Time series analysis and long short-term memory neural network to predict landslide displacement

Beibei Yang [a, b], Kunlong Yin [a] *, Suzanne Lacasse [b], and Zhongqiang Liu [b]

[a] Faculty of Engineering, China University of Geosciences, Wuhan 430074, China

[b] Norwegian Geotechnical Institute, Oslo 0806, Norway

Abstract: A good prediction of landslide displacement is an essential component for implementing an early warning system. In the Three Gorges Reservoir Area (TGRA), many landslides deform distinctly and in steps from April to September each year under the influence of seasonal rainfall and periodic fluctuation in reservoir water level. The sliding becomes more uniform again from October to April. This landslide deformation pattern leads to accumulated displacement versus time showing a step-wise curve. Most of the existing predictive models express static relationships only. However, the evolution of a landslide is a complex nonlinear dynamic process. This paper proposes a dynamic model to predict landslide displacement, based on time series analysis and long short-term memory neural network (LSTM). The accumulated displacement was decomposed into a trend term and a periodic term in the time series analysis. A cubic polynomial function was selected to predict the trend displacement. By analyzing the relationships between landslide deformation, rainfall and reservoir water level, a LSTM model was used to predict the periodic displacement. The LSTM approach was found to more properly model the dynamic characteristics of landslides than static models, and make full use of the historical information. The performance of the model was validated with the observations of two step-wise landslides in the TGRA, the Baishuihe landslide and Bazimen landslide. The application of the model to those two landslides demonstrates that the LSTM model provides a good representation of the measured displacements and gives a more reliable prediction of landslide displacement than the static support vector machine (SVM) model. It is concluded that the proposed model can be used to effectively predict the displacement of step-wise landslides in the TGRA.

Keywords: Displacement prediction; Step-wise landslide; Time series; Long short-term memory neural network; Three Gorges Reservoir

1. Introduction

Landslides are one of the most damaging geological disasters in many areas of the world and can cause a large number of casualties and very important property losses every year (Haque et al. 2016; Froude and Petley 2018; Hong et al. 2016; Zhang and Huang 2018). In China, for instance, an annual report (China Institute of Geo-Environment Monitoring 2017) shows that 9710 geological hazards occurred in 2016 and 7403 of them were landslides. Other geological hazards in 2016 included 1484 avalanches, 584 debris flows, 221 ground collapses, 12 ground fissures and 6 cases of ground subsidence. The Xinmo landslide on 24 June 2017 caused 10 fatalities, 73 persons were reported missing and 64 houses were damaged. It was one of the most catastrophic landslide events in recent years in China (Fan et al. 2017, 2018; Scaringi et al. 2018).

Reliable early warning systems are a reasonable approach for landslide risk reduction. Such systems can be successfully implemented if one can predict landslide displacement (Yao et al., 2015). As an example, at the 1985 Xintan landslide, located 26 km upstream of the Three Gorges Dam (TGD), the effective forecasting of the landslide displacement helped reduce significantly the economic losses and casualties (Zhou et al., 2016).

Since Saito (1965) proposed its empirical formula for landslide prediction, numerous landslide prediction models have been developed. Landslide prediction models can be categorized into two groups: physical models and data-driven models (Huang et al., 2017). The physical models primarily use general creep theory, large-scale laboratory experiments and physical parameters monitored in situ to predict landslide events (Helmstetter et al. 2004; Mufundirwa et al. 2010). The

models can provide clear physical explanations of landslides, but the implementation of sufficiently accurate models can be complex, time consuming and expensive (Corominas et al. 2005; Thiebers, 2014; Ma et al. 2017). Compared with physical models, the data-driven models are often more popular because of simple process, accurate prediction and lower costs (Corominas et al. 2005; Zhou et al. 2018a).

A variety of data-driven models have had significant developments in recent years (Kayacan et al. 2010; Liu et al. 2014; Zhou et al. 2016; Liu et al. 2018; Reichenbach et al. 2018). Among data-driven models, machine learning (ML) methods are widely applied for landslide displacement predictions and have achieved good performances (Zhou et al., 2018a). Du et al. (2013) decomposed the cumulative displacement into trend and periodic components, using a back-propagation neural network (BPNN) to predict the periodic displacement. Lian et al. (2015) applied artificial neural network (ANN) to model landslide displacement and achieved a good prediction performance. Cao et al. (2016) developed an extreme learning machine (ELM) considering the response of influencing factors, and achieved promising results with their model. Zhou et al. (2016) and Miao et al. (2018) developed displacement prediction models based on the time series analysis and the support vector machine (SVM) model and obtained highly accurate results. To overcome some deficiencies of the classical ML methods, Li et al. (2018) proposed a novel data-driven framework with an ensemble-based ELM and copula models. The effectiveness and utility of the proposed model have been confirmed by applying it to back-calculate the displacement prediction of a landslide in the Three Gorges Reservoir Area (TGRA).

Most of the existing prediction models are static models, which consider landslide displacement prediction as a static regression problem (Yao et al., 2015). However, landslide evolution is a complex nonlinear dynamic process (Qin et al., 2002). The influencing factors and deformation conditions at one time affect the deformation and stability conditions at the next time (Xu and Niu 2018). Dynamic modeling approaches, which can include relationships between different times of landslide displacement time series, were therefore used in this study to develop a landslide displacement prediction model.

A major part of the research on dynamic modeling has concentrated on recurrent neural network (RNN), which is a special type of ANN. RNN can remember historical information and apply it to the current output, which leads to the achievement of state feedback in the networks (Han et al. 2004; Chen and Chou 2012). Since conventional RNN fails to capture the long temporal dependency for the input sequence, an improved version of RNN, named long short-term memory neural network (LSTM), was developed (Hochreiter and Schmidhuber 1997). Their long-range dependencies are more accurate than conventional RNN (Hasim et al., 2014). The LSTM model was used earlier to predict the displacement of the Baijiabao landslide in the TGRA and achieved better performance than the BPNN and SVM models (Xu and Niu 2018).

In this paper, a dynamic model based on time series analysis and LSTM was proposed to predict landslide displacement. Through time series analysis, the accumulated displacement was decomposed into a trend term and a periodic term. The trend term displacement was predicted using a cubic polynomial function. The periodic term displacement was predicted by a LSTM model based on the analysis of the relationship between landslide displacement and key influencing factors. The total displacement forecast was obtained by adding the trend and periodic displacements. The proposed model was used to predict the displacements of two step-wise landslides in the TGRA, the Baishuihe and Bazimen landslides. For comparison, the SVM model was also used to predict the periodic term displacement.

2. Machine learning models for landslide displacement prediction

2.1. Recurrent neural network and long short-term memory neural network

2.1.1. Recurrent neural network

An artificial neural network (ANN) is typically composed of an input layer, a hidden layer and an output layer. The input layer is the first layer of the neural network and brings the initial data into the system for further processing by subsequent layers of artificial neurons. The hidden layer is a layer between the input and output layers, in which artificial neurons receive a set of weighted inputs and produce an output through an activation function. The output layer is the last layer of

neurons that produces the outputs from the ANN modeling.

In traditional ANN, the units are related to each other from one layer to the other (input layer, hidden layer and output layer), but independent in one given layer, as illustrated in Fig. 1. Recurrent neural network (RNN) is an expansion of ANN. RNN contains recurrent connections, and each unit in the hidden layer is related to the others for different time steps. For example, in the hidden layer, the unit at time step t-1 is connected to the unit at time step t in the RNN, while this connection does not exist in the ANN (Fig. 1). The recurrent connections in the RNN implement passing information from one step of the network (time step t-1) to the next one (time step t). For a given input sequence $x=(x_1, x_2,\ldots, x_T)$, the hidden state vector sequence is $h=(h_1, h_2,\ldots, h_T)$ and the output sequence is $y=(y_1, y_2,\ldots, y_T)$. The output sequence y can be obtained by iterating the following equations from time t=1 to T (Fan et al., 2014):

$$h_t = \Re(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \tag{1}$$

$$y_t = W_{hy}h_t + b_y \tag{2}$$

where x_t and y_t are the input and output at time step t; h_t and h_{t-1} are the hidden states at time step t and t-1, respectively; T is the prediction period; W_{xh} is the weight matrix between input and hidden vectors; W_{hh} represents the weight matrix between different time steps of the hidden vectors; W_{hy} represents the weight matrix that connects the hidden information to the output vectors; b_h and b_y are the corresponding bias vectors of W_{hh} and W_{hy}, respectively; and $\Re$ is the nonlinear activation function for hidden nodes and is usually a sigmoid function or hyperbolic tangent function.

RNN shares the same parameters (W_{xh}, W_{hh} and W_{hy}) across all steps, which reduces the total number of learning parameters and the learning workload of networks. The hidden state (h_t) contains information about the history of the past elements of a sequence. Because of this special design, RNN exhibits a superior capability of modeling nonlinear time series problems effectively (Graves et al., 2013).

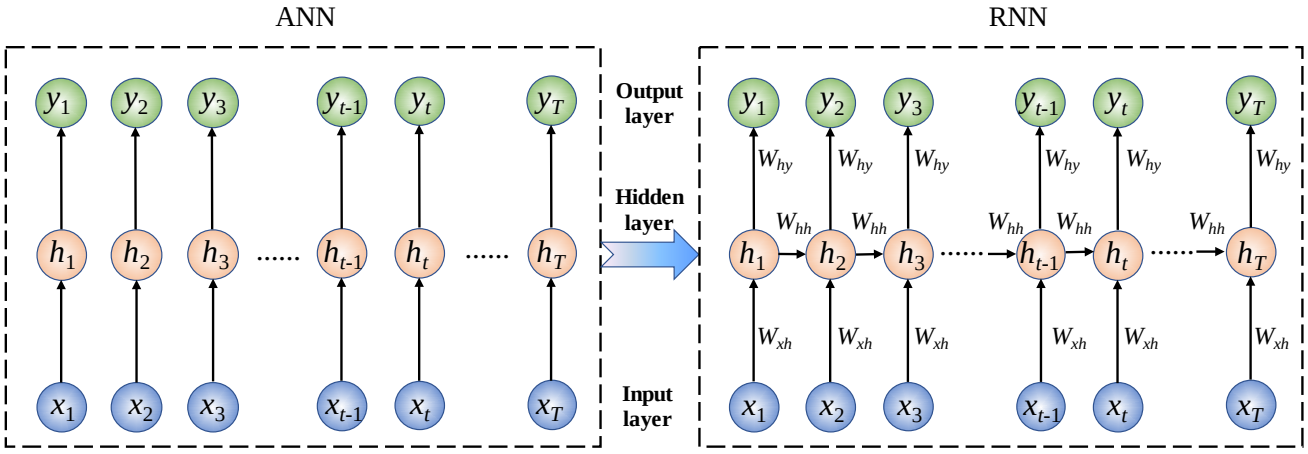

Fig. 1 Comparison between artificial neural network (ANN) and recurrent neural network (RNN)

2.1.2. *Long short-term memory neural network*

In theory, RNN should be capable of handling arbitrarily long sequences (Vincent et al., 2010). In practice however, RNN can not do so because of the vanishing gradient or exploding gradient problems (Bengio et al., 1994). To overcome such disadvantage of conventional RNN, Hochreiter and Schmidhuber (1997) proposed an improved version of RNN named long short-term memory neural network (LSTM). Compared with the conventional RNN, the basic unit of the hidden layer in LSTM is a memory block (Fig. 2). The memory block contains memory cell and three gates, named 'Input gate', 'Forget gate' and 'Output gate' (Felix and Jürgen 2000). These three gates regulate the flow of information into and out of the memory cell. The 'Input gate' controls the flow of input activations into the memory cell, meaning that the 'Input gate' decides to feed the input data. The 'Forget gate' controls whether the information from the previous time step is remembered or forgotten. When the 'Forget gate' is open, information from the previous time step is passed along to the next time step;

when the 'Forget gate' is closed, all the information from the previous time step is forgotten and cannot be passed along to the next time step. This means that the 'Forget gate' can filter information, keeping useful information and throwing away useless information. The 'Output gate' controls the flow of output activations into other blocks or to the final results (Xu and Niu 2018). The key difference between RNN and LSTM lies in the formulation of the hidden vector (h_t). The hidden vector (h_t) in LSTM neural network can be obtained as follows:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (3)$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right) \quad (4)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (5)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_0) \quad (6)$$

$$h_t = o_t \tanh(c_t) \quad (7)$$

where i_t, f_t, o_t and c_t are the values of the input gate, forget gate, output gate and memory cell in the memory block; b_i, b_f, b_o and b_c are their corresponding bias values; σ is the sigmoid function; W_x represents the weights between input nodes and hidden nodes; W_h represents the weights between hidden nodes and memory cell; and W_c represents the weights that connect memory cell to output nodes.

