SCIENTIFIC REPORTS
nature research

OPEN

Forecasting of landslide displacements using a chaos theory based wavelet analysis-Volterra filter model

Yuanyao Li[1*], Ronglin Sun[2], Kunlong Yin[1], Yong Xu[3], Bo Chai[2] & Lili Xiao[4]

Landslide displacement time series can directly reflects landslide deformation and stability characteristics. Hence, forecasting of the non-linear and non-stationary displacement time series is necessary and significant for early warning of landslide failure. Traditionally, conventional machine learning methods are adopted as forecasting models, these forecasting models mainly determine the input and output variables experientially and does not address the non-stationary characteristics of displacement time series. However, it is difficult for these conventional machine learning methods to obtain appropriate input-output variables, to determine appropriate model parameters and to acquire satisfied prediction performance. To deal with these drawbacks, this study proposes the wavelet analysis (WA) to decompose the displacement time series into low- and high-frequency components to address the non-stationary characteristics; then proposes thee chaos theory to obtain appropriate input-output variables of forecasting models, and finally proposes Volterra filter model to construct the forecasting model. The GPS monitoring cumulative displacement time series, recorded on the Shuping and Baijiabao landslides, distance measuring equipment monitoring displacements on the Xintan landslide in Three Gorges Reservoir area of China, are used as test data of the proposed chaotic WA-Volterra model. The chaotic WA-support vector machine (SVM) model and single chaotic Volterra model without WA method, are used as comparisons. The results show that there are chaos characteristics in the GPS monitoring displacement time series, the non-stationary characteristics of landslide displacements are captured well by the WA method, and the model input-output variables are selected suitably using chaos theory. Furthermore, the chaotic WA-Volterra model has higher prediction accuracy than the chaotic WA-SVM and single chaotic Volterra models.

The safety of local people' life and property are threatened seriously by the reservoir landslides distributing along the Three Gorges Reservoir[1–3]. Forecasting of Landslide displacements is considered as an important part of an operational early warning system[2,4–7]. Hence, it is crucial and significant to forecast the landslide displacements for early warning[8].

Over the past decades, Global Position System (GPS) has been widely and successfully used to monitor the landslide displacement time series[9]. GPS technology makes it possible to real-timely track the deformation processes of a landslide. The landslide displacement prediction models mainly include physically based models and data-based models[6,10]. The data-based models are established through training and testing the input-output variables using linear and/or non-linear models. The modeling processes of data-based models are more simple and accurate than those of physically based models[6,11]. Hence, this study propose data-based models to forecast the landslide displacement time series.

Some data-based models namely the classical black box models (Auto Regressive, Moving Average, Multiple Linear Regression, *et al.*) have been proposed for time series prediction since 1970[12]. These classical black box models are not accurate enough for nonlinear time series prediction because they are linear models. In the past three decades, many Artificial Intelligence (AI) methods, which provide a way to solve the problems of

[1]Institute of Geological survey, China University of Geosciences, Wuhan, 430074, China. [2]School of Environmental Studies, China University of Geosciences, Wuhan, 430074, China. [3]Wuhan Center, China Geological Survey, Wuhan, 430205, China. [4]School of Highway, Chang'an University, Xi'an, 710064, China. *email: yuanyaoli2007@126.com

complexity, dynamism and nonlinear characteristics in nonlinear time series, had been used to predict landslide displacement time series. These models include artificial neural networks (ANNs)[13–16], extreme learning machine[17], fuzzy logic approach[18] and support vector machine (SVM)[19–22], *et al.*

However, the nonlinear time series prediction performances of these AI methods are limited. This is because that, (1) the nonlinear and non-stationary characteristics of landslide displacement time series are not captