Neural Comput & Applic
DOI 10.1007/s00521-013-1446-3

ICONIP 2012

Ensemble of extreme learning machine for landslide displacement prediction based on time series analysis

Cheng Lian · Zhigang Zeng · Wei Yao · Huiming Tang

Received: 25 March 2013 / Accepted: 20 June 2013

Abstract Landslide hazard is a complex nonlinear dynamical system with uncertainty. The evolution of landslide is influenced by many factors such as tectonic, rainfall and reservoir level fluctuation. Using a time series model, total accumulative displacement of landslide can be divided into the trend component displacement and the periodic component displacement according to the response relation between dynamic changes in landslide displacement and inducing factors. In this paper, a novel neural network technique called ensemble of extreme learning machine (E-ELM) is proposed to investigate the interactions of different inducing factors affecting the evolution of landslide. Grey relational analysis is used to sieve out the more influential inducing factors as the inputs in E-ELM. Trend component displacement and periodic component displacement are forecasted, respectively; then, total predictive displacement is obtained by adding the calculated predictive displacement value of each sub. Performances of our model are evaluated by using real data from Baishuihe landslide in the Three Gorges Reservoir of China, and it provides a good representation of the measured slide displacement behavior.

Keywords Extreme learning machine · Artificial neural networks · Ensemble · Grey relational analysis · Landslide · Displacement prediction

C. Lian (✉) · Z. Zeng
School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China
e-mail: liancheng@hust.edu.cn

C. Lian · Z. Zeng
Key Laboratory of Image Processing and Intelligent Control of Education Ministry of China, Wuhan 430074, China

W. Yao
School of Computer Science, South-Central University for Nationalities, Wuhan 430074, China

H. Tang
Faculty of Engineering, China University of Geosciences, Wuhan 430074, China

1 Introduction

Landslides are a recurrent problem throughout the Three Gorges Reservoir Area in China, which often cause significant damage to people and property. It is very important for us to improve the technology of landslides forecasting to prevent and reduce the loss caused by landslides. It is well known that landslide hazard is a complex nonlinear dynamical system with uncertainty, in which tectonic, rainfall and reservoir level fluctuation and so on all influence the evolution of landslide. These factors can be divided into the trigger and the primary cause [4, 13]. Studies on the interactions of the different factors affecting landslide occurrence are very important for the prediction of landslide. A time series decomposable model was proposed to establish the response relation between dynamic changes in landslide displacement and inducing factors [7, 22]. Total displacement of landslide can be divided into the trend component displacement and the periodic component displacement. The trend component displacement is determined by the potential energy and constraint condition of the slope. The periodic component displacement is affected by the periodic dynamic functioning of inducing factors such as rainfall, difference in temperature of day and night. Rainfall is one of the dominant exterior factors triggering the landslides, which has the lagged effect on the evolution of landslide. However, the optimized variables

Published online: 03 July 2013

associated with rainfall that played the significant role used in the past landslide models are mainly empirically chosen [7]. Grey relational analysis (GRA) is one of the most important contents of grey theory [5, 17]. It has been successfully applied in many fields such as management, economics and engineering [20, 24]. In this paper, GRA is used to expediently sieve the more influential factors associated with rainfall out.

In recent years, artificial neural networks (ANNs) have been widely applied in the area of landslide forecasting [3, 18]. To compare with logistic regression analysis, ANNs give a more optimistic evaluation of landslide susceptibility [18]. However, most ANN-based landslide forecasting methods used gradient-based learning algorithms such as back-propagation neural network (BPNN), which are relatively slow in learning and may easily get stuck in a local minimum [1, 6]. Recently, a novel learning algorithm for single-hidden-layer feedforward neural networks (SLFNs) called extreme learning machine (ELM) has been proposed [12, 25]. ELM not only learns much faster with higher generalization performance than the traditional gradient-based learning algorithms but also avoids many difficulties faced by gradient-based learning methods such as stopping criteria, learning rate, learning epochs and local minimum [9–11, 21]. Up to now, lots of applications involving ELM appeared in the pattern recognition and static function regression [15, 16, 19]. Although ELM has many advantages, a disadvantage is that its output is usually different from time to time because the input weights and biases are randomly chosen, which is also a common existing problem in ANNs that parameters are initialized randomly. So we do not know exactly on which time the initiation will give a good result. The idea of ELM ensembles has been proposed, which can significantly improved the generalization ability of learning systems through training a finite number of ELMs and then combining theirs results [14, 20]. The final output of E-ELM is the average of the outputs of each ELM network. A case study of Baishuihe landslide in the Three Gorges Reservoir Area is presented to illustrate the capability and merit of our model.

