ORIGINAL PAPER

Prediction of landslide displacement using multi-kernel extreme learning machine and maximum information coefficient based on variational mode decomposition: a case study in Shaanxi, China

Qing Ling[1,2] · Qin Zhang[2] · Jing Zhang[2] · Lingjie Kong[1] · Weiqi Zhang[3] · Li Zhu[4]

Abstract
Prediction of landslide movement is an efficient approach in the reduction in landslide risk. However, it is also a tough task due to the scientific challenges in forecasting a sophisticated natural disaster. This paper proposes a VMD-MIC-M-KELM (variational mode decomposition-maximum information coefficient-multi-kernel extreme learning machine) technique for prediction of landslide movements. The original displacement is first decomposed into a predefined number of components by VMD. Then, the triggers of each component are selected based on MIC between subseries and influencing factors. The decomposed terms are predicted by M-KELM respectively via k-fold cross-validation. Finally, predicted total displacement is achieved by summing up all forecasting subseries. A case study of Miaodian landslide (China) is presented for validation of the developed model. The verification results demonstrate the higher ability of the approach to forecast monthly displacement for periods up to 12 months as compared to the Poly-KELM and SVR models. Thus, improved monthly predictions may be achieved with constantly updated datasets from the monitoring system, which would offer reliable information for early warning of landslide.

Keywords Kernel extreme learning machine · Kernel functions · Variational mode decomposition · Maximum information coefficient · Displacement prediction

✉ Qin Zhang
dczhangq@chd.edu.cn

[1] School of Civil Engineering, Lanzhou University of Technology, Lanzhou Gansu, China

[2] College of Geology Engineering and Geomatics, Chang' an University, Xian Shaanxi, China

[3] The First Geodetic Team of the Ministry of Natural Resources, Xian Shaanxi, China

[4] Information Engineering University, Zhengzhou Henan, China

1 Introduction

Landslide is a very destructive natural disaster and has posed severe threat to humans, assets and the environment. Generally, high-precision monitoring data of landslide displacement are capable of representing the complicated deformation and failure characteristics of landslide and thus are an important index to evaluate landslide stability (Shihabudheen et al. 2017). Hence, predicting landslide deformation is regarded as one of the most effective ways for landslide disaster prevention and is also vital in avoiding or at least minimizing devastating influences on human lives and infrastructures. However, due to the adverse impacts of numerous factors, such as intrinsic geotechnical triggers (variations in landslide geometry and stress conditions, changes in material rheology, etc.) and external environmental factors (precipitation, widespread irrigation and snowmelt, etc.) (Guzzetti et al. 2005; Kawabata and Bandibas 2009; Bernardie et al. 2015; Intrieri and Gigli 2016), accurate prediction of landslide displacement is still a challenging task and has been a hot research topic in geological hazard worldwide.

At present, with respect to the study of landslide displacement prediction, researchers first decompose the original displacement into several subcomponents using methods like wavelet analysis (WA), empirical mode decomposition (EMD) and ensemble empirical mode decomposition (EEMD). Then, all the subseries are predicted separately. At last, the forecasting displacement is obtained by summation of all the predicted subcomponents, namely the almost reconstruction of the measurements (Du et al.2009; Xu et al.2011; Zhang et al.2015; Zhou et al. 2016; Shihabudheen et al. 2017; Deng et al.2017). As for prediction models, due to the influence of internal geological conditions and external factors (rainfall, temperature, reservoir water level, etc.), the cumulative displacement is a monotonically nonlinear time series, and its evolution contains multilevel information. Consequently, satisfactory results are not always achieved using creep theory-based physical models in numerical simulation or displacement–time-based statistical techniques in mathematics (Calvello et al. 2008; Federico et al 2012; Krkač 2015; Corominas et al. 2005; Sassa et al. 2010). With the development of artificial intelligence (AI) and machine learning (ML), many nonlinear intellegent models have wide application in prediction of landslide displacement, such as artificial neural network (ANN), the grey model, the support vector regression (SVR) model and the kernel extreme learning machine (KELM) model(Wu et al. 2007; Melchiorre et al. 2008; Alimohammadlou et al. 2014; San 2014; Liu et al. 2016; Pham et al. 2016; Zhou et al.2018).

Although outstanding achievements have been made in displacement–time series analysis and prediction models, there are still three problems to be addressed. First, the decomposed components divided by WA, EMD and EEMD are too more (more than five components) to realize the relationship between displacement components and factors due to the ignorance of physical meaning of each component. Second, the frequently used prediction models are constructed based on single factor or multiple factors selected by empirical methods. These models do not take into account the joint constraints of multiple influencing factors or the actual evolution of landslide deformation, resulting in low reliability of predictions. Finally, due to the influence of intrinsic factors and environmental triggers, landslide displacement exhibits several properties like complexity, randomness and uncertainty, leading to low universality of the developed model.

Therefore, to counteract the abovementioned issues, we propose a novel model, i.e., VMD-MIC-M-KELM, for the purpose of forecasting landslide movements which has not yet been studied or put forward before. This new model adopts VMD to decompose the

original displacement into three sub-components with physical meaning of each term, and selects the corresponding influence factors by MIC (Reshef et al. 2011). Moreover, the M-KELM model is developed to predict subcomponents by linearly integrating the linear and nonlinear kernel function-based ELMs. Then, the predictions are achieved by summation of all the predicted subcomponents. The Miaodian landslide of Jingyang in Shaanxi Province (China) is adopted to validate the ability of the developed technique. The results show that the proposed model is capable of realizing the relationship between displacement components and factors with artificially defined number of decomposition and highlighting the temporal evolution of landslide displacement with representative triggers. Further, the developed approach also achieves well applicability in prediction with multi-kernel functions. Besides, traditional models such as the Poly-KELM and SVR are applied for comparison. The results indicate that the VMD-MIC-M-KELM model performs better as compared with the other two techniques.

2 Methodology

2.1 Variational mode decomposition (VMD) algorithm

VMD decomposes the original displacement $f(t)$ into a given number (k) of band-limited subsignals or modes (u_k) (Dragomiretskiy and Zosso 2014). Each mode (u_k) is mostly compact around a center frequency ω_k. On dividing the displacement into components, valuable information concealed in the displacement can be retrieved (Dragomiretskiy and Zosso 2014).

Assuming that the landslide displacement $f(t)$ is equal to the summation of k decomposed components, the constrained problem is expressed by:

$$\begin{cases} \min_{u_k, \omega_k} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k(t)} \right\|_2^2 \right\} \\ s.t. \sum_{k=1}^{K} u_k = f \end{cases} \tag{1}$$

where $\delta(t)$ is the Dirac distribution and t is time script. The solution to the original minimization problem (1) is achieved by a chain of iterative steps as explained in Dragomiretskiy and Zosso (2014) and Dragomiretskiy (2015).

2.2 Multi-kernel-based extreme learning machine (M-KELM)

ELM is an efficient SLFN (single-hidden-layer feed-forward network) with randomly produced hidden nodes (Huang et al 2006). Despite its attractive attributes such as higher generalization performance, better accuracy, faster speed in contrast to traditional neural network, there are still some deficiencies to be addressed. For instance, it is difficult to determine the number of hidden layers. The solution equation may be ill-conditioned due to the singular output matrix of hidden layers. In addition, there still exists model overfitting since the optimum output weight of ELM is obtained by the Moore–Penrose inverse without adding regularization parameter (Ranjeeta et al.2018; Fang et al. 2020). Therefore,

to further enhance the generalization ability and stability of ELM, a kernel-based ELM (KELM) was proposed by Huang et al. in 2012, which is inspired by the kernel functions of support vector machine. The formula of KELM is as follows:

$$f(x) = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T (\frac{I}{C} + \Omega_{ELM})^{-1} Y \tag{2}$$

where $\frac{I}{C}$ is a positive value added to the diagonal of HH^T based on ridge regression, and Y is the output vector. Ω_{ELM} is the kernel matrix for KELM which can be defined as follows:

$$\begin{cases} \Omega_{ELM} = HH^T \\ \Omega_{ELM}(i, j) = h(x_i)h(x_j) = K(x_i, x_j) \end{cases} \tag{3}$$

The performance of KELM depends on the kernel function which transforms low-dimensional nonlinear data into high-dimensional linear data. Hence, proper kernel function can significantly improve model generalization capability. At present, the frequently used kernel functions satisfying Mercer's conditions and its parameters are shown in Table 1.

