

Landslides (2018) 15:2047–2059
 DOI 10.1007/s10346-018-1020-2
 Received: 5 November 2017
 Accepted: 21 May 2018
 Published online: 29 May 2018
 © Springer-Verlag GmbH Germany
 part of Springer Nature 2018

Huajin Li · Qiang Xu · Yusen He · Jiahao Deng

Prediction of landslide displacement with an ensemble-based extreme learning machine and copula models

Abstract Research on the dynamics of landslide displacement forms the basis for landslide hazard prevention. This paper proposes a novel data-driven approach to monitor and predict the landslide displacement. In the first part, autoregressive moving average time series models are constructed to analyze the autocorrelation of landslide triggering factors. A linear ensemble-based extreme learning machine using the least absolute shrinkage and selection operator is applied in predicting the displacement of landslides. Five benchmarking data-driven models, the support vector machine, neural network, random forest, k -nearest neighbor, and the classical extreme learning machine, are considered as baseline models for validating the ensemble-based extreme learning machines. Numerical experiments demonstrated that the proposed prediction model produces the smallest prediction errors among all the algorithms tested. In the second part, parametric copula models are fitted on the predicted displacement, to investigate the relationship between the triggering factors and landslide displacement values. The Gumbel-Hougaard copula model performs best, which indicates strong upper tail correlation between the triggering factors and displacement values. Thresholds for the triggering factors can be obtained by monitoring the landslide moving patterns with large displacement values. The effectiveness and utility of the proposed data-driven approach have been confirmed with the landslide case study in the region of the Three Gorges Reservoir.

Keywords Landslide displacement prediction · Extreme learning machine · LASSO · Copula theory · Value-at-Risk

Introduction

Landslides are a recurrent geological phenomenon which pose serious threats to the local community. In the Three Gorges Reservoir region in central China, more than 4200 landslides have occurred, which are attributed to the complex geological environment and to heavy rainfall (Yin et al. 2010). This region has the highest frequency of landslide occurrence in China (Zhou et al. 2016) and the majority of landslides in this area shows multiple reactivations with large displacement (Cao et al. 2016). Hence, predictive modeling of landslide displacement is essential for prevention of landslide hazards.

Methods for predicting the cumulative displacements of landslides have been studied by researchers in previous years. The methods can be briefly categorized into two groups: physical models and data-driven models (Huang et al. 2017; Li et al. 2012). Due to the complexity of geological parameters for reservoir landslides, physical models are less used in practice (Huang et al. 2017).

Recently, data-driven models which contain the triggering factors as input and the moving response of the landslide as

output of the prediction model are widely applied in displacement prediction. The input part refers to the earthquake, precipitation, and reservoir water level fluctuations while the output response denotes the landslide displacement. Xu et al. (2011) introduced the autoregressive time series model to predict landslide displacement. Krkač et al. (2017) presented a methodology for prediction of landslide movements based on random forest (RF). Du et al. (2013) also applied the time series model to decompose the cumulative displacement into a trend component and a seasonal component. Lian et al. (2015) developed artificial neural network (ANN) to model the displacement with highly accurate results. Zhou et al. (2016) selected particle swarm algorithm (PSO) to optimize the parameters of the support vector machine (SVM) and achieved promising results as well. Among data-driven models, machine learning algorithms provide more accurate results than classical statistical models.

However, the classical machine learning algorithms such as ANN and SVM have two major constraints: (1) The level of appropriateness of parameter initialization (e.g., number of hidden layers, C values) influences the performance of the algorithm, (2) overfitting is a common problem when less training data are available. At the same time, the study regarding the relationship between the triggering factors and the landslide displacement of landslides is still insufficient. Under extreme severe conditions such as heavy precipitation, a single factor may trigger landslide movement easily. Therefore, it is necessary to develop a predicting data-driven framework to solve this problem.

Extreme learning machine (ELM) is one of the most widely applied algorithms for predicting time series data. As a single-hidden layer feedforward neural network, ELM overcomes the challenge of appropriate parameter initialization. Its advantages include a simple theoretical basis, global minimum optimization, and powerful generalization, strengthening the capacity in prediction modeling (Huang 2003; Huang et al. 2006a, b). ELM has been successfully applied in many other fields with promising prediction results (Cao et al. 2016; Lian et al. 2013; Miche et al. 2015; Ouammi et al. 2010; Zong and Huang 2011).

The use of the single extreme learning machine may produce more prediction errors when a small dataset is used for training. An ensemble-based method is a commonly used solution to mitigate the problem (He and Kusiak 2018; Tramèr et al., 2017). A novel linear ensemble-based method named least absolute shrinkage and selection operator (LASSO) has been proposed to resolve this problem (Tibshirani 1996). This ensemble-based method minimizes the sum of squared errors by placing a boundary on the sum of the absolute values of the coefficients. Through assigning different weights to different ELMs, based on their prediction performances, LASSO is capable to reduce prediction errors. It also exhibits the stability of ridge regression

(Sun et al. 2017). The assembly of different ELMs based on LASSO may further improve the prediction accuracy in comparison with any single ELM.

Beside the displacement prediction, advanced monitoring, and identification of landslides with seasonal reactivations are also highly demanded. The predicted slope failures of such seasonal landslide reactivations can prevent severe damages to local communities. Indicated by previous research, tail correlation exists between the triggering factors and displacement values (Cao et al. 2016). Therefore, in this research, parametric copula models are constructed to model the relationship between the triggering factors and landslide displacement values. Thresholds for the triggering factors can be computed to predict landslide reactivations.

In this paper, a novel data-driven framework with an integrated extreme learning machine containing copula models is proposed to monitor and predict the uncertainties of landslide displacement. Considering the computational expensiveness and less accurate results of the classical algorithms, the ELM is selected in this study. To overcome the deficiencies of the single ELM, a LASSO algorithm is utilized to assemble the ELMs. For monitoring the landslide reactivations with large displacement, tail correlations between displacement values and triggering factors including precipitation and reservoir water level fluctuations are investigated through parametric copula models.

Methodology

In this research, a novel data-driven predictive modeling and monitoring framework is proposed; two datasets are selected for this research: time series of monthly precipitation and monthly average reservoir water level (Fig. 1). First, the autoregressive moving average (ARMA) time series model is selected to analyze the trend and seasonality of the average monthly reservoir water level and monthly precipitation. Autocorrelation factors (ACFs) of the ARMA model are computed to illustrate their trend and seasonality. Second, Pearson's correlation coefficients have been used to map the relationship between the instant displacements and historic precipitations or reservoir water level fluctuations. All significant positive values (with p values < 0.05) of historic precipitation or reservoir water level are considered as indicators of landslide displacement. Hence, they are selected as inputs in the prediction model based on LASSO-ELM. Next, four parametric copula models are constructed to fit the predicted displacement data. Parameters for each model are derived by the maximum likelihood estimation. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) are calculated to evaluate the performance of the four parametric copula models. Copula models with the smallest values of AIC and BIC perform best. Last, thresholds indicating large displacements can be extracted from the best copula models. Value-at-Risks (VaRs) of both triggering factors (e.g., precipitation and reservoir water level fluctuations) are obtained as thresholds to monitor the landslides for landslide hazard prevention.

Time series analysis

Autoregressive moving average models are widely utilized in analyzing time series data. The ARMA models are a generic class of time series models capable of predicting the current values of a

variable from its past values and the past error terms. The parameters of an ARMA model produce general statistical inferences regarding the temporal dynamics and long-term memory of a variable (McLeod and Li 1983). A classical ARMA time series model is expressed in (1).

$$x_k = \sum_{i=1}^p \phi_i x_{k-i} + \alpha_k + \sum_{j=1}^q \theta_j \alpha_{k-j} \quad (1)$$

where x_k denotes the values of the current value of the target variable, α_k is the white noise, ϕ denotes the autoregressive parameter, θ denotes the moving average parameter, and p and q are order parameters. In this research, x_k represents the instant value of the triggering factors.

