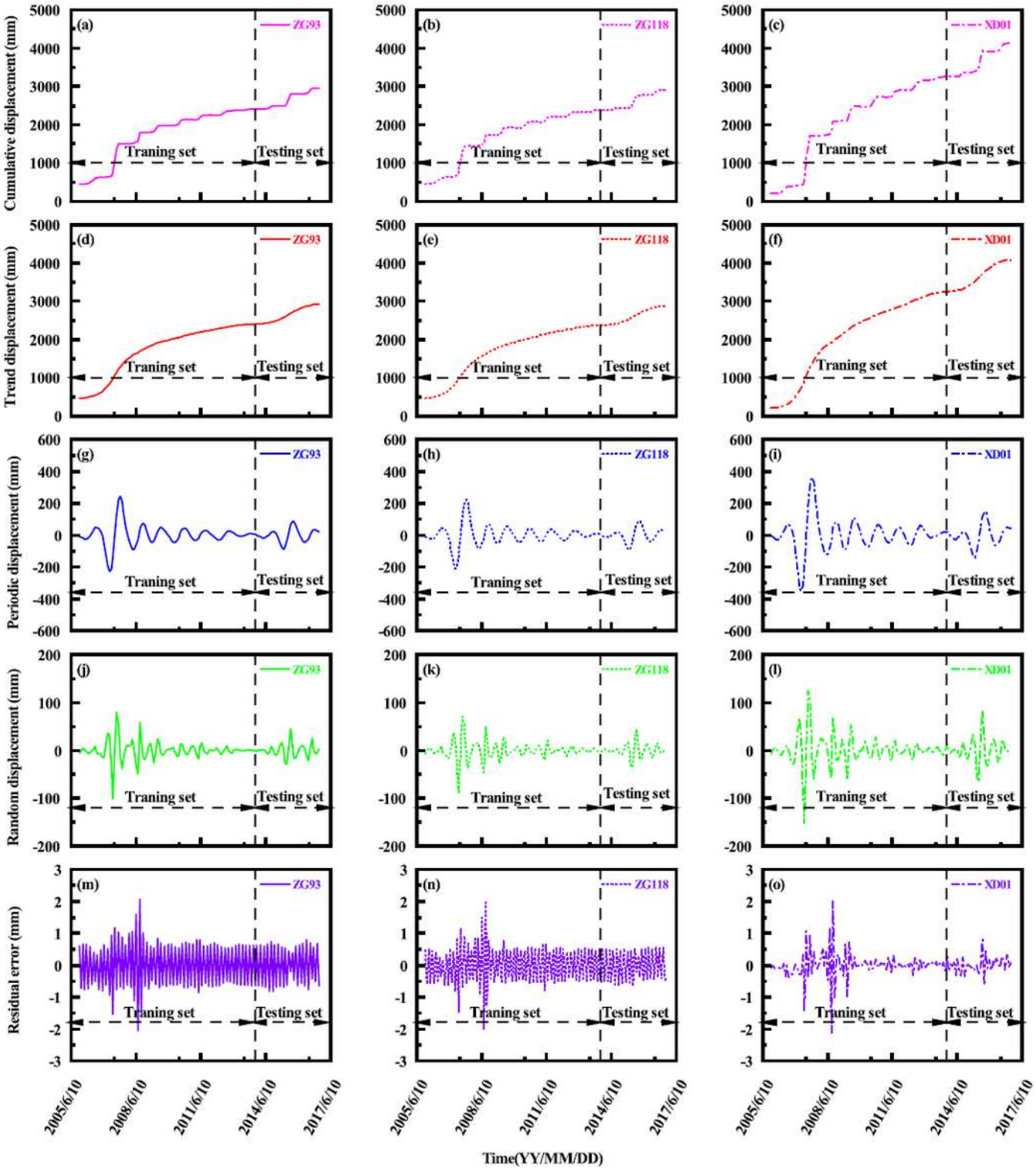

Figure 10

Decomposition results of displacement time series for monitoring sites ZG93, ZG118 and XD01