The kernel functions in Table 1 can be grouped in three varieties. The first category is the global kernel function with strong generalization and weak learning ability like polynomial kernel function and sigmoid kernel function. The second type is the local kernel with weak learning ability and strong generalization, such as Gaussian radial basis kernel function. The last is the wavelet kernel (Morlet) with multilevel and multiresolution properties which can approximate any function accurately (Fig. 1, test point $x = 0.3$, parameter $c = -1.5$). Thus, researchers can select suitable kernel-based ELM to predict landslide displacement. However, the intrinsic factors and external factors of landslides have various characteristics such as complexity, diversity and randomness, which lead to poor performance in prediction by adopting one single kernel based ELM. Consequently, a new multi-kernel ELM (M-KELM) model is developed by linearly integrating all the kernel functions in Table 1 for accurately predicting future evolution and development trend of landslide.

The newly developed multi-kernel function is as follows:

$$K(x, x_j) = \mu_1 K_P + \mu_2 K_g + \mu_3 K_s + \mu_4 K_w \tag{4}$$

Subject to $\sum_{i=1}^{4} \mu_i = 1$.

Table.1 Various kernel functions and their parameters

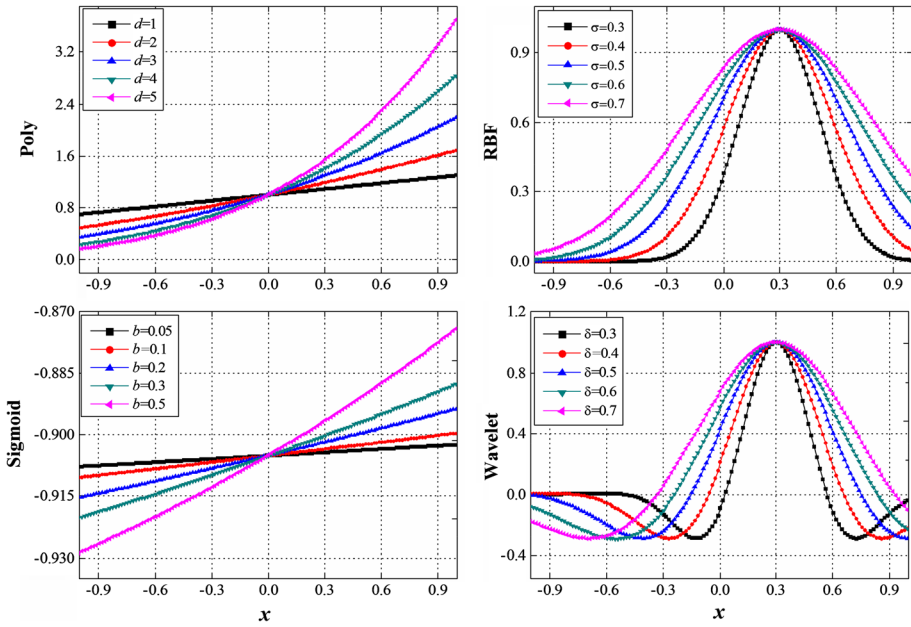| Kernel functions	Equations	Parameters
Polynomial kernel (Poly)	$K_P(x, x_j) = (xx_j + 1)^d$	d
Gaussian radial basis kernel function (RBF)	$K_g(x, x_j) = \exp\left(-\left\|x - x_j\right\|^2 / \sigma^2\right)$	σ
Sigmoid kernel (Sigmoid)	$K_s(x, x_j) = \tanh(bxx_j + c)$	b, c
Wavelet kernel (Morlet)	$K_w(x, x_j) = \prod_{j=1}^{d}\left(\cos\left(1.75 \times \frac{\|x - x_j\|}{\delta}\right) \exp\left(-\frac{\|x - x_j\|^2}{2\delta^2}\right)\right)$	δ

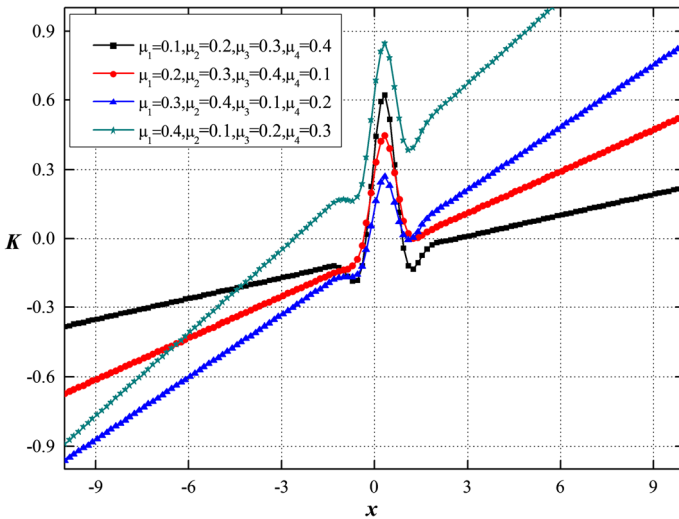**Fig. 1** Characteristic curve of single-kernel function

Fig. 2 Characteristic curve of multi-kernel function with different weighting factors

Figure 2 shows the characteristic curve of multi-kernel function with predefined parameters (test point $x = 0.3$). The multi-kernel values with various weighted factors are greatly affected regardless of the distance from the test point as depicted in Fig. 2. Further, the approximation performance is also significantly improved. Accordingly, the

developed multi-kernel function integrates the advantages of all single kernels as shown in Table 1.

In addition, eight parameters, including four kernel parameters (σ,b,c and d) and four weight factors (e.g. μ_1,μ_2,μ_3,μ_4), should be appropriately selected to enhance performance of predicting via k-fold cross-validation. Thus, in this paper artificial bee colony algorithm (ABC) is applied for optimization as explained in Karaboga and Bandibas (2009) and Karaboga et al. (2014).

2.3 The developed coupling technique and evaluation metrics

This paper proposes a model, i.e., VMD-MIC-M-KELM, for predicting of landslide movements. Poly-KELM and SVR are adopted for comparison. All hyperparameters related to the models discussed in this paper are optimized by ABC algorithm via nine-fold cross-validation. The flowchart of the developed approach is shown in Fig. 3.