2 Methodology

The change in landslide displacement is determined by its own geological conditions and dynamic functioning of inducing factors. The displacement of landslide sequence is an instability time series. Based on the time series analysis, total displacement of landslide can be broken down into different corresponding components according to the different influential factors. Total displacement of landslide can be divided into trend component displacement, which is determined by the potential energy and constraint condition of the slope, and periodic component displacement, which is affected by the periodic dynamic functioning of inducing factors such as rainfall, difference in temperature of day and night. Trend component displacement nearly increases under large time scales, and periodic component displacement fluctuated increases under small time scales. Based on the above analysis, the displacement of landslide sequence can be described in terms of 4 basic classes of components: trend component, periodic component, impulse component and random component. It can be expressed as follows [7, 22]:

$$A_t = T_t + P_t + I_t + R_t, \quad t = 1, 2, \ldots, N \tag{1}$$

where A_t is the time series value (the accumulative change in landslide displacement) at time t. T_t is the trend component revealed the long-term trend of the sequence, which is determined by the potential energy and constraint condition of the slope. P_t is the periodic component, which is influenced by changes in the natural environment cycle like rainfall and difference in temperature of day and night. I_t is the impulse component that responds to the abrupt events like reservoir level fluctuation, and R_t is the random component, which reflects the impacts of random factors like earthquake. We can treat I_t as P_t when the impulse factors show characteristic of periodicity. In this paper, we present the prediction model without taking into account of influences of random component. Then, the model can be simplified as follows:

$$A_t = T_t + P_t, \quad t = 1, 2, \ldots, N \tag{2}$$

It has been found that SLFNs can approximate any continuous nonlinear function with arbitrary precision [8]. Based on this, E-ELM model is used to forecast the trend component displacement. For the forecasting of the periodic component displacement, grey relational analysis will sieve the more influential factors about rainfall out. Then, the input variables of the E-ELM are chosen though the combination of the GRA and the expert knowledge. The periodic component displacement is forecasted by E-ELM model. Total predictive displacement is obtained by adding the calculated predictive displacement value of each sub.

2.1 Double moving average method (DMAM)

Double moving average method can be used to separate the trend component displacement and periodic component displacement of landslide [23]. $a_t (t = 1, 2, \ldots, N)$ is the time series observed value, the total accumulative displacement of landslide. n is the number of step ($n < N$), so the period is $2n$. The trend component displacement can be extracted using the following equation:

$$M_t(1) = (a_t + a_{t-1} + \cdots + a_{t-n+1})/n$$
$$M_t(2) = [M_t(1) + M_{t-1}(1) + \cdots + M_{t-n+1}(1)]/n \tag{3}$$

where $M_t(1)$ and $M_t(2)$ represent a moving average and double moving average values, respectively. $M_t(2)$ is the trend component displacement extracted. The periodic component displacement is obtained by removing the trend component displacement from the total displacement of landslide. DMAM can eliminate some of the randomness in the data and leave a smooth trend-cycle component if the period can be selected accurately.

2.2 GRA

Grey relational analysis is proposed by Deng [5]. The influence degree of a compared series on the reference series is known as the grey relational grades (GRG), which describes the relative variations between one major factor and all other factors in a given system. If the relative variations between two factors are basically consistent during their development process, then the grey relational grade is regarded as large, or otherwise as small. The GRG between two sequences, the compared series and the reference series, is called relational coefficient $\xi_{0i}(k)$. Before calculating the grey relational coefficients, each data series must be normalized by dividing the respective data from the original series with their averages.