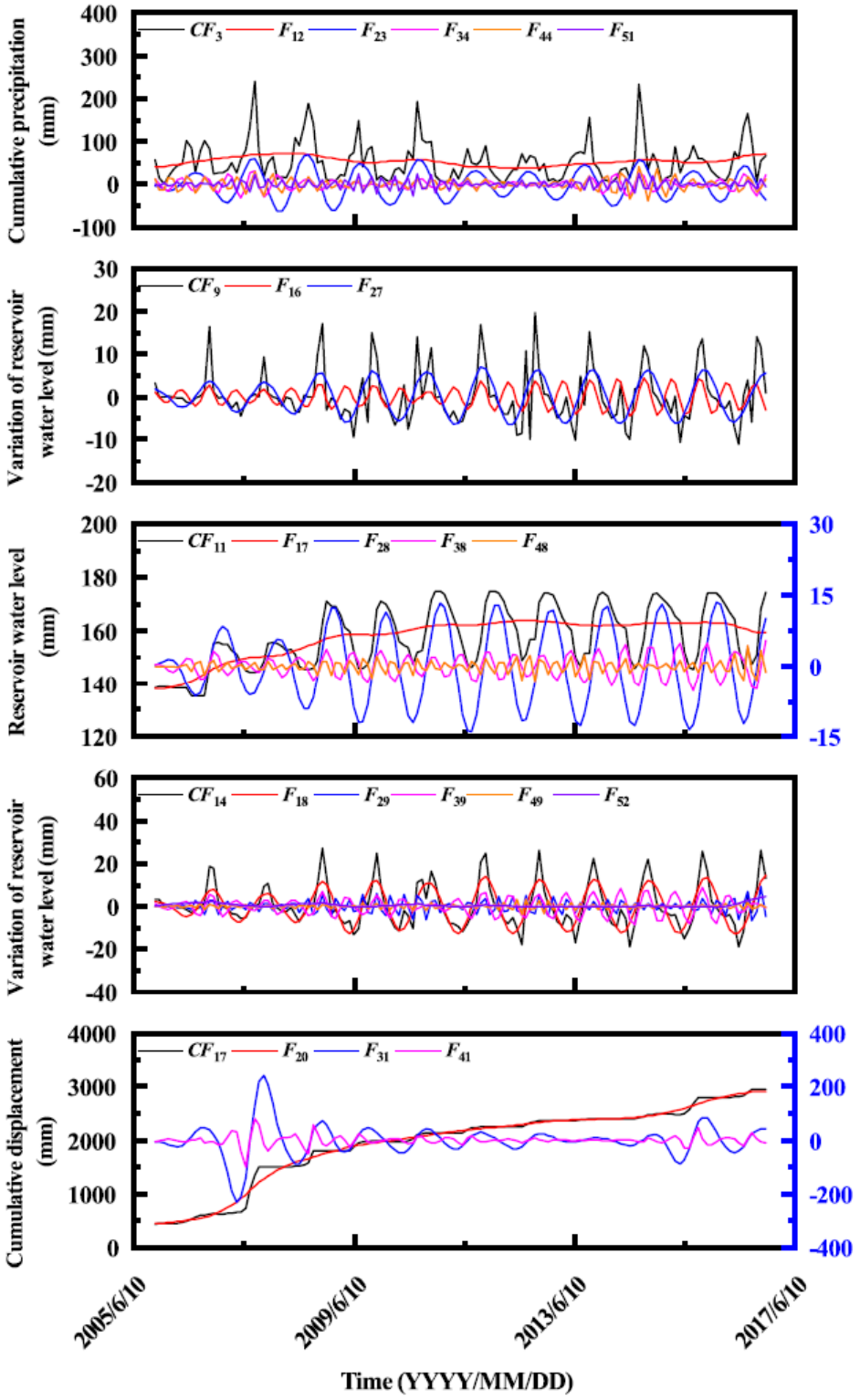

Figure 11

Representative decomposition results of decomposed factors

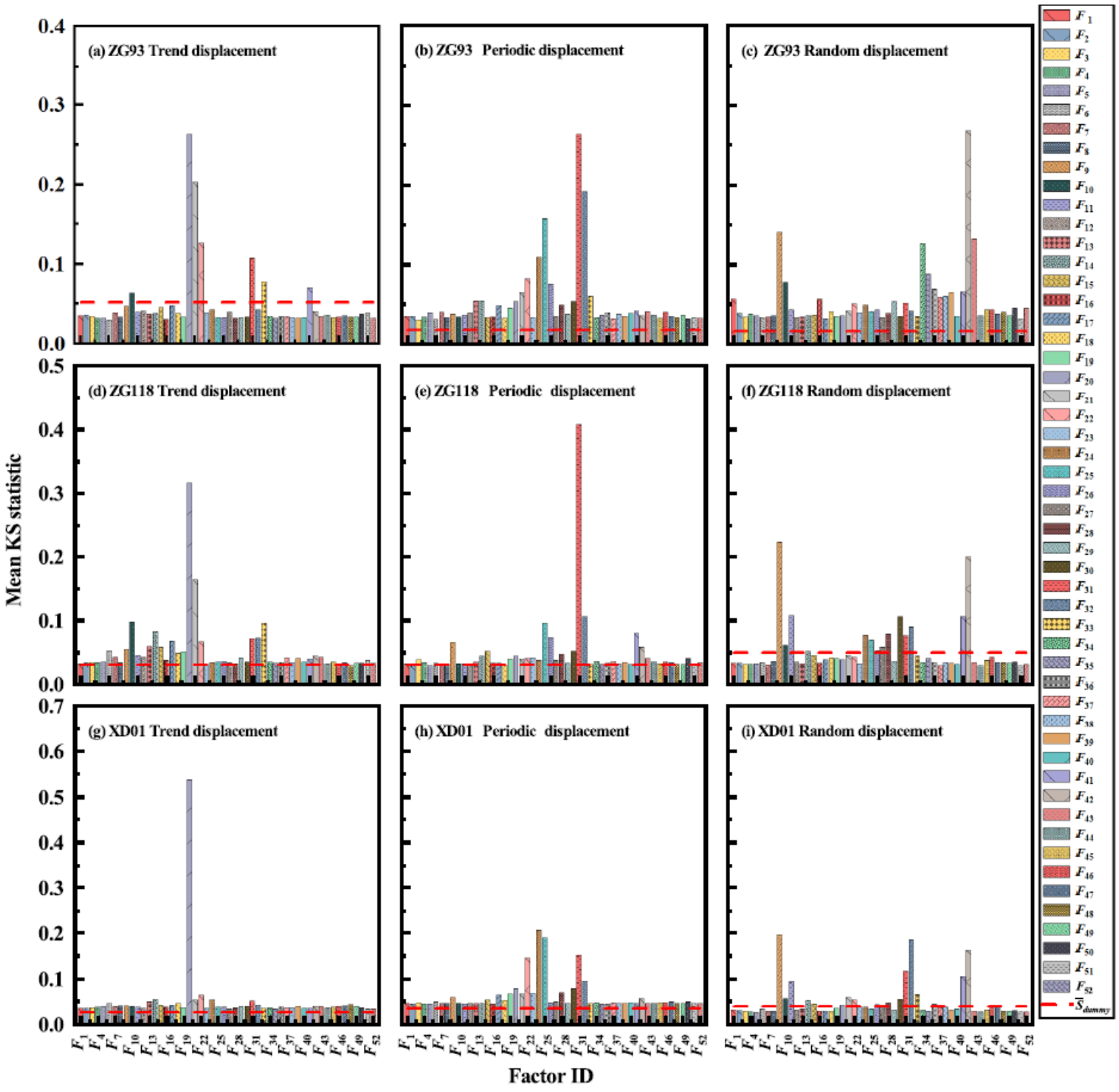

Figure 12

GSA results of the input inducing factors for the prediction of the different displacement components of monitoring sites ZG93, ZG118, and XD01