Four measures, i.e., the root-mean-square error (RMSE), absolute percentage error (APE), mean absolute error (MAE) and goodness of fit (R^2), are applied for evaluating the performance of the proposed approach. The detailed mathematical expressions are as follows:
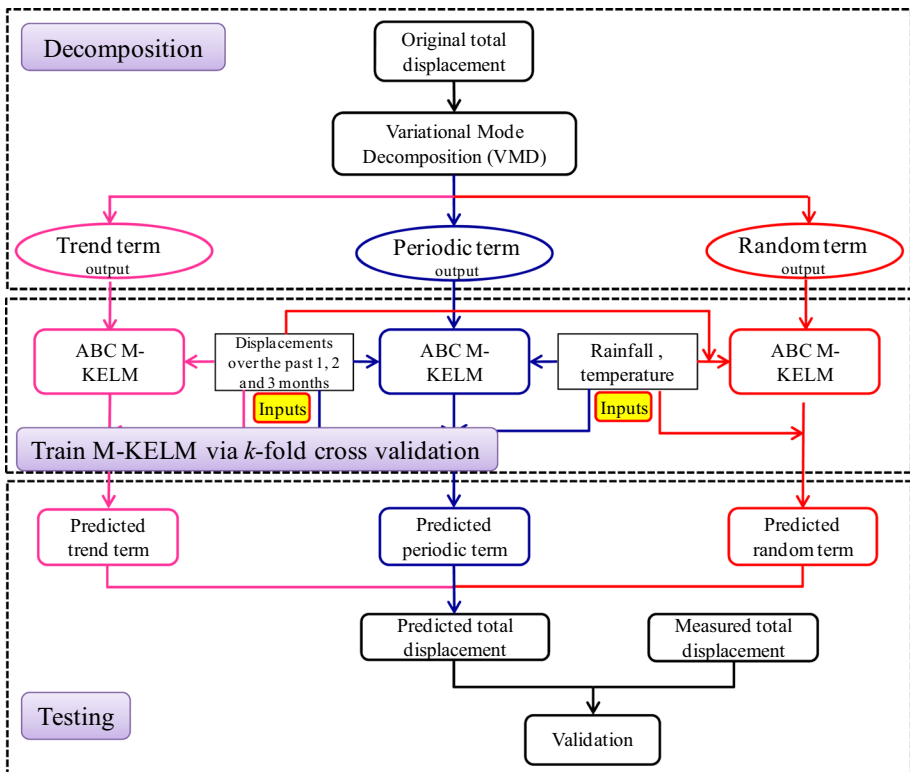

Fig. 3 Proposed approach for landslide displacements prediction

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{L}_i - L_i)^2} \tag{5}$$

$$APE = \left| \frac{\hat{L}_i - L_i}{L_i} \right| \tag{6}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| \hat{L}_i - L_i \right| \tag{7}$$

$$R^2 = \frac{(n \sum_{i=1}^{n} \hat{L}_i L_i - \sum_{i=1}^{n} \hat{L}_i \sum_{i=1}^{n} L_i)^2}{\left(n \sum_{i=1}^{n} \hat{L}_i^2 - (\sum_{i=1}^{n} \hat{L}_i)^2 \right) \left(n \sum_{i=1}^{n} L_i^2 - (\sum_{i=1}^{n} L_i)^2 \right)} \tag{8}$$

where n is the number of measurements, L_i is the observations and $\hat{L}_i$ is the predictions.

3 Case study

3.1 Geological conditions

The study region is in the lower reaches of Jing River, the central part of Shaanxi Province belonging to the Northwest Loess Plateau region of China. The landslide covers an area of $6.2 \times 10^4 m^2$ with a maximal width and length of ~267 m and ~227 m, respectively. The topographic inclination follows the SW direction. Both sides of the slope are high and abrupt, and the rear is nearly vertical. The abrupt terrain conditions are favorable for the development of landslides.

The Jingyang landslide comprises three layers, i.e., Malan loess, Paleosol, and Lishi loess. Vertical joints are well developed in the loess layer which has high permeability, offering favorable conditions for groundwater seepage. Paleosol has low permeability and is thus regarded as an impermeable layer. Geographically, the region is a temperate continental monsoon climate. Two distinct seasons are reported in this region, namely hot and rainy summer, and cold and dry winter. Heavy and concentrated precipitation of 100–200 mm/month has been registered from July to September. In the rainy season, the highest daily precipitation of over 80 mm has been recorded. During dry period, average precipitation less than 20 mm/ month has been reported in five months (i.e., November to March), where November and December have the lowest and even no rains. The temperature ranges from − 20.8 to 41.8 °C. Further in one year there are frost-free period of 213 days and frost duration of 51 days with a maximum frozen soil depth of 44 cm. Besides, the groundwater level changes with seasonal precipitation, which is also one of the main factors that poses adverse influence on landslide displacement (Cao et al. 2016).

3.2 The evolution of landslide deformation

To assess the stability of Miaodian landslide, we constructed a deformation monitoring network in June 2015 (Fig. 4d). Twelve monitoring stations were built in the study region, including nine monitoring stations (MD01-MD09) and three datum stations (JZ01, JZ02 and JZ03) established on stable area. The Leica TS30 measurement robot (Switzerland), with a nominal accuracy of 0.5″ (angle surveying), 0.6 mm+1 ppm×d (precise distance surveying) and 0.6 mm+1 ppm×d (standard distance surveying) (d is the measured distance, unit km), was adopted to monitor landslide displacement with high precision. A local independent coordinate system was applied with coordinate origin outside monitoring region. The original coordinate of JZ02 was set as (500,500,300). The length and geodetic azimuth angle from JZ01-JZ02 were achieved by GNSS relative positioning technique. Further, for validation the high-precision total station was employed to measure the
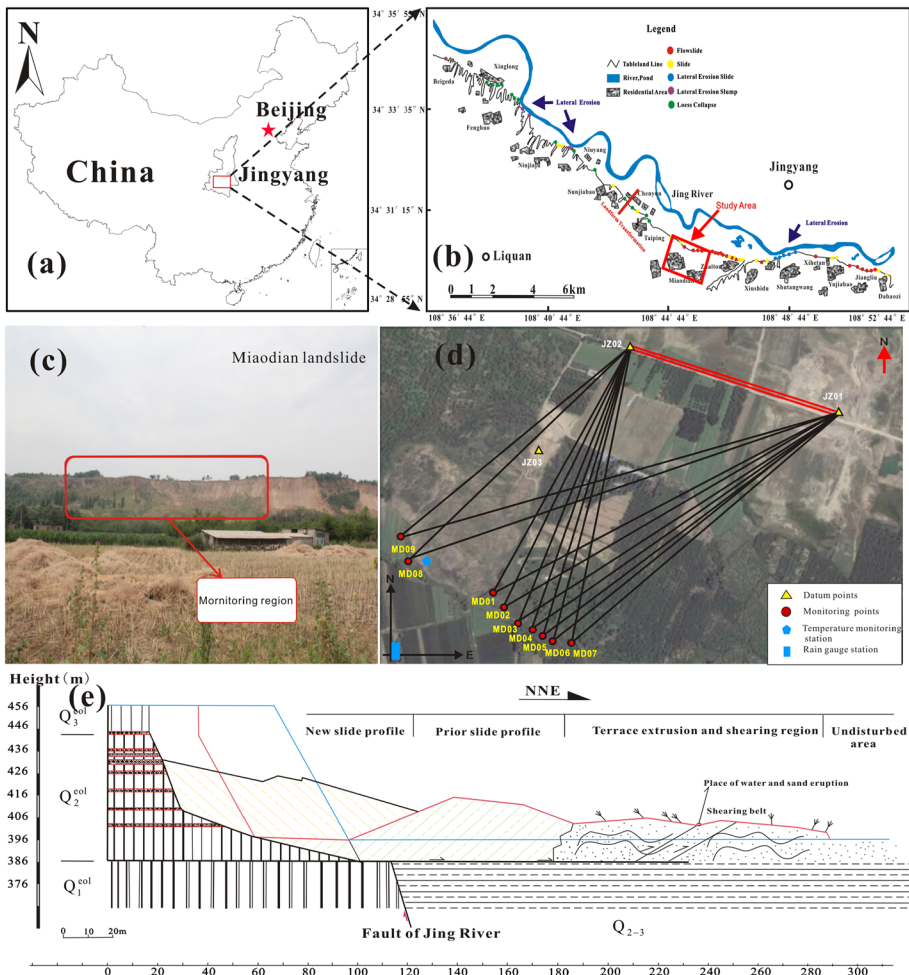
Fig. 4 **a** Map of China; **b** geographic locations of Miaodian landslides; **c** photograph of the Miaodian landslide; **d** deformation monitoring network of Miaodian landslide; **e** geological profiles of Miaodian landslide

length between the two datum points. Accordingly, the coordinate of JZ01 in independent coordinate system was obtained. Due to regular deformation monitoring since May 2015, measurements of the landslide (plane precision better than ± 5 mm, height precision better than ± 3 mm) from May 2015 to February 2018 were available. In addition, the data of temperature and rainfall were gathered from the nearby meteorological stations and temperature and rain gauge stations established by surveyors in Jingyang (Fig. 4).

