

# SCIENTIFIC REPORTS



OPEN

## A novel method for landslide displacement prediction by integrating advanced computational intelligence algorithms

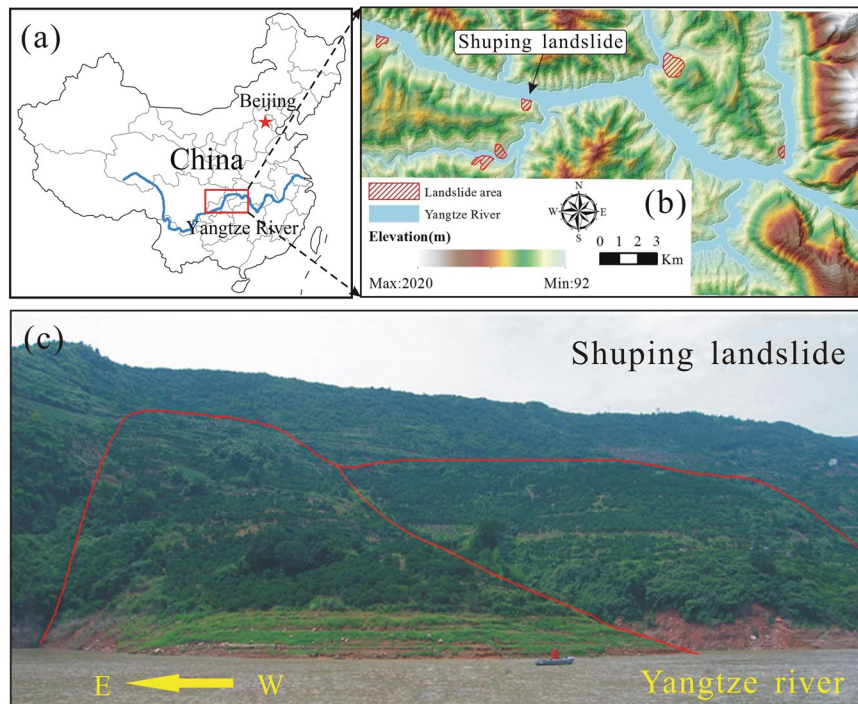
Chao Zhou<sup>1</sup>, Kunlong Yin<sup>1</sup>, Ying Cao<sup>1</sup>, Bayes Ahmed <sup>2</sup> & Xiaolin Fu<sup>3</sup>

Landslide displacement prediction is considered as an essential component for developing early warning systems. The modelling of conventional forecast methods requires enormous monitoring data that limit its application. To conduct accurate displacement prediction with limited data, a novel method is proposed and applied by integrating three computational intelligence algorithms namely: the wavelet transform (WT), the artificial bees colony (ABC), and the kernel-based extreme learning machine (KELM). At first, the total displacement was decomposed into several sub-sequences with different frequencies using the WT. Next each sub-sequence was predicted separately by the KELM whose parameters were optimized by the ABC. Finally the predicted total displacement was obtained by adding all the predicted sub-sequences. The Shuping landslide in the Three Gorges Reservoir area in China was taken as a case study. The performance of the new method was compared with the WT-ELM, ABC-KELM, ELM, and the support vector machine (SVM) methods. Results show that the prediction accuracy can be improved by decomposing the total displacement into sub-sequences with various frequencies and by predicting them separately. The ABC-KELM algorithm shows the highest prediction capacity followed by the ELM and SVM. Overall, the proposed method achieved excellent performance both in terms of accuracy and stability.

Landslides are a common natural hazard and cause extensive losses in mountainous areas. For example, the Maoxian landslide that occurred in the southwest of China in June 2017, damaged 62 houses and more than 100 people were buried<sup>1–3</sup>. In the north of Vietnam, a massive amount of soil and rocks, triggered by heavy rainfall, rolled down in an adjacent hill and caused 8 fatalities in October 2017<sup>4</sup>. In order to reduce such landslide disaster risks, developing an effective early warning system (EWS) should be prioritized in landslide research<sup>5,6</sup>.

Landslide displacement prediction is considered as an essential component for developing EWSs. It can be used to set warning thresholds and to recognize when a landslide undergoes an unpredicted acceleration<sup>6,7</sup>. Since Saito (1965) proposed the empirical formula for landslide prediction<sup>8</sup>, numerous landslide prediction models have been developed<sup>9–16</sup>. They can be grouped into two categories: physical-based models and data-based models. Physical-based models provide a good understanding and prediction of landslide failure. Jiang *et al.* (2011) investigated and predicted landslide deformation by establishing a three-dimensional geological model and performing a numerical simulation with the saturated–unsaturated fluid–solid coupling theory<sup>17</sup>. The physical-based models are complex and time consuming<sup>18,19</sup>, as they need to consider many conditions, such as complex landslide geometries, spatial variations in soil properties and three-dimensional groundwater flow etc.<sup>19</sup>. The data-based models have been applied recently because of their good prediction accuracy and simplicity. Du *et al.* (2012) analyzed the deformation characteristics of colluvial landslides and conducted the displacement prediction using a Back Propagation Neural Network (BPNN) method<sup>20</sup>. Zhou *et al.* (2016) integrated time series

<sup>1</sup>Engineering Faculty, China University of Geosciences, Wuhan, 430074, China. <sup>2</sup>Institute for Risk and Disaster Reduction, University College London (UCL), London, WC1E 6BT, UK. <sup>3</sup>Administration of Prevention and Control of GeoHazards in the Three Gorges Reservoir of China, Yichang, 443000, China. Correspondence and requests for materials should be addressed to K.Y. (email: [yinkl@126.com](mailto:yinkl@126.com)) or Y.C. (email: [caoying@cug.edu.cn](mailto:caoying@cug.edu.cn))



**Figure 1.** (a) Map of China; (b) Map showing the location of the Shuping landslide; and (c) Photograph of the Shuping landslide (The two maps was created by Chao Zhou using ArcGIS 9.3, <http://www.esri.com/>).

decomposition and support vector machine (SVM) to establish a displacement prediction method by considering the response relationship between triggering factors and landslide deformation<sup>21</sup>.

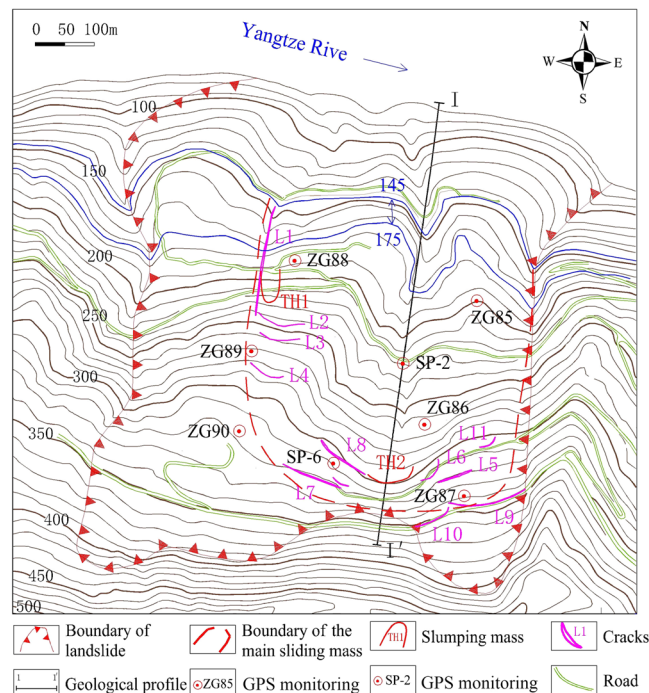
Although these methods perform well and provide crucial parameters for EWS, they require a huge amount of monitoring data to analyze the deformation mechanism. The least developed landslide-prone regions suffer the most<sup>22</sup> because they cannot afford or do not have the available resources to monitor the parameters for landslide deformation analysis and prediction, such as information about groundwater, shear tension, pore water pressure etc. To overcome the limitations, a variety of computational intelligence methods have been applied in landslide study to achieve scientifically valid and accurate results<sup>23–28</sup>. These methods can maximize the extraction of useful information from limited data for landslide displacement prediction.

Landslide evolution is a complex nonlinear dynamic process, and is influenced by geological conditions of landslide mass and external triggering factors (e.g. precipitation, earthquake, human engineering activities, etc.). The displacement data includes multiple deformation information with various rules<sup>29</sup>. Wavelet transform (WT) is an effective signal processing algorithm which can decompose time series signals into several components with different frequencies. Many researchers have tried to predict landslide using artificial neural networks (ANN) methods<sup>30–34</sup>. The conventional ANN models can process nonlinear problems, but the selection of network structure and parameters is mainly dependent on experience. In addition, some other drawbacks of the traditional ANNs models are associated with excessive tunable parameters, slow learning rate, high possibility of entrapment in local minima, long computational time, and over-tuning<sup>35</sup>. Extreme learning machine is an algorithm based on single-hidden layer feed-forward neural network, which has been reported with good prediction capacity<sup>36–40</sup>. However, the random connection weights between the input and hidden layers make the output of ELM fluctuant. Consequently, ELM was introduced into Kernel learning and the kernel-based ELM (KELM) was proposed by Huang *et al.*<sup>41</sup>. Considering that the performance of KELM is affected by its parameters<sup>38,41</sup>, it is significant to select appropriate parameters. The artificial bee colony (ABC) is a swarm intelligence-based global optimization algorithm, and has been utilized to search the optimal parameters of KELM due to its high accuracy and fast convergence characteristic<sup>42</sup>.

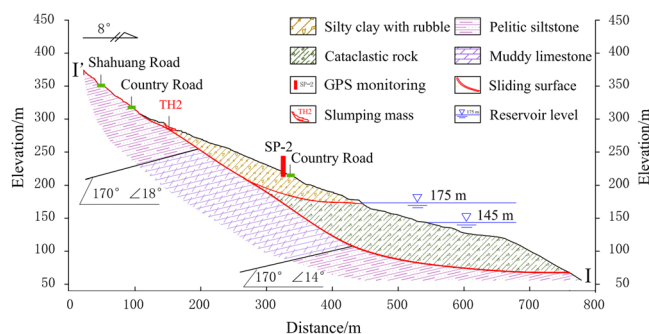
In this study, the advanced computational intelligence algorithms namely the WT, ABC and the KELM methods (WT-ABC-KELM) were integrated to establish a novel landslide displacement prediction method. The Shuping landslide in the Three Gorges Reservoir area (TGRA) of China was undertaken as a case study. The ABC-KELM, WT-ELM, ELM, and SVM methods were applied for performance comparison.

## Case study

**Geological conditions.** The Shuping landslide occurred in the Zigui county that is located in the TGRA (30°59′37″N, 110°37′0″E). The landslide location is near the Yangtze River and 47 kilometers from the Three Gorges Dam (Fig. 1). The landslide is fan-shaped in plane with a main sliding direction of N11°E. The left and right boundaries are defined by two gullies, while the elevation of the upper boundary varies between 340 m and 400 m a.s.l., and the elevation of the toe is 60 m a.s.l. The landslide covers an area of  $5.5 \times 10^5 \text{ m}^2$ , with a mean



**Figure 2.** Topographical map of the Shuping landslide, with the location of the monitoring networks (Drawn by Ying Cao using AutoCAD 2014, <https://www.autodesk.com.cn/>), the cross-section can be seen in Fig. 3.



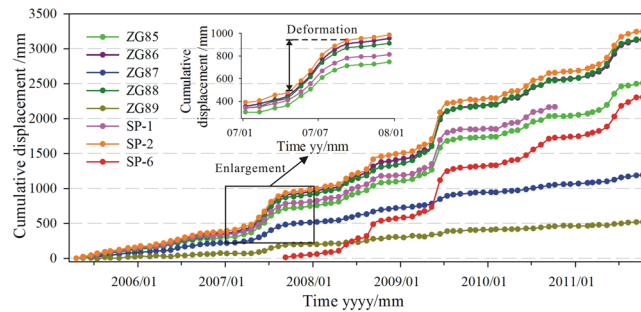
**Figure 3.** Schematic geological cross-section (I-I') of the Shuping landslide (Drawn by Ying Cao using AutoCAD 2014, <https://www.autodesk.com.cn/>), the track can be seen in Fig. 2.

longitudinal dimension of 800 m and a width of 700 m (Fig. 2). The Shuping landslide is an ancient slumping deposit. It developed in the reverse syncline slope which is composed of mudstone and siltstone with muddy limestone of Triassic Badong Formation, with the dip direction of 120~173° and the dip angle of 9~38°. The mean depth of the sliding surface is about 50 m (Fig. 3). The Shuping landslide is a large colluvial mass movement with an estimated volume of  $2750 \times 10^4 \text{ m}^3$ . The eastern and middle part of the landslide is the main sliding mass with greater deformation. The area and volume of which are  $35 \times 10^4 \text{ m}^2$  and  $1575 \times 10^4 \text{ m}^3$ , respectively.

**Deformation characteristics.** The Shuping landslide is an ancient landslide with complex geological conditions and a large volume. The deformation was significantly slow before the construction of the Three Gorges Dam (see boundary of landslide in Fig. 2). After the construction of the dam, the water level of Yangtze River raised from 75 m to 145 m a.s.l. Under the influence of such periodic reservoir fluctuations and heavy precipitation, the Shuping landslide revived. The landslide deformed partially which is shown as the main sliding mass in Fig. 2. The deformation monitoring of the sliding body is mainly based on the global positioning system (GPS) and macroscopic geological inspection. Eight effective GPS monitoring stations were installed surrounding the Shuping landslide (ZG85-ZG90, SP-2 and SP-6) (Fig. 2), with the layout of two vertical and three horizontal directions to monitor the deformation of the whole landslide.