Based on the constructed ARMA model, the correlation between the current values of the variable and its past values may be computed. The autocorrelation function is widely employed to display the correlation. The ACF based on an ARMA model is expressed in (2) (Bustos and Yohai 1986).

$$r_l = \frac{\frac{1}{N-1} \sum_{i=1}^{N-l} (x_i - \bar{x})(x_{i+l} - \bar{x})}{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (2)$$

where r_l is the ACF of the l th lag of the variable, x_i is the i th data sample, \bar{x} is the mean of all data samples, and N is the total number of samples in the dataset.

Extreme learning machine and LASSO regularization

ELM in Fig. 2 is a single-hidden layer feedforward neural network (SLFN). The basic ELM algorithm consists of three layers: the input layer, the hidden layer, and the output layer. The input layer weight matrix and the hidden layer biases are randomly assigned to compute the hidden layer output matrix. Based on that, the output layer weight matrix can be computed via least-square linear regression by using the hidden layer output matrix and target output. The learning model expressed in (3)–(4) contains the training set (x_j, t_j) , the hidden node output function $G(w, b, x)$, and the number of hidden nodes L (Huang et al. 2006a, b).

$$f_L(x_j) = y_j, \forall j \quad (3)$$

$$\sum_{i=1}^L \beta_i G(w_i, b_i, x_j) = t_j, j = 1, 2, \dots, N \quad (4)$$

where x_j represents the input parameters, w_i is the weight vector connecting the i th hidden node and the input nodes, b_i is the bias of the i th hidden node, and β_i is the weight vector connecting the i th hidden node and the output nodes.

The training strategy of an ELM includes three steps. First, the hidden node parameters a_i and b_i are randomly assigned. Second, the hidden layer output matrix H is computed from (5). Last, the output weight β is computed as $\beta = H^+ T$, where H^+ is the

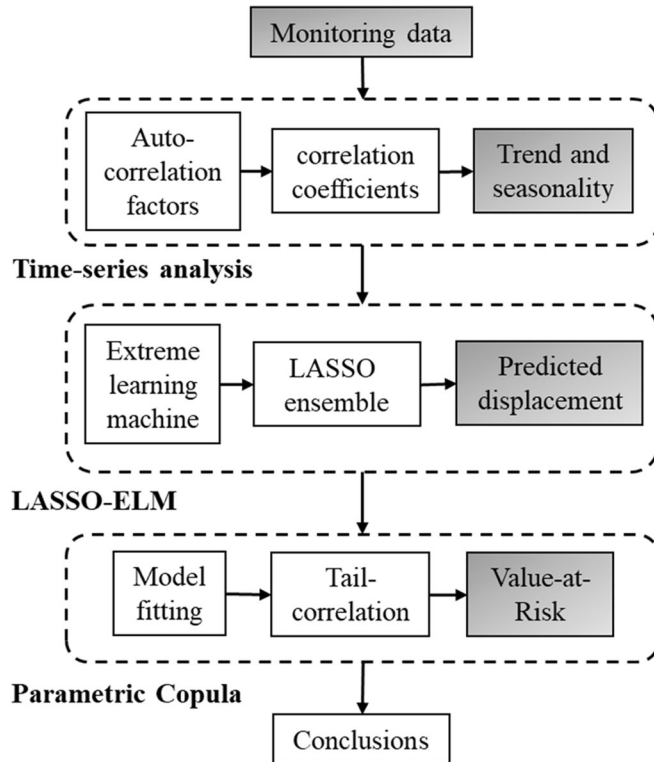


Fig. 1 Process of the data-driven framework

Moore-Penrose generalized inverse of the hidden layer output matrix H .

$$H(w_1, \dots, w_L; b_1, \dots, b_L; x_1, \dots, x_N) = \begin{bmatrix} G(w_1, b_1, x_1) & \dots & G(w_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ G(w_1, b_1, x_N) & \dots & G(w_L, b_L, x_N) \end{bmatrix}_{N \times L} \quad (5)$$

The single basic extreme learning machine randomly assigns hidden nodes and incrementally updates the output weights of the hidden layer nodes. With a small training dataset, the prediction error may be large. A linear ensemble-based ELM can overcome this challenge. Yu et al. (2013) proposed an ensemble of ELMs with least absolute shrinkage and selection operator (LASSO) regularization, which provides less prediction errors than any single ELM. The ensemble of ELMs is expressed in (6) and the LASSO regularization is expressed in (7).

$$\bar{y}_i = B_0 + \sum_{p=1}^8 y_i^p B_p \quad (6)$$

$$\min_{B_0, B} \left\{ \sum_{i=1}^N \left(\bar{y}_i - B_0 - \sum_{p=1}^8 y_i^p B_p \right)^2 \right\} \quad (7)$$

where B_0 and B_p are estimated intercept and weights of the linear ensemble regression model, \bar{y}_i denotes the target output vector, and y_i^p represents the vector of the predicted output from the P th ELM. In this study, our ensemble model is a regularized least-square regression model using LASSO. With LASSO regularization, we can estimate the intercept and weights B_0 and B_p of the linear ensemble model. The scheme of the linear ensemble-based ELMs can be illustrated in Fig. 3.

Benchmarking data-driven models

In this research, five benchmarking data-driven models including the support vector machine (SVM), neural network (NN), random forest (RF), k -nearest neighbor (kNN), and classical extreme learning machine (ELM) are compared with the proposed LASSO-ELM.