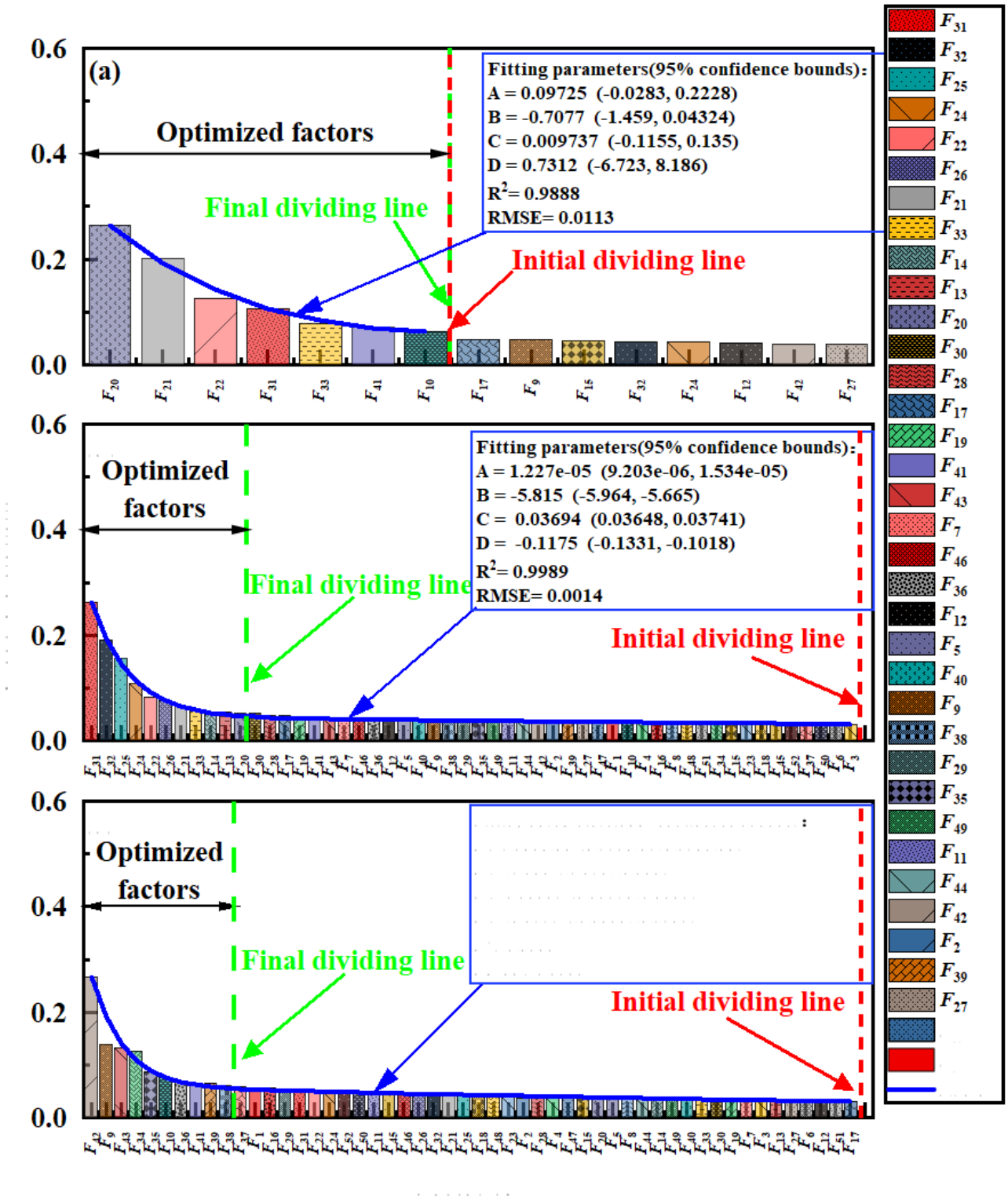

Figure 13

Division of the valid and invalid factors for the (a) trend displacement, (b) periodic displacement, and (c) random displacement prediction models