The grey relational coefficient $\xi_{0i}(k)$ of compared series $x_i(t)$ to reference series $x_0(t)$ at time $t = k$ is defined as follows:

$$\xi_{0i}(k) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|} \tag{4}$$

where $x_0 = (x_0(1), x_0(2), \ldots, x_0(n))$ and $x_i = (x_i(1), x_i(2), \ldots, x_i(n))$, $i = 1$ to m, are the reference sequences and compared sequences after normalization; $|x_0(k) - x_i(k)|$ denotes the absolute difference between two sequences; $\min_i \min_k |x_0(k) - x_i(k)|$ and $\max_i \max_k |x_0(k) - x_i(k)|$ denote the distance for each time in all compared sequences; ρ is a distinguishing coefficient ($0 < \rho \leq 1$) that is used to adjust the range of the comparison environment. According to the sensitivity analysis, the suggested value of ρ is 0.5, due to the moderate distinguishing effects and good stability of outcomes [2]. Then, the mean also is known as the GRG which is calculated as:

$$\xi_{0i} = \frac{1}{n} \sum_{k=1}^{n} \xi_{0i}(k), \quad i = 1, 2, \ldots, m \tag{5}$$

The purpose of GRA is to realize the relationship between two sets of time series data in relational space. If the data for two series at all individual time points are the same, then all the relational coefficients would equal one. The higher GRG compared with all factors are chosen to delegate the input data of the neural network model.

2.3 ELM

Extreme learning machine is a single-hidden-layer feedforward neural network with randomly generated hidden nodes independent of the training data. Input weights and biases can be randomly chosen, and the output weights can be analytically determined using the Moore−Penrose (MP) generalized inverse. Compared with traditional popular gradient-based learning algorithms for SLFNs, ELM not only learns much faster with higher generalization ability but also avoids many difficulties, such as the stopping criteria, learning rate, learning epochs and local minima.

For N distinct samples $(\mathbf{x}_i, \mathbf{t}_i)$, where $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^T \in \mathbf{R}^n$ and $\mathbf{t}_i = [t_{i1}, t_{i2}, \ldots, t_{im}]^T \in \mathbf{R}^m$, standard SLFNs with $\tilde{N}$ hidden neurons and activation function $g(x)$ are mathematically modeled as

$$\sum_{i=1}^{\tilde{N}} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{o}_j, \quad j = 1, \ldots, N. \tag{6}$$

where $\mathbf{w}_i = [w_{i1}, w_{i2}, \ldots, w_{in}]^T$ is the weight vector connecting the ith hidden neuron and the input neurons, $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]^T$ is the weight vector connecting the ith hidden neuron and the output neurons, $\mathbf{o}_j = [o_{j1}, o_{j2}, \ldots, o_{jm}]^T$ is the jth output vector of the SLFN and b_i is the threshold of the ith hidden neuron. $\mathbf{w}_i \cdot \mathbf{x}_j$ denotes the inner product of $\mathbf{w}_i$ and $\mathbf{x}_j$. The above N equations can be written compactly as:

$$\mathbf{H}\beta = \mathbf{O} \tag{7}$$

where

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_1 + b_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \quad \text{and} \quad \mathbf{O} = \begin{bmatrix} \mathbf{o}_1^T \\ \vdots \\ \mathbf{o}_N^T \end{bmatrix}_{N \times m}$$

where $\mathbf{H}$ is called the hidden layer output matrix of the neural network. The ith column of $\mathbf{H}$ is the ith hidden neuron's output vector with respect to inputs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$.

Extreme learning machine theories claim that the input weights $\mathbf{w}_i$ and hidden biases b_i can be randomly generated instead of tuned. To minimize the cost function $\|\mathbf{O} - \mathbf{T}\|$, where $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_N]^T$ is the target value matrix, the output weights is as simple as finding the least−square (LS) solution to the linear system $\mathbf{H}\beta = \mathbf{T}$, as follows:

$$\hat{\beta} = \mathbf{H}^{\dagger}\mathbf{T} \tag{8}$$

where $\mathbf{H}^{\dagger}$ is the MP generalized inverse of the matrix $\mathbf{H}$. The minimum norm LS solution is unique and has the smallest norm among all the LS solutions.

As a result of the input weights and biases are randomly chosen, the output of ELM is usually different from time to time. And we do not know exactly on which time the initiation will give a good result. In order to improve the generalization ability of ELM, E-ELM has been proposed. E-ELM consists of λ ELM networks with same structure, including the number of hidden nodes and same activation function. The final output of E-ELM is the average of the outputs of each ELM network. It is the same to repeat run the ELM for λ times with the same training data and calculate as follows:

$$\bar{y}_i = \frac{1}{\lambda}\sum_{j=1}^{\lambda} y_i^j, \quad j = 1, 2, \ldots, \lambda \tag{9}$$

where y_i^j is the output of jth ELM network with the input of $\mathbf{x}_i$. Obviously, the outputs obtained by E-ELM will become more stable when the parameter λ becomes larger, but the computation time also increases.