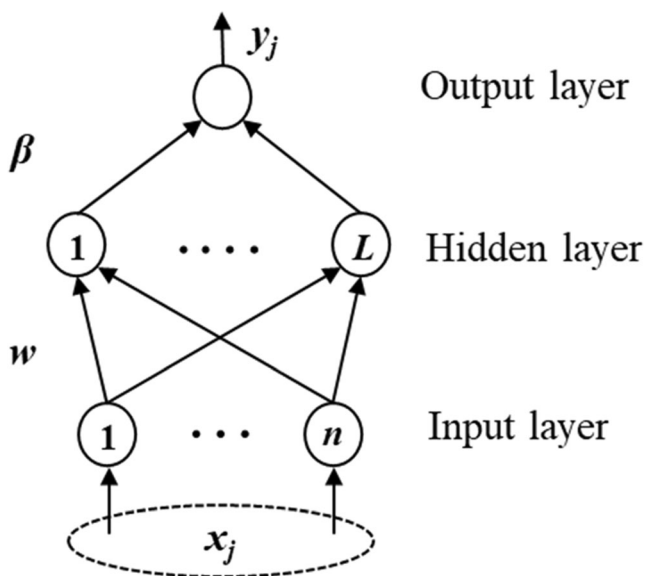


Fig. 2 Extreme learning machine

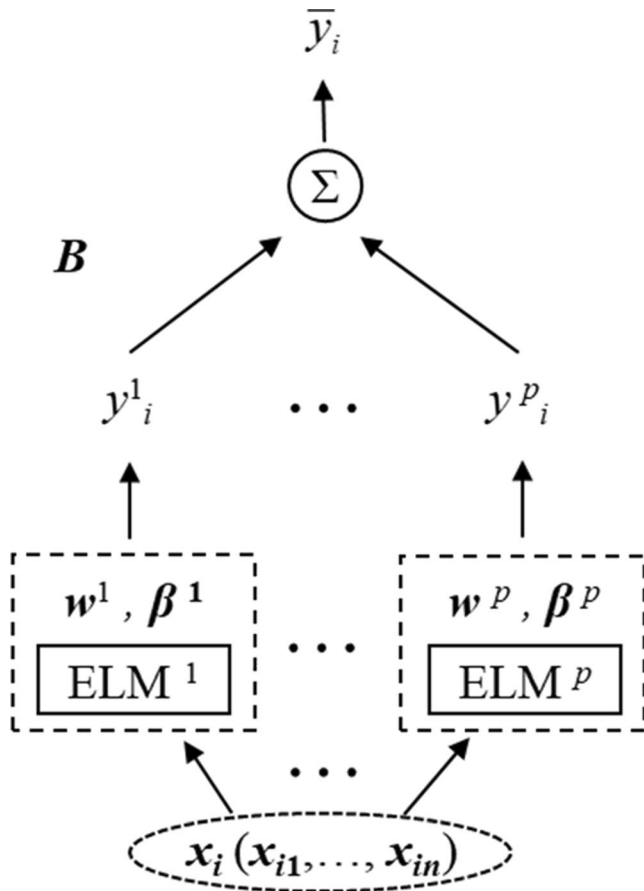


Fig. 3 Scheme of LASSO ensemble-based extreme learning machines

The support vector machine model (Abdi and Giveki 2013; Tong and Koller 2001) is applied in this study with the Gaussian kernel function expressed in (8).

$$K(X, X') = \exp\left(\frac{-\|X-X'\|}{2\sigma^2}\right) \quad (8)$$

where X is the vector of the input data and σ denotes the standard deviation of the input data. The capacity factor C of the SVM model with values 1, 10, 100, and 1000 is considered in the training of the SVM model. In addition, the parameter $\gamma = 1 \square 2\sigma^2$ with values of 0.001, 0.01, 0.1, and 1 is also considered during the training process. The optimal parameter settings of SVM are evaluated through 10-fold cross validation.

The neural network model (LeCun et al. 1990) in this research applies back-propagation (BP) to optimize its performance by adjusting the weight of each neuron. The structure of the NN model includes the input layer, the hidden layer, and the output layer. The Sigmoid function is selected as the activation function in this study. The number of hidden neurons in each hidden layer with values of 10, 20, 30, 40, and 50 and the number of hidden layers with values of 1, 2, 3, 4, and 5 are all evaluated via 10-fold cross validation to determine the optimal parameter setting.

The random forest model (Breiman 2001) assembles multiple decision trees, which are generated based on the values of an independent set of random variables. For classification and regression tasks, the best split is used among a subset of randomly selected predictors at the split node. The maximum depth of random forest with values of 10, 30, 50, 70, and 90 and the number of trees with values of 1000, 1500, 2000, and 2500 are all evaluated via 10-fold cross validation.

The k -nearest neighbor model (Denoeux 1995) for predicting the landslide displacement is also presented in this research. The Euclidean distance is used in the kNN model. Various values of k from 1 to 20 are evaluated through a 10-fold cross validation to determine the optimal k value.

The classical extreme learning machine model introduced in the “Extreme learning machine and LASSO regularization” section has been applied with the radial basis function (RBF) as the activation function. The numbers of hidden neurons within the hidden layer with values of 2, 4, 6, 8, ..., 98, and 100 are all evaluated via 10-fold cross validation to determine the optimal parameter setting.

Model evaluation

To assess the landslide displacement prediction performance of the data-driven models, four widely applied metrics, namely mean absolute error (MAE (9)), mean absolute percentage error (MAPE (10)), mean square error (MSE (11)), and root mean square error (RMSE (12)), are selected in this study.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\bar{y}_i - y_i| \quad (9)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\bar{y}_i - y_i}{y_i} \right| \quad (10)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\bar{y}_i - y_i)^2 \quad (11)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{y}_i - y_i)^2} \quad (12)$$

where y_i denotes the actual displacement value and \bar{y}_i represents predicted displacement value.

Copula theory

The triggering factors of landslides (e.g., reservoir water level fluctuations, precipitation) are uncertain and dynamic. Analyzing the correlation structure between the triggering factors and the landslide displacement is essential for landslide hazard monitoring and prevention. The thresholds of these triggering factors are valuable for monitoring of reactivation of landslides. Applying copula models for tail correlation analysis, the thresholds for precipitation and reservoir water level can be computed. Hence, in this research, copula models are constructed to study the relationship between the triggering factors and displacement values.

The theory of copula was firstly proposed by Sklar (1996) for nonlinear and asymmetric multivariate analysis. Nonlinearity and tail correlation between the variables can be fully investigated through the copula models. The general formula of the copula model expressed in (13) is an N -dimensional joint distribution function composed of N univariate marginal distribution functions.

$$F(x_1, x_2, \dots, x_N) = C[F_{x_1}(x_1), F_{x_2}(x_2), \dots, F_{x_N}(x_N)] \quad (13)$$

where x_N represents the N th variable; $F_{x_N}(x_N)$ denotes the marginal cumulative density function of the N th variable; $F(x_1, \dots, x_N)$ is the N -dimensional joint distribution; and $C[F_{x_1}(x_1), \dots, F_{x_N}(x_N)]$ is the copula function.

In practice, two copula families, namely the Archimedean copula family and the elliptical copula family, are widely used. In the elliptical copula family, the Gaussian copula and the Student t copula are most frequently used. In the Archimedean copula family, the Frank copula, the Joy-Clayton copula, the Gumbel-Hougaard copula, and the Ali-Mikhail-Haq copula are frequently used in practice (Reboredo 2011). All copula models in the elliptical family are symmetric and are less sensitive to tail correlations between variables. In the Archimedean copula family, the Frank copula evaluates the co-movement of highly associated variables (Hao and AghaKouchak 2013). The Gumbel-Hougaard copula is sensitive to the upper tail correlation and the Joe-Clayton copula fits the lower tail-correlated variables well. The Ali-Mikhail-Haq copula is a modified form of Gumbel-Hougaard copula and it is also performing well with upper correlated variables (Onken et al. 2009).

Previous research indicates that the landslides were mostly triggered by heavy precipitation and large fluctuation of reservoir water level (Du et al. 2013; Keefer et al. 1987; Yao et al. 2015). In this study, the Archimedean copula models which are more sensitive to the tail-correlated dataset are applied to model the data. To select the copula model that fits the data best, two performance evaluation metrics, namely Akaike information criterion (AIC (14)) and Bayesian information criterion (BIC (15)) (Chen and Gopalakrishnan, 1998; Posada and Buckley, 2004), are computed in this study.

$$AIC = -2\ln(L) + 2m \quad (14)$$

$$BIC = -2\ln(L) + m\ln(n) \quad (15)$$

where m is the number of the estimated variables, n is the number of data samples, and L represents the maximum log-likelihood function. In this research, the copula model with the minimum AIC and BIC values performs best.

For landslide hazard prevention, thresholds of triggering factors can be obtained as the indicator of large landslide displacements. Since the copula model studies the correlation between triggering factors (e.g., reservoir water level fluctuations, precipitation) and displacement values, the thresholds can be computed through Value-at-Risk (VaR) (He et al. 2017) expressed in (16).

$$VaR_p(x_1, \dots, x_N) = C^{-1}(x_1, \dots, x_N) \quad (16)$$

where x_N denotes the N th triggering factor, C^{-1} represents the reverse function of the copula function, and p is the confidence level to compute the thresholds. In this paper, p is set to 0.95 and only two-dimensional copula models are applied for our research respectively.