The output y_t can then be obtained by Eq.2. The output sequence $y=(y_1, y_2,\ldots,y_T)$ will be obtained by iterating Eq.2 to Eq.7 from times t=1 to T.

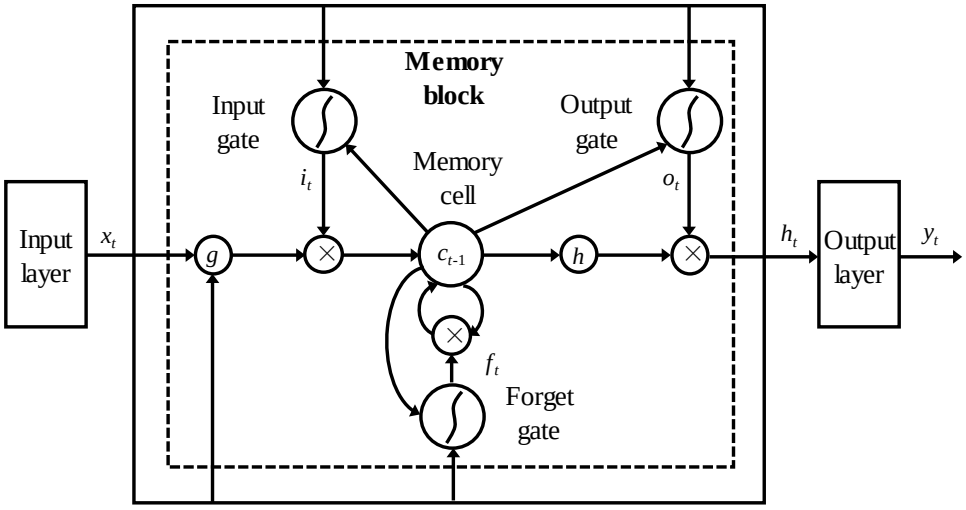

Fig. 2 Long short-term memory neural network (LTSM) architecture (Ma et al., 2015)

2.2. Support vector machine model

The support vector machine (SVM) model is a non-linear regression forecasting method proposed by Cortes and Vapnik (1995). The essence of the SVM model is to map data into a high-dimensional feature space through a non-linear transformation and then perform linear regression within this space (Smola and Scholkopf 2004). The SVM model is a static model, which has performed well earlier for predicting the displacement of step-wise landslides in the TGRA (Zhou et al. 2016; Miao et al. 2018). Particle swarm optimization (PSO) is applied to search for the optimal parameters of the SVM model. Inspired from the feeding behavior characteristic of a bird flock, the PSO is used for solving optimization problem (Eberhart and Kennedy 1995), where each particle in the algorithm is regarded as a potential solution to the optimization problem. A fitness function is defined to search for the optimal solution in the entire solution space.

2.3. Proposed model

2.3.1. Decomposition of the displacement time series into trend and periodic terms

The displacement will be modeled with two components: a trend component and a periodic (i.e. changing with time) component (Du et al., 2013). Many factors can affect landslide displacement (Zhou et al., 2018a). On the long-term, the deformation trend is controlled by internal geological conditions, such as lithology, geological structure and progressive weathering. The landslide displacement due to these factors increases approximately monotonically with time, and usually on a long-time scale. Long term creep, degradation and weathering are not included explicitly in the model. This long-term displacement will be the trend component in the model.

The components that affect landslide displacement on the slopes in the TGRA over a short time period, are on the other hand, considered explicitly. The short-term deformation in the TGRA is influenced by two external triggering factors (rainfall and reservoir water level). This short-term displacement is the periodic component in the model.

The accumulated displacement time series can therefore be decomposed as follows:

$$S(t) = \phi(t) + \eta(t) \tag{8}$$

where t is time, $S(t)$ is the accumulated displacement, $\phi(t)$ is the trend displacement, and $\eta(t)$ is the periodic displacement.

2.3.2. Prediction process with the proposed model

The process of the proposed model, both the creation of the model and its validation through machine learning techniques, is illustrated in Fig. 3. Using the available data for a landslide, the model first creates a model of the accumulated displacement. It uses the majority of the data to model as well as possible the observations. A volume of the available data is used to fit the data, and to train the model. The training consists of data processing, fitting of trend term and building the LSTM model for periodic term. The models are then validated with the rest of the volume of data available.

The two components of the accumulated displacement, the trend term and the periodic term are predicted separately. As stated in Section 2.3.1, the trend displacement is taken as approximately monotonically increasing under the influence of internal geological conditions (lithology, geological structure, progressive weathering, etc.). The prediction model of the trend displacement can be constructed by fitting the shape of the growth curve of displacement. A polynomial function was found to give the best results and thus was used for the description of trend displacement in this study.

The periodic displacement fluctuates under the influence of two external triggers, rainfall and reservoir water level. By analyzing the response relationships between the periodic displacement, rainfall and reservoir water level, multi-layer LSTM models were developed to predict the periodic displacements for a selected location on two landslides in the TGRA. The total accumulated displacement was obtained by adding the predicted trend and periodic displacements.

The purpose of the paper is to illustrate the techniques of LSTM model for landslide displacement prediction only. The SVM model was used to predict landslide periodic displacement for the purpose of comparison with the LSTM model. Zhou et al. (2016) applied the SVM model to predict the displacement of the Bazimen landslide in the TGRA (one of the two cases studied in this paper), and achieved an improved prediction with the SVM model compared with the BPNN model. Therefore, this paper did not consider a BPNN model for comparison purpose.

To verify the goodness of the predictive model, the root mean square error (RMSE) and mean absolute percentage error (MAPE) were adopted in this study (Eqs. 9 and 10). They are two common measures of forecast accuracy, and each have advantages and disadvantages. The RMSE is a scale-dependent measure. It is useful when comparing different methods applied to the same data set, but should not be used when comparing across data sets with different scales. The MAPE has the advantage of being scale-independent and so is frequently used to compare forecast performance across different data sets, but has the disadvantage of being infinite or undefined if the observed values are zero. A complete overview of advantages and disadvantages for these two measures can be found in Hyndman and Koehler (2006). A lower RMSE and MAPE-value (minimum values of 0) reflects an improved accuracy (Huang et al., 2016).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \hat{x}_i\right)^2} \tag{9}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - \hat{x}_i}{x_i}\right| \tag{10}$$

where x_i is the measured value; $\hat{x}_i$ is the predicted value; N is the number of predicted values; $\bar{x}$ is the mean of measured values; and $\bar{\hat{x}}$ is the mean of predicted values.

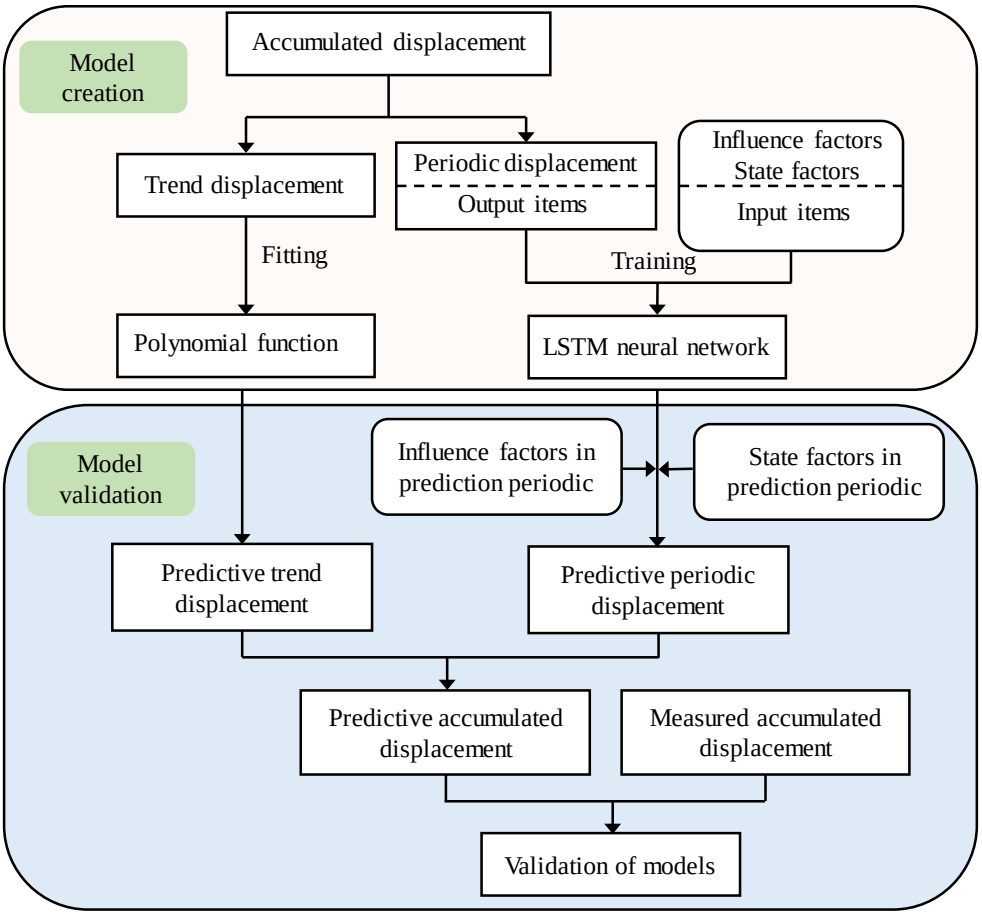

Fig. 3 Flowchart of the proposed forecast model

3. Case studies

The TGRA is well-known as a landslide-prone area in China. The construction of the dam dramatically changed the hydrogeological conditions of the entire region. The water level in this area was raised to the designed highest elevation of 175 m by 2009. Since then, a 30 m water level variation in the reservoir (from 145 m to 175 m elevation) is imposed through annual reservoir regulation. The rapid and relatively large water level variation affects negatively the stability of slopes along the shores, resulting in massive slope failures, and increasing the risk of potential landslide disasters (Yin et al. 2016; Zhou et al. 2018b).

Two well-studied examples of step-wise landslides, the Baishuihe landslide and Bazimen landslide in the TGRA, each with different deformation characteristics, were used as cases to develop and validate the proposed model. The Baishuihe landslide occurred in Zigui town on the south side of the Yangzi River, 56 km from the TGD. The Bazimen landslide is also located in Zigui town on the west side of the Xiangxi River, a major tributary of the Yangtze River, 0.8 km from the estuary. The locations of the two landslides are shown in Fig. 4. These two landslides were selected because of the quantity

and quality of the available data. The two landslides are typical and yet different:

- In the TGRA, a large number of landslides deform sharply from April to September each year, and become more stable again from October to April. This step-wise accumulated displacement versus time is observed for many landslides in the TGRA. The Baishuihe and Bazimen landslides are two step-wise deformation landslides, and are representative to test the proposed model.
- The Baishuihe and Bazimen landslides began to deform more noticeably after the first impoundment of the Three Gorges Reservoir (TGR) in 2003. GPS stations were then installed to monitor slope deformation, with 1-month data collection frequency. These two landslides have complete displacement dataset over several years.
- These two landslides have different deformation patterns. The Baishuihe landslide is a retrogressive landslide, where deformations occurred first at the bottom of the slope and retrogressed upwards (Du et al. 2013; Miao et al. 2018). The Bazimen landslide is an "advancing" landslide, where the landslide was initiated from the upper part and developed downwards (Du et al. 2013; Zhou et al. 2016). It was important to test the new model for the two types of landslide progression.