Under the influence of periodic reservoir fluctuations and seasonal precipitations, the displacement speed of the Shuping landslide is found to be increasing between May and September in 2007 (Fig. 4). In this period, the landslide deformed 453.5 mm while the total deformation of 2007 was 593 mm. Several transverse arc-shaped





**Figure 4.** The GPS monitoring data of the Shuping landslide.



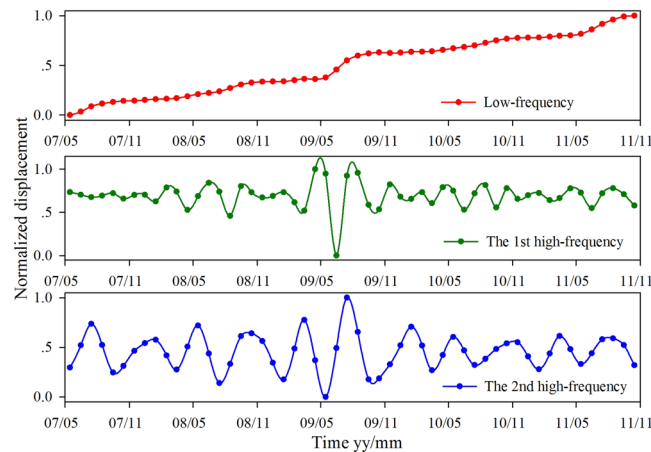
**Figure 5.** The deformation phenomena on the landslide mass: (a) slump of TH1, (b) crack of L2, (c) crack of L4, and (d) crack of L9.

tensile cracks occurred on the eastern upper part of the landslide (L5-L11), with the length of about 20 m and the width of 5–8 m. From June 13 to July 11 in 2007, the landslide experienced severe deformation. The maximum deformation of SP-2 and ZG86 in the middle part was about 130 mm (Fig. 4), and the maximum daily speed reached around 5.10 mm.

Due to heavy rainfall in the late August 2008, the cracks on the eastern part showed further development, and were connected with the cracks of L9 and L10. The direction of all the cracks was basically consistent with the upper landslide boundary (Fig. 2). Meanwhile, a small-scale slump (TH1, Fig. 5a) occurred in the west of the landslide with a length and width of 50 m and 30 m respectively, and a volume larger than 1000 m<sup>3</sup>. In August 2008, the GPS monitoring stations of SP-2 in the middle part and SP-6 in the upper part showed larger deformation than the other monitoring stations. Their monthly deformation was found 143.7 mm and 169.1 mm respectively.

The deformation speed of the Shuping landslide severely increased between May and August in 2009. The most serious deformation occurred in June with crack extension and road subsidence on the upper landslide (L2-L4) (Fig. 5b,c). All the cracks on the eastern part and most of the echelon cracks on the western part were connected. The displacements monitored by the GPS stations (GPS85, GPS86, GPS88, SP-2, SP-6) were over 300 mm within a month (Fig. 4). Probably it formed the boundaries of the main sliding zone of the Shuping landslide.

During the year of 2010 and 2011, the deformation of the main sliding mass increased since May, and more surface cracks developed and increased into larger scales on the boundaries. The cracks on the eastern boundary were of compression-shearing character (L11), the cracks on the upper boundary were formed by tensile force



**Figure 6.** The decomposition result of the total displacement of SP-2.

(L9 and L10) (Fig. 5d), and the cracks on the western boundary are of shearing-tension character with echelon distribution (L2 -L4) (Fig. 2).

## Results and Discussion

**The decomposition of landslide displacement.** Landslide total displacement can be decomposed into multi-level sub-sequences with various frequencies using the WT. In order to select an appropriate number of decomposition levels, Equation (1)<sup>43–47</sup> was used in the proposed method:

$$L = \text{int}[\log(N)] \quad (1)$$

where  $L$  is the decomposition level, and  $N$  is the number of time series data. Here,  $N$  is 54. The total displacement sequence was decomposed into 2 levels (i.e.  $L$ ), which can be expressed as follows:

$$D_s = D_a + D_{h1} + D_{h2} \quad (2)$$

where  $D_s$  is the original total displacement;  $D_a$  is the low-frequency component; and  $D_{h1}$  and  $D_{h2}$  are the first and second high-frequency components, respectively. As the computational intelligence algorithms are more sensitive in their desired range, each sub-sequence of displacement was normalized into the desired range between [0,1] based on Equation (3) before modelling.

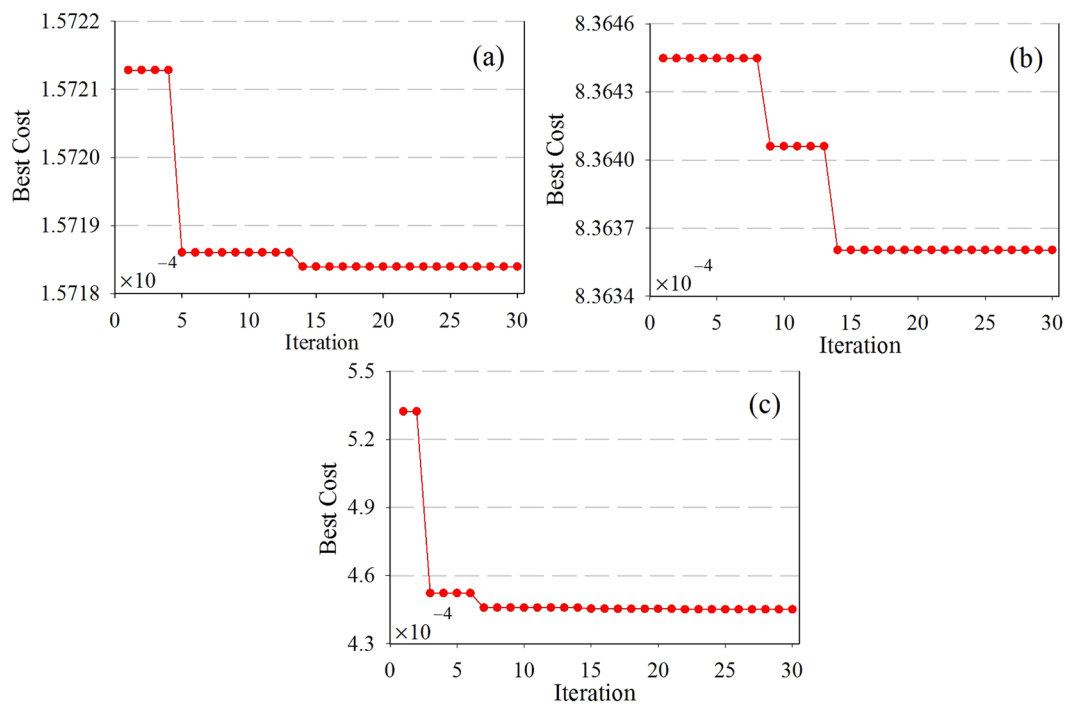
$$D_i = \frac{D_{ori,i} - D_{ori,min}}{D_{ori,max} - D_{ori,min}} \quad (3)$$

where  $D_i$  are the normalized values,  $D_{ori,i}$  are the original values,  $D_{ori,min}$  is the minimum value of an original sub-sequence, and  $D_{ori,max}$  is the maximum value of an original sub-sequence. As shown in Fig. 4, the GPS station of SP-2 experienced the largest deformation. In this study, the displacement data of SP-2 from May 2007 to October 2010 was used for model training, and the data after October 2010 was used for performance test. After data normalization, the decomposition results of SP-2 are shown in Fig. 6.