Field investigation

Study area

Case study was focused on the Baishuihe landslide, which is located on the south side of the Yangtze River within Zigui County, China. Its exact location is presented in Fig. 4. This landslide has a maximal length of 780 m from north to south and a maximum width of 700 m from east to west. The total area of the Baishuihe landslide is 0.42 km² and it contains a total volume of 1260 × 10⁴ m³. The current slide ranges in elevation between 75 and 390 m from the toe to the main scarp respectively (Du et al. 2013). The central part of the Baishuihe landslide is relatively flat while the gradients in the upper and lower parts of landslide are larger. The bedrock geology of the landslide area consists mainly of low-strength sandstones and mudstones (Lian et al. 2013). In recent years, the slope deformation of the Baishuihe landslide is more intense, which is attributed to fluctuations of the reservoir water level and heavy precipitation in the flood season. Baishuihe landslide is a retrogressive landslide, with failure surface which propagates from the Yangtze River in the direction of the slope. The cumulative displacement of the landslide has reached a distance of more than 2500 mm during the year 2003 to 2015.

Data collection

The Baishuihe landslide shows repeated reactivations since the first activation in 2003. After first reactivation, slope deformation monitoring has been conducted by configuring 11 global positioning system (GPS) points, with 1-month data collection frequency, on the landslide surface area as illustrated in Fig. 5. The deformation rate in the bottom part is significantly higher than that in the upper part. Thus, more GPS monitoring points were assigned in this area. Among those, ZG93 and ZG118 have a complete displacement dataset over several years and are hence utilized for detailed numerical analysis in this study. Previous research indicated that the main triggering factors of landslides in the Three Gorges Reservoir are precipitation and reservoir water level fluctuations (Cao et al. 2016; Du et al. 2013; Zhou et al. 2016). Hence, the historic data of these two factors are also utilized in this study. The dataset of reservoir water level has been collected through on-site investigation, and the dataset of precipitation has been obtained from a monitoring site 9.5 km away from the landslide location. The raw dataset is presented in Fig. 6 respectively.

In this study area, based on the GPS time series data (see Fig. 6), two significant types of displacements occurred: seasonal faster displacement and slower displacement. The seasonal faster displacement which illustrates “step-like” patterns has steep positive gradients or steps in cumulative horizontal displacement plots (Massey et al. 2013). In addition, the average cycle length for seasonal faster displacement is approximately 1 year. The long period of slower motion comprises semi-constant displacement rates over several months or most parts of the year. Considering the seasonal patterns of reservoir water level fluctuations and precipitation illustrated in Fig. 6, the collected dataset suggests a

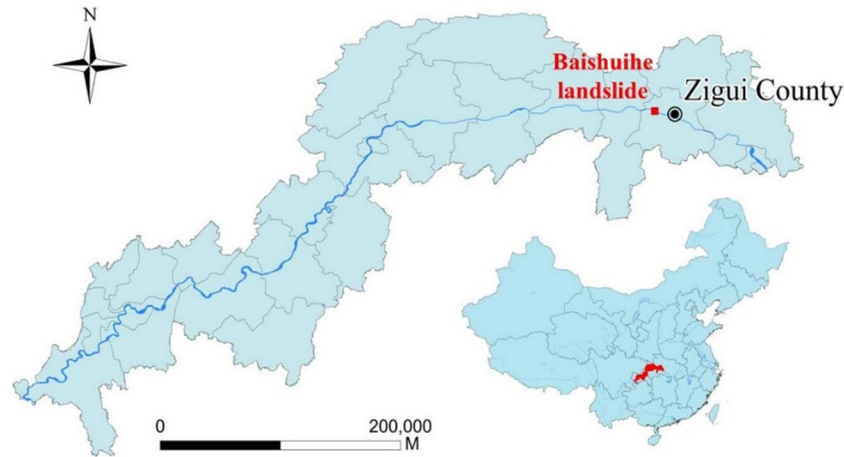


Fig. 4 Location of the Baishuihe landslide

complex relationship between them and the landslide displacements.

Experimental results

The landslide displacement data utilized in this paper were collected between 2003 and 2015 from the GPS monitoring points on the Baishuihe landslide in the Three Gorges Reservoir, China. Several reactivations of the landslide are recorded during this monitoring period. Before analyses of landslide displacement and triggers, all monitoring data are pre-processed, including outlier removal and missing value imputation.

Time series analysis

Previous studies (Du et al. 2013; Keefer et al. 1987; Yao et al. 2015) indicate that the reservoir water level fluctuations and precipitation have a seasonal trend that generates a landslide with a seasonal faster displacement. In this paper, the precipitation and reservoir water level fluctuations, illustrated in Fig. 6, are analyzed independently by constructing autoregressive moving average time series models as expressed in (1). The autocorrelation function (expressed in (2)), which indicates the correlation between

the current value and the historic values, is computed and illustrated in Fig. 7 with a time window of 30 months of historic data.

In the realm of time series modeling, the seasonality pattern matters to the future prediction accuracy. Significant seasonality patterns of the two major triggering factors are observable in Fig. 7. The monthly precipitation and monthly average reservoir water level are showing strong yearly seasonal behavior. The significance of the autocorrelations is also evaluated through Durbin-Watson (DW) statistic tests (Granger and Newbold 1974) and the factors with a DW statistic value less than 1.5 are determined to have significant impact on the current value. In this study, the DW test statistics of all autocorrelations of selected monthly lags are less than 1.5.

As shown in Fig. 8, the first two monthly lags for precipitation have the highest absolute values of correlation coefficient and are hence selected for predicting the monthly instant displacement in the forthcoming month. For reservoir water level fluctuations, the selected lags that contain the positive correlation coefficient are the third to eighth monthly lagged data. These historic precipitation or reservoir water level fluctuations are selected as inputs in the ARMA time series models to forecast the current value. Consequently, the trend patterns of the time series models can be