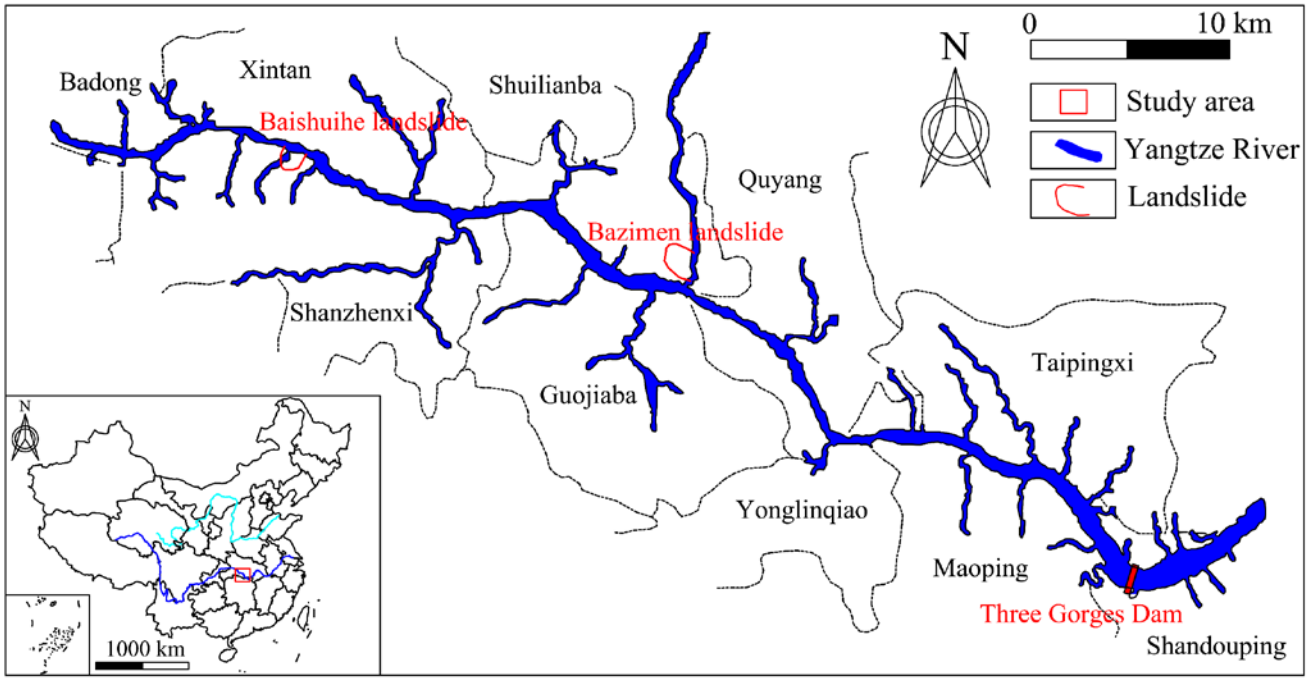

Fig. 4 Locations of Baishuihe and Bazimen landslides in the Three Gorges Reservoir Area

3.1. Baishuihe landslide

3.1.1. Geological conditions

The Baishuihe landslide is fan-shaped (Fig. 5) and its main sliding direction is oriented at N20°E. The upper boundary of the landslide is the interface between bedrock and soil, at the elevation of 400 m above the sea level. The left and right boundaries are bedrock and a gully, respectively. The frontal part of the landslide moves down to the bed of the Yangzi River and the elevation of the toe ranges from 120 m to 130 m. The landslide extends 780 m from north to south, has a width of 700 m from east to west and covers an area of 42×10^4 m^2. The average thickness of the moving masses is 30 m and the estimated volume is 1260×10^4 m^3 (Fig. 5).

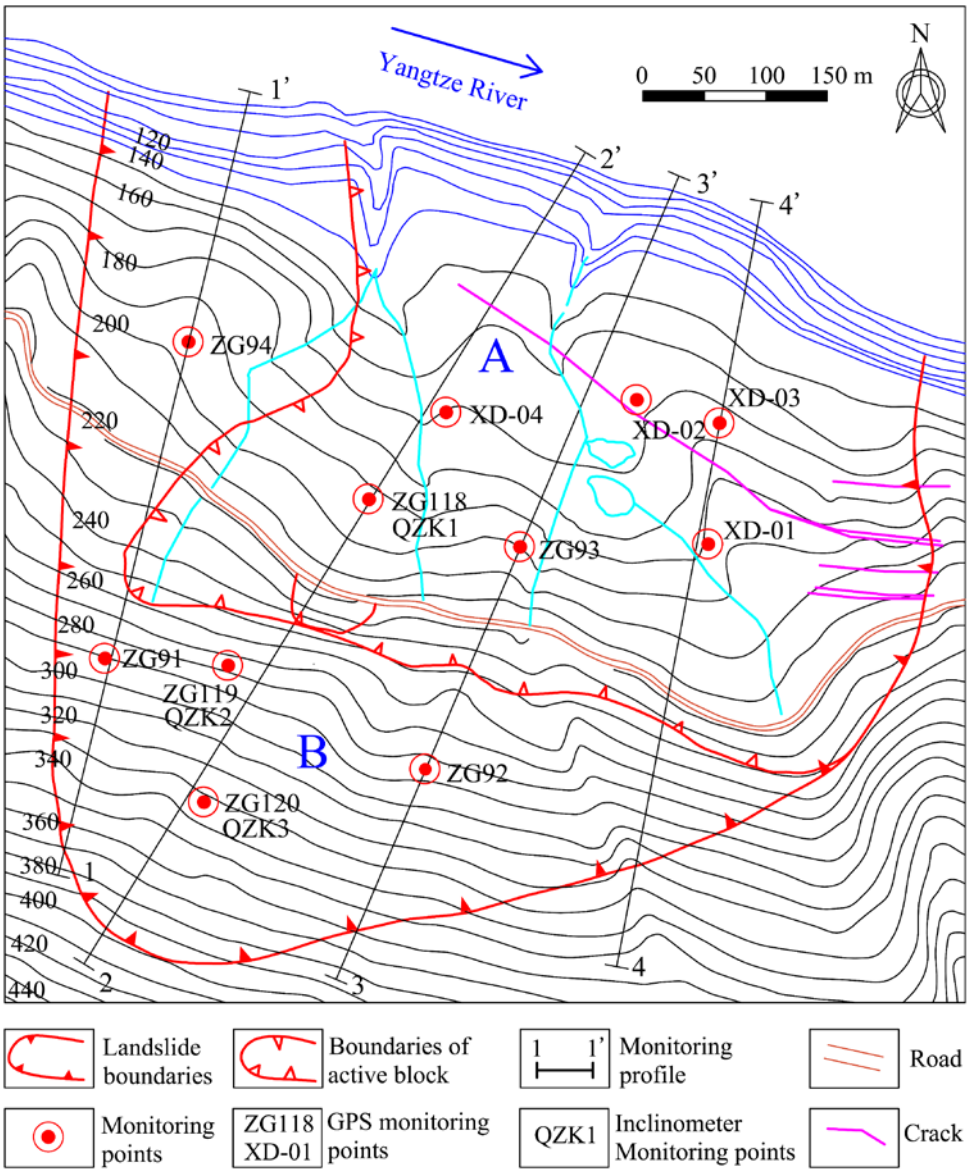

Fig. 5 Monitoring arrangement in the Baishuihe landslide and displacement blocks A and B

The moving masses in the Baishuihe landslide consist mainly of Quaternary deposits, including silty clay and fragmented rubbles with a loose and disorderly structure. The lithology of the bedrock is mainly Jurassic siltstone, silty mudstones and muddy siltstones, with dip directions of 15° and dip angles of 36°, as shown in Fig. 6a (Miao et al., 2014). Two sliding surfaces were found in the landslide, occurring at different depths. The secondary sliding surface is located at the interface between soil and the cataclastic rock at a depth varying from 12 m to 21.5 m. The initial sliding surface develops at the interface between the bottom of the cataclastic rock and bedrock, with a depth deeper than 30 m (Fig. 6b).

The Baishuihe landslide reactivates frequently and have had several intense deformation periods since 2003. A 300-m long transversal tension crack occurred at the east frontal part of the landslide in June 2003. Twenty-one residential houses were destroyed due to landslide deformations in 2004. During August 2005 and August 2006, many transversal cracks were observed on the ground surface. Approximately 10×10^4 m^3 of debris piled up on the road on June 30, 2007 (Fig. 5). The cracks in the western boundary of the landslide spread sporadically in plume form in August 2009.

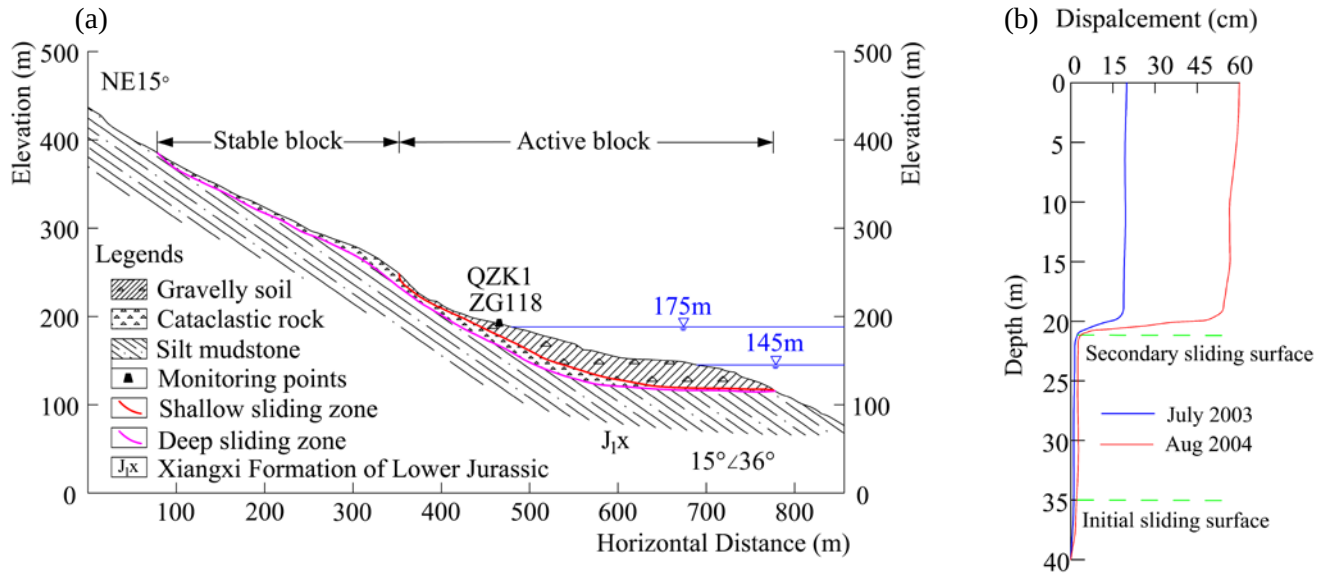

Fig. 6 a Schematic geological cross-section 2–2′ of the Baishuihe Landslide. b Lateral displacement versus depth from inclinometer QZK1, Baishuihe landslide

3.1.2. Analysis of the monitoring data

The Baishuihe landslide movements between 2003 and 2013 are typical of a retrogressive landslide, where the deformations occurred first at the bottom of the slope and then developed upwards (Du et al. 2013; Miao et al. 2018). From the site investigation and monitoring data, the Baishuihe landslide can be divided into two blocks (Fig. 5), the active block A and the relatively stable block B (Huang et al. 2017; Zhou et al. 2018a). Eleven GPS monitoring points were installed on the landslide, six on the active block and five on the relatively stable block (Fig. 5). Fig. 7 shows the measured displacement at the six GPS stations on the active block A. The accumulated displacement of the active block varied spatially. The accumulated displacements at Stations XD1 and XD3, on the east part of the landslide, were higher than at other stations. The accumulated displacements at ZG118 and ZG93 were higher than at XD2 and XD4. Fig. 7 shows that the displacement time series of Stations XD1, XD2, XD3 and XD4 either started later or were shorter than for Stations ZG118 and ZG93, while Stations ZG118 and ZG93 have similar displacement trend. Station ZG118, with longer displacement time series and relatively high displacement, was selected as an example for the analysis of the displacement with the new model.