The observations of nine stations are depicted in Fig. 5. The results indicate that each point had certain degree of deformation, where MD09 had an apparent S-shaped curve. The total displacement of MD09 was up to − 608 mm until February 2018. Initially, the displacement of MD09 increased slowly, and then it became faster before gradually slowing down. MD09 tended to deform on favorable circumstances (e.g., precipitation and temperature variations) due to its special location as depicted in Fig. 5. Accordingly, we choose MD09 station for detailed analysis.

Further, to explore deformation mechanisms and the triggers of Miaodian landslide, data on monthly displacement, largest daily rainfall and monthly cumulative precipitation at MD09 are plotted in Fig. 6, and the associations among these triggers are summarized as follows.

Initially, the precipitation was heavy, whereas small deformation was developed from July to October 2015. Due to significant transformation of stress field, seepage field and strata structure after landside occurred in June 2015, the slip mass demonstrated gentle response to the variations of external triggers. Hence, the development of landslide deformation had little correlation with precipitation changes during this period.

In the second phase, rainfall was gentle but large deformation occurred from November 2015 to May 2016, which indicated precipitation exerted little impact on landslide movement. According to collected data in the study region, winter (Early November–Late December) and summer irrigation (Early March–Middle April) was implemented during this period. Widespread irrigation may have generated rapid increase in groundwater level and pore pressure. Further, high saturation at the initial position of irrigation would have resulted in obvious vertical deformation of the area. As depicted in Fig. 6, two peak values had been registered during this period, one of which is bigger with displacement of 58 mm. In addition, precipitation also retards impact on landslide deformation. For instance, due to heavy rainfall in the previous month, high deformation was recorded in November 2015.

Fig. 5 Monitoring curves of accumulated displacement of landslides for nine stations
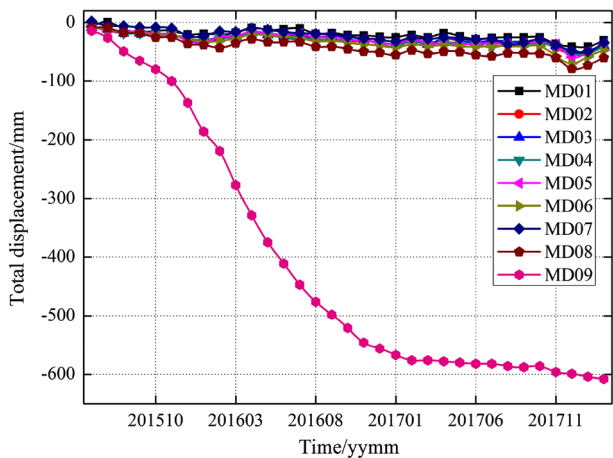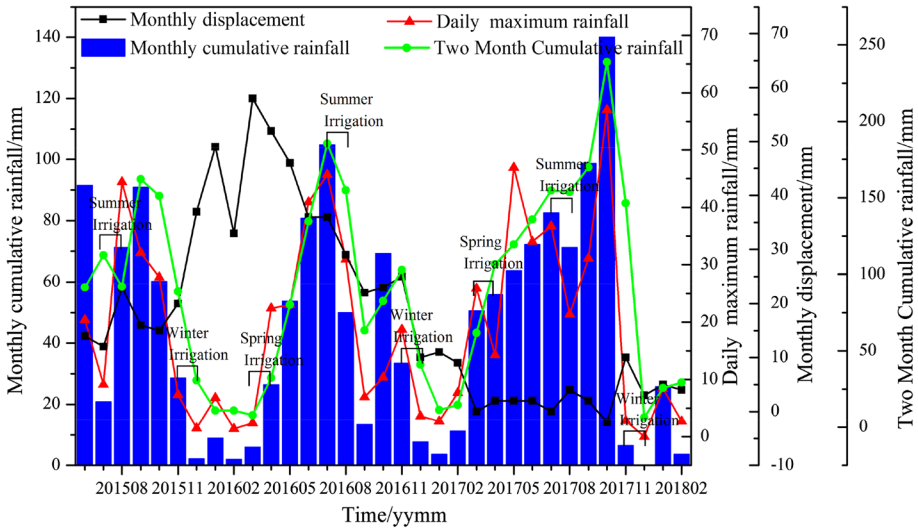
Fig. 6 Monthly cumulative rainfall, daily maximum precipitation and monthly displacement monitoring data of station MD09

Similar patterns can also be found in November 2016 (Fig. 6). Accordingly, the influence of widespread irrigation and lagged impact of rainfall accounted for the deformation in this period.

In the third phase, from June 2016 to February 2017, observations in Fig. 6 show that the monthly displacement exhibited close correlation with rainfall intensity. Precipitation impacts the soil properties such as reduction in soil anti-shearing strength, soil liquefaction as well as even flow of soil or debris mass, thus advancing the susceptibility of soil materials to landslides (Highland and Bobrowsky 2008). Besides, large displacements in August and November were also attributable to the hysteresis effect of rainfall since heavy precipitation was registered in the previous two months. Further, continuous rainfall also aggravated landslide movement. Thus, rainfall was the major trigger that affected landslide deformation and also exerted a "retarded impact" on landslide movement during this stage.

In the last phase (March 2017–February 2018), since the field of stress and seepage as well as geotechnical characteristics in front of the landslide underwent a significant adjustment in the previous stage, the response of the landslide to the external periodic factors became stable. Further, due to the accomplishment of canal renovation, widespread irrigation exerted minimal impact on landslide deformation (Fig. 7). So landslide displacement increased slowly regardless of the heavy rainfall.

3.3 Decomposition of total displacement by VMD

Observations of MD09 for the period from July 2015 to February 2017 were adopted for training via ninefold cross-validation (Li et al. 2018a, b; Li et al. 2020), and those from March 2017 to February 2018 were employed for testing. All data were normalized within the range [0, 1] using Eq. (9):
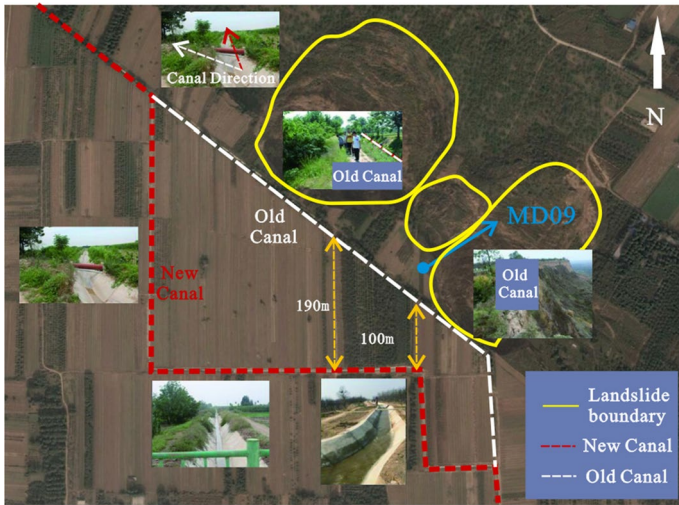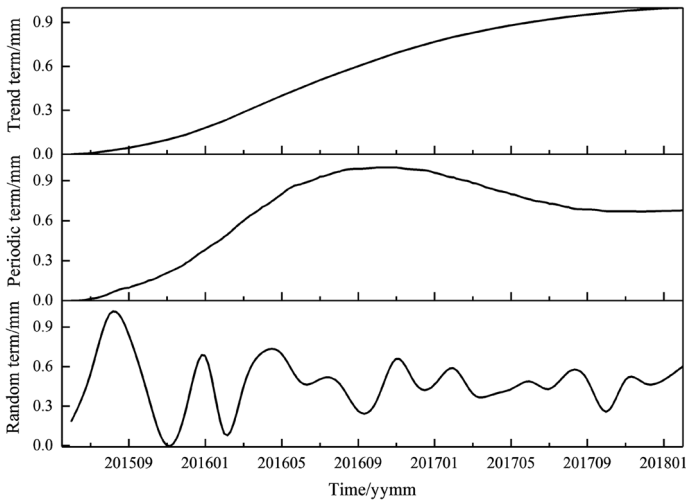
Fig. 7 Position of new and old irrigation canal