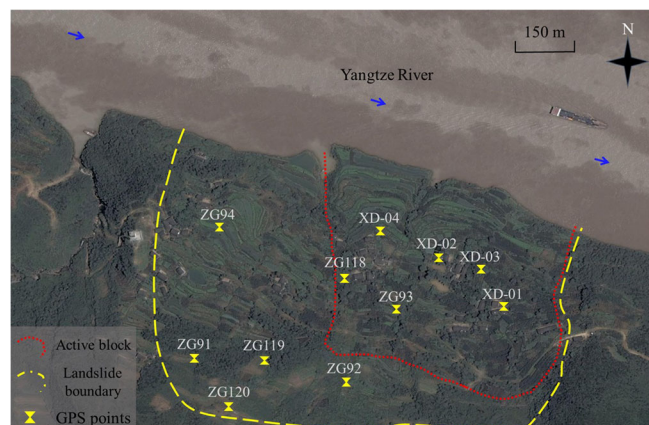


Fig. 5 Configuration of GPS monitoring points in Baishuihe landslide

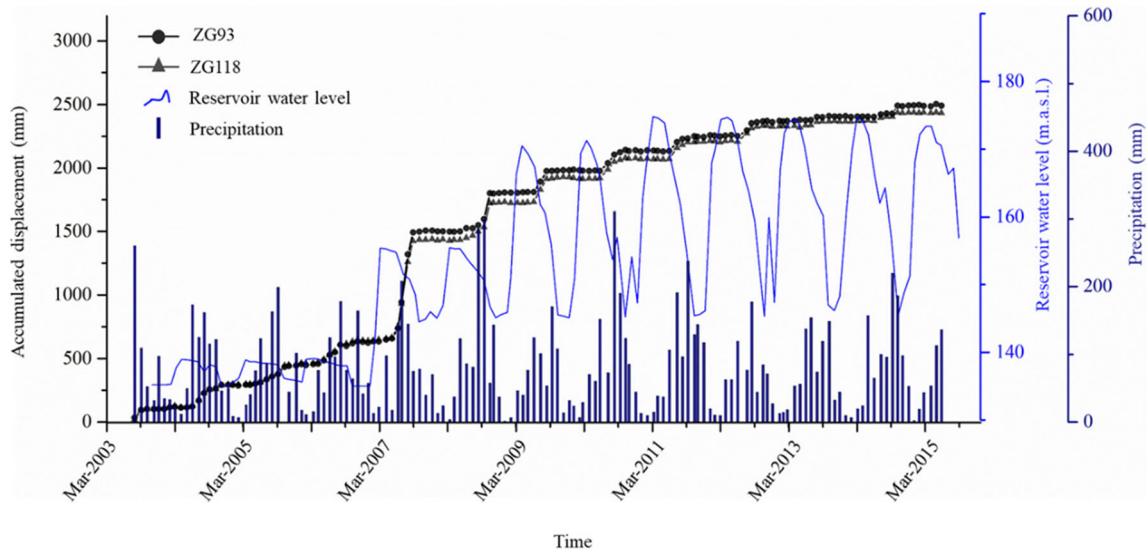


Fig. 6 Historic landslide cumulative displacement, reservoir water level, and precipitation during 2003 to 2015

represented by the selected lags. Hence, both the trend patterns and the seasonality patterns are embedded in the ARMA models to predict the future precipitation and reservoir water level fluctuations in the region of the Baishuihe landslide.

Displacement prediction

In this research, our proposed framework containing multiple extreme learning machines assembled by LASSO regularization (LASSO-ELM) is applied to predict the monthly displacement of

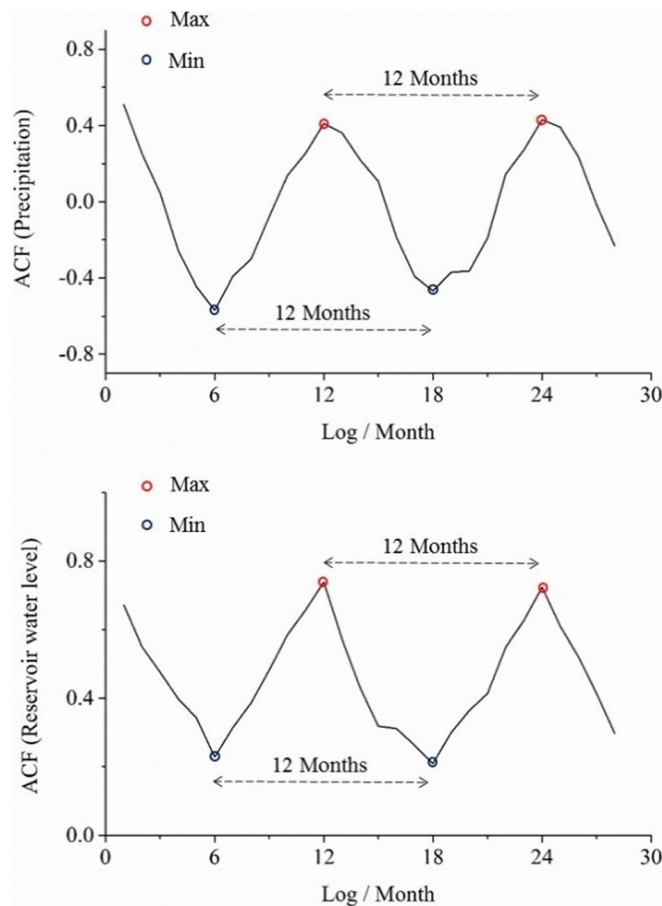


Fig. 7 Autocorrelation function of precipitation and reservoir water level fluctuations

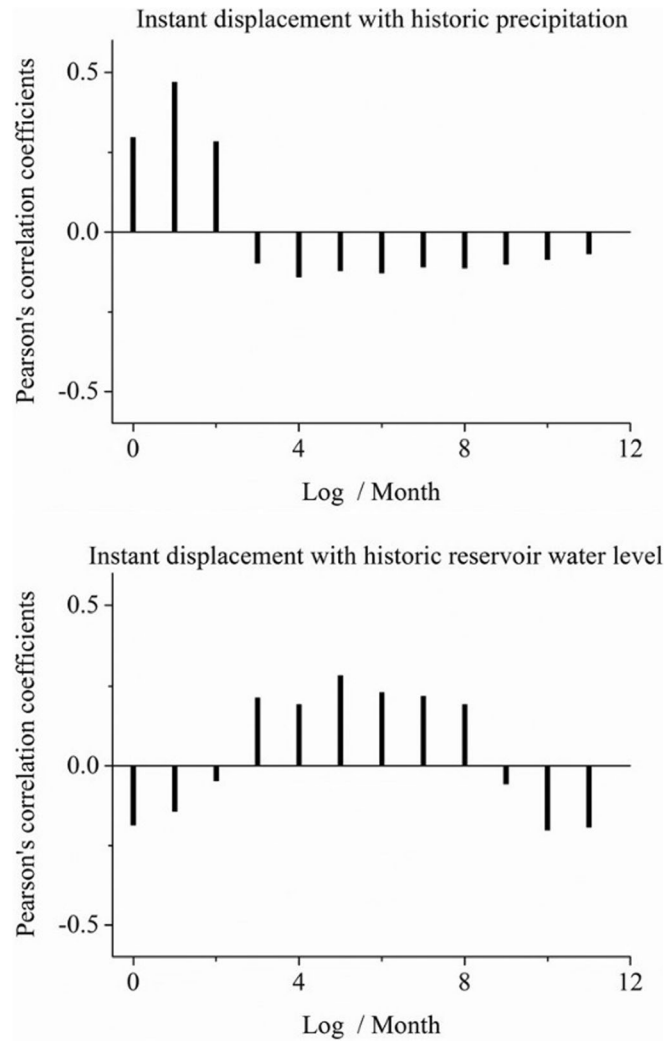


Fig. 8 Pearson's correlation coefficient between instant displacement values, precipitation, and reservoir water level fluctuations

the Baishuihe landslide. Data collected from January 2004 to February 2012 from monitoring points are applied as training dataset and data collected between March 2012 and May 2015 are used for model validation. Benchmarking prediction algorithms including the support vector machine, neural networks, random forest, k -nearest neighbor, and the classical extreme learning machine are compared with our proposed framework. Four widely applied metrics such as given in Eqs. (9) to (12) are used to assess the performances with respect to landslide displacement prediction.

To achieve the best prediction performance, the parameters of our framework are examined. Based on previous studies (Fernández-Navarro et al. 2011; Huang and Siew 2005), the best performing activation function for the extreme learning machine is the radial basis function (RBF), which is selected in this study. By testing 10% of the randomly sampled data, the optimal number of hidden neurons in the single hidden layer of each ELM is 58. Eight extreme learning machines are combined in this study using LASSO regularization to optimize the weights based on the training performance. The prediction performance of the proposed

LASSO-ELM approach and other benchmarking methods is presented in Table 1.

From Table 1, the SVM and NN produced better results in comparison with RF and kNN. However, the major constraint of utilizing SVM and NN is the challenge to locate the global optimal solution by trying all model parameters. On the other hand, the extreme learning machine contains one single hidden layer and the controlling parameter is the number of hidden neurons which can be searched easily through testing on a small sample dataset. Furthermore, the linear ensemble structure with LASSO regularization offers more weights on the best performing ELM and increases the prediction accuracy further. Therefore, it is clear that the LASSO-ELM has the best performance with the lowest MAPE and RMSE values. It generally indicates that LASSO-ELM outperforms the other models in predicting landslide displacement values.