2.4 Normalization and unnormalization

The normalized method for the input and output data set is given as follows:

$$X_{\text{normalize}} = \frac{(X_{ij} - \max\{X_{ij}\}) + (X_{ij} - \min\{X_{ij}\})}{(\max\{X_{ij}\} - \min\{X_{ij}\})}, \tag{10}$$

$$i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, N$$

The unnormalized method for the input and output data set is given as follows:

$$P_{\text{unnormalize}} = \frac{P_{ij}(\max\{X_{ij}\} - \min\{X_{ij}\}) + \max\{X_{ij}\} + \min\{X_{ij}\}}{2} \tag{11}$$

$$i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, N$$

2.5 Steps of constructing the landslide displacement prediction model

Step 1 Separate the trend component displacement and the periodic component displacement.

(a) Extract the trend component displacement of landslide using (3).
(b) The periodic component displacement is obtained by removing the trend component displacement from the total displacement of landslide.

Step 2 Forecast the trend component displacement.

(a) The input and output data are divided into training data and predicting data.
(b) Normalize the training data using (10).
(c) Select the activation function and the number of hidden nodes of ELM based on the cross-validation method.
(d) Use the regression ability of ELM to obtain the predictive trend component displacement.
(e) Repeat Step 2(d) for λ times for the same data and then to compute the average predictive trend component displacement and its forecasting error.

Step 3 Forecast the periodic component displacement.

(a) Divide the input and output data into training data and predicting data.
(b) Calculate the GRG to select the factors about rainfall that have a significant effect on the periodic component displacement.
(c) Normalize the training data using (10).
(d) Select the activation function and the number of hidden nodes of ELM based on the cross-validation method.
(e) Use the regression ability of ELM to obtain the predictive further periodic component displacement.
(f) Repeat Step 3(e) for λ times for the same data and then to compute the average predictive periodic component displacement and its forecasting error.

Step 4 The predictive values of trend component displacement and periodic component displacement are superposed to obtain the total displacement prediction.

The scheme of the E-ELM is shown in Fig. 1.

3 Simulation studies

3.1 Date collection

Baishuihe landslide is located on the south bank of Yantze River and its 56 km away from the Three Gorges Dam. The bedrock geology of the study area consists mainly of sandstone and mudstone, which is an easy slip stratum. The slope is of the category of bedding slopes. Figure 2 shows the monitoring data of landslide accumulative displacement at ZG118 monitoring point, and Fig. 3 shows the monitoring data of rainfall and reservoir level elevation. The total number of data is 38 groups from June 2004 to July 2007. The data between June 2004 to December 2006