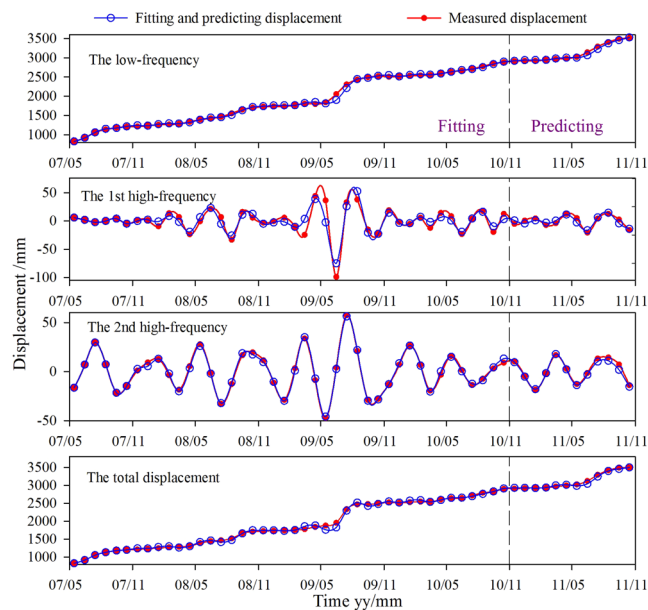
**The modelling of displacement prediction.** In the modelling of the proposed method, the radial basis function was adopted as the kernel function of KELM, and the parameters of the KELM, i.e. regularization coefficient  $c$  and kernel parameter  $\gamma$ , were optimized through the ABC algorithm and the training process. In the hybrid model of ABC-KELM, the parameters of  $c$  and  $\gamma$  were set as solution target, the KELM was set as solution function, and the root mean square error (REMS) of the fitting result of KELM was set as fitness (best cost). In the optimization processing, the problem solution ( $c$  and  $\gamma$ ) was selected through the strategy of ABC algorithm, and the corresponding REMS was calculated for fitness comparison (see the Methods Section). The initial parameters of the ABC algorithm, namely the maximum number of iterations, the number of population, and the acceleration threshold were set as 30, 50, and 20, respectively.

Each displacement sub-sequence was predicted separately using the ABC-KELM model. The best cost in the parameters optimization processes of the three sub-sequences is shown in Fig. 7. The optimized parameters and inputs for each sub-sequence modelling are shown in Table 1. The fitting and prediction results of each sub-sequence can be calculated (Fig. 8) based on Equation (3). As per the decomposition principle of the displacement time series (Equation (2)), the predicted total displacement was obtained by summing up the predicted values of all the sub-sequences (Fig. 8).

To verify the performance of the proposed method, four methods, namely, the ABC-KELM, WT-ELM, ELM and SVM<sup>21,48</sup>, were executed with the same data. Regarding the ABC-KELM method, the parameters of KELM were optimized using the ABC algorithm. The optimal regularization coefficient and the kernel parameter were  $1.0 \times 10^6$  and 1492.9397, respectively. The number of the neurons in the hidden layer (i.e. an integer) is the sole parameter of the ELM algorithm. It was selected by the trial-and-error approach within the range from 1 to 50.



**Figure 7.** The parameters searching process for the three sub-sequences predictions: (a) the low-frequency, (b) the 1st high-frequency, and (c) the 2nd high-frequency.



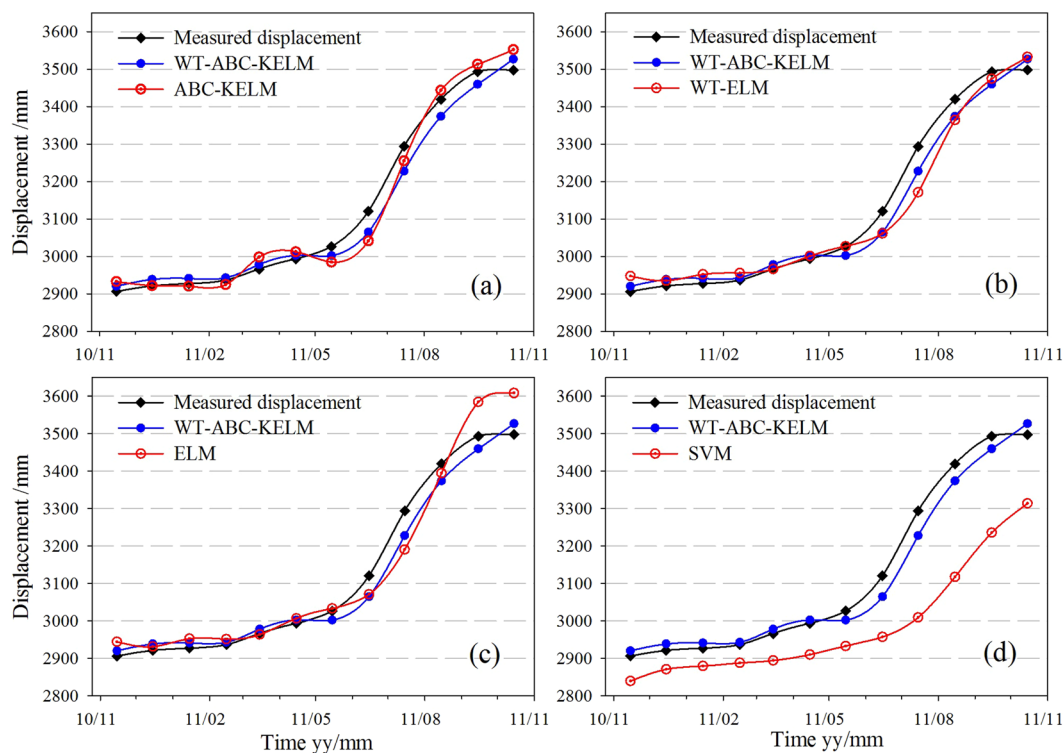
**Figure 8.** The results of fitting and prediction values of SP-2.

Sub-sequences	Parameters	Inputs
The low-frequency	$c = 1.0 * 10^6, \gamma = 1727.0881$	$[d(t-3), d(t-2), d(t-1)]$
The 1st high-frequency	$c = 3.35 * 10^4, \gamma = 929.1308$	$[d(t-6), d(t-5), d(t-4), d(t-3), d(t-2), d(t-1)]$
The 2nd high-frequency	$c = 1.0 * 10^6, \gamma = 38.1863$	$[d(t-6), d(t-5), d(t-4), d(t-3), d(t-2), d(t-1)]$

**Table 1.** The parameters and inputs of the prediction models for sub-sequences. Note:  $d(t-n)$  means the displacement of  $n$  months ago.