To validate the prediction accuracy, the predicted accumulative displacement values produced by all tested methods are presented in Fig. 9. The predicted displacement from the proposed LASSO-ELM is the closest to the measured displacement values between March 2012 and March 2015. Hence, the predicted increment

Table 1 Performance evaluation of prediction models

Algorithm	MAE (mm)	MAPE (%)	MSE (mm)	RMSE (mm)
SVM	53.61	4.90	4.58	78.04
NN	44.94	2.45	4.28	58.41
RF	67.12	5.24	4.75	86.84
kNN	74.75	5.60	5.40	103.79
ELM	37.65	2.59	2.47	47.54
LASSO-ELM	26.44	1.30	2.28	35.63

displacement values from LASSO-ELM are used for tail correlation analysis in the “Tail correlation analysis” section.

Tail correlation analysis

With predicted displacement values, the performance of the proposed LASSO-ELM approach has been validated. In order to identify the seasonal faster displacement patterns of the landslides, the computation of the upper thresholds of the triggering factors is essential for the hazard prevention. In this section, tail correlation analysis is conducted to investigate the relationship between the predicted displacement values and the major triggering factors. Copula models are constructed to study the co-movement between the landslide increment displacement and triggering factors. Based on the computed results from the “Time series analysis” section, the past monthly precipitation and the monthly average reservoir water level from 5 months ago have the most significant autocorrelation and are hence selected to construct the copula models.

In this study, two families of copula models (i.e., elliptical family, Archimedean family) are selected to fit the predicted displacement and the two major triggering factors. The normality of the selected variables matters to the fitting performance of copula models. Table 2 expresses the summary of the normality tests of the selected variables for the copula models. Landslide displacement increments refer to the predicted increments in the current month. Reservoir water level refers to the average level 5 months ago. Precipitation is the precipitation in the past month. Anderson-Darling (A-D) test is the A^2 statistics for the test of normality. And

the Jarque-Bera (J-B) test is the χ^2 test for the normality. The Kolmogorov-Smirnov (K-S) test is a Goodness-of-Fit test for normality (Wilcox 2005). All normality tests with p values less than 0.05 refer to the non-normal distribution of all variables. The computational results indicate the existence of tail correlations between the variables. Hence, the Archimedean family of copula models is selected to fit the data.

Table 3 presents the fitted parameters of the copula models of predicted displacements versus reservoir water level and predicted displacements versus precipitation. Using the data collected from March 2012 to March 2015, the Archimedean copula model parameters are estimated through the maximum likelihood estimation. The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are computed for the calculation of model fitting performances. The formulas to compute AIC and BIC have been defined in the “Copula theory” section (see Eq. (14)–(15)). The smaller values of AIC and BIC indicate better performances of the Archimedean copula models.

The performances of the Archimedean copula models are illustrated in Table 3. The AIC and BIC values are computed to evaluate the fitting performance of the copula models. In comparison with the others, the Gumbel-Hougaard copula models show the smallest values of AIC and BIC, and hence they outperform in both datasets. It demonstrates strong upper tail correlations among predicted displacement values and the two triggering factors. The constructed Gumbel-Hougaard copula models for reservoir water level versus predicted displacements and precipitation versus predicted displacements are illustrated in Fig. 10. The

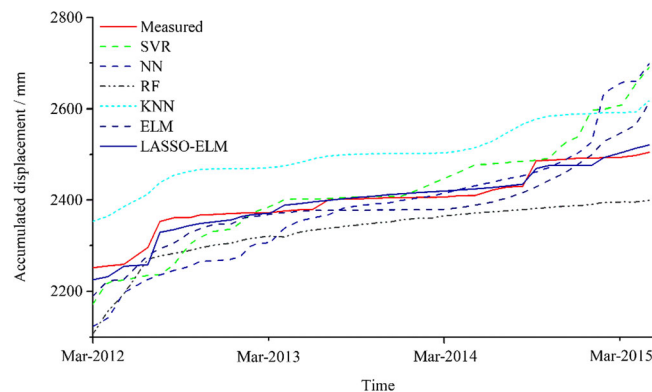
**Fig. 9** Performance comparison of benchmarking algorithms and the proposed framework

Table 2 Normality test results

Parameter		Anderson-Darling test Test stat	<i>p</i> value	Jarque-Bera test Test stat	<i>p</i> value	Kolmogorov-Smirnov Test stat	<i>p</i> value
March 2012 to March 2013	Landslide displacement increments	1.4870	0.0005	15.2730	0.0010	0.3379	0.0010
	Reservoir water level	4.0523	0.0005	12.2426	0.0010	0.2737	0.0010
	Precipitation	0.6725	0.0434	98.3975	0.0010	0.2197	0.0010
March 2013 to March 2014	Landslide displacement increments	2.3964	0.0005	35.8083	0.0010	0.3840	0.0010
	Reservoir water level	1.2481	0.0027	9.2258	0.0168	0.1758	0.0024
	Precipitation	2.0965	0.0005	105.7399	0.0010	0.2423	0.0010
March 2014 to March 2015	Landslide displacement increments	2.4824	0.0005	18.8034	0.0010	0.4341	0.0010
	Reservoir water level	1.4317	0.0008	11.1637	0.0096	0.2310	0.0010
	Precipitation	0.8905	0.0117	68.2922	0.0010	0.1912	0.0127

computed upper tail parameters and the upper tail thresholds, namely Value-at-Risks (see expression (16)), have been provided in Table 4.

Based on Fig. 10 and Table 4, the upper tail correlations among precipitation and reservoir water level are strong. It indicates that the seasonal faster displacements are likely to be triggered by large precipitation from last month and large reservoir water level fluctuation from the periods 5 months ago. Based on the constructed copula models, the Value-at-Risks (VaRs) as the upper tail thresholds are computed through (15) and presented in Table 4. The VaR of monthly precipitation is 201.17 mm and the VaR of average monthly reservoir water level is 170.49 m. In practice, large displacement values have high probability of occurrence as the triggering factors exceed the VaR values.

Model validation

In this research, the reservoir water level fluctuations and precipitation are major triggering factors of seasonal faster displacements and their VaRs are valuable for predicting the displacement patterns. To validate the effectiveness of the proposed monitoring framework, the historical data of these two triggering factors are utilized for seasonal faster displacement prediction. A confusion matrix is selected to assess the performance of the proposed framework (see Table 5).

According to Table 5, all seasonal faster displacements in the Baishuihe landslide have been identified prior to its occurrence based on monitoring of the historic reservoir water level and precipitation data. Only one false identification has been produced. Hence, the effectiveness and accuracy of using the thresholds of reservoir water level and precipitation have been confirmed.

Overall, the proposed framework has successfully identified all periods of seasonal faster by monitoring the triggering factors. Based on the prediction by LASSO-ELM, the seasonal faster displacements can be predicted and monitored in advance. Hence, the proposed framework is applicable to the real geological investigations in this area.