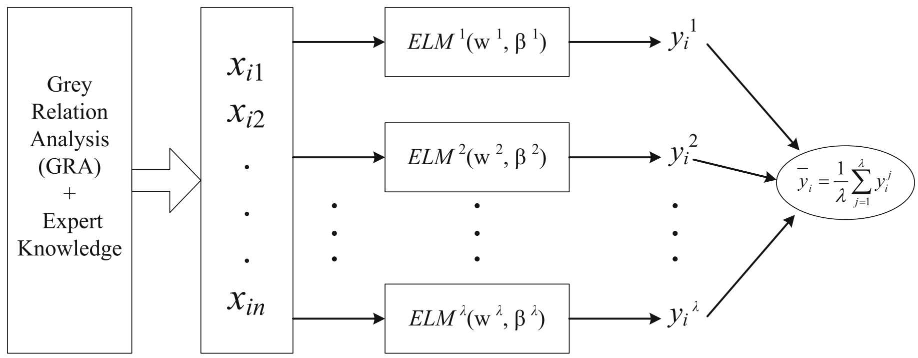

Fig. 1 The scheme of the E-ELM integration system

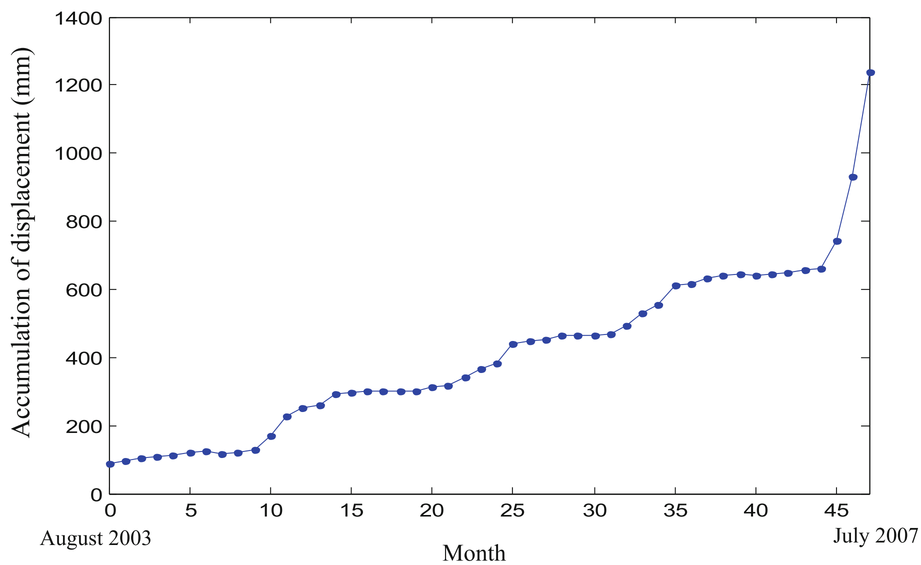

Fig. 2 Monitoring curves of landslide accumulative displacement

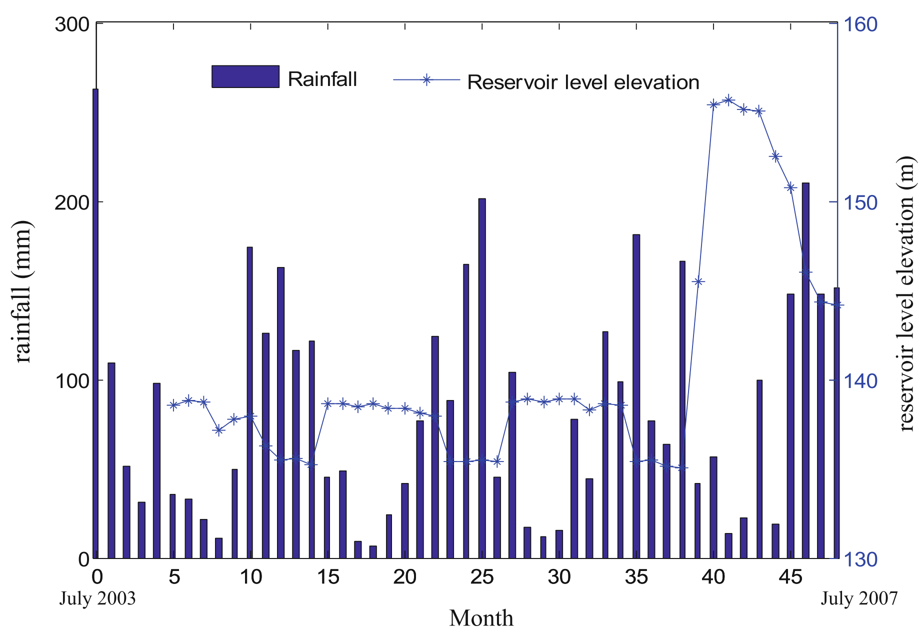

Fig. 3 Monitoring curves of rainfall and reservoir level elevation

were selected as training data in order to construct the forecasting model, and the rest of 7 groups of data from January 2007 to July 2007 were selected as predicting data.

DMAM is used to separate the trend component displacement and periodic component displacement of landslide. We choose $n = 6$, so the period is 12 months. $M_t(2)$

is the trend component displacement extracted which is calculated using Eq. (2). Then, the periodic component displacement can be obtained by removing the trend component displacement from the total displacement of landslide. The results are as shown in Fig. 4.

3.2 Analysis and prediction of trend component displacement

Trend component displacement is determined by the potential energy and constraint condition of the slope, which nearly increases under large time scales. ELM can approximate any continuous function which is used to forecast the trend term displacement. The activation function of E-ELM is the sigmoidal function $g(x) = 1/(1 + e^{-x})$. The number of hidden nodes is 3. Considering both computation time and the stability of E-ELM, the parameter λ is selected 1,000. The predicted values of trend component displacement are shown in Table 1.