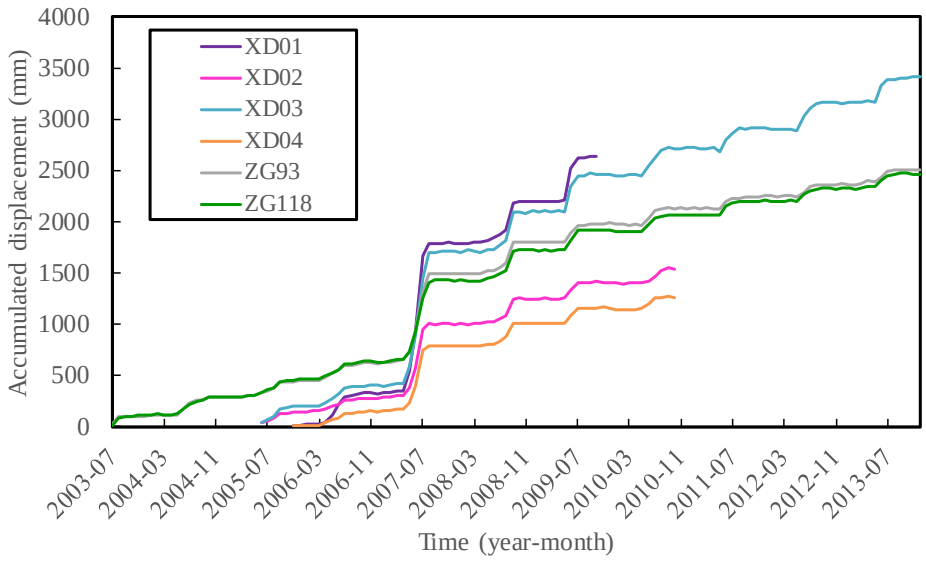

Fig.7 Accumulated displacement in the Baishuihe landslide

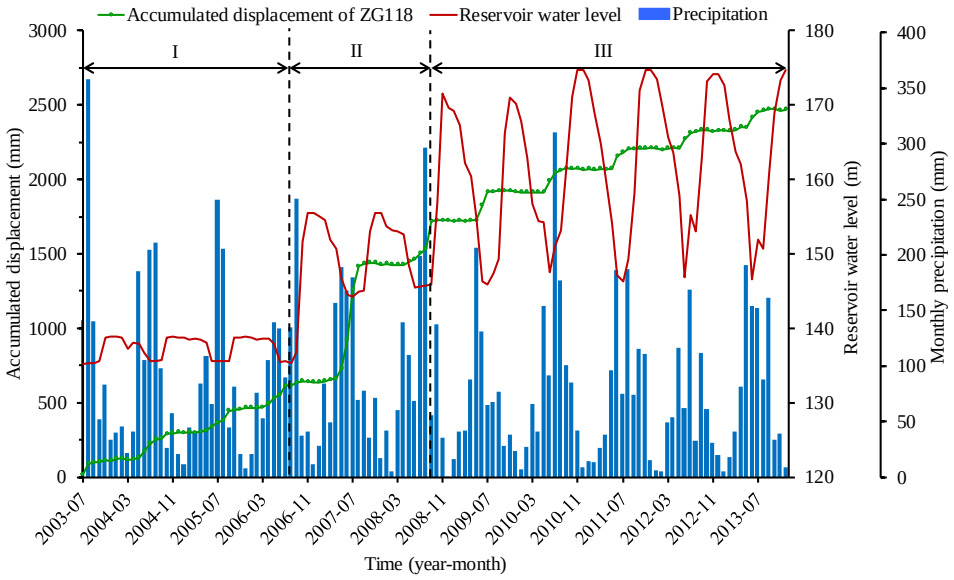

Fig.8 Rainfall, reservoir water level and accumulated displacement at ZG118 in the Baishuihe landslide

Fig. 8 shows the measured curves of accumulated displacement, reservoir water level and rainfall at Location ZG118 from August 2003 to December 2013. The fluctuation of reservoir water level can be divided into three periods.

(1) From August 2003 to August 2006, the reservoir water fluctuated between 135 m and 139 m. As only a small part of the slope was submerged in water and the water level fluctuated within a range of 4 m, the effect of reservoir water level on the landslide deformation was limited. The accumulated displacement exhibited a step-wise behavior from May to September, corresponding to rainfall in the annual flood season. The maximum monthly displacements in the flood season of 2004, 2005 and 2006 were 52.6 mm, 63.8 mm and 62.2 mm, respectively. In the dry season from October to April in each of the following year, the increase of the accumulated displacement tended to be smaller, with some maximum monthly displacements of 31.9 mm, 22.3 and 12.4 mm. From this subset of data, it was inferred that rainfall during the flood season was the main factor contributing to the step-wise landslide deformations during this period.

(2) From September 2006 to September 2008, the reservoir water fluctuated between 145 m and 155 m. Compared with the first period, a wider volume of the landslide was submerged in water, and the fluctuation of reservoir water level became larger. During the period of reservoir water rise, the displacement increased slowly, and the maximum monthly displacement was merely 19 mm. However, when the reservoir water level fell from 155 m to 145 m for the first time under a rainfall of 685.1 mm, the largest deformation with the monthly displacement of 334 mm occurred in July 2007. This phenomenon was due to the joint influence of strong rainfall and large reservoir water drawdown rate, which changed the stress field, seepage field and structure of the slope. After that, as saturation of the materials increased, the slope gradually adapted to the new regulation of the reservoir level and tended to be stable again.

(3) From October 2008 to December 2013, the reservoir water fluctuated between 145 m and 175 m. The slope was again affected over a larger volume, and the fluctuation range of reservoir water level increased to 30 m. When the reservoir water level fell from 175 m to 145 m for the first time under a rainfall of 626.4 mm, the maximum monthly displacement was only 98.9 mm, which was much less than the monthly displacement of 334 mm in July 2007 when the reservoir level fell from 155 m to 145 m. The annual accumulated displacement maintained a step-wise behavior during the flood season. The steps were 74 mm, 84 mm, 63 mm and 63 mm in the subsequent four years. The reason for this phenomenon was that the landslide underwent a long-term stress adjustment and adapted again to the new schedule of reservoir water fluctuation.

3.2. Bazimen landslide

3.2.1. Geological conditions

The Bazimen landslide is also fan-shaped. The initial sliding direction of the landslide is oriented at N120°E. Its upper boundary is at the interface between bedrock and soil. The left and right boundaries are homologous gullies. The front part of the landslide extends to the bed of the Xiangxi River. The elevation of the landslide ranges from 110 m to 250 m, and covers an area of 13.5×10^4 m^2, with a maximum longitudinal dimension of 380 m and width of 100 m to 350 m. The estimated volume is 200×10^4 m^3 (Fig. 9).

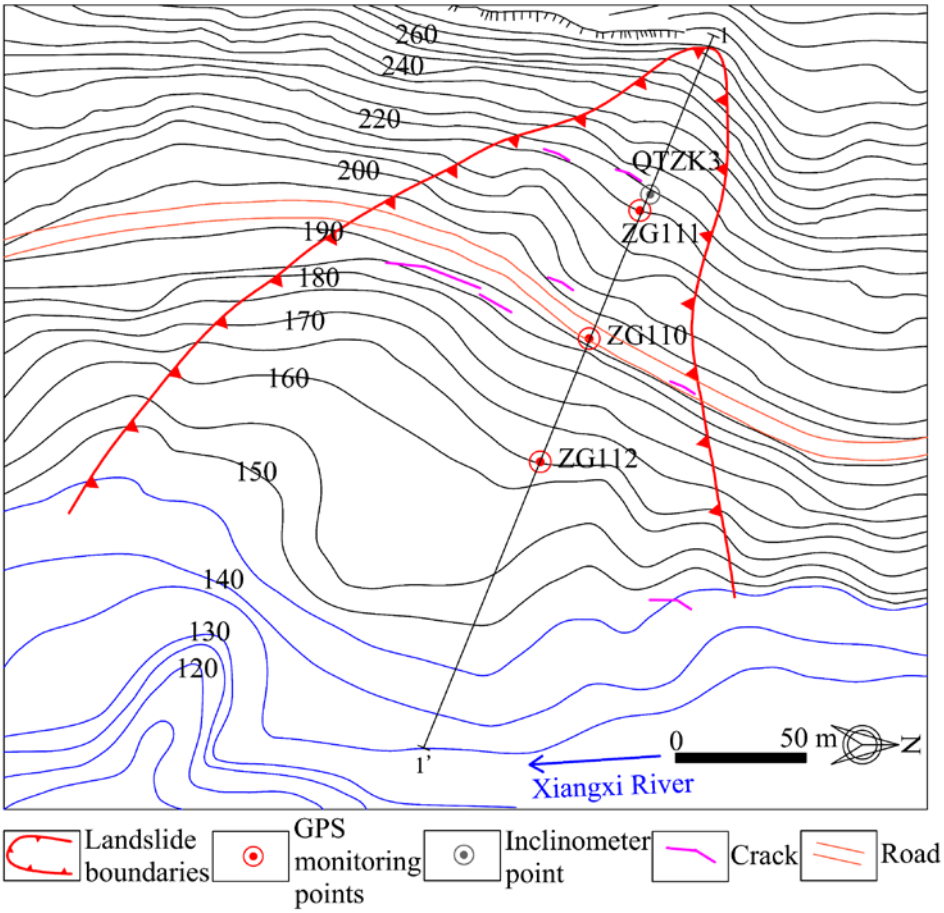

Fig. 9 Monitoring arrangement in the Bazimen landslide

The sliding masses of the Bazimen landslide consist of loose Quaternary deposits and fragmented rubbles. The bedrock of the landslide is formed by clastic sediments with siltstone and sandstone, with dip directions of 30°-40° (Fig. 10a). There are two main sliding surfaces, with varying depths in different parts of the landslide (Fig. 10b). The depth of the initial sliding surface is 27 m to 33 m and the depth of the upper secondary sliding surface is 8 m to 20 m.

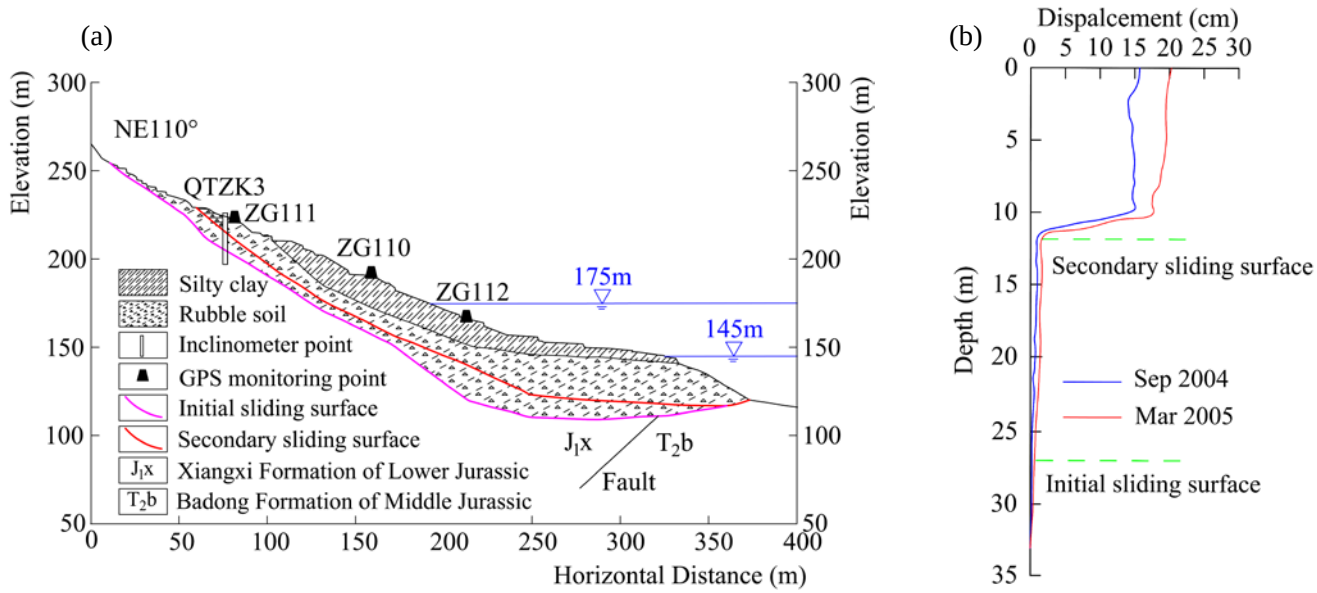

Fig. 10 a Schematic geological cross-section 1–1′ of the Bazimen Landslide b Lateral displacement versus depth from inclinometer QTZK3, Bazimen landslide

Since the impoundment of the TGR, the Bazimen landslide has reactivated and begun to deform more noticeably. In August 2003, two large cracks appeared on each side of the road within the sliding area (Fig. 9), with the length of 45 m and 30 m, respectively. During the May to August 2004 flood season and as the reservoir water level dropped from 139 m to 135 m, various cracks were observed within the trailing edge and middle part of the landslide. From June to July 2007, several cracks appeared on the nearby provincial road and buildings at the frontal part and upper part of the landslide.

3.2.2. Analysis of the monitoring data

Three GPS monitoring points were installed on the Bazimen landslide (Fig.9). The elevations at Stations ZG112, ZG110 and ZG111 are 165 m, 191 m and 215 m, respectively. The data on the accumulated displacement from August 2003 to December 2012 are shown on Fig. 11. The displacements in the upper part of the landslide (ZG111) were larger than that in the middle (ZG110) and lower (ZG112) areas. It can then be inferred that the Bazimen landslide is an "advancing" landslide, where the landslide was initiated from the upper part and developed downwards with increasing displacement (Du et al. 2013; Zhou et al. 2016). The monitoring station ZG111, with the largest deformation and the longest series of monitoring data, was selected for the analysis of the deformation characteristics of the Bazimen landslide.

