Engineering Geology 204 (2016) 108–120

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo

Application of time series analysis and PSO–SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir, China

Chao Zhou [a], Kunlong Yin [a,*], Ying Cao [a], Bayes Ahmed [b]

[a] Engineering Faculty, China University of Geosciences, Wuhan 430074, China
[b] Institute for Risk and Disaster Reduction, Department of Earth Science, University College London (UCL), WC1E 6BT, UK

ARTICLE INFO

Article history:
Received 29 September 2015
Received in revised form 6 January 2016
Accepted 21 February 2016
Available online 24 February 2016

Keywords:
Step-like landslide
Particle Swarm Optimization (PSO)
Support Vector Machine (SVM)
Displacement prediction
Three Gorges Reservoir area

ABSTRACT

The landslide displacement in the Three Gorges Reservoir, China, experiences step-like deformation that is influenced by rainfall and the periodic scheduling of the reservoir. In view of the step-like characteristic, the Particle Swarm Optimization and Support Vector Machine (PSO–SVM) coupling model based on the response of the induced factors was proposed to predict the landslide displacement. The moving average method was adopted to divide the total displacement into trend term and periodic term. The trend displacement was controlled by the geological conditions and predicted by polynomial function, while the periodic displacement was under the combined control of the triggers and the evolution state of the landslide. Therefore, the PSO–SVM model, based on the factors of the precipitation, the variation range of the reservoir and the displacements of the prior-periods, was proposed to predict the periodic displacement. The typical step-like landslide in the Three Gorges Reservoir, which is known as the Bazimen landslide, was taken as a case study to verify the prediction results. The values of the root mean square error and the mean absolute percentage error were 13.28 and 25.95, respectively. The results showed that rainfall and reservoir water level were the dominant factors for the step-like landslide deformation. The evolution state of the landslide was also significant in reflecting the response relationship between the displacement and inducing factors. In conclusion, the proposed PSO–SVM model can better represent the response relationship between the factors and the periodic displacement, which made the predicted values of the total displacement fit with the measured values greatly.

1. Introduction

Landslides and associated slope failure phenomena are one of the major geological hazards in the world, particularly in China and South Asia (Miyagi et al., 2011; Yin et al., 2009; Kirschbaum et al., 2010; Ahmed, 2014). Landslides can cause massive casualties, losses and damages to properties. The recent earthquakes in Nepal (i.e. on 25 April 2015) generated around 3000 landslides (ICIMOD, 2015). Moreover, the rainfall-induced Abe Barek landslides on 2 May 2014 caused 2700 casualties in Ago district, Afghanistan (Zhang et al., 2015). China suffers severely from landslide hazards, where on an average 700 to 900 people are killed every year (Chen et al., 2008). Therefore, it has become an urgent worldwide problem to reduce property damages and casualties caused by landslide hazard.

However, for many cases, engineering measure to stabilize dangerous slopes is not the best option for the local administration and the affected people, as it needs very high cost. Monitoring, prediction and early warning are the most economical landslide risk reduction measures which are applicable for both developed and developing countries. The prediction accuracy of landslide movement is one of the critical parameters for early warning (Sassa et al., 2009; Du et al., 2013). The accurate prediction of landslide displacement can reduce the damage to a certain extent. For example, the forecasting of the Xintan and Qianjiangping landslide in China helped reducing the economic losses and casualties significantly (Demir et al., 2015).

Since 1960s when Saito proposed the empirical formula for landslide prediction, the theory and method have obtained greater development (Cao et al., 2015; Sassa et al., 2009; Saito, 1965). Numerous theories and models have been applied until now to predict landslide, which fall into three main categories: deterministic model, statistical model and non-linear model. With the rapid development of theories and technologies related to pattern recognition and artificial intelligence, the non-linear models have become popular in predicting the movement of landslide, including the Gaussian process model (Liu et al., 2014), the neural network model (Gelisli et al., 2015; Pradhan et al., 2014), the Support Vector Machine (SVM) model (Goetz et al., 2015; Kavzoglu et al., 2015a), and the Extreme Learning Machine (ELM) model (Cao et al., 2015), and so on.

The neural network technology has provided a powerful tool for the research on landslide displacement prediction (Lian et al., 2015; Liu et al., 2014; Du et al., 2013; Ran et al., 2010). But the defects of faulty

* Corresponding author.
E-mail addresses: zhouchao_rs@163.com (C. Zhou), yinkl@126.com (K. Yin), caoying_sophis@126.com (Y. Cao), bayesahmed@gmail.com (B. Ahmed).

http://dx.doi.org/10.1016/j.enggeo.2016.02.009
0013-7952/© 2016 Elsevier B.V. All rights reserved.

theory foundation, local minimum and over fitting weakened the ability of prediction (Sangheum et al., 2014). The SVM is a machine learning method established based on the statistical learning theory for small sample and the principle of structural risk minimization. SVM advances a lot in dealing with the problems of nonlinear regression prediction (Vapnik, 2000), and it is useful in predicting landslide displacement (Liu et al., 2014). However, it is important to obtain optimal model parameters through an effective and intelligent method, because SVM is sensitive to the selection of the parameters (Kavzoglu et al., 2015b). The Particle Swarm Optimization (PSO) algorithm is a swarm intelligence-based global optimization method proposed by Kennedy and Eberhart, which is applicable to the selection and optimization of the model parameters and has been widely used (Vandenbergh and Engelbercht, 2006; Trelea, 2003; Kennedy and Eberhart, 1995).