Fig. 8 Decompositions of the total displacement of MD09

$$L_{\mathrm{norm}} = \frac{L_i - L_{\mathrm{min}}}{L_{\mathrm{max}} - L_{\mathrm{min}}} \tag{9}$$

where L_{norm} is the normalized data, L_i are the observed data, L_{min} is the minimum value of observations, and L_{max} is the maximum value of observations.

To ensure that each component decomposed by the VMD has practical physical meaning, the decomposition level was set to 3. The penalty parameter and rising step were set to 0.3 and 0.1, respectively after several trials. Thus, there were three subseries

generated, where high frequency was the random term, low frequency was the periodic term and the residual was the trend term (Fig. 8).

3.4 Association analysis between decomposed components and influencing factors

Whether the choice of triggers is proper or not will directly affect the performance of the developed technique. Generally, the grey relation grade (GRG) is adopted to analyze the correlation between the decomposed components and their triggers to verify the reasonability of selected triggers (Du et al. 2009; Lian et al. 2013; Zhang et al. 2015; Miao et al. 2018). If the correlation between the influencing factor and the component is greater than 0.6, the two factors are closely related to the resolution coefficient set to 0.5. In this study, MIC is also introduced to analyze the associations. There are obvious differences between the results of the two approaches (Table 1). Accordingly, based on analysis as discussed in Sect. 3.3, and the values of GRG and MIC, the following factors were selected as the triggers of each component.

Trend term is mainly affected by intrinsic geotechnical conditions during the evolution of landslide deformation, and changes gradually with time. The MIC listed in Table 1 also shows high associations between displacement in the previous 1, 2 and 3 months and trend term. So the movements during the previous 1, 2 and 3 months were chosen as the triggers of trend component.

Periodic term is not only affected by the precipitation and temperature variations over a period of time, but also related to its own changes. Therefore, the following seven terms were adopted as the main influencing factors, namely daily largest rainfall, monthly cumulative rainfall, bimonthly cumulative precipitation, monthly mean temperature and displacements during the previous 1, 2 and 3 months (Table 1).

Random term is influenced by environmental factors such as rainfall, temperature, and wind load. Due to limitations of monitoring conditions, certain data such as groundwater level variations contributed by irrigation and seasonal freeze–thaw were not available. Hence, in this study we selected six factors as enlisted in Table 1 to predict random displacement.

4 Results and analysis

The prediction performances of trend component, periodic component and random component using M-KELM, Poly-KELM and SVR models by k-fold cross-validation are presented in the next sections.

4.1 Trend term prediction

As shown in Table 1, the trend term is mainly triggered by inherent geotechnical conditions. Therefore, the factors, i.e., deformation for the previous 1, 2, and 3 months, are chosen as model inputs and the trend term is used as output (Cao et al. 2016; Zhou et al. 2018). Measurements from July 2015 to February 2017 are adopted for training by using ninefold cross-validation, and those from March 2017 to February 2018 are employed for testing. M-KELM is adopted to forecast the trend term. Poly-KELM and SVR models are applied for comparison. Moreover, optimization of all hyperparameters in abovementioned models

Table 2 Correlation analysis between displacements and impact factors

Displacement	Influencing factors													
	1-month displacement		2-month displacement		3-month displacement		Monthly precipitation		Bimonthly precipitation		Daily largest rainfall		Monthly mean temperature	
	GRG	MIC	GRG	MIC	GRG	MIC	GRG	MIC	GRG	MIC	GRG	MIC	GRG	MIC
Trend term	0.789	1	0.664	1	0.544	1	–	–	–	–	–	–	–	–
Periodic term	1	0.933	1	0.771	0.588	0.706	1	0.329	1	0.350	1	0.329	1	0.489
Random term	0.756	0.160	0.769	0.260	0.745	0.204	0.765	0.258	–	–	0.765	0.237	0.765	0.276

Table 3 Performance of M-KELM, Poly-KELM and SVR for trend term

Model	Training			Validating			Testing		
	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2
M-KELM	0.044	0.054	1	0.303	0.305	1	0.509	0.636	1
Poly-KELM	0.897	0.945	1	0.739	0.739	1	0.700	0.919	1
SVR	1.169	1.604	1	0.305	0.318	1	0.647	0.779	1

The unit of RMSE and MAE is mm
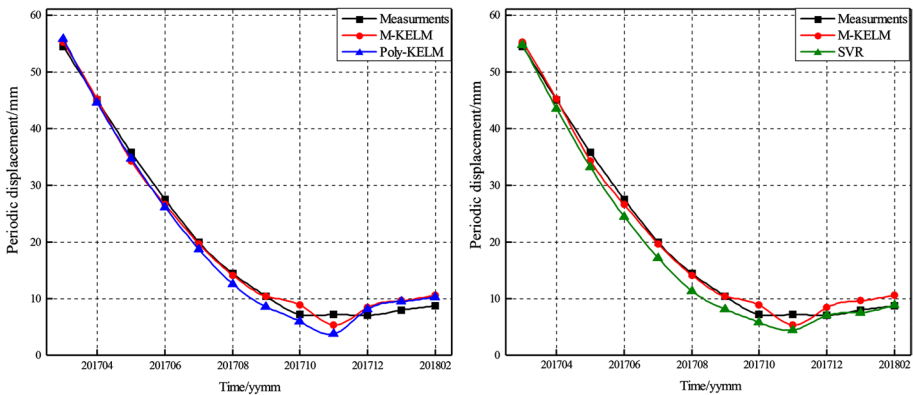
Fig. 9 Predictions and measurements of periodic displacement

is performed by ABC algorithm. The performance of all discussed techniques in terms of R^2, RMSE and MAE is enlisted in Table 2.

As shown in Table 2, it is observed that the M-KELM performs better as compared with the other predictive techniques. For instance, in model testing, the values of RMSE and MAE for the developed model are found to be 0.636 mm and 0.509 mm, respectively; for Poly-KELM, the values are 0.919 mm and 0.700 mm, respectively; for SVR, the values are 0.779 mm and 0.647 mm, respectively. The trend term accuracy of the proposed technique is improved by 44.5% and 18.4% on RMSE, 37.5% and 21.3% on MAE, respectively, in comparison with single kernel-based ELM and SVR. Similar pattern can also be found in model training and validating. Accordingly, the developed approach, which integrates the merits of four kernel functions, produces a more satisfactory result.