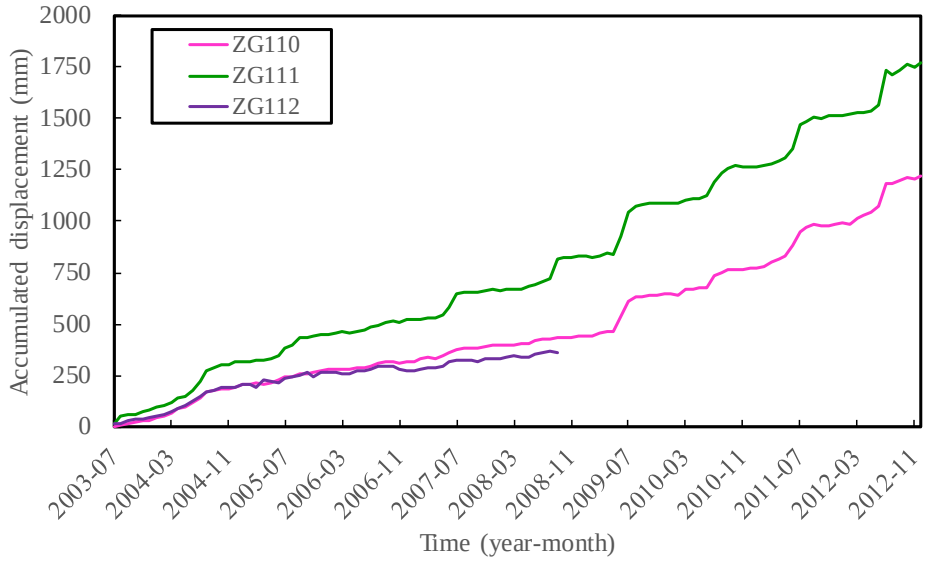

Fig.11 Accumulated displacement in the Bazimen landslide

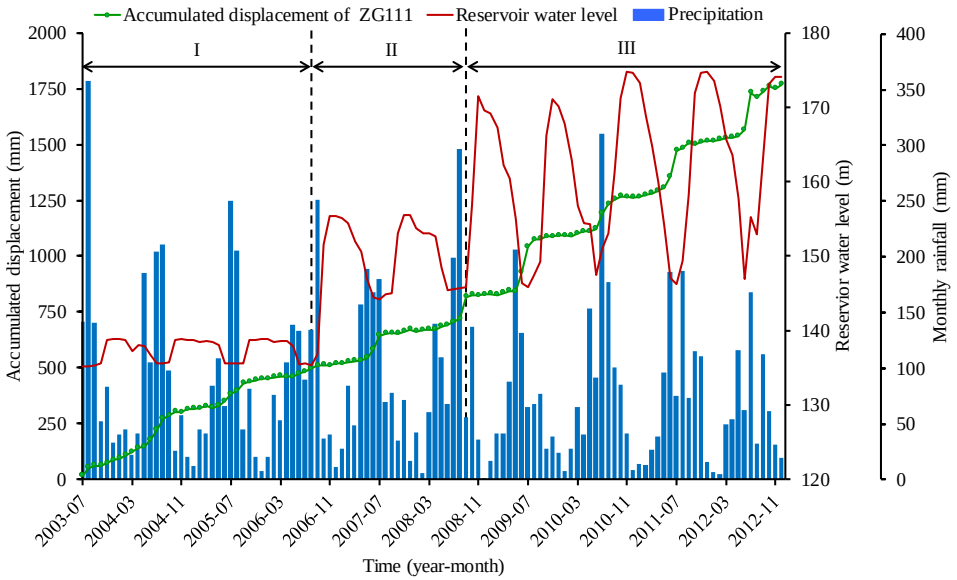

Fig.12 Rainfall, reservoir water level and accumulated displacement at Station ZG111 in the Bazimen landslide

Fig. 12 shows the measured curves of accumulated displacement, reservoir water level and rainfall at Location ZG111 from August 2003 to December 2012. The fluctuation of reservoir water level can be divided into three periods.

From August 2003 to August 2006, when the reservoir water fluctuated between 135 m and 139 m, the displacement showed trends similar to those of the Baishuihe landslide. The maximum monthly displacement occurred in the flood season with the values of 50.2 mm, 36.1 mm and 13.7 mm from 2004 to 2006. The displacement curve tended to be gentle in dry season.

Between September 2006 and September 2008, the reservoir water fluctuated between 145 m and 155 m. The landslide remained in a quasi-stable state between September 2006 and April 2007, and began to deform intensely from May 2007. When the reservoir water level fell from 155 m to 145 under a rainfall of 685.1 mm during the flood season, a maximum monthly displacement of 61.6 mm occurred, which was much smaller than for the Baishuihe landslide (334 mm). The Bazimen landslide then deformed more intensely and the maximum monthly displacement was 97 mm during the flood season of 2008 under the rainfall of 729.4 mm and a reservoir water level drawdown of 6.84 m.

From October 2008 to December 2012, the reservoir water level fluctuated between 145 m and 175 m. Between October 2008 and April 2009, the landslide was virtually stable. From May to September 2009, with the decrease of 11.1 m in reservoir water level and the precipitation of 547.4 mm, the maximum monthly displacement increased to 114.2 mm. The annual accumulated displacement curves kept a step-wise characteristic during the flood season. The maximum monthly displacements in the following years were 68 mm, 118 mm, and 166.8 mm, which were a trend different from the displacement of the Baishuihe landslide.

4. Calculation and results

4.1. Displacement decomposition

As mentioned, Location ZG118 on the Baishuihe landslide and Location ZG111 on the Bazimen landslide were selected to establish and validate the prediction model:

(1) For Location ZG118, the monitoring data from August 2003 to December 2012 was selected as training dataset, and the remaining data from January 2013 to December 2013 was used to test the model.

(2) For Location ZG111, the monitoring data from August 2003 to December 2011 was selected as training dataset,

while the data from January 2011 to December 2012 was used to test the model.

Considering the accumulated displacement curve (Figs. 8 and 12), the displacement during the flood season exhibits step-wise growth each year. Miao et al. (2018) showed that using a moving average method can eliminate the influence of the step in accumulated displacements and reflect the long-term trend. The water level of the TGR fluctuates between 145 m and 175 m every year as shown in Fig. 8 and Fig. 12. Considering the scheduling cycle of the TGR, the moving average cycle was selected as 12 months (Zhou et al. 2016; Miao et al. 2018).

The moving average method was used to extract the trend term of the displacement. The original displacement time series was expressed as $S(t)= \{s_1,s_2,\ldots,s_t,\ldots,s_n\}$. The trend term was calculated from follows:

$$\phi(t) = \frac{s_t + s_{t-1} + \cdots s_{t-n+1}}{k}, (t = k, k+1, k+2, \cdots, n) \tag{11}$$

where $\phi(t)$ is the trend term at time t and t is time; s_t is the accumulated displacement at time t; n is the number of measured accumulated displacement points; and k is the moving average cycle and was set to 12.

The periodic term was obtained after removing the trend term from the accumulated displacement. The values of the trend and periodic terms at Locations ZG118 and ZG111 are shown in Fig. 13.

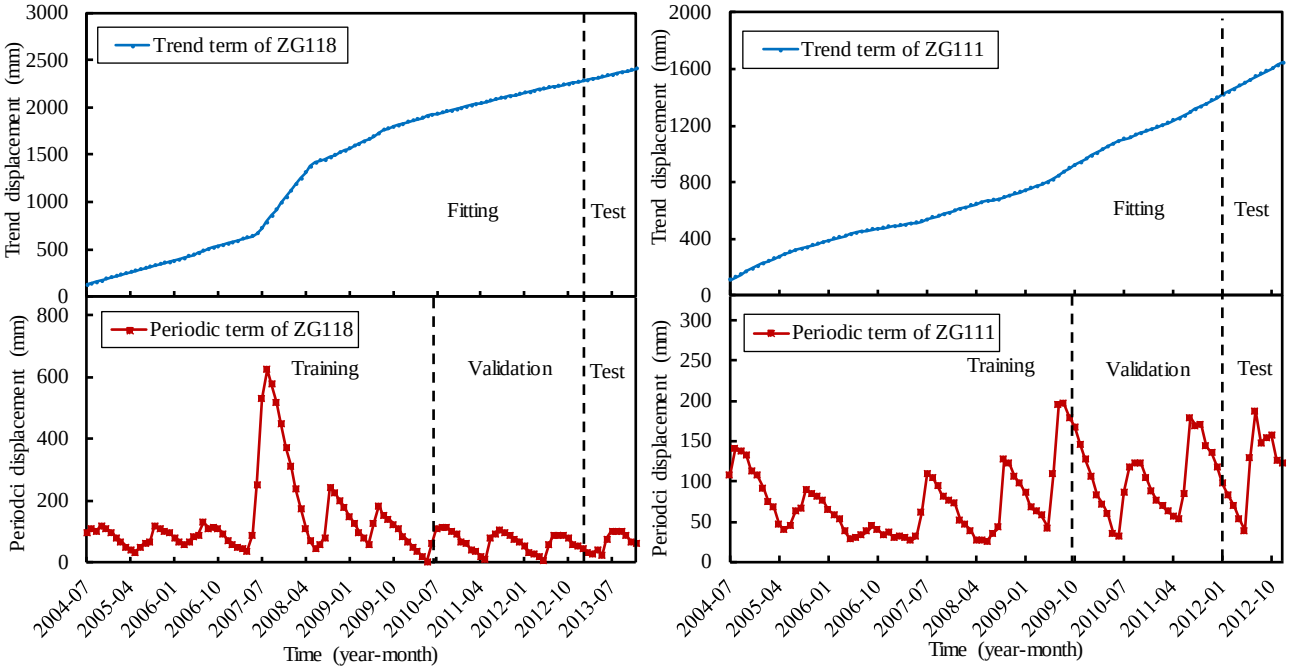

Fig. 13 Trend and periodic terms of ZG118 measurements in Baishuihe landslide and ZG111 measurements in Bazimen landslide

4.2. Trend displacement prediction

The trend curve of Location ZG118 was divided into three periods. For Location ZG111, two segments were considered (Fig. 13). The least squares method of polynomial fitting with a cubic polynomial was used (Eq.12). The results of the fitting calculation are shown in Table 1. The predicted trend terms were close to those measured, as illustrated in Fig. 14.

$$\phi(t) = at^3 + bt^2 + ct + d \tag{12}$$

where $\phi(t)$ is the trend displacement at the time t and t *is* time; and a, b, c and d are the coefficients, where a cannot be zero.

Table1 Regression coefficients of the trend displacement term based on polynomial fitting for the two landslides

Case	Period	*a*	*b*	*c*	*d*	Regression coefficient (R^2)

ZG118	July 2004 to May 2007	0.001	-0.0065	14.055	114.74	0.9992
ZG118	Jun 2007 to Aug 2009	0.0441	-3.91	98.947	529.5	0.9956
ZG118	Sep 2009 to Dec 2012	0.0009	-0.1303	16.382	1766.3	0.9998
ZG111	July 2004 to May 2005	-0.0378	0.4454	17.569	93.323	1
ZG111	June 2005 to Dec 2011	-0.0002	-0.1269	4.6793	335.22	0.9966

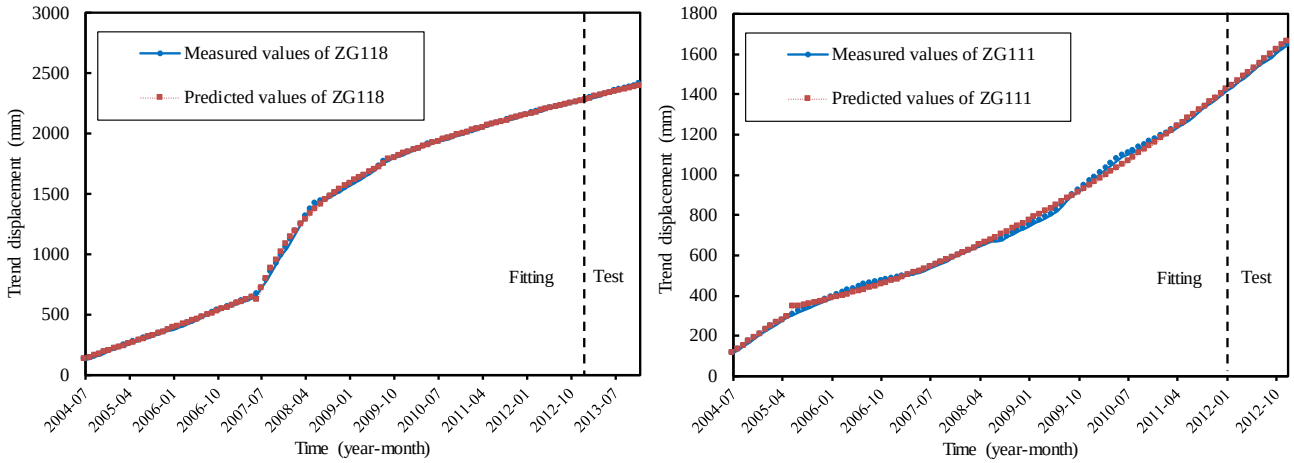

Fig. 14 Comparison of predicted and measured trend displacement term at Location ZG118 and Location ZG111

4.3. Periodic displacement prediction

4.3.1. Main controlling factors

The factors that influence the periodic displacement were used as input sequences and the periodic displacement was the output sequence in the model. As stated in Section 2.3.1, the periodic displacement is believed to be mainly affected by external triggering factors. This selection of influencing factors will affect directly the training of the model.