Discussion

As described in the “Time series analysis” section, the past monthly precipitation and the monthly reservoir water level from 5 months ago had the highest correlation coefficient with the instant displacement values respectively. Based on previous research (Cao et al. 2016; Du et al. 2013; Lian et al. 2015; Matsuura et al. 2008; Miao et al. 2018), the instant displacement has a delayed response to the triggering factors. Miao et al. (2014) carried out a research to study a delayed response of the Baishuihe landslide movement to triggering factors. Based on

Table 3 Parameter estimation of Archimedean copula models

		Copula Joe-Clayton	Gumbel-Hougaard	Frank	Ali-Mikhail-Haq
Reservoir water level vs displacement	ρ	0.208	1.662	2.375	0.506
	Log-likelihood	− 39.707	− 50.325	− 24.474	− 29.818
	AIC	− 77.413	− 98.651	− 46.947	− 57.636
	BIC	− 75.830	− 97.067	− 45.364	− 56.052
Precipitation vs displacement	ρ	0.223	1.208	3.458	0.488
	Log-likelihood	− 28.563	− 40.248	− 29.600	− 36.999
	AIC	− 55.126	− 78.496	− 57.201	− 71.998
	BIC	− 53.543	− 76.912	− 55.617	− 70.415

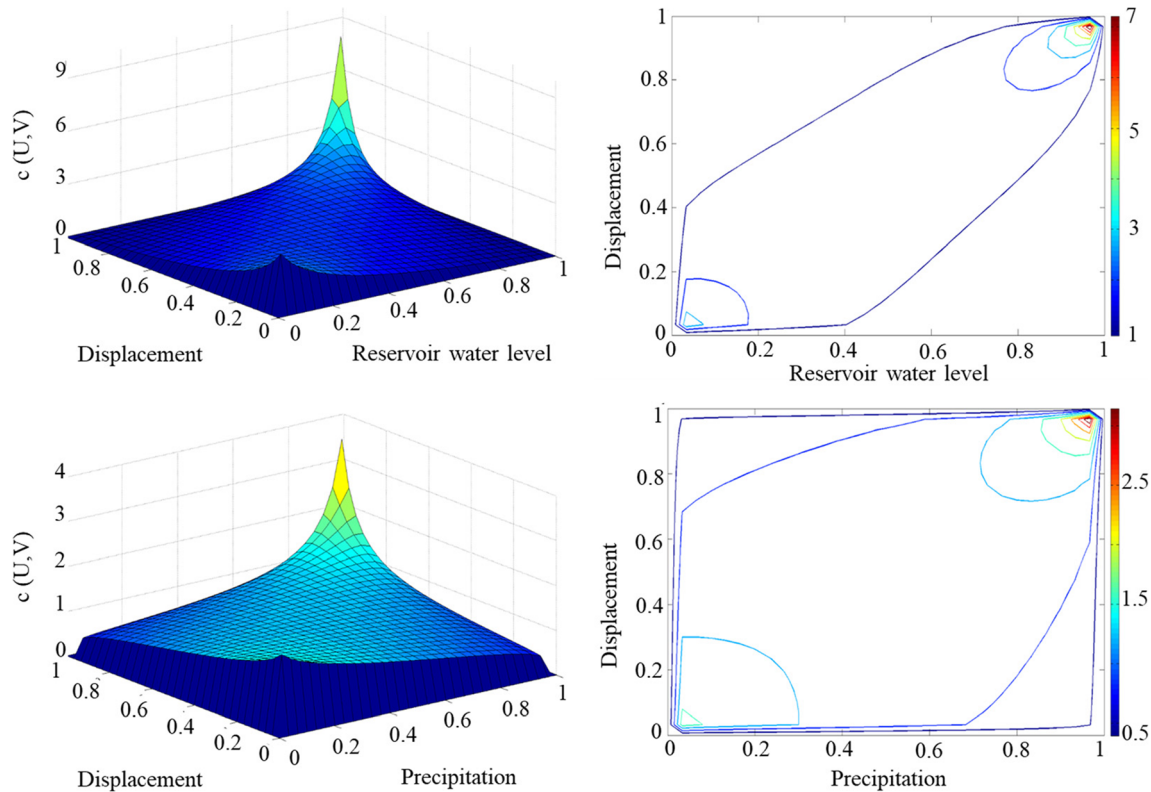


Fig. 10 Fitted Gumbel-Hougaard copula models

his work, the incremental pore water pressure which triggers the landslide displacement is caused by precipitation and periodic fluctuations of reservoir water level. For ancient deep-seated reservoir landslide in the state of limit equilibrium, there still exists a time lag for the transmission of the incremental water pressure to the sliding surface of the landslide. The length of the time lag is closely associated with the soil permeability and the thickness of landslide mass. In our case, since the Baishuihe landslide is a deep-seated landslide, it exhibits a significant time lag between the triggering factors and the instant displacement values.

Periods of seasonal faster displacement and slower displacement comprise the displacement patterns of the Baishuihe landslide as discussed in the “Data collection” section. The reservoir water level fluctuations and precipitation both illustrate seasonal patterns and have a complex relationship with the landslide displacements. In this paper, we propose a framework using LASSO-ELM and Gumbel-Hougaard copula models to quantify the effect of reservoir water level and precipitation on the displacements. Displacement prediction models are constructed using the historic reservoir water level data and the precipitation data. Then, the

Table 4 Tail correlation parameters and Value-at-Risk

Correlation coefficient	Precipitation vs displacement	Reservoir water level vs displacement
Upper tail	0.378	0.647
Kendall's tau	0.336	0.332

copula models are created to analyze the tail correlation between the displacement and triggering factors. There are three major advantages of our proposed approach. First, the extreme learning machine used in this research contains only one hidden layer and can be trained easily and rapidly. Second, the ensemble of the ELMs strengthens the model performance and improves the displacement prediction accuracy. Based on the experimental results presented in Table 1, the ensemble approach outperforms all algorithms tested. Third, the copula model provides more significant upper tail correlations compared to other approaches. We are enabled to construct the threshold to monitor and forecast the seasonal fast displacement based on historic data.

The data used in this research was monitored and collected at monthly resolution and it resulted in an appropriate size of the dataset. It enables many state-of-the-art machine learning algorithms such as ELM to offer sufficient prediction performance. However, when it comes to large displacement dataset that contains highly frequent data and complex nonlinear characteristics,

Table 5 Confusion matrix of model validation results

		Predicted Seasonal faster displacement	Slower displacement
Actual	Seasonal faster displacement	7	0
	Slower displacement	1	28

the ELM which has only one hidden layer may not be capable to fully recognize these patterns. This is one of the major constraints of the utility of ELM-based algorithms. In such cases, the recently proposed deep learning algorithms may be a feasible approach to tackle such complexity and may be applied in the near future research.

Conclusions

In this research, a data-driven approach for monitoring and predicting of landslide displacements has been proposed. Data collected from the Baishuihe landslide in the Three Gorges Reservoir between 2003 and 2015 have been utilized in this case study. A numerical analysis of the capacity of an assembly of extreme learning machines and copula models in modeling landslide seasonal faster displacement is performed.

To confirm the effectiveness and accuracy of the assembly of ELMs based on LASSO regularization, a comparative analysis of the LASSO-ELM, classical ELM, SVM, RF, and NN is conducted in the first part of this paper. Four evaluation metrics, the MAE, MAPE, MSE, and RMSE are selected to assess the prediction performance. Based on the predicted displacement, Gumbel-Hougaard copula models are constructed to investigate the tail correlation between precipitation, reservoir water level fluctuations, and the predicted landslide displacement. Thresholds for monitoring cyclic faster displacements are computed based on the tail correlation analysis with copula models.

Results of computational studies validated the predictive modeling capacity of LASSO-ELM in forecasting landslide seasonal faster displacement. The LASSO-ELM provided the lowest MAE, MAPE, MSE, and RMSE while the SVM produced the highest prediction error. Based on the predicted displacements, a significant upper tail correlation between the displacement and past month precipitation is assessed through the copula model. Upper tail correlation between the displacement and the level of the reservoir from 5 months ago is also significant. Hence, the reactivation landslide hazards can be predicted and prevented in advance through monitoring of the precipitation and reservoir water level respectively.