As shown in Table 1, E-ELM model shows a good extrapolation capability. The predictive values and measurement values are very close for every calculation, and the relative error all falls into 5 %, the predicting precision is high enough which can reveal the long-term trend of the evolution of landslide.

3.3 Analysis and prediction of periodic component displacement

As we know, the periodic component displacement of landslide will be affected by many factors. Baishuihe landslide is located at the Three Gorges Reservoir Area; based on the expert knowledge, the periodic component is mainly affected by rainfall, and the impulse component is mainly affected by reservoir level fluctuation. Since the reservoir level adjustment shows characteristic of periodicity with one year cycle in Three Gorges Reservoir Area, we can treat reservoir level fluctuation as the periodic component

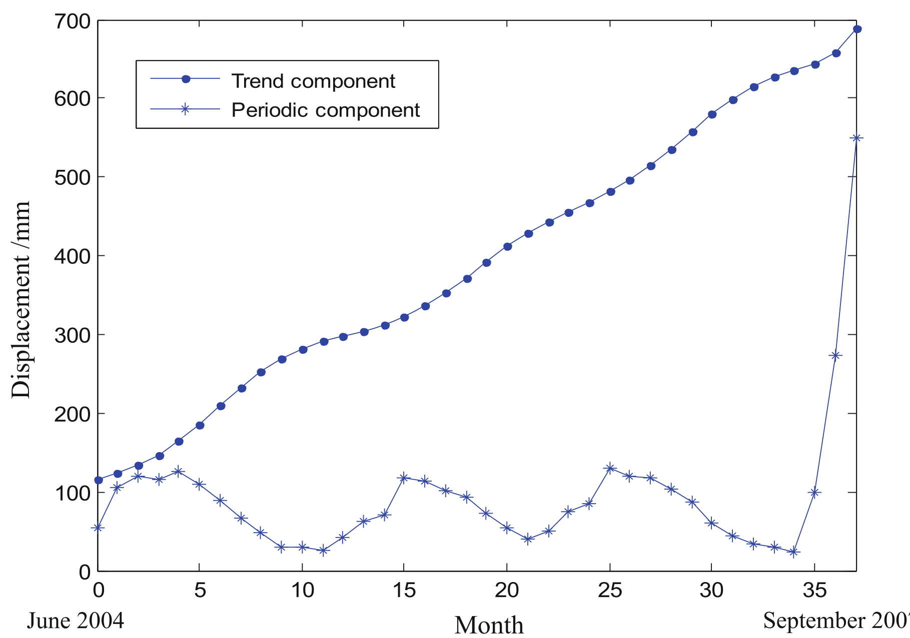

Fig. 4 Extracted values of trend component displacements and periodic component displacements

Table 1 Trend component displacement comparison between predicted values and measured values

Time	Measured value (mm)	Predicted value (mm)	Absolute error (mm)	Relative error (%)
01/01/2007	599.6	580.6	19.0	3.17
02/01/2007	615.5	596.2	19.3	3.14
03/01/2007	627.6	611.7	15.9	2.53
04/01/2007	636.3	627.3	9.0	1.41
05/01/2007	644.6	642.8	1.8	0.28
06/01/2007	658.6	658.4	0.2	0.03
07/01/2007	688.1	673.9	14.2	2.06

inducing factors. As a result, rainfall has the lagged effect on the evolution of landslide. To select the most influential indices about rainfall by using the GRA method to be the input data of the E-ELM model, rainfall time series is processed into 6 kinds, and the GRG of each factor is shown in Table 2. Where R_1 represents rainfall of current month, R_2 represents rainfall of last month, R_3 represents rainfall of the month before last, R_4 represents cumulative of rainfall of current month and last month, R_5 represents cumulative of rainfall anterior two month, R_6 represents cumulative of rainfall of current month and anterior two month. The factors about rainfall with higher GRG are shown in bold in Table 2. From the result of GRA, two selected factors about rainfall with higher GRG are R_5 and R_6. So here are three inputs of ELM: R_5, R_6 and reservoir level fluctuation. The reservoir level fluctuation is calculated as the difference between the water level at this month and last month.