In this paper, the typical step-like landslide (i.e. the Bazimen landslide) in the Three Gorges Reservoir area, China, is taken as an example. A time series model was adopted to divide the total displacement into the trend component determined by geological conditions and the periodic components controlled by influencing factors. The trend of displacement was predicted by the polynomial function. By analyzing the response relationship between the deformation characteristics and the factors, the PSO–SVM based on the factors of the reservoir level, the precipitation and the displacements of the prior-periods was proposed to predict the periodic displacement. The total forecast displacement was obtained by adding the trend and periodic displacements together.

2. Proposed displacement forecast model for landslide

2.1. Decomposition of the displacement time series

Landslide displacement shows approximately monotonic increasing in the long term under the influence of its geological condition, while it behaves fluctuant increasing in the short term because of the triggers. Based on the related researches (Du et al., 2013; Ren et al., 2015), in a time series model, the total displacement can be divided into two parts, expressed in Eq. (1):

$$x_t = s_t + v_t \tag{1}$$

where, x_t is the total displacement; s_t is the trend term; v_t is the periodic term.

2.2. Principles of the PSO–SVM model

2.2.1. Support Vector Machine

The SVM is a non-linear regression forecasting method proposed by Vapnik et al. in 1995. The input variables are mapped into a high-dimensional linear feature space through a non-linear transformation. Then the optimal decision function is constructed. The dot product operation in the higher dimensional feature space is replaced by the kernel function in original space, and the global optimal solution is obtained by the training of the finite sample. The regression function for SVM is:

$$f(x) = \langle W \cdot \Phi(x) \rangle + b \tag{2}$$

where, W is the weight vector, $\Phi(x)$ is nonlinear mapping from the input space to the output space, b is bias.

Transforming the estimation function into function minimization problem by the ε insensitive loss function:

$$\mathrm{R}_{\min} = \frac{1}{2}\|\mathrm{W}\|^2 + C\sum_{i=1}^{m}(\xi_i + \xi_i{}^*) \tag{3}$$

The constraint conditions are as follows:

$$\begin{cases} W^T\phi(x_i) + b_i - y_i \le \varepsilon + \xi_i; \\ y_i - W^T\phi(x_i) - b_i \le \varepsilon + \xi_i^*; \\ \xi_i, \xi_i^* \ge 0, i = 1, \ldots, l. \end{cases} \tag{4}$$

where C is penalty factor; ξ_i and ξ_i^* are relaxation factors. The Lagrange multiplier is introduced at last, and converted into an equivalent dual problem based on the Wolf duality theory:

$$\min\frac{1}{2}(\alpha-\alpha^*)^T Q(\alpha-\alpha^*) + \varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) \tag{5}$$

The constraint conditions are expressed in Eq. (6):

$$\begin{cases} \sum_{i=1}^{l}(\alpha_i + \alpha_i^*) = 0; \\ 0 \le \alpha_i, \alpha_i^* \le C, i = 1, 2 \ldots l. \end{cases} \tag{6}$$

where $Q_{ij} = k(x_i, x_j) = \phi(x_i)^T\phi(x_j)$.

The SVM regression prediction model can be obtained by quadratic programming:

$$f(x, \alpha_i^*, \alpha_i) = \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)K(x_i, x) + b \tag{7}$$

where $K(x_i, x)$ is the kernel function of SVM (Vapnik, 2000).

There are four commonly used kernel functions available: linear kernel, polynomial kernel, radial basis kernel function (RBF) and sigmoid function (Elbisy, 2015). With wide convergence domain, the RBF is widely used. Therefore, the RBF is adopted as kernel function for SVM model in this paper (Altınel et al., 2015; Farzan et al., 2015).

2.2.2. Particle Swarm Optimization

The PSO is an evolutionary algorithm developed in recent years (Vandenbergh and Engelbercht, 2006; Trelea, 2003; Eberhart and Shi, 2001). Inspired from the feeding behavior characteristic of a bird flock, it is used for solving optimization problem. Each particle in the algorithm indicates a potential solution to the problem. The common feature is represented by position, speed and fitness value. The fitness value can be calculated through the fitness function, which can estimate the merit of the particles. The velocity of the particle determines the movement direction and distance, and adjusts dynamically with the movement experience of all the particles, so as to optimize the individual in the search-space (Kennedy and Eberhart, 1995). Considering the sensitivity of SVM to the model parameters, the PSO method was proposed to establish the PSO–SVM prediction model in this paper.

2.3. Proposed model for landslide displacement prediction

Landslide is a multi-dimensional nonlinear dynamic system influenced by various factors such as geological conditions, reservoir water level and rainfall during evolutionary process (Eid, 2014). The long-term deformation trend is determined by the geological conditions, while the triggers control the short-term mutation (Glade et al., 2005). In the Three Gorges Reservoir area, under the influence of periodic rainfall and the scheduling of the reservoir, the deformations of some landslides increase suddenly. Then, the landslides become steady again when the influence disappear. Hence, the cumulative displacements often showed step-like growth from May to September every year (Fig. 1). Therefore, the response relationship between the displacement and the influencing factors is the key to predict the step-like landslide accurately (Li et al., 2010; Miao et al., 2014).

The adopted PSO–SVM model helps predicting the periodic displacement, and the trend term is predicted by polynomial function in this