Acknowledgements

We are immensely grateful to the suggestions and guidance from Prof. T. W. J. van Asch from Utrecht University.

Funding information This research was supported by the Science Fund for Creative Research Groups of the National Natural Science Foundation of China (Grant No. 41521002) and the Key Program of National Natural Science Foundation of China (Grant No. 41630640).

References

- Abdi M, Giveli D (2013) Automatic detection of erythematous-squamous diseases using PSO-SVM based on association rules. *Eng Appl Artif Intell* 26(11):603–608
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
- Bustos O, Yohai V (1986) Robust estimates for ARMA models. *J Am Stat Assoc* 81(393):155–168
- Cao Y, Yin K, Alexander DE et al (2016) Using an extreme learning machine to predict the displacement of step-like landslides in relation to controlling factors. *Landslides* 13(Suppl 2):1–12
- Chen S, Gopalakrishnan P (1998) Speaker, environment and channel change detection and clustering via the Bayesian information criterion. In: *Proc. darpa broadcast news transcription and understanding workshop*, vol 8. pp 127–132
- Denoeux T (1995) A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE Transactions on Systems, Man, and Cybernetics* 25(5):804–813
- Du J, Yin K, Lacasse S (2013) Displacement prediction in colluvial landslides, Three Gorges Reservoir, China. *Landslides* 10(2):203–218
- Fernández-Navarro F, Hervás-Martínez C, Sánchez-Monedero J, Gutiérrez PA (2011) MELM-GRBF: a modified version of the extreme learning machine for generalized radial basis function neural networks. *Neurocomputing* 74(16):2502–2510
- Granger C, Newbold P (1974) Spurious regressions in econometrics. *J Econ* 2(2):111–120
- Hao Z, AghaKouchak A (2013) Multivariate standardized drought index: a parametric multi-index model. *Adv Water Resour* 57:12–18
- He Y, Kusiak A (2018) Performance assessment of wind turbines: data-derived quantitative metrics. *IEEE Transactions on Sustainable Energy* 9(1):65–73
- He Y, Kusiak A, Ouyang T, Teng W (2017) Data-driven modeling of truck engine exhaust valve failures: a case study. *J Mech Sci Technol* 31(6):2747–2757
- Huang G (2003) Learning capability and storage capacity of two-hidden-layer feedforward networks. *IEEE Trans Neural Netw* 14(2):274–281
- Huang G, Siew C (2005) Extreme learning machine with randomly assigned RBF kernels. *Int J Inf Technol* 11(1):16–24
- Huang G, Chen L, Siew C (2006a) Universal approximation using incremental constructive feedforward networks with random hidden nodes. *IEEE Trans Neural Netw* 17(4):879–892
- Huang G, Zhu Q, Siew C (2006b) Extreme learning machine: theory and applications. *Neurocomputing* 70(1):489–501
- 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
- Keefer D, Wilson R, Mark R et al (1987) Real-time landslide warning during heavy rainfall. *Science* 238(4829):921–925
- Krkač M, Špoljarić D, Bernat S, Arbanas SM (2017) Method for prediction of landslide movements based on random forests. *Landslides* 14(3):947–960
- LeCun Y, Boser B, Denker J et al (1990) Handwritten digit recognition with a back-propagation network. In: *Advances in neural information processing systems*. pp 396–404
- Li X, Kong J, Wang Z (2012) Landslide displacement prediction based on combining method with optimal weight. *Nat Hazards* 61(2):635–646
- Lian C, Zeng Z, Yao W, Tang H (2013) Displacement prediction model of landslide based on a modified ensemble empirical mode decomposition and extreme learning machine. *Nat Hazards* 66(2):759–771
- Lian C, Zeng Z, Yao W, Tang H (2015) Multiple neural networks switched prediction for landslide displacement. *Eng Geol* 186:91–99
- Massey C, Petley D, McSaveney M (2013) Patterns of movement in reactivated landslides. *Eng Geol* 159:1–19
- Matsuura S, Asano S, Okamoto T (2008) Relationship between rain and/or meltwater, pore-water pressure and displacement of a reactivated landslide. *Eng Geol* 101(1–2):49–59
- McLeod A, Li W (1983) Diagnostic checking ARMA time series models using squared-residual autocorrelations. *J Time Ser Anal* 4(4):269–273
- 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 et al (2018) Prediction of landslide displacement with step-like behavior based on multialgorithm optimization and a support vector regression model. *Landslides* 15(3):475–488
- Miche Y, Akusok A, Veganzones D, Björk KM, Séverin E, du Jardin P, Termenon M, Lendasse A (2015) SOM-ELM—self-organized clustering using ELM. *Neurocomputing* 165:238–254
- Onken A, Grünewälder S, Munk M et al (2009) Analyzing short-term noise dependencies of spike-counts in macaque prefrontal cortex using copulas and the flashlight transformation. *PLoS Comput Biol* 5(11):e1000577
- Ouammi A, Dagdougui H, Sacile R, Mimet A (2010) Monthly and seasonal assessment of wind energy characteristics at four monitored locations in Liguria region (Italy). *Renew Sust Energ Rev* 14(7):1959–1968
- Posada D, Buckley T (2004) Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Syst Biol* 53(5):793–808
- Reboredo J (2011) How do crude oil prices co-move?: A copula approach. *Energy Econ* 33(5):948–955

-
- Sklar A (1996) Random variables, distribution functions, and copulas: a personal look backward and forward. Lecture notes-monograph series 1–14
- Sun K, Huang S, Wong D et al (2017) Design and application of a variable selection method for multilayer perceptron neural network with lasso. *IEEE T Neur Net Lear* 28(6):1386–1396
- Tibshirani R (1996) Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B Methodol* 58:267–288
- Tong S, Koller D (2001) Support vector machine active learning with applications to text classification. *J Mach Learn Res* 2(Nov):45–66
- Tramèr F, Kurakin A, Papernot N et al (2017) Ensemble adversarial training: attacks and defenses arXiv preprint arXiv: 1705.07204
- Wilcox R (2005) Kolmogorov–Smirnov test. *Encyclopedia of biostatistics*. <https://doi.org/10.1002/0470011815.b2a15064>
- Xu F, Wang Y, Du J et al (2011) Study of displacement prediction model of landslide based on time series analysis. *Chin J Rock Mech Eng* 4:014
- Yao W, Zeng Z, Lian C, Tang H (2015) Training enhanced reservoir computing predictor for landslide displacement. *Eng Geol* 188:101–109
- Yin Y, Wang H, Gao Y, Li X (2010) Real-time monitoring and early warning of landslides at relocated Wushan Town, the Three Gorges Reservoir, China. *Landslides* 7(3):339–349
- Yu Q, Miche Y, Eirola E, van Heeswijk M, Séverin E, Lendasse A (2013) Regularized extreme learning machine for regression with missing data. *Neurocomputing* 102:45–51
- 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
- Zong W, Huang G (2011) Face recognition based on extreme learning machine. *Neurocomputing* 74(16):2541–2551
-
- H. Li · Q. Xu** 
- State Key Laboratory of Geo-hazard Prevention and Geo-environment Protection, Chengdu University of Technology, No. 1 Erxianqiao East Rd., Chengdu, 610059, Sichuan, China
Email: xுqiang_68@126.com; xq@cduet.edu.cn
- Y. He**
Department of Mechanical and Industrial Engineering,
The University of Iowa,
Iowa City, IA 52242, USA
- J. Deng**
College of Computing and Digital Media,
DePaul University,
Chicago, IL 60604, USA