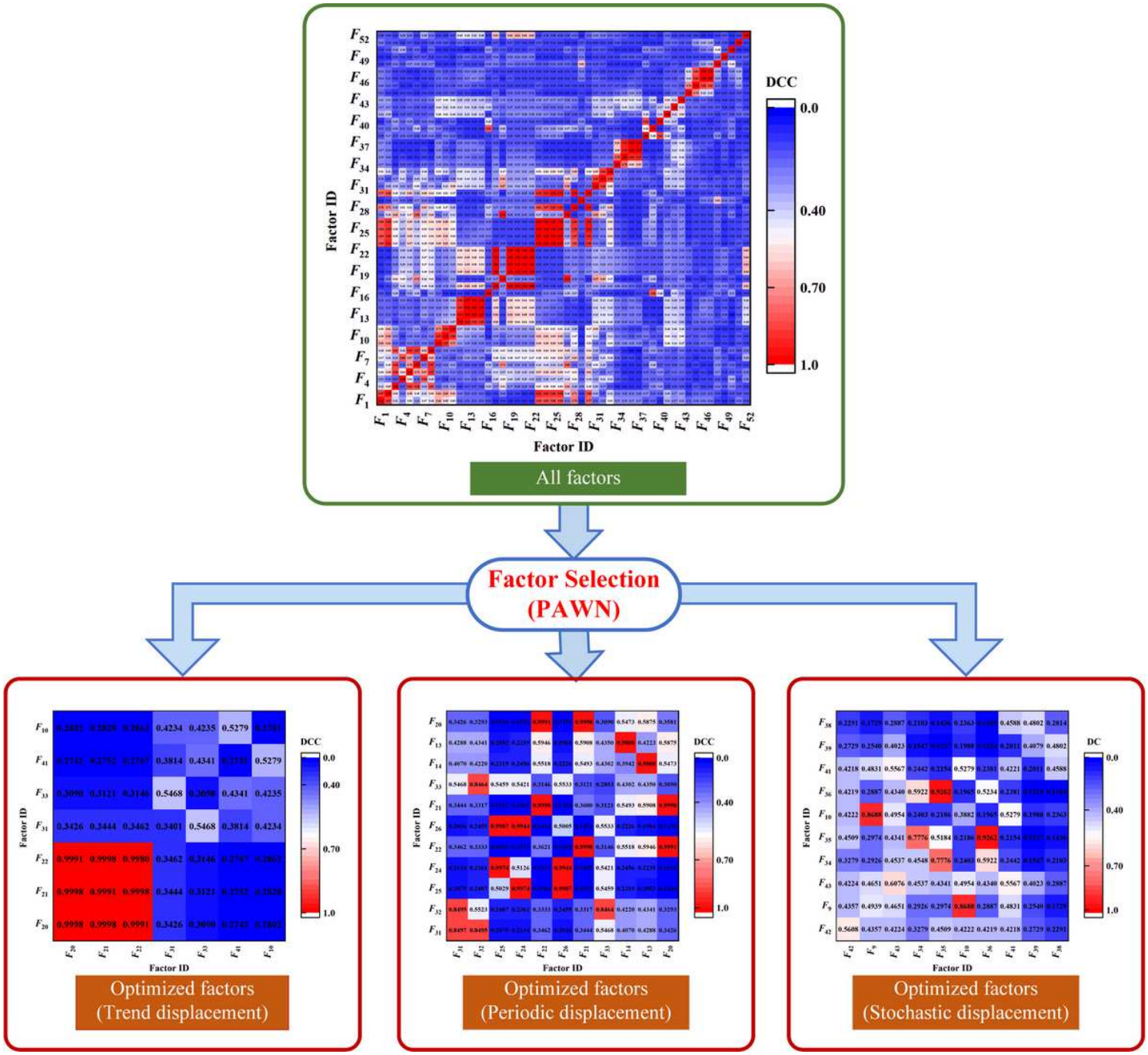

Figure 14

DCC matrixes of the reduced factors for the different displacement components

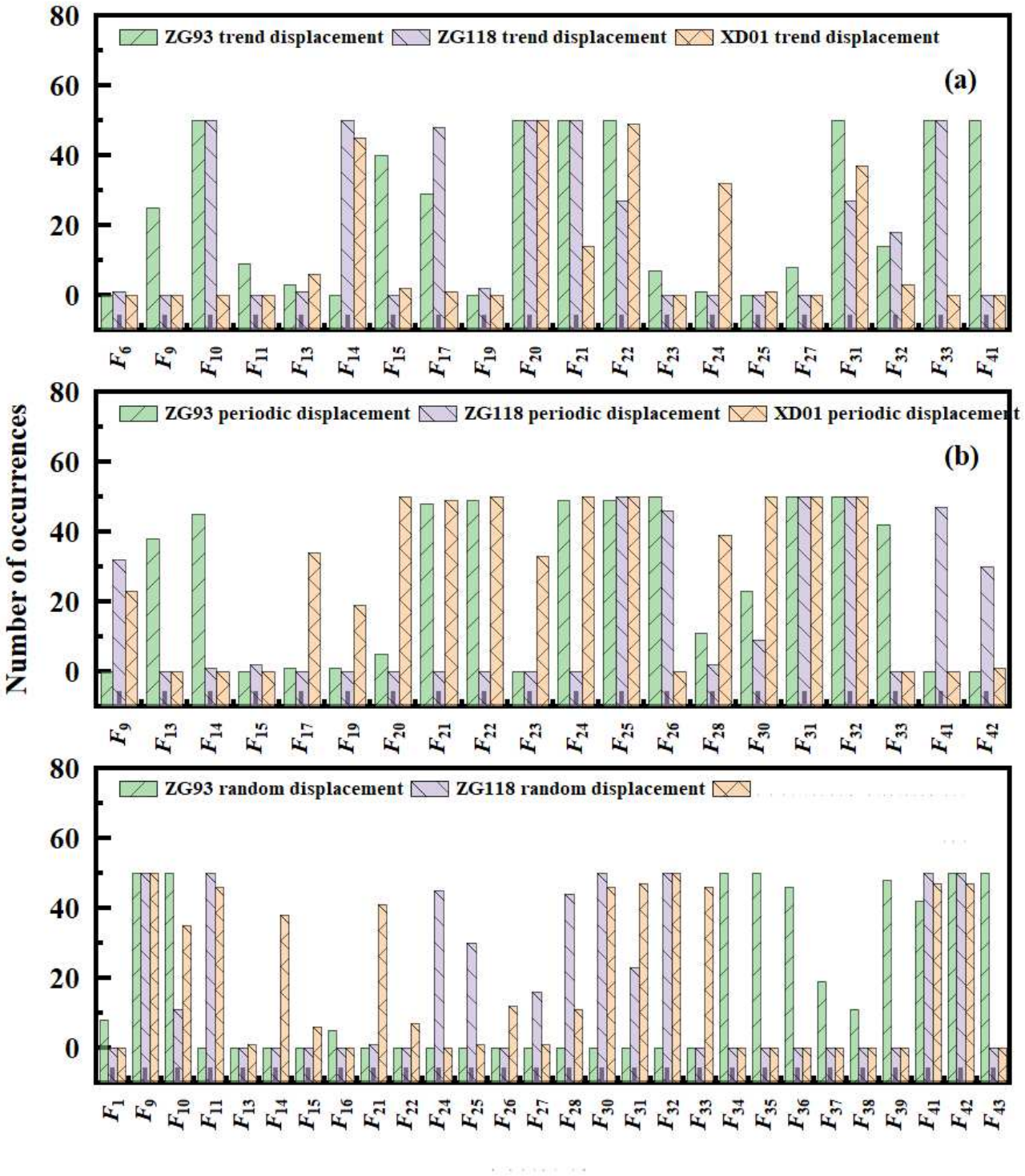

Figure 15

Reduction results of the input inducing factors for the prediction models after multiple calculations: (a) trend displacement prediction model; (b) periodic displacement prediction model; (c) random displacement prediction model