Rainfall is one of the major external forces that trigger landslides in the TGRA (Tomas et al. 2014; Yang et al. 2017). Combining the rainfall with the monitoring data, the deformations of the Baishuihe landslide and Bazimen landslide were affected significantly by heavier rainfall. Earlier research on the relationship between landslides and rainfall (Roering et al. 2015; Bogaard and Greco 2018) suggests that high precipitation over the last one-month or over the last two-months before failure can strongly increase landslide deformation. In this study, the periodic displacement follows quite well with the 1- and 2-month antecedent rainfalls at both Locations ZG118 and ZG111 (Fig.15).

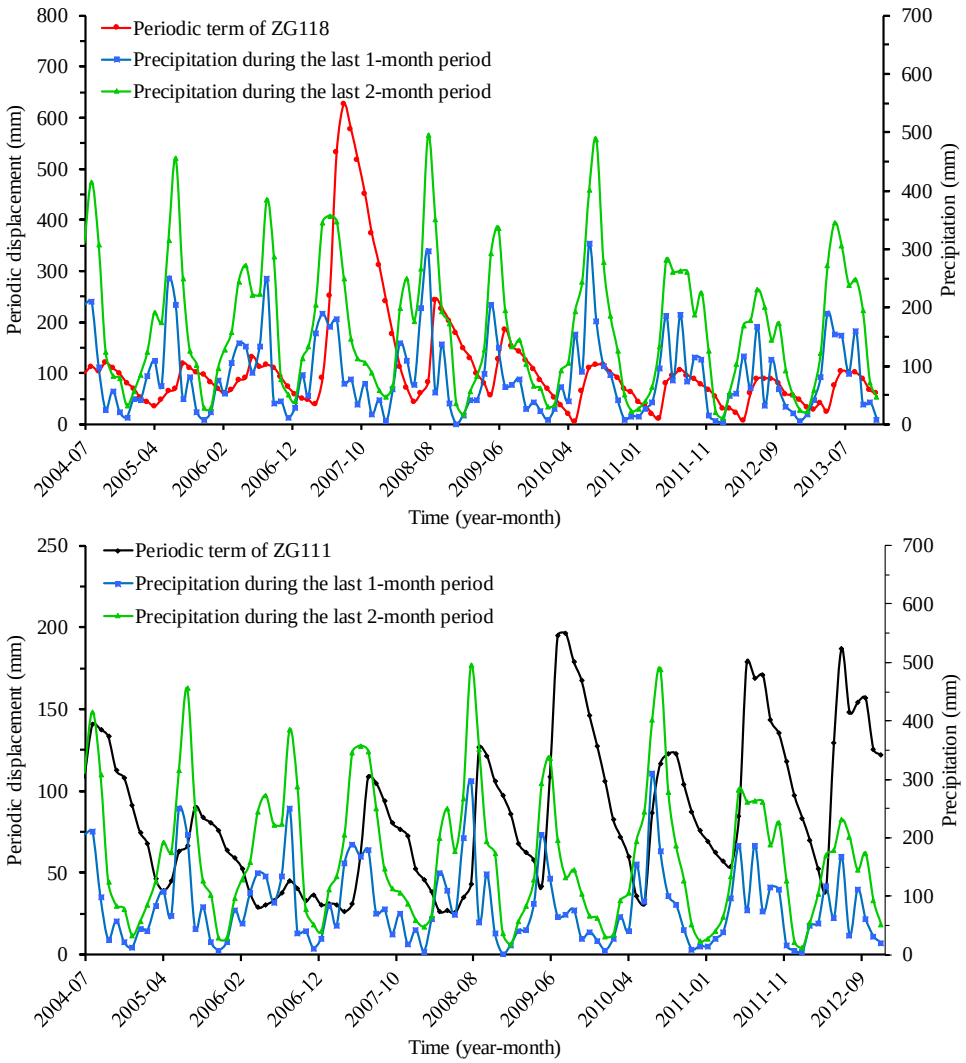

Fig. 15 Illustration of rainfall and periodic displacements at a Location ZG118 and Location ZG111

The analysis of the monitoring data of the Baishuihe landslide and Bazimen landslide suggests that the periodic variation of reservoir water level is another dominant exogenetic trigger inducing the step-wise deformation. The influence of the reservoir water level fluctuation on landslide deformation is strongest during the drawdown period. A faster drawdown rate of the reservoir water level leads to a larger deformation. Fig. 16 shows that there is a close relationship between the periodic displacement and the monthly change in reservoir water level. The reservoir level shows a “delay effect” on the deformation of the Baishuhe landslide while it is not noticeable on the Bazimen landslide (Miao et al. 2018; Zhou et al. 2016). The reservoir water level had a greater effect on the deformation of the Baishuhe landside, comparing with the Bazimen landslide. For this reason, the reservoir water level in the current month was adopted as an influence factor for the Baishuhe landside and not for the Bazimen landslide (Du et al., 2013).

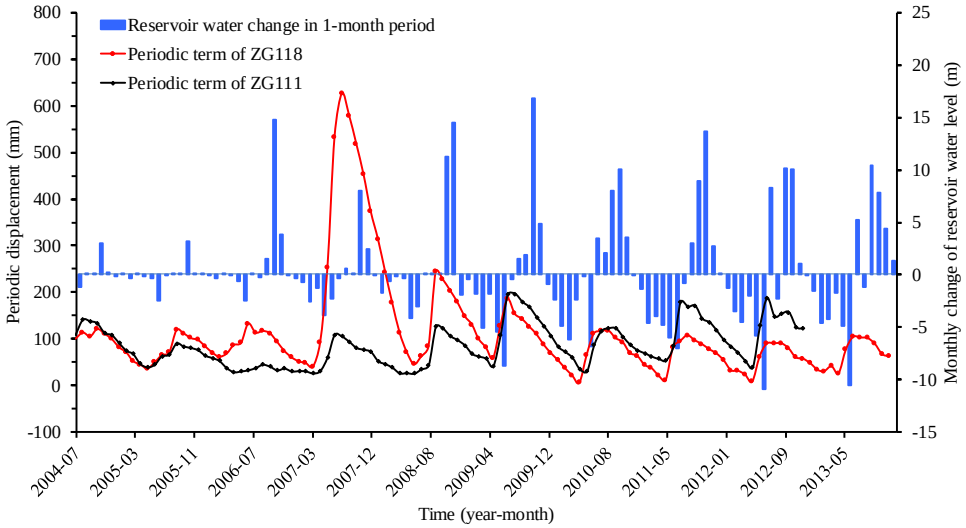

Fig. 16 Monthly change in reservoir level and periodic displacements at Locations ZG118 and ZG111

Slopes in different evolutionary states may respond differently to the same external trigger factors. When the slope is stable, even strong influence factors are usually not enough to cause large deformation. In contrast, if a slope is marginally stable or even unstable, the slightest increase in external "loads" may cause disequilibrium and large deformation (Glade et al., 2005). Therefore, the displacements over the preceding one, two and three months were adopted to represent the current state of the slope in this study (Cao et al. 2016; Zhou et al. 2016, 2018a). Table 2 summarizes the input items for the analysis of periodic displacement.

Based on the gray relational analysis, the grey relational grade (GRG) can be obtained, which is useful to evaluate the degree of correlation between the periodic displacement and influencing factors (Tan et al., 2018). The GRG ranges from 0 to 1, and a value of 0.6 or higher designates a strong correlation. The GRG between periodic displacement and the input parameters are shown in Table 2. The GRG-values are everywhere higher than 0.6, suggesting that the selected influencing factors have a strong correlation with the periodic displacement.

Table 2 Inputs for the modeling of the periodic displacement, and the grey relational grade (GRG) between influencing factors and periodic displacement

Input	Baishuihe landslide	GRG	Bazimen landslide	GRG
Precipitation	Input 1, the 1-month cumulative antecedent rainfall	0.78	Input 1, the 1-month cumulative antecedent rainfall	0.69
	Input 2, the 2-month cumulative antecedent rainfall	0.78	Input 2, the 2-month cumulative antecedent rainfall	0.67
Reservoir level	Input 3, reservoir level change in 1-month period	0.86	Input 3, reservoir level change in 1-month	0.78
	Input 4, reservoir level change in 2-month period	0.85		
	Input 5, the average elevation of reservoir level in the current month	0.84		
Evolution state	Input 6, the displacement over the past 1 month	0.86	Input 4, the displacement over the past 1 month	0.87
	Input 7, the displacement over the past 2 months	0.84	Input 5, the displacement over the past 2 months	0.86

	Input 8, the displacement over the past 3 months	0.81		Input 6, the displacement over the past 3 months	0.85

4.3.2. *Establishment of the LSTM model*

The training dataset was split into two parts: 70% of the data was used for training the LSTM model and 30% for validation (Fig. 13)[1]. The periodic term displacements and controlling factors were normalized to [-1, 1]. The LSTM model was implemented in Python by the Keras package and used TensorFlow as a backend. The one-step-ahead prediction was adopted in this study. The controlling factors were selected as input sequences of the LSTM model. As stated in 2.1.2, the weights of the input sequences keep changing in the process of training the model and are therefore not assigned fixed values.

The grid search method was used to search for the optimal parameters of the LSTM model. Its process can be described as follows: plausible values for each parameter are combined, and then all the combination values are used in the training process of the model. The performance metric of each combination is typically measured by cross-validation on the training set (Hsu et al., 2010). When all the parameter combinations have been tried, an optimal parameter combination with the best performance is returned automatically.

Both models at Locations ZG118 and ZG111 had three layers. The first two layers were LSTM layers and the last one was a dense layer. The length of the input sequence was of great importance and controlled the number of historical data point fed as an input (Xu and Niu 2018). The optimal length of the input sequence was also determined by the grid search method and was finally set to 12. Fig. 17 showed that the predicted values fitted well with the measured values in the establishment of the model and the model could then be used to predict the future behavior.

To verify the forecasting ability, the SVM model (Section 2.2) was also used to predict the periodic displacement for comparison with the LSTM model.

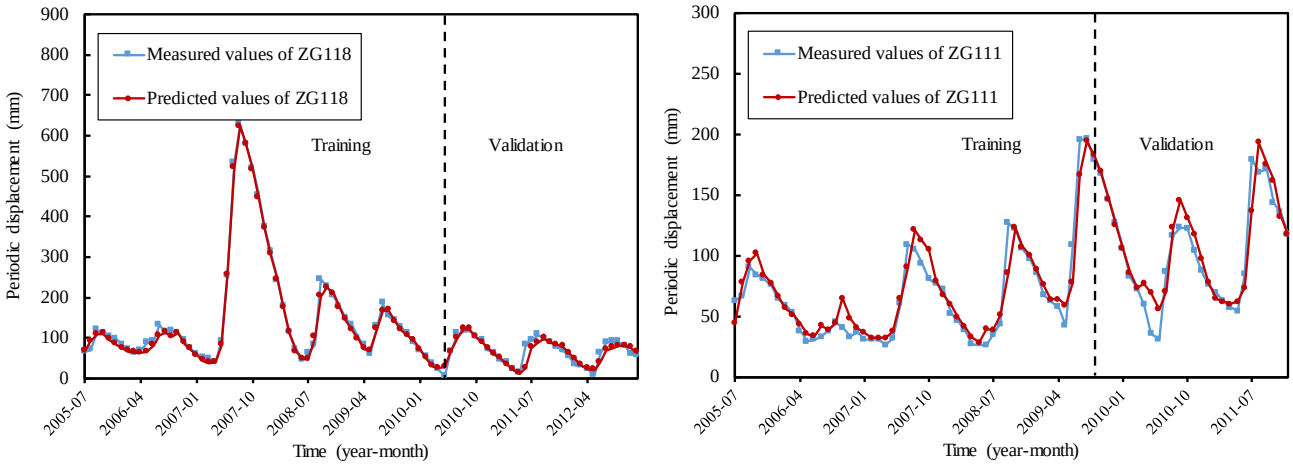

Fig. 17 The establishment of the LSTM model for Location ZG118 and Location ZG111

4.3.3. Predicted periodic displacement

Fig. 18, Table 3 and Table 4 present the comparisons between measured and predicted periodic displacement at Locations ZG118 and ZG111 using the LSTM and SVM models. For ZG118, the values of RMSE and MAPE with the LSTM model

[1] In supervised machine learning, the percentages of train set and validation set can usually be 50% and 50%. But when the dataset is not large, this division may be not reasonable. As the displacement is measured on a monthly cycle, only 125 and 113 measured values can be obtained from the GPS monitoring points of the two landslides. For similar size of database (73 data points), Xu and Niu (2018) divided the trainset into two part: 70% for training and 30% for validation, and got a well-trained LSTM model. This division was selected for this study.

were 7.53 mm and 10.52%, while the accuracy factors of the SVM model were 15.43 mm and 17.81%. For ZG111, the values of RMSE and MAPE were 13.83 mm and 8.60 %, and the accuracy of the SVM model was 28.32 mm and 16.52%.