Methods	Parameters	Inputs
ABC-KELM	$c = 1.0 * 10^6, \gamma = 1492.9397$	$[d(t-2), d(t-1)]$
ELM	$n = 10$	$[d(t-2), d(t-1)]$
SVM	$\Theta = 85.4376, g = 0.01$	$[d(t-2), d(t-1)]$
WT-ELM		
The low-frequency	$n = 10$	$[d(t-3), d(t-2), d(t-1)]$
The 1st high-frequency	$n = 15$	$[d(t-6), d(t-5), d(t-4), d(t-3), d(t-2), d(t-1)]$
The 2nd high-frequency	$n = 21$	$[d(t-6), d(t-5), d(t-4), d(t-3), d(t-2), d(t-1)]$

**Table 2.** The parameters and inputs of the four compared methods.



**Figure 9.** The predicted results of the compared methods: (a) ABC-KELM, (b) WT-ELM, (c) ELM, and (d) SVM.

For the WT-ELM method, the number of the neurons of the low-frequency, the 1st high-frequency, and the 2nd high-frequency were 10, 15, and 21, respectively. For the ELM method, the number of the neurons was set as 10. The radial basis function was applied as the kernel function of SVM, and the penalty factor  $\Theta$  and the kernel function  $g$  were set as 85.4376 and 0.01, respectively. The parameters and inputs of the four models used for comparison are shown in Table 2. Their predicted total displacements are shown in Fig. 9.

**Accuracy comparison and discussion.** Validation is an essential component in testing the effectiveness and scientific significance of the method applied for landslide displacement prediction study. Three statistical indexes –the root mean square error (RMSE), the mean absolute percentage error (MAPE), and the correlation coefficient (R)– were used for assessing the prediction accuracy of the five prediction methods (Table 3).

As shown in Fig. 8, the proposed method i.e. WT-ABC-KELM achieved high accuracy both in model fitting and prediction stages. To compare the performance of the ABC-KELM and ELM algorithms, the accuracy of the WT-ABC-KELM and ABC-KELM methods were compared with the WT-ELM and ELM, respectively. The accuracy statistic parameters (REMS, MAPE and R) of the WT-ABC-KELM are calculated 33.13, 0.85%, and 0.991, respectively (Table 3), which are found better than the WT-ELM. The REMS, MAPE and R of the ABC-KELM are estimated 46.79, 1.04%, and 0.980, respectively. It indicates the ABC-KELM algorithm can simulate the actual situation more accurately than the ELM. The predicted values of the ABC-KELM algorithm are closer to the measured values. The REMS, MAPE and R of the SVM are calculated 166.39, 4.25%, and 0.959, which indicates the prediction capacity of the SVM algorithm is worse than the ABC-KELM and ELM algorithms. It can be stated that the ABC-KELM algorithm has the strongest prediction capacity followed by the algorithm of ELM and SVM.

As reported in the previous literature<sup>38,41,49,50</sup>, the ELM outperforms the conventional computational intelligence algorithms. However, the connection weights between the input and hidden layers of the ELM algorithm



Methods	REMS	MAPE	R
WA-ABC-KELM	33.13	0.85%	0.991
WA-ELM	36.23	0.93%	0.989
ABC-KELM	46.79	1.04%	0.980
ELM	53.00	1.25%	0.977
SVM	166.39	4.25%	0.959

**Table 3.** The accuracy comparison of the prediction methods.

Model	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	Mean	Variance
ABC-KELM	46.79	46.79	46.79	46.79	46.79	46.79	46.79	46.79	46.79	46.79	46.79	0
ELM	53.00	63.76	76.72	61.72	78.12	63.15	96.65	60.11	90.04	81.84	72.51	184.12

**Table 4.** The comparison of the REMS values between the ABC-KELM and ELM in 10 trial.

are randomly determined and it leads to random fluctuation of the predicted results. In KELM algorithm, a fixed kernel function is applied to replace the output function to avoid the randomly generating process of the connection weights in the ELM algorithm. Under the same condition (Table 2), ten trials were performed using the ABC-KELM and ELM methods, and the REMS values of predicted results are shown in Table 4. Apparently, the predicted results of the ELM method are fluctuant, and the RMSE variance of the 10 predictions is as high as 184.12. On the contrary, the predicted results of the ABC-KELM method in different trials are the same without fluctuations. For generating landslide early warnings, the fluctuations of the predicted results as found in the case of ELM may produce false warnings and mislead disaster managers into making erroneous decisions. Overall, the ABC-KELM outperforms the ELM slightly. The prediction of the KELM is fixed and that is crucial in practice.

Landslide deformation is a complex nonlinear process affected by many factors and shows highly nonlinear total cumulative displacement sequence. The WT can decompose the total displacement into sub-sequences with various frequencies for easier prediction. As shown in Table 3, the REMS, MAPE and R of the ABC-KELM are calculated 46.79, 1.04%, and 0.980, respectively. Meanwhile, the three statistic indexes of the ELM are 53.00, 1.25%, and 0.977, respectively. Notably, the prediction accuracy of both methods is found lower than the corresponding hybrid methods using the WT (Table 3 and Fig. 9). The results of performance comparison demonstrate that decomposing the total displacement into several sub-sequences and predicting them separately can effectively improve the prediction accuracy.

The computational intelligence methods can accurately predict landslide displacement as demonstrated in the current case study (Table 3 and Fig. 9). The ANN algorithms, including the KELM, require past data for training. The past data is the foundation for forecasting and the prediction accuracy depends on the training data (past displacement). When the deformation pattern in prediction period is similar to the training period, for example, the two periods are under the same evolution stage, the landslide displacement can usually be predicted accurately. Conversely, when the deformation pattern in predicting period is different from the training period, the performance of these training-based models will be limited. It is important to select the monitoring data as training samples whose deformation pattern is similar to the prediction period.

## Conclusions

The conventional methods for landslide displacement prediction require enormous professional monitoring data. To overcome such limitations, a novel method is proposed by integrating the WT, ABC and KELM algorithms (WT-ABC-KELM). The Shuping landslide in the TGRA in China is taken as a case study. The WT-ELM, ABC-KELM, ELM, and SVM methods were implemented for comparison purpose. The results suggest:

- The prediction methods applying the WT algorithm performed better. The WT algorithm can decompose a highly nonlinear series of landslide displacement into sub-sequences with different frequencies to reduce the difficulty of prediction and improve the prediction accuracy.
- The ABC algorithm can efficiently optimize the parameters of the KELM, and the hybrid model of ABC-KELM outperformed the ELM and SVM models. The predicted value of KELM is fixed compared to the ELM algorithm.

Overall, the proposed WT-ABC-KELM method integrates the advantages of the three computational intelligence algorithms and produces better prediction accuracy than the WT-ELM, ABC-KELM, ELM, and SVM methods.