All the data sets should be normalized into the range of $[-1, 1]$. The activation function of E-ELM is the sigmoidal function $g(x) = 1/(1 + e^{-x})$. The number of hidden nodes is 18. Considering both computation time and the stability of E-ELM, the parameter λ is selected 1,000. The predicted values of periodic component displacement are shown in Table 3.

Table 2 The GRG of factors about rainfall

Data	R_1	R_2	R_3	R_4	R_5	R_6
GRG	0.5902	0.6557	0.6929	0.6718	**0.7442**	**0.7121**

Rainfall and reservoir level fluctuation are the major factors that affect the stability of landslide, but there are many other factors such as temperature difference between day and night and other random factors such as human project activities. So the predictive values in February 2007 and March 2007 are not very precise, but that lack of precision may not matter in engineering. The predictive values successfully reflect the evolution tendency of landslide from January 2007 to June 2007; especially the model successfully predicts the obvious deformation from April 2007 to June 2007 which is able to provide early warnings. The landslide collapsed in July 2007. Once the collapse happens, the displacement will increase exponentially without constrain. Then, the landslide is at the unsteady state. Obviously, without any training samples belong to the same balancing system, neural networks model is not suited to forecast the evolution of displacement in this condition. Actually, the impact factors of neural network model are also changed in this condition.

3.4 Analysis and prediction of the total displacement of landslide

The total displacement prediction is obtained by adding the predictive values of trend component displacement and periodic component displacement. The predicted values are shown in Table 4. In addition, the traditional E-ELM is also used to predict landslide displacement for comparison purposes, the activation function is selected the sigmoidal

Table 3 Periodic component displacement comparison between predicted values and measured values

Time	Measured value (mm)	Predicted value (mm)	Absolute error (mm)	Relative error (%)
01/01/2007	44.6	48.2	3.6	8.07
02/01/2007	32.8	55.9	23.1	70.43
03/01/2007	29.0	52.7	23.7	81.72
04/01/2007	24.5	30.0	5.5	22.45
05/01/2007	99.2	101.5	2.3	2.32
06/01/2007	272.3	216.0	56.3	20.68
07/01/2007	550.3	122.4	427.9	77.76

Table 4 Total displacement of landslide comparison between predicted values and measured values

Time	Measured value (mm)	E-ELM based on the time series analysis		Traditional E-ELM	
		Predicted value (mm)	Relative error (%)	Predicted value (mm)	Relative error (%)
01/01/2007	644.2	628.8	2.39	625.3	2.93
02/01/2007	648.3	652.1	0.59	632.1	2.50
03/01/2007	656.6	664.4	1.19	659.7	0.47
04/01/2007	660.8	657.3	0.53	645.8	2.27
05/01/2007	743.8	744.3	0.07	690.6	7.15
06/01/2007	930.9	874.4	6.07	791.4	14.99
07/01/2007	1238.4	796.3	35.70	967.5	21.88

function, the parameter λ is selected 1,000, and the number of hidden nodes is selected 6 by trial and error.

As shown in Table 4, E-ELM based on the time series analysis method shows a better prediction precision than the traditional E-ELM. The predictive values and measurement values are very close for every calculation except in July 2007. The simulation results illustrate the proposed method has well extrapolated ability for prediction, and successfully predicts the obvious deformation from April 2007 to June 2007. On the basis of the mentioned analysis in Sect. 3.3, the total displacement of landslide is unpredictable in July 2007 using our model. But the forecasting ability of our model is good enough to provide early warnings.

4 Conclusion

Based on the time series analysis, total accumulative displacement of landslide can be divided into the trend component displacement and the periodic component displacement according to the response relation between dynamic changes in landslide displacement and inducing factors. Trend component displacement and periodic component displacement are forecasted, respectively; then, total predictive displacement is obtained by adding the calculated predictive displacement value of each sub. A relatively novel neural network technique, E-ELM, is used to analyze and forecast the trend component displacement and the periodic component displacement. A good prediction of the periodic component displacement is essential to ensure an accurately predicted total displacement. So the inputs of E-ELM with respect to the periodic component are selected based on the combination of GRA and expert knowledge. Based on the expert knowledge, rainfall and reservoir level fluctuation are the major factors that affect the stability of landslides in the Three Gorges Reservoir Area, and GRA method is used to expediently sieve the more influential factors associated with rainfall out.