4.2 Periodic term prediction

Seven triggers as enlisted in Table 1 are applied for the inputs of all discussed techniques, while the periodic displacements are used for output. Similarly, in the training stage, nine-fold cross-validation is adopted to achieve the optimum hyperparameters for M-KELM, Poly-KLEM and SVR with datasets of the period from July 2015 to February 2017. The performances of all discussed models are shown in Table 3. Predictions of periodic term are depicted in Fig. 9. As shown in Table 3, the accuracy of M-KLEM is nearly equal to that of Poly-KELM and SVR in the stage of model training. However, in model testing, the

Table 4 Performance of M-KELM, Poly-KELM and SVR for periodic term

Model	Training			Validating			Testing		
	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2
M-KELM	0.866	1.040	1	0.620	0.866	1	1.017	1.223	0.995
Poly-KELM	0.944	1.159	1	0.700	0.725	1	1.512	1.645	0.991
SVR	0.819	1.461	0.999	2.110	2.134	1	1.689	2.041	0.994

The unit of RMSE and MAE is mm
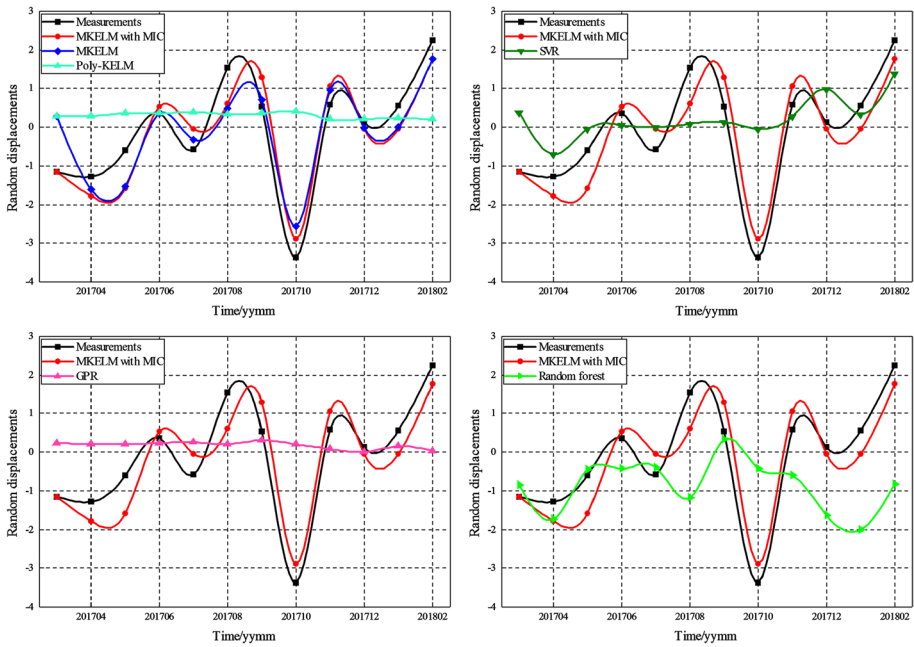
Fig. 10 Predictions and measurements of random displacement

proposed technique performs better than the other two approaches. For example, the lowest RMSE and MAE values of 1.223 mm and 1.017 mm are reported in the developed model. Also the R^2 value (0.995) in M-KELM is found to be higher in comparison with those values (0.991 and 0.994, respectively) in Poly-ELM and SVR. Thus, the M-KELM is regarded as the best technique in this section.

4.3 Random term prediction

Six factors as shown in Table 1 are selected for prediction of random displacement. Since all datasets are normalized within the range [0, 1] using Eq. (9), six input factors have the same weight. However, each factor has different influence on landslide deformation. Therefore, for better performance, different impact factors are given the same weight as their MIC values shown in Table 1(Li et al. 2018a, b) after normalization. In addition, Poly-KELM,

SVR, Gaussian progress regression (GPR) and random forest are also adopted for comparison. The performance evaluation indices (R^2, RMSE and MAE) of all discussed models are enlisted in Table 4. For clear visualizing the compared results, the predictions of all the discussed models by using ninefold cross-validation are shown graphically in Fig. 10.

As shown in Table 4 and Fig. 10, it is obvious that during model training, M-KELM with and without MIC has similar performance. However, in model testing, the M-KELM model with MIC is superior to the M-KELM model without MIC. For instance, in Mar 2017, the absolute error for M-KELM model with MIC is 0.013 mm, while the value for M-KELM model without MIC is 1.446 mm. Hence, the developed technique, which gives different impact factors the same weight as their MIC values, performs better than the M-KELM without MIC.

In addition, better performance is also achieved by the proposed approach as compared with the other models. For example, the lowest RMSE value is recorded in the M-KELM with MIC during model testing strategy, where RMSE value showed a variation from 1.543 mm(random forest), 1.483 mm(Poly-KELM),1.459 mm (GPR), 1.247 mm (SVR), 0.683 mm (M-KELM) to a minimum of 0.585 mm (proposed approach). The MAE values show similar patterns. Also the highest R^2 value of 0.834 is reported in M-KELM with MIC. Therefore, the proposed approach, which integrates the advantages of multi-kernel functions and MIC, has better performance among all the considered models in this section.

4.4 Total displacement prediction

The cumulative displacement is achieved by summation of predicted trend, periodic and random displacements from ninefold cross-validation. The performance evaluation indices such as R^2, RMSE and MAE of M-KELM, Poly-KELM and SVR models are included in Table 5. The comparison of total predictions is plotted in Fig. 11. Model training and testing results produced by the developed approach are depicted in Fig. 12.

As shown in Table 5 and Fig. 11, the best performance in model fitting and predicting phases is registered in the developed model. Considering the results of predictions, the lowest RMSE and MAE values are reported in the proposed methodology. The RMSE and MAE are found to be 1.573 mm and 1.309 mm, respectively, with the M-KELM with MIC technique; 2.494 mm and 2.185 mm, respectively, with the Poly-KELM model; and

Table 5 Performance of M-KELM with MIC, M-KELM, SVR, GPR, Random Forest and Poly-KELM for random term

| Model	Training			Validating			Testing		
	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2
M-KELM with MIC	0.002	0.003	1	0.581	0.581	1	0.514	0.585	0.834
M-KELM	0.002	0.003	1	0.401	0.411	1	0.542	0.683	0.765
SVR	0.025	0.026	1	0.414	0.501	1	0.921	1.247	0.295
GPR	0.065	0.082	1	0.171	0.242	1	1.087	1.459	0.302
Random Forest	1.375	1.690	0.959	0.534	0.546	1	1.359	1.543	0.225
Poly-KELM	1.208	1.538	0.566	1.600	1.725	0.346	1.079	1.483	0.291