## Methods

**Wavelet transform.** The wavelet transform (WT) is an effective algorithm in signal processing, which provides good localization features in both time and frequency domains. The original signal can be decomposed by a set of wavelet functions and are obtained by stretching and translating the mother wavelet. The mother wavelet can be defined as follows:



$$\begin{cases} \int_{-\infty}^{+\infty} \varphi(t) dt = 0 \\ \varphi_{a,b}(t) = |a|^{-0.5} \varphi\left(\frac{t-b}{a}\right) \end{cases} \quad (4)$$

where  $\varphi_{a,b}(t)$  is a successive wavelet function,  $a$  is a frequency factor,  $b$  is a time factor. The WT algorithms can be divided into two classes: the continuous wavelet transformation (CWT) and discrete wavelet transformation (DWT). Compared to the CWT, which requires complex computation and massive data, the DWT requires less time and is easy to implement.

The DWT algorithm proposed by Mallat (1989)<sup>51</sup> has been widely used. It applies high-pass and low-pass filters to extracting approximation and detail sequence from the original signal. From a given signal  $S(t)$ , one approximation series  $a$  and  $n$  detail series ( $d_1, \dots, d_n$ ) can be extracted by Mallat's algorithm. There are many choices for the wavelet functions, such as the Haar<sup>52</sup>, Meyer<sup>53</sup>, Daubechies<sup>54</sup> and so on. The Daubechies wavelet (db  $N$ , where  $N$  is the degree of wavelet) is smooth, orthogonal and compactly supported. In the proposed method, the wavelet function of db 4 was utilized to decompose landslide displacement sequence.

**Kernel-based extreme learning machine.** The extreme learning machine (ELM), proposed by Huang *et al.* (2006), is a single-hidden layer feed-forward neural network (SLFN) with randomly generated hidden nodes<sup>36</sup>. Because of its excellent generalization ability and fast learning speed, the ELM has been adopted in various fields recently<sup>55-57</sup>. For  $N$  arbitrary samples ( $x_i, t_i$ ), where  $x_i = [x_{i1}, x_{i2}, \dots, x_{im}]^T \in R^m, t_i = [t_{i1}, t_{i2}, \dots, t_{in}]^T \in R^n$ , the standard SLFNs with  $M$  hidden nodes and activation function  $g(x)$  are mathematically written as:

$$\sum_{i=1}^M \beta_i g(W_i \cdot X_j + b_i) = O_j, j = 1, 2, \dots, N \quad (5)$$

where  $W_i = [W_{1i}, W_{2i}, \dots, W_{mi}]^T$  is the weight vector connecting the  $i$ th hidden node and the input nodes;  $\beta_i = [\beta_{1i}, \beta_{2i}, \dots, \beta_{ni}]^T$  is the weight vector connecting the  $i$ th hidden node and the output nodes;  $b_i$  is the threshold of the  $i$ th hidden node,  $W_i \cdot X_j$  is the inner product of  $W_i$  and  $X_j$ ; and  $O_j = [O_{j1}, O_{j2}, \dots, O_{jn}]^T$  is the output vector of SLFNs.

The standard SLFNs with  $M$  hidden nodes and activation function  $g(x)$  can approximate these  $N$  samples with zero error, which means that  $\sum_{j=1}^M \|O_j - t_j\| = 0$ , i.e., there exist  $\beta_i, w_i$  and  $b_i$  such that

$$\sum_{i=1}^M \beta_i g(w_i \cdot x_j + b_i) = y_j, j = 1, 2, \dots, N \quad (6)$$

The above equation can be expressed compactly as follows:

$$H\beta = Y \quad (7)$$

where  $H$  is the output matrix of the hidden layer. As the randomly selected input weight and the hidden-layer threshold are determined, the network training is equivalent to finding the least squares solution of  $\hat{\beta}$ . The equation can be written as:

$$\hat{\beta} = H^+Y \quad (8)$$

where  $H^+$  is the Moore-Penrose generalized inversion of matrix  $H$ .

In order to overcome the randomness of the ELM algorithm, and to improve its generalization capability and stability, Huang *et al.* extended the ELM into kernel learning and proposed the kernel-based ELM (KELM)<sup>39</sup>. Based on the orthogonal projection method and ridge regression theory, the output weight  $\beta$  can be calculated by adding a positive constant  $1/C$  as:

$$\beta = H^T(1/C + HH^T)^{-1}Y \quad (9)$$

Hence, the output function of ELM is shown as follows:

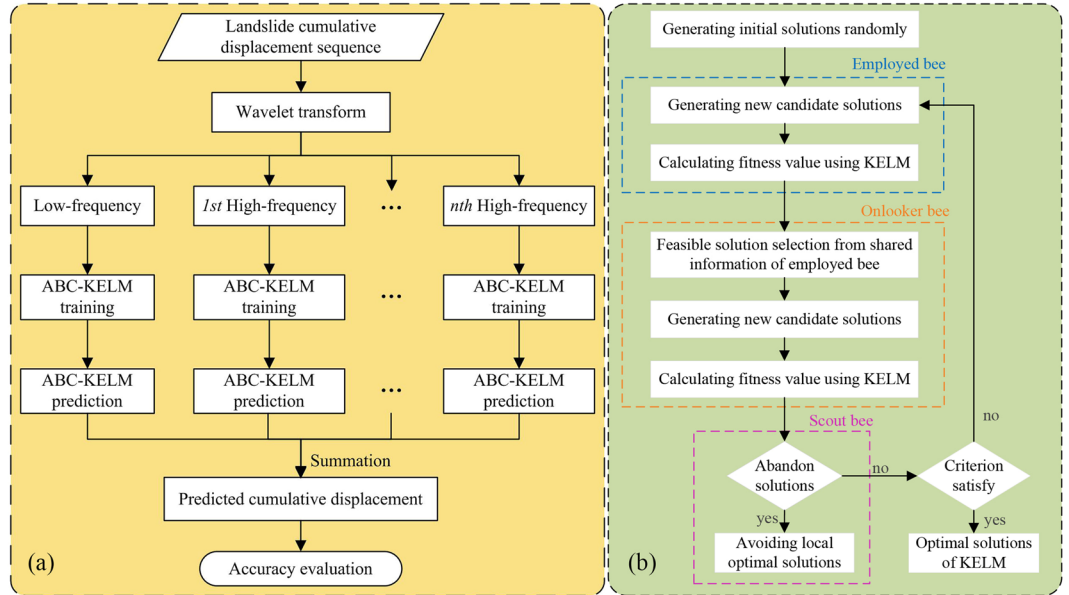
$$f(x) = h(x)H^T(1/C + HH^T)^{-1}Y \quad (10)$$

Usually, the output function  $h(x)$  is unknown for users, a kernel matrix for the ELM can be utilized to replace  $h(x)$ . The output function of the KELM can be written as follows:

$$f(x) = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T (1/C + HH^T)^{-1}Y \quad (11)$$

where  $K(x, x_i)$  is the kernel function.