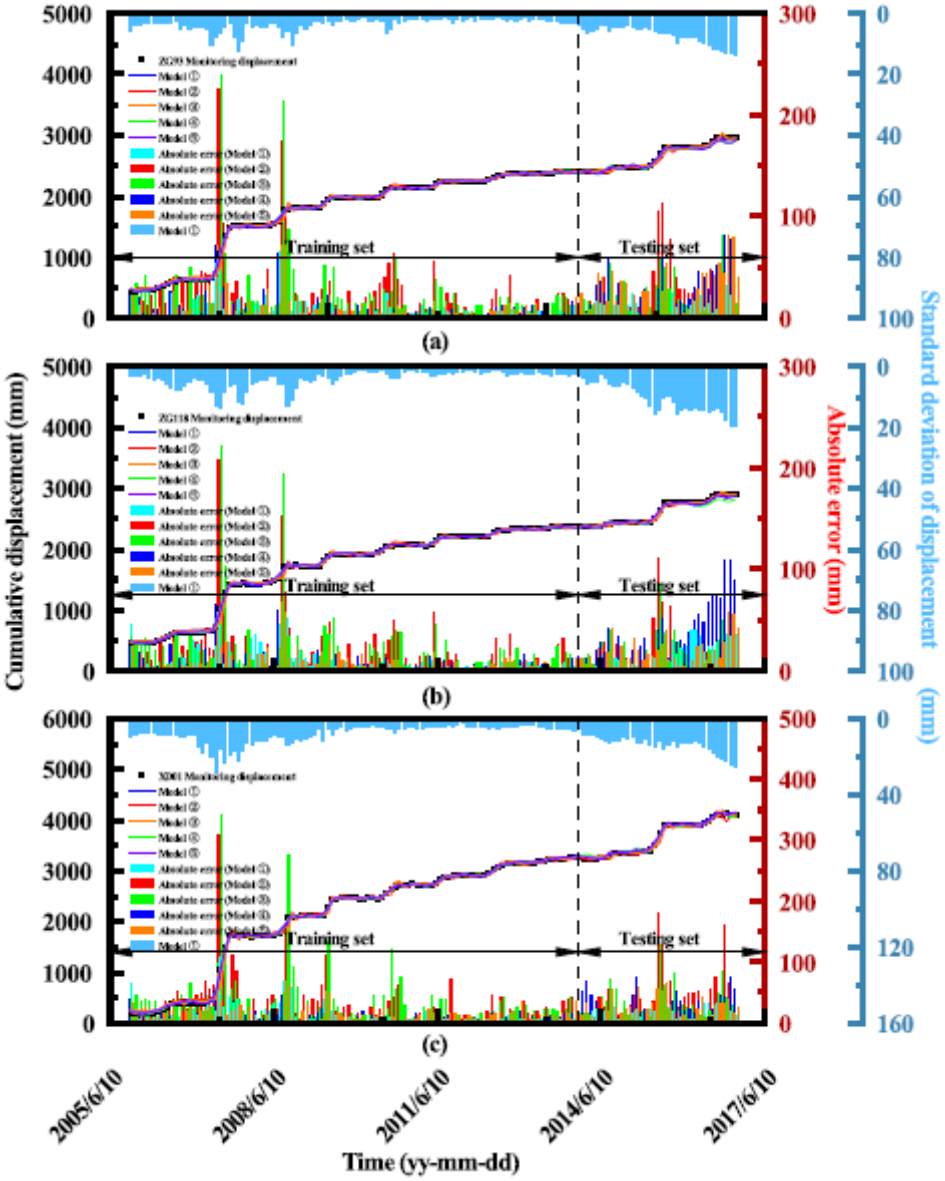

Figure 16

Comparison of the displacement prediction results and their prediction errors

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Tables.docx