The unit of RMSE and MAE is mm
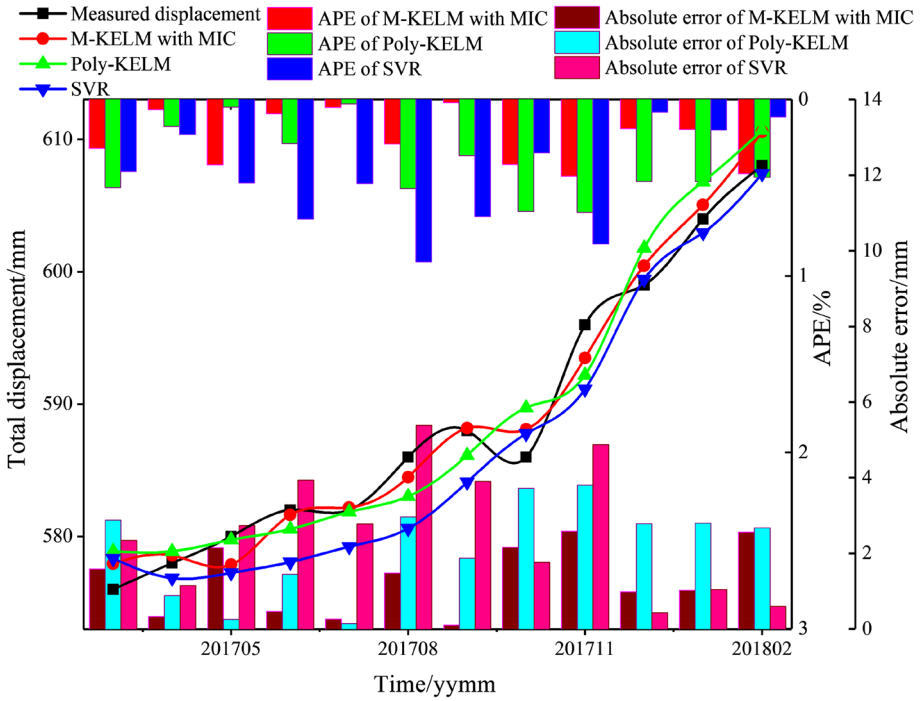
Fig. 11 Total displacement prediction of MD09 using M-KELM, Poly-KELM and SVR
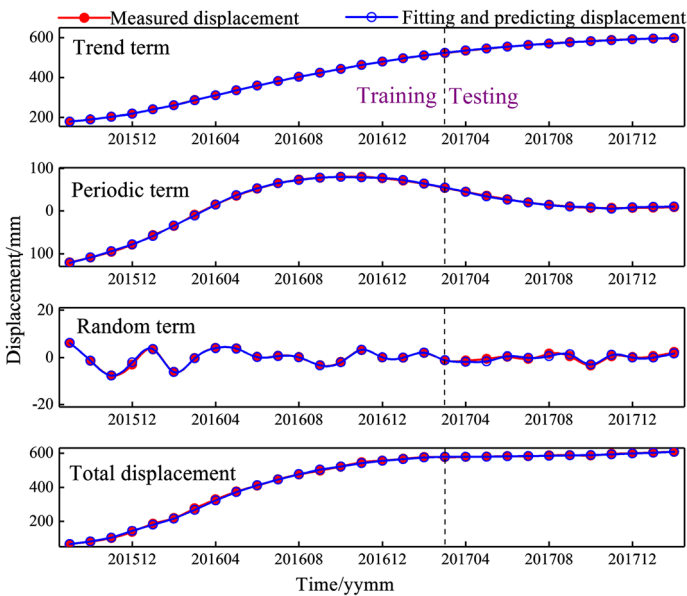

Fig. 12 Fitting and predicting displacements of MD09 using M-KELM

3.031 mm and 2.579 mm, respectively, with the SVR approach. Also R^2 value of 0.980 is found to be higher in M-KELM with MIC in comparison with the two approaches. The proposed model produced an APE varying from 1E-3% to 0.434% as depicted in Fig. 11. Besides, the developed technique also yields better performance than other discussed approaches during highly displacement fluctuating stages (Fig. 11, from September to December 2017). For instance, in September 2017, the absolute error and APE are found to be 0.101 mm and 0.018%, respectively, with the proposed approach; 1.874 mm and 0.312%, respectively, with the Poly-KELM model; and 3.901 mm and 0.663%, respectively, with the SVR technique. Further, higher MAE values are also found in the Poly-KELM and SVR techniques. The lower the performance measures, the higher the predictive capability. Similar patterns can also be found in November 2017.Accordingly, the results in this paper revealed that VMD-MIC-M-KELM model is the optimum approach among all the models considered.

4.5 Discussion

In this study, six models for comparison with the developed approach are adopted to predict decomposed displacement, respectively. The results show that accurate predictions of trend term are achieved by adopting all six models, whereas inaccurate predictions of periodic term and random term occur in models such as Wavelet-KELM, Sigmoid-KELM, RBF-KELM and ELM. For the predictions of cumulative displacement, only three models, i.e., M-KELM, Poly-KELM and SVR, perform better in comparison with the other models. Accordingly, in this paper we only exhibit the performances of the M-KELM, Poly-KELM and SVR approaches.

KELM and other statistical approaches have similar limitations, e.g., the requirement of long and successive data time series, the similar natural/anthropogenic conditions during the training and testing period and the choice of appropriate influencing factors. Moreover, poor predictions may be attained provided that an occurrence happened in the testing stage that did not take place in the training stage. In this paper, the developed model eliminates some limitations mentioned above. The time series employed for the development of an accurate model in Miaodian landslide are more than 3 years and continuous. Results of the model fitting and prediction indicate that the selected parameters are relatively representative. Refined predictions would be achieved by continually updating the developed technique with newly monitoring datasets from the monitoring system.

To explore further, Table 6 shows the total displacement forecasting performance of M-KELM with and without MIC. In fact, the same experiments are also conducted on period term prediction. However, the results show that the addition of MIC has little effect on model performance improvement. Hence, we only present the results of random term

Table 6 Performance of M-KELM, Poly-KELM and SVR in model training and testing

Model	Training			Testing		
	MAE	RMSE	R^2	MAE	RMSE	R^2
M-KELM with MIC	0.985	1.199	1	1.309	1.573	0.980
Poly-KELM	2.452	3.236	1	2.185	2.494	0.951
SVR	2.217	2.918	1	2.579	3.031	0.946

The unit of RMSE and MAE is mm

improved by MIC. As shown in Table 6, it is obvious that the performance of M-KELM with MIC for total displacement forecasting is improved by 13.77% on MAE, 12.95% on RMSE and 0.8% on R^2, as compared with the MKELM approach. The results reveal that it is effective to add MIC value in random displacement prediction, which contributes a lot to the model improvement in total displacement prediction.

Despite the better accuracy of our method in landslide prediction, inaccurate predictions occasionally occur without proper triggers. For instance, it is obvious that predictions in August and November 2017 for the three models are relatively smaller than the measurements (Fig. 11). The highest absolute errors for all discussed models are reported in August and November 2017. As depicted in Fig. 9, the absolute errors in the two months are found to be 1.481 mm and 2.588 mm, respectively, with the M-KELM model with MIC; 2.961 mm and 3.808 mm, respectively, with the Poly-KELM approach; and 5.386 mm and 4.872 mm, respectively, with the SVR technique. According to our collections, some farmlands near the landslide are far away from the canal and farmers have to irrigate the land broadly by traditional method for spring and autumn farming annually. Hence, groundwater level variations contributed by summer and winter irrigation which were not considered in this study account for the higher absolute errors even though the canal renovation had been completed. Therefore, the influence of groundwater level due to widespread irrigation should be further considered in the model for better performance (Table 7).

In addition, the penalty parameter and rising step of VMD were obtained by continuously trials to insure the fidelity of decomposed time series. This procedure is very laborious and time-consuming since the subseries varies with the change in penalty parameter and rising step. Hence, determination of these two parameters, i.e., penalty parameter and rising step, should be further studied to enable researchers free from selection.

Further, although random component is available by VMD in this study, it is still unable to achieve better performance since we are lack of monitoring data of relevant triggers owing to limited monitoring conditions. Accordingly, data with high quality should be required for better performance in the future.

5 Conclusion

Due to the impact of geological factors and environmental triggers, it remains a tough task to predict landslide displacement accurately for landslide prevention. Thus, this paper presents a predictive technique, i.e., VMD-MIC-M-KELM, to achieve better performance for forecasting long-term landslide movements. The Miaodian loess landslide (China) was taken as a case study. With more than 3-year observations in the study region, the developed model is able to accurately predict landslide movements for periods up to 12 months. The technique developed in this study proved to consider the physical meaning of subcomponents with manually set number of decomposition. Besides, the triggers selected were also representative. Moreover, the M-KELM approach was capable of producing

Table 7 the total displacement forecasting performance of M-KELM with MIC and M-KELM	Model	Testing		
		MAE/mm	RMSE/mm	R^2
	M-KELM with MIC	1.309	1.573	0.980
	M-KELM	1.518	1.807	0.972

satisfactory results in model prediction by employing a weighted integration of individual kernel functions. Model validation also showed that the technique developed in this paper achieved the best performance in comparison with the Poly-KELM and SVR.