**Artificial Bee Colony algorithm.** Artificial bee colony (ABC) is one of the latest optimization algorithms, and is being widely studied and successfully applied to multiple contexts<sup>42</sup>. The ABC algorithm was inspired by



**Figure 10.** (a) The flowchart of the proposed forecast method; (b) the hybrid model of ABC-KELM.

the intelligent foraging behaviour of honey bee swarms. In this algorithm, three types of bees are defined: the employed bees are associated with specific food sources (problem solutions), the onlooker bees in the hive watch the dance of employed bees and aim to choose a good food source (suboptimal solutions), and the scout bees search for new food sources (candidate solutions) in a random way.

The employed bees search for food sources in their memory and meanwhile they share the information of these food sources with the onlooker bees. The onlooker bees tend to select better food sources from those found by the employed bees. The food source with higher quality (lower best cost) will have a better chance to be selected by the onlooker bees than the one with lower quality. The scout bees are translated from a few employed bees, who abandon their food sources and search for new ones. In general, an iteration of ABC optimization algorithm contains four steps: the food source initialization, employed bee phase, onlooker bee phase, and the scout bee phase (Fig. 10b).

**The proposed coupling model and accuracy evaluation.** Landslide displacement is caused by various factors and contains multi-level information of deformation. In the proposed method, the original total displacement is decomposed into several sub-sequences with different frequencies. Subsequently, each displacement sub-sequence is predicted separately by the KELM, while the parameters of KELM were optimized using the ABC. The predicted total displacement is obtained by summing up all the predicted displacement sub-sequences. The flowchart of the proposed prediction method is shown in Fig. 10.

Accuracy evaluation is an important component for performance test of the proposed model. The statistical indexes, such as the root mean square error (RMSE), mean absolute percentage error (MAPE) and correlation coefficient (R), were utilized to verify the efficiency of the proposed method. Smaller values of RMSE and MAPE, and larger value of R indicate higher prediction accuracy. The equations of the three indexes are shown as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{D}_i - D_i)^2} \tag{12}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{D}_i - D_i}{D_i} \right| \tag{13}$$

$$R = \frac{\sum_{i=1}^N (D_i - \bar{D})(\hat{D}_i - \bar{\hat{D}})}{\sqrt{\sum_{i=1}^N (D_i - \bar{D})^2} \sqrt{\sum_{i=1}^N (\hat{D}_i - \bar{\hat{D}})^2}} \tag{14}$$

where  $N$  is the number of displacement values;  $D_i$  is the observed displacement value;  $\hat{D}_i$  is the predicted displacement value;  $\bar{D}$  is the mean of observed values;  $\bar{\hat{D}}$  is the mean of predicted values.

**Data Availability Statement.** The relevant datasets used in this study are available from the corresponding author on reasonable request.

## References

- Fan, X. *et al.* Failure mechanism and kinematics of the deadly June 24th 2017 Xinmo landslide, Maoxian, Sichuan, China. *Landslides* **14**, 203–218 (2017).
- Intrieri, E. *et al.* The Maoxian landslide as seen from space: detecting precursors of failure with Sentinel-1 data. *Landslides* **15**, 123–133 (2017).
- Qiu, J. *et al.* The catastrophic landslide in Maoxian County, Sichuan, SW China, on June 24, 2017. *Nat Hazards* **89**, 1485–1493 (2017).
- Petley, D. The Landslide Blog. American Geophysical Union Blog Network, <http://blogs.agu.org/landslideblog/> (2017).
- Baum, R. L. & Godt, J. W. Early warning of rainfall-induced shallow landslides and debris flows in the USA. *Landslides* **7**, 259–272 (2010).
- Intrieri, E., Gigli, G., Mugnai, F., Fanti, R. & Casagli, N. Design and implementation of a landslide early warning system. *Eng. Geol.* **147**, 124–136 (2012).
- Crosta, G. B. & Agliardi, F. How to obtain alert velocity thresholds for large rockslides. *Phys. Chem. Earth* **27**, 1557–1565 (2012).
- Saito, M. Proceedings of the 6th international conference on soil mechanics and foundation engineering (ed. Montre al, Que.) 537–541 (Oxford, 1965).
- An, H. *et al.* Development of time-variant landslide-prediction software considering three-dimensional subsurface unsaturated flow. *Environ Modell Softw* **85**, 172–183 (2016).
- Calvello, M., Cuomo, S. & Ghasemi, P. The role of observations in the inverse analysis of landslide propagation. *Comput. Geotech.* **92**, 11–21 (2017).
- Conte, E., Donato, A. & Troncone, A. A simplified method for predicting rainfall-induced mobility of active landslides. *Landslides* **14**, 35–45 (2017).
- Goetz, J. N., Brenning, A., Petschko, H. & Leopold, P. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. *Comput. Geosci.* **81**, 1–11 (2015).
- Krkač, M., Špoljarić, D., Bernat, S. & Arbanas, S. M. Method for prediction of landslide movements based on random forests. *Landslides* **14**, 947–960 (2016).
- Miao, S., Hao, X., Guo, X., Wang, Z. & Liang, M. Displacement and landslide forecast based on an improved version of Saito's method together with the Verhulst-Grey model. *Arab J Geosci* **10**, 53 (2017).
- Wang, W. *et al.* A new approach for modeling landslide movement over 3D topography using 3D discontinuous deformation analysis. *Comput. Geotech.* **81**, 87–97 (2017).
- Sasahara K. *Geotechnical hazards from large earthquakes and heavy rainfalls* (eds Hazarika, H., Kazama, M. & Lee, W.) 223–233 (Japan, 2017).
- Jiang, J. *et al.* Numerical simulation of Qiaotou Landslide deformation caused by drawdown of the Three Gorges Reservoir, China. *Environ. Earth Sci.* **62**, 411–419 (2011).
- van Asch, T. W. J., Malet, J. P., van Beek, L. P. & Amitrano, D. Techniques, issues and advances in numerical modelling of landslide hazard. *Bull. Soc. Géol. Fr.* **178**, 65–88 (2007).
- Corominas, J., Moya, J., Ledesma, A., Lloret, A. & Gili, J. A. Prediction of ground displacements and velocities from groundwater level changes at the Vallcebre landslide (Eastern Pyrenees, Spain). *Landslides* **2**, 83–96 (2005).
- Du, J., Yin, K. & Lacasse, S. Displacement prediction in colluvial landslides, three Gorges reservoir, China. *Landslides* **10**, 203–218 (2013).
- Zhou, C., Yin, K., Cao, Y. & Ahmed, B. Application of time series analysis and PSO-SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir, China. *Eng. Geol.* **204**, 108–120 (2016).
- Petley, D. Global patterns of loss of life from landslides. *Geology* **40**, 927–930 (2012).
- Cheng, G. *et al.* Automatic landslide detection from remote-sensing imagery using a scene classification method based on BoVW and pLSA. *Int. J. Remote Sens.* **34**, 45–59 (2013).
- Cheng, G. & Han, J. A survey on object detection in optical remote sensing images. *ISPRS J. Photogramm.* **117**, 11–28 (2016).
- Cracknell, M. J. & Reading, A. M. Geological mapping using remote sensing data: a comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Comput. Geosci.* **63**, 22–33 (2014).
- Nejad, F. P. & Jaksa, M. B. Load-settlement behavior modeling of single piles using artificial neural networks and CPT data. *Comput. Geotech.* **89**, 9–21 (2017).
- Wang, X., Li, Y., Chen, T., Yan, Q. & Ma, L. Quantitative thickness prediction of tectonically deformed coal using extreme learning machine and principal component analysis. *Comput. Geosci.* **101**, 38–47 (2017).
- Zhou, C. *et al.* Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. *Comput. Geosci.* **102**, 23–37 (2018).
- Crozier, M. J. *Landslides: causes, consequences and environment* (ed. Crozier, M. J.) 30–50 (London, 1986).
- Cao, Y., Yin, K., Alexander, D. E. & Zhou, C. Using an extreme learning machine to predict the displacement of step-like landslides in relation to controlling factors. *Landslides* **13**, 725–736 (2016).
- Liu, Z., Shao, J., Xu, W., Chen, H. & Shi, C. Comparison on landslide nonlinear displacement analysis and prediction with computational intelligence approaches. *Landslides* **13**, 725–736 (2014).
- Yao, W., Zeng, Z., Lian, C. & Tang, H. Training enhanced reservoir computing predictor for landslide displacement. *Eng. Geol.* **188**, 101–109 (2016).
- Zhou, C. & Yin, K. Landslide displacement prediction of WA-SVM coupling model based on chaotic sequence. *Electr. J. Geol. Eng.* **19**, 2973–2987 (2014).
- Zhu, X. *et al.* Comparison of two optimized machine learning models for predicting displacement of rainfall-induced landslide: A case study in Sichuan Province, China. *Eng. Geol.* **218**, 213–222 (2017).
- Taylor, J. A., Mathieu, J. L., Callaway, D. S. & Poolla, K. Price and capacity competition in balancing markets with energy storage. *Energy Syst.* **8**, 169–197 (2017).
- Huang, G., Zhu, Q. & Siew, C. Extreme learning machine: theory and applications. *Neurocomputing* **70**, 489–501 (2006).
- Vasu, N. N. & Lee, S. R. A hybrid feature selection algorithm integrating an extreme learning machine for landslide susceptibility modeling of Mt. Woomeyon, South Korea. *Geomorphology* **263**, 50–70 (2016).
- Yang, Z., Ce, L. & Lian, L. Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. *Appl. Energy* **190**, 291–305 (2017).
- Shihabuddeen, K. V., Pillai, G. N. & Peethambaran, B. Prediction of landslide displacement with controlling factors using extreme learning adaptive neuro-fuzzy inference system (ELANFIS). *Appl. Soft Comput.* **61**, 892–904 (2017).
- Barzegar, R., Fijani, E., Moghaddam, A. A. & Tziritis, E. Forecasting of groundwater level fluctuations using ensemble hybrid multi-wavelet neural network-based models. *Sci. Total Environ.* **599**, 20–31 (2017).
- Huang, G., Zhou, H., Ding, X. & Zhang, R. Extreme learning machine for regression and multiclass classification. *IEEE T. Syst. Man Cy. B* **42**, 513–529 (2012).
- Karaboga, D. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *J. Global Optim.* **39**, 459–471 (2007).
- Aussem, A., Campbell, J. & Murtagh, F. Wavelet-based feature extraction and decomposition strategies for financial forecasting. *J. Comp. Intel. Fin.* **6**, 5–12 (1998).