The application shows that our method can achieve a good prediction result. In particular, our model successfully predicts the obvious deformation of landslide, which is able to provide early warnings. Therefore, this method has a good perspective in application and further development. It must be pointed out that landslide hazard has its own characteristics which varied with geological environment; landslide forecast model should be established according to the concrete types of landslides. Actually, expert judgment still should be taken into account in practical applications.

Acknowledgments The work is supported by the Natural Science Foundation of China under Grant 61125303 and 61203286, the 973 Program of China under Grant 2011CB710606, the Specialized Research Fund for the Doctoral Program of Higher Education of China under Grant 20100142110021.

References

1. Amro EJ, John M (1990) A new error criterion for posterior probability estimation with neural nets. In: Proceedings of international joint conference on neural networks, pp 185–192
2. Chang TC, Lin SJ (1999) Grey relation analysis of carbon dioxide emissions from industrial production and energy uses in Taiwan. J Environ Manage 56(4):247–257
3. Chen HQ, Zeng ZG (2013) Deformation prediction of landslide based on improved back-propagation neural network. Cogn Comput 5(1):56–62
4. Cubito A, Ferrara V, Pappalardo G (2005) Landslide hazard in the Nebrodi Mountains (Northeastern Sicily). Geomorphology 66(1–4):359–372
5. Deng JL (1982) Control problems of gray systems. Syst Control Lett 1(4):288–294
6. Drucker H, Cun LY (1992) Improving generalization performance using double backpropagation. IEEE Trans Neural Netw 3(6):991–997
7. Du J, Yin KL, Chai B (2009) Study of displacement prediction model of landslide based on responses analysis of inducing factors. Chin J Rock Mech Eng 28(9):1783–1789
8. Hornik K (1991) Approximation capabilities of multilayer feedforward networks. Neural Netw 4(2):251–257
9. Huang GB (2003) Learning capability and storage capacity of two-hidden-layer feedforward networks. IEEE Trans Neural Netw 14(2):274–281
10. Huang GB, Babri HA (1998) Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions. IEEE Trans Neural Netw 9(1):224–229
11. Huang GB, Chen L, Siew CK (2006) Universal approximation using incremental constructive feedforward networks with random hidden nodes. IEEE Trans Neural Netw 17(4):879–892
12. Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. Neurocomputing 70(1–3):489–501
13. Kawabata D, Bandibas J (2009) Landslide susceptibility mapping using geological data, a DEM from ASTER images and an Artificial Neural Network (ANN). Geomorphology 113(1): 97–109
14. Lan Y, Chai Y, Huang GB (2009) Ensemble of online sequential extreme learning machine. Neurocomputing 72(13–15): 3391–3395
15. Lan Y, Hu ZJ, Soh YC, Huang GB (2013) An extreme learning machine approach for speaker recognition. Neural Comput Appl 22(3–4):417–425
16. Li GQ, Niu PF (2013) An enhanced extreme learning machine based on ridge regression for regression. Neural Comput Appl 22(3–4):803–810
17. Liu SF, Dang YG, Fang ZG et al (2004) Grey system theory and its application, (third version), Science Press, Beijing
18. Nefeslioglu HA, Gokceoglu C, Sonmez H (2008) An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. Eng Geol 97:171–191
19. Nian R, He B, Lendasse A (2013) 3D object recognition based on a geometrical topology model and extreme learning machine. Neural Comput Appl 22(3–4):427–433. http://link.springer.com/article/10.1007%2Fs00521-012-0892-7
20. Sun ZL, Choi TM, Au KF, Yu Y (2008) Sales forecasting using extreme learning machine with applications in fashion retailing. Decis Support Syst 46(1):411–419
21. Tamura S, Tateishi M (1997) Capabilities of a four-layered feedforward neural network: four layers versus three. IEEE Trans Neural Netw 8(2):251–255

22. Wang JF (2003) Quantitative prediction of landslide using S-curve. Chin J Rock Mech Eng 14(2):1–8
23. Xu F, Wang Y, Du J, Ye J (2011) Study of displacement prediction model of landslide based on time series analysis. Chin J Rock Mech Eng 30(4):746–751
24. Yin XG, Yu WD (2007) The virtual manufacturing model of the worsted yarn based on artificial neural networks and grey theory. Appl Math Comput 185(1):322–332
25. Zhu QY, Qin AK, Suganthan PN, Huang GB (2005) Evolutionary extreme learning machine. Pattern Recogn 38(10):1759–1763