However, there are also some limitations that require to be addressed in the future: (1) groundwater information contributed by broad irrigation should be modeled into landslide prediction model; (2) parameters of VMD should be optimized by potential algorithm like ANNs; and (3) prediction performance of random component should be improved with efficient monitoring datasets. The developed model can be employed to forecast landslide displacements influenced by factors such as precipitation, temperature, groundwater, but be inapplicable to predict movements triggered by artificial evacuation, earthquakes, or other emergency events.

Accordingly, given the reliability of predictions and robustness, VMD-MIC-M-KELM presented in this study proves to be applicable for prediction of landslide behavior with high coverage of data sources such as displacement, intrinsic geotechnical factors and environmental factors. And the technique may be constantly updated on a monthly basis from the monitoring system, thus providing reliable information for early warning of landslide.

References

Alimohammadlou Y, Najafi A, Gokceoglu C (2014) Estimation of rainfall-induced landslides using ANN and fuzzy clustering methods: a case study in Saeen Slope, Azerbaijan province. Iran Catena 120:149–162

Bernardie S, Desramaut N, Malet JP, Gourlay M, Grandjean G (2015) Prediction of changes in landslide rates induced by rainfall. Landslides 12(3):481–494

Calvello M, Cascini L, Sorbino G (2008) A numerical procedure for predicting rainfallinduced movements of active landslides along pre-existing slip surfaces. Int J Numer Anal Meth Geomech 32(4):327–351

Cao Y, Yin KL, Alexander DE, 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

Corominas J, Moya J, Ledesma A, Lloret A, Gili JA (2005) Prediction of ground displacements and velocities from groundwater level changes at the Vallcebre landslide (Eastern Pyrenees, Spain). Landslides 2:83–96

Deng DM, Liang Y, Wang LQ, Sun ZH, Wang C, Huang MM (2017) PSO-SVR prediction method for landslide displacement based on reconstruction of time series by EEMD: a case study of landslides in Three Gorges Reservoir area. Rock and Soil Mechanics 38(12):1001–1009

Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. IEEE Trans Signal Process 62(3):531–544

Dragomiretskiy K (2015) Variational methods in signal decomposition and image processing. Ph.D. thesis

Du J, Yin KL, Chai B (2009) Study of displacement prediction model of landslide based on response analysis of inducing factors. Chin J Rock Mechan Eng 28(9):1783–1789

Fang YM, Zhao XD, Zhang P, Liu L, Wang SY (2020) Prediction modeling of silicon content in liquid iron based on multiple kernel extreme learning machine and improved grey wolf optimizer. Control Theory Appl 37(7):1644–1654

Federico A, Popescu M, Elia G, Fidelibus C, Internò G, Murianni A (2012) Prediction of time to slope failure: a general framework. Environ Earth Sci 66:245–256

Guzzetti F, Reichenbach P, Cardinali M, Galli M, Ardizzone F (2005) Probabilistic landslide hazard assessment at the basin scale. Geomorphology 72:272–299

Highland LM, Bobrowsky P (2008) The Landslide Handbook— A Guide to Understanding Landslides. Us Geological Survey

Huang GB, Zhou HM, Ding XJ, Zhang R (2012) Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems Man and Cybernetics Part B(Cybernetics) 42(2): 513–529

Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. Neurocomputing 70:489–501

Intrieri E, Gigli G (2016) Landslide forecasting and factors influencing predictability. Nat Hazards Earth Syst Sci 16(12):2501–3251

Karaboga D, Gorkemli B, Ozturk C, Karaboga N (2014) A comprehensive survey: artificial bee colony (ABC) algorithm and applications. Artif Intell Rev 42(1):21–57

Kawabata D, Bandibas J (2009) Landslide susceptibility mapping using geological data, a DEM from ASTER images and an artificial neural network (ANN). Geomorphology 113:97–109

Krkač M (2015) A phenomenological model of the Kostanjek landslide movement based on the landslide monitoring parameters. Dissertation, University of Zagreb (in Croatian)

Li LW, Wu YP, Miao FS, Liao K, Zhang LF (2018a) Displacement prediction of landslide based on variational mode decomposition and GWO-MIC-SVR model. Chin J Rock Mechan Eng 37(06):100–111

Li H, Xu Q, He Y, Deng J (2018b) Prediction of landslide displacement with an ensemble-based extreme learning machine and copula models. Landslides 15:2047–2059

Li H, Xu Q, He Y, Fan X, Li S (2020) Modeling and predicting reservoir landslide displacement with deep belief network and EWMA control charts: a case study in Three Gorges Reservoir. Landslides 17(3):693–707

Lian C, Zeng ZG, Yao W, Tang HM (2013) Displacement prediction of landslide based on PSOGSA-ELM with mixed kernel. Sixth Interational Conference on Advanced Computational Intelligence China 52–57

Liu Y, Liu D, Qin ZM, Liu FB, Liu LB (2016) Rainfall data feature extraction and its verification in displacement prediction of Baishuihe landslide in China. B Eng Geol Environ 75(3):897–907

Melchiorre C, Matteucci M, Azzoni A, Zanchi A (2008) Artificial neural networks and cluster analysis in landslide susceptibility zonation. Geomorphology 94:379–400

Miao FS, Wu YP, Xie Y, Li Y (2018) Prediction of landslide displacement with step-like behavior based on multi algorithm optimization and a support vector regression model. Landslides 15:475–488

Pham BT, Bui DT, Prakash I, Dholakia M (2016) Evaluation of predictive ability of support vector machines and naive Bayes trees methods for spatial prediction of landslides in Uttarakhand state (India) using GIS. J Geomatics 10:71–79

Ranjeeta B, Dash PK, Das PP (2018) Short-term electricity price forecasting and classification in smart grids using optimized multi-kernel extreme learning machine. Neural Comput Appl 1–24

Reshef DN, Reshef YA, Finucane HK, Grossman SR, McVean G, Turnbaugh PJ, Lander ES, Mitzenmacher M, Sabeti PC (2011) Detecting novel associations in large data sets. Science 334(6062):1518–1524

San BT (2014) An evaluation of SVM using polygon-based random sampling in landslide susceptibility mapping: the Candir catchment area (western Antalya, Turkey). Int J Appl Earth Obs Geoinforma 26:399–412

Sassa K, Osamu N, Solidum R, Yamazaki Y, Ohta H (2010) An integrated model simulating the initiation and motion of earthquake and rain induced rapid landslides and its application to the 2006 Leyte landslide. Landslides 7:219–236

Shihabudheen KV, Pillai GN, Peethambaran B (2017) Prediction of landslide displacement with controlling factors using extreme learning adaptive neuro-fuzzy inference system (ELANFIS). Appl Soft Comput 61:892–904

Wu YP, Teng WF, Li YW (2007) Application of grey-neural network model to landslide deformation prediction. Chin J Rock Mechan Eng 26(03):632–636

Xu YQ, Tang YQ, Li XY, Ye JM (2011) The landslide deformation prediction with improved Euler method of gray system model GM(1,1). Hydrogeology Engineering Geology 38(1):110–113

Zhang J, Yin KL, Wang JJ, Huang FM (2015) Displacement prediction of Baishuihe Landslide based on time series and PSO-SVR model. Chin J Rock Mechan Eng 34(2):382–391

Zhou C, Yin KL, 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 Eng Geol 204:108–120

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