The displacements predicted by LSTM model were in better agreement with the measured values than those with the SVM model, especially in the critical early warning period of the step-wise landslides (from May to July) as highlighted by the red rectangles in Fig. 18. The predicted accuracy of each model during the critical early warning period is summarized in Table 5. Taking the monitoring point ZG118 as an example, during May to July 2013, the reservoir water level decreased and the cumulative precipitation rose to 347 mm, which caused the displacement to increase sharply. The relationship between the influencing factors and the displacement was represented effectively by the LSTM model, and the values of RMSE and MAPE were 7.11 mm and 10.43%.

During May to July 2013, the values predicted by the SVM model were far from the measured values, with RMSE and MAPE-values of 21.83 mm and 22.02%. During this period, a Maximum Relative Error (MRE) of 48.49% obtained by the SVM occurred in June, while the MRE of the LSTM model was14.61%. Similarly, for the Location ZG111, the MRE from the SVM model and the LSTM model were 60.93 % and 22.17% respectively in June 2012.

However, the LSTM model showed some poorer predictions of local peak displacements (April 2013 of ZG118 and August 2012 of ZG111), as highlighted by the green rectangles in Fig. 18. The main reason for this is that the LSTM model is a dynamic model and the predicted result of the current time step is affected significantly by the previous information. When a local peak occurs, the model still learns from the former trend, and this leads to some large errors. Therefore, the curves of LSTM-predicted displacement are smoother than the measured values and cannot reflect local peaks well. It should be noted that the LSTM model shows reasonably good results after the local peak, because of the structure of the gates that can forget the useless information from the previous time step.

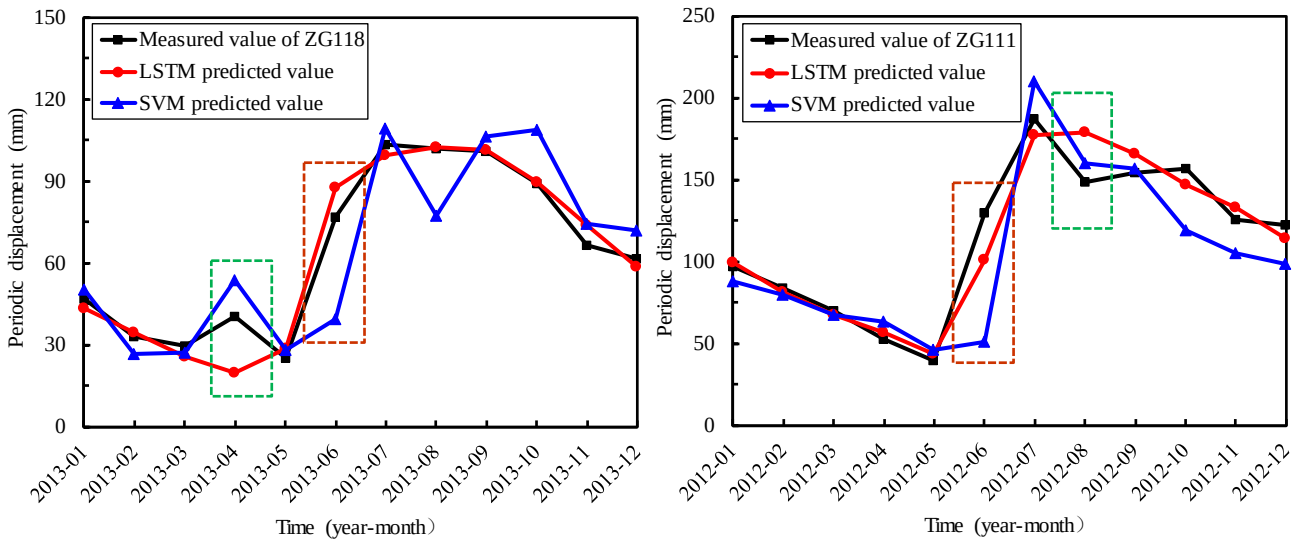

Fig. 18 Predicted and measured periodic displacements at Location ZG118 and Location ZG111 with the LSTM and SVM models

Table 3 Accuracy of the predicted periodic displacement with the LSTM and VSM models at Location ZG118

Time	Measured value (mm)	LSTM model			SVM model		
		Predicted value (mm)	Absolute error (mm)	Relative error (%)	Predicted value (mm)	Absolute error (mm)	Relative error (%)
2013-01	47.08	43.57	3.51	7.46	50.32	3.24	6.88
2013-02	33.04	34.76	1.72	5.21	26.48	6.56	19.85
2013-03	29.82	25.85	3.97	13.31	27.40	2.42	8.12
2013-04	40.69	19.86	20.83	51.19	53.89	13.20	32.44
2013-05	25.35	28.62	3.27	12.90	28.29	2.94	11.60

Time	Measured value (mm)	LSTM model			SVM model		
		Predicted value (mm)	Absolute error (mm)	Relative error (%)	Predicted value (mm)	Absolute error (mm)	Relative error (%)
2013-06	76.68	87.88	11.20	14.61	39.50	37.18	48.49
2013-07	103.44	99.51	3.93	3.80	109.62	6.18	5.97
2013-08	102.13	102.60	0.47	0.46	77.56	24.57	24.06
2013-09	100.83	101.48	0.66	0.65	106.25	5.43	5.39
2013-10	88.99	89.78	0.79	0.89	108.81	19.82	22.27
2013-11	66.63	73.98	7.35	11.03	74.67	8.04	12.07
2013-12	61.82	58.90	2.92	4.72	72.04	10.22	16.53
Min			0.47	0.46		2.42	5.39
Max			20.83	51.19		37.18	48.49
Mean			5.05	10.52		11.65	17.81
RMSE		7.53			15.43		

Table 4 Accuracy of the predicted periodic displacement with the LSTM and VSM models at Location ZG111

Time	Measured value (mm)	LSTM model			SVM model		
		Predicted value (mm)	Absolute error (mm)	Relative error (%)	Predicted value (mm)	Absolute error (mm)	Relative error (%)
2012-01	97.16	99.18	2.02	2.08	88.10	9.06	9.33
2012-02	83.42	81.31	2.11	2.53	79.90	3.52	4.22
2012-03	69.87	66.95	2.92	4.18	67.07	2.80	4.00
2012-04	52.60	56.35	3.75	7.13	63.13	10.53	20.02
2012-05	39.14	43.94	4.80	12.26	45.96	6.82	17.41
2012-06	129.38	100.69	28.69	22.17	50.55	78.83	60.93
2012-07	187.05	177.65	9.40	5.03	210.33	23.28	12.44
2012-08	148.18	179.29	31.11	20.99	160.11	11.93	8.05
2012-09	154.33	166.03	11.70	7.58	156.50	2.17	1.41
2012-10	157.11	147.15	9.96	6.34	119.11	38.00	24.19
2012-11	125.53	133.27	7.74	6.17	104.74	20.79	16.56
2012-12	122.34	114.15	8.19	6.69	98.19	24.15	19.74
Min			2.02	2.08		2.17	1.41
Max			31.11	22.17		78.83	60.93
Mean			10.20	8.60		19.32	16.52
RMSE		13.83			28.32		

Table 5 Prediction accuracy of each model during the critical early warning period (from May to July)

Model	RMSE		MAPE	
	ZG118	ZG111	ZG118	ZG111
LSTM	7.11	17.65	10.43	13.15
SVM	21.83	30.27	22.02	47.62

4.3.4. Accumulated displacement prediction

Following the decomposition principle of displacement time series (Eq. 8), the accumulated displacement was determined

by making the sum of the predicted trend and periodic displacements. Fig. 19 shows that the predicted accumulated displacements at the two monitoring points compare very well with the measured displacements. Between January and December 2013, at Location ZG118 on the Baishuihe landslide, the RMSE- and MAPE-values were 10.89 mm and 0.36 %. Between January and December 2012, at Location ZG111 on the Bazimen landslide, the RMSE-value was 16.93 mm and the MAPE-value was 0.73 %. Furthermore, Fig. 19 shows that the MRE of the accumulated displacement and periodic displacement occurred at the same time steps (April 2013 for the Baishuihe landslide and June 2012 for the Bazimen landslide), which proves that the accurate prediction of the periodic displacement is the key to the prediction of the accumulated displacement for step-wise landslides.

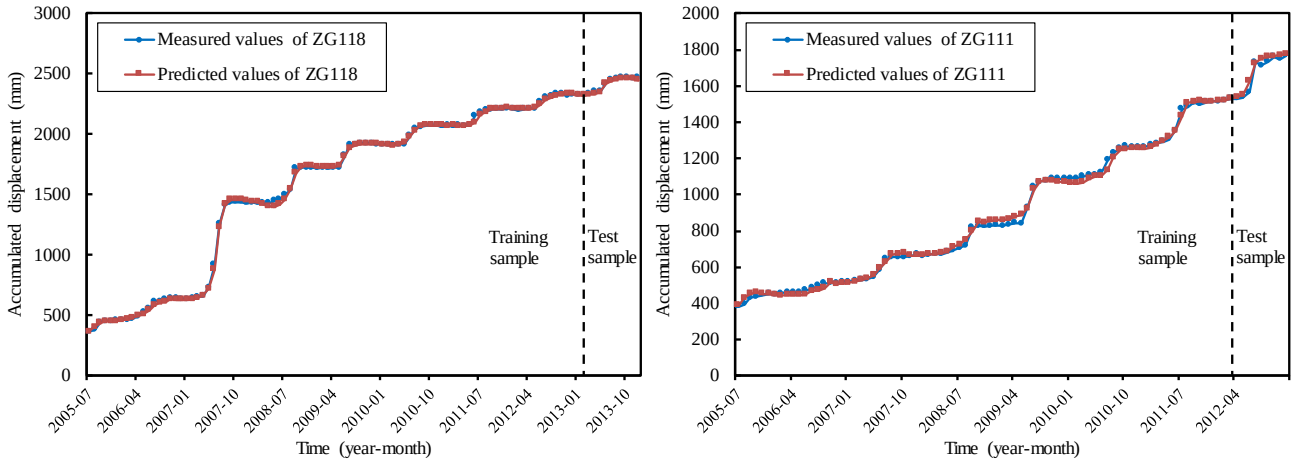

Fig. 19 Observed and predicted accumulated displacement at Location ZG118 and Location ZG111

5. Discussion

The LSTM model showed a more satisfactory performance than the SVM model, using the same training dataset, especially during the critical early warning period of the step-wise landslides (from May to July). In the architecture of LSTM models, nodes of hidden layers relate to others at different time steps. The connections to the previous time step in the LSTM model are set as inputs, so the hidden state contains a dynamic history of the input features sequence. The model can learn rules from a sequence that contains historical information and thus can make full use of the historical information. The SVM model, on the other hand, with no connections between different time steps, can only learn rules from a point in time, and the response between outputs and inputs cannot be learned completely (Xu and Niu 2018).