44. Adamowski, J. & Chan, H. F. A wavelet neural network conjunction model for groundwater level forecasting. *J. Hydrol.* **407**, 28–40 (2011).
45. Barzegar, R., Adamowski, J. & Moghaddam, A. A. Application of wavelet-artificial intelligence hybrid models for water quality prediction: a case study in AjiChay River, Iran. *Stoch. Env. Res. Risk A.* **30**, 1797–1819 (2016).
46. Moosavi, V., Vafakhah, M., Shirmohammadi, B. & Behnia, N. A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods. *Water Resour. Manag.* **27**, 1301–1321 (2013).
47. Wang, W. & Ding, J. Wavelet network model and its application to the prediction of hydrology. *Nat. Sci.* **1**, 67–71 (2003).
48. Vapnik, V. N. *The nature of statistical learning theory* (ed. Vapnik, V. N.) 138–146 (New York, 2000).
49. Deo, R. C. & Şahin, M. Application of the extreme learning machine algorithm for the prediction of monthly effective drought index in eastern Australia. *Atmos. Res.* **153**, 512–525 (2015).
50. Lian, C., Zeng, Z., Yao, W. & Tang, H. Extreme learning machine for the displacement prediction of landslide under rainfall and reservoir level. *Stoch. Env. Res. Risk A.* **28**, 1957–1972 (2014).
51. Mallat, S. G. A theory for multi-resolution signal decomposition: the wavelet representation. *IEEE T. Pattern Anal.* **11**, 674–693 (1989).
52. Haar, A. Zur theorie der orthogonalen funktionensysteme. *Math Ann.* **69**, 331–371 (1910).
53. Meyer, Y. *Ondelettes et opérateurs* (1990).
54. Daubechies, I. *Ten lectures on wavelets*, Vol. 61. (SIAM, 1992).
55. Lima, A. R., Cannon, A. J. & Hsieh, W. W. Nonlinear regression in environmental sciences using extreme learning machines: a comparative evaluation. *Environ. Modell. Softw.* **73**, 175–188 (2015).
56. Ortín, S. *et al.* A unified framework for reservoir computing and extreme learning machines based on a single time-delayed neuron. *Sci. Rep.* **5** (2015).
57. Zhu, H. *et al.* Hyperspectral Imaging for Presymptomatic Detection of Tobacco Disease with Successive Projections Algorithm and Machine-learning Classifiers. *Sci. Rep.* **7** (2017).

## Acknowledgements

This paper was prepared as part of the projects “Studies on the predictive criteria and mechanism of landslides in gently-dipped strata in the Three-Gorges reservoir” (No. 41572292) and “Study on the hydraulic properties and the rainfall infiltration law of the ground surface deformation fissure of colluvial landslides” (No. 41702330). The National Natural Science Foundation of China funds both the projects. The comments from the two anonymous reviewers and the editor have significantly improved the quality of this article. The authors thank Dr. Gao Lin for his assistance in model implementation. The first author would like to thank the China Scholarship Council for funding his research at the University of Florence, Italy.

## Author Contributions

Chao Zhou designed the experiments, analyzed the data and contributed designing the analysis tools; Kunlong Yin and Ying Cao performed the experiments and wrote the first draft; and Bayes Ahmed and Xiaolin Fu commented on the scientific integrity and edited the manuscript.

## Additional Information

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018