For the two Three Gorges landslides analyzed, rainfall and the reservoir water level fluctuation were two major triggering factors for deformation. The relative influence of the rainfall and the reservoir water level fluctuation on landslide deformation may not, however, remain the same during the entire monitoring period. For the Baishuihe landslide, the analysis of the monitoring data showed that rainfall was the dominant influence factor in the early stage of the monitoring period, while both rainfall and reservoir water fluctuation contributed significantly to the displacement at the end stage. The change in the displacement response as a function of the external influence factors may lead to a change in learned rules during the process of training the model. When one point is predicted based on the previous learned rules between inputs and outputs, the LSTM model can assess the learned rules, forgetting invalid ones and/or remembering useful ones. Thus, the prediction accuracy of the LSTM model would not be affected by this change in the influence factors with time. For static models like the SVM model, all the rules learned from the earlier information will be used to predict the next stage throughout the whole training process. As a result, the goodness of the prediction will be affected. To establish a more accurate causal SVM relationship, the earlier information should be removed and the latest monitoring data should gradually be added to the static models (Du et al. 2013; Miao et al. 2018). However, in this case, it is not easy to judge how much earlier information should be removed, because of the complex relationships between landslide deformation and the

landslide triggering factors. Such drawbacks represent a significant obstacle for the future development of static models.

Other models, such as an ELM model, also showed an improved performance than the SVM model for landslide displacement prediction (Cao et al. 2016; Huang et al. 2016; 2017). To further evaluate the proposed forecast model, the comparison of the prediction accuracy with the LSTM model and other data-driven methods can be done in the future. On the other hand, Zhou et al. (2016) showed that the SVM model behaved better than a BPNN model to predict the displacement of the Bazimen landslide in the TGRA. Therefore, an LSTM model will behave better than a BPNN model.

In each landslide studied, only one monitoring point was modelled. The point with the most data versus time and large displacement was selected for analysis. However, a single point can only represent local displacements, which is insufficient to characterize the entire landslide. An analysis of the deformation characteristics at several monitoring points in different parts of the landslide are necessary to do a prediction for the entire landslide. Physical models are also another useful method to help model and predict landslide movements.

Although the LSTM model achieved a good performance in landslide displacement prediction, it has some drawbacks. As a deep ML method, a LSTM model needs to tune more parameters than classical ML methods (Reimers and Gurevych 2017). It is always difficult to obtain the optimal of all parameters at the same time. Another drawback is that the LSTM model can be quite demanding with respect to the sizes of the dataset. The neural networks cannot be fully trained if training samples are insufficient, and thus the prediction accuracy of the model would be affected. In the case of the two landslides modelled in this paper, the displacement was measured by GPS with a monthly frequency, it may take years to obtain enough displacement series. Some researchers have used interpolation method to get more training data (e.g. Yao et al. 2015). Interpolated points may change the trend of the displacement series at some time steps, thus leading to large errors (Xu and Niu 2018). One should, in the future, consider means to develop larger datasets in the case of limited monitoring periods.

6. Conclusions

Most classical ML methods, such as the BPNN and the SVM, do landslide displacement prediction as a static regression problem. Since landslide evolution is a complex nonlinear dynamic (varying in time) process, dynamic modeling approaches are more suitable to construct predictors. In this study, a dynamic model for landslide displacement prediction was established, using time series analysis and the deep ML "long short-term memory neural network" (LSTM) model. The static SVM model was also used to predict landslide displacement for comparison. The superiority of the LSTM model over the SVM model was verified with the comparison of predicted and measured displacements for two typical step-wise landslides in the TGRA.

The LSTM model can establish connections between landslide conditions at different time, and learn rules from previous deformation time steps. Thus, a LSTM model can make full use of the historical information. With its structure of memory block, the LSTM model can judge whether the learned rules from the previous time step are useful or not, and then determine the learned rules should be passed along to the next time step or be abandoned. The prediction accuracy is thus not affected by the errors in some previous points. These characteristics of the LSTM model contribute to improving the predictions.

Overall, the proposed dynamic modeling approach, based on time series analysis and LSTM, can achieve accurate prediction in case of the slow and step-wise deformation periods. This dynamic method has the potential for broad application to predict landslide displacement in the TGRA and other landslide-prone regions.

Acknowledgment

This research was supported by the National Natural Sciences Foundation of China (No. 41572292) and the Research Council of Norway (Klima 2050). The authors wish to thank Dr. Du Juan, Dr. Huang Faming, Dr. Miao Fasheng and Dr. Zhou Chao for their assistance in collecting the data. The first author wishes to thank the China Scholarship Council (CSC)

and the Norwegian Geotechnical Institute (NGI) for funding her research period at NGI.

Reference

Bengio Y, Simard P, Frasconi P (1994) Learning long-term dependencies with gradient descent is difficult. Neural Netw IEEE Trans 5(2):157-166

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

Cao Y, Yin K, Alexander 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

Chen S, Chou W (2012) Short-term traffic flow prediction using EMD-based recurrent hermite neural network approach. The 15th International IEEE Conference on Intelligent, pp. 1821-1826

China Institute of Geo-Environment Monitoring (2017) Bulletin of geologic hazards from January to December in 2016. China Institute of Geo-Environment Monitoring, Beijing

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(2):83-96

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20: 273-297

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

Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. Proceedings of the 6th International Symposium on Micro Machine and Human Science. Nagoya, Japan, IEEE, pp 39-43

Fan X, Xu Q, Scaringi G, Dai L, Li W, Dong X, Zhu X, Pei X, Dai K, Havebith HB (2017) Failure mechanism and kinematics of the deadly June 24th 2017 Xinmo landslide, Maoxian, Sichuan, China. Landslides 14(6):2129-2146

Fan X, Xu Q, Scaringi G (2018) Brief communication: Post-seismic landslides, the tough lesson of a catastrophe. Nat Hazards Earth Syst Sci 18:397-403

Fan Y, Qian Y, Xie F, Soong FK (2014) TTS Synthesis with Bidirectional LSTM based Recurrent Neural Networks. Proceedings of the 15th Annual Conference of the International Speech Communication Association (INTERSPEECH), pp. 1964-1968

Felix A, Jürgen S (2000) Recurrent nets that time and count. Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, pp. 3:189-194

Froude M, Petley D (2018) Global fatal landslide occurrence from 2004 to 2016. Nat Hazards Earth Syst Sci 18(8):2161-2181

Glade T, Anderson M, Crozier MJ (2005) Landslide hazard and risk: issues, concepts and approach. In: Landslide hazard and risk. Wiley, pp 1-40

Graves A, Mohamed A, Hinton G (2013) Speech recognition with deep recurrent neural networks. Proceedings of the International Conference on Acoustics, Speech and Signal Processing Acoustics, pp. 6645-6649

Han M, Xi J, Xu S, Yin F (2004) Prediction of chaotic time series based on the recurrent predictor neural network. IEEE Trans Signal Process 52(12):3409-3416

Haque U, Blum P, da Silva PF, et al (2016) Fatal landslides in Europe. Landslides 13(6):1545-1554

Hasim S, Andrew S, Francoise B (2014) Long short-term memory recurrent neural network architectures for large scale acoustic modeling. Proceedings of the 15th Annual Conference of the International Speech Communication Association (INTERSPEECH), pp. 338-342

Helmstetter A, Sornette D, Grasso J, et al (2004) Slider block friction model for landslides: application to Vaiont and La Clapiere landslides. J Geophys Res 109(B2):1-15

Hochreite S, Schmidhuber J (1997) Long short -term memory. Neural Comput 9(8): 1735-1780

Hong H, Pourghasemi H, Pourtaghi Z (2016) Landslide susceptibility assessment in Lianhua County (China): a comparison between a random forest data mining technique and bivariate and multivariate statistical models. Geomorphology 259:105-118

Huang F, Yin K, Zhang G, Gui L, Yang B, Liu L (2016) Landslide displacement prediction using discrete wavelet transform and extreme learning machine based on chaos theory. Environ Earth Sci 75:1376-1394

Huang F, Huang J, Jiang S, Zhou C (2017) Landslide displacement prediction based on multivariate chaotic model and extreme learning machine. Eng Geol 218: 173-186

Hsu C, Chang C, Lin C (2010). A practical guide to support vector classification. Technical Report, National Taiwan University

Hyndman RJ, Koehler AB (2006). Another look at measures of forecast accuracy. Int J Forecast 22(4): 679-688

Kayacan E, Ulutas B, Kaynak O (2010) Grey system theory-based models in time series prediction. Expert Syst Appl 37: 1784-1789

Lian C, Zeng Z, Yao W, Tang H (2015) Multiple neural networks switched prediction for landslide displacement. Eng Geol 186: 91-99

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

Liu B, Shao J, Xu W, Chen H, Shi C (2014) Comparison on landslide nonlinear displacement analysis and prediction with computational intelligence approaches. Landslides 11(5):889-896

Liu Y, Qin Z, Hu B et al (2018) State fusion entropy for continuous and site-specific analysis of landslide stability changing regularities. Nat Hazards Earth Syst Sci 18:1187-1199

Ma J, Tang H, Liu X, Hu X, Sun M, Song Y (2017) Establishment of a deformation forecasting model for a step-like landslide based on decision tree C5.0 and two-step cluster algorithms: a case study in the Three Gorges Reservoir area, China. Landslides 14(3):1275-1281

Ma X, Tao Z, Wang Y, Yu H, Wang Y (2015) Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transp Res Part C Emerg Technol 54: 187-197

Miao H, Wang G, Yin K, Kamai T, Li Y (2014) Mechanism of the slow-moving landslides in Jurassic red-strata in the Three Gorges Reservoir, China. Eng Geol 171:59–69

Miao F, Wu Y, Xie Y, Li Y (2018) Prediction of landslide displacement with step-like behavior based on multialgorithm optimization and a support vector regression model. Landslides 15(3): 475-488

Mufundirwa A, Fujii Y, Kodama J (2010) A new practical method for prediction of geomechanical failure time. Int J Rock Mech Min Sci 47(7):1079-1090

Qin S, Jiao J, Wang S (2002) A nonlinear dynamical model of landslide evolution. Geomorphology 43(1-2): 77-85

Reichenbach P, Rossi M, Malamud B, Mihir M, Guzzetti F (2018). A review of statistically-based landslide susceptibility models. Earth-Sci Rev 180:60-91

Reimers N, Gurevych I (2017). Optimal hyperparameters for deep LSTM-networks for sequence labeling tasks. arXiv preprint arXiv:1707.06799

Roering J, Mackey B, Handwerger A et al (2015) Beyond the angle of repose: a review and synthesis of landslide processes in response to rapid uplift, Eel River, Northern California. Geomorphology 236:109-131

Saito M (1965) Forecasting the time of occurrence of a slope failure. In: Proceedings of the 6th International Mechanics and Foundation Engineering pp: 537-541

Scaringi G, Fan X, Xu Q et al (2018) Some considerations on the use of numerical methods to simulate past landslides and possible new failures: the case of the recent Xinmo landslide (Sichuan, China). Landslides 15:1359-1375

Smola A, Scholkopf B (2004) A tutorial on support vector regression. Stat Comput 14(3):199-222

Tan F, Hu X, He C, Zhang Y, Zhang H, Zhou C, Wang Q (2018) Identifying the main control factors for different deformation stages of landslide. Geotech Geo Eng 36(1): 469-482

Thiebers (2014) Integration of a limit-equilibrium model into a landslide early warning system. Landslides 11(5):859-875

Tomas R, Li Z, Liu P, Singleton A, Hoey T, Cheng X (2014) Spatiotemporal characteristics of the Huangtupo landslide in the Three Gorges region (China) constrained by radar interferometry. Geophys J Int 197(1):213-232

Vincent P, Larochelle H, Lajoie I et al (2010) Stacked denoising autoencoders: learning useful representation in a deep network with a local denoising criterion. J Mach Learn Res 11(12):3371-3408

Xu S, Niu R (2018) Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorge area, China. Comput Geosci 111: 87-96

Yao W, Zeng Z, Lian C, Tang H (2015) Training enhanced reservoir computing predictor for landslide displacement. Eng Geol 188: 101-109

Yang B, Yin K, Xiao T, Chen L, Du J (2017) Annual variation of landslide stability under the effect of water level fluctuation and rainfall in the Three Gorges Reservoir, China. Environ Earth Sci 76: 564-580

Yin Y, Huang B, Wang W, Wei Y, Ma X, Ma F, Zhao C (2016) Reservoir-induced landslides and risk control in three Gorges project on Yangtze river, China. J Rock Mech Geotech Eng 8 (5):577-595

Zhang H, Huang X (2018) Trend and spatiotemporal distribution of fatal landslides triggered by non-seismic effects in China. Landslides 15:1663-1674

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

Zhou C, Yin K, Cao Y, Intrieri E, Ahmed B, Catani F (2018a) Displacement prediction of step-like landslide by applying a novel kernel extreme learning machine method. Landslides DOI 10.1007/s10346-018-1022-0

Zhou C, Yin K, Cao Y, Ahmed B, Li Y, Catani F, Pourghasemie HR (2018b) Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. Comput Geosci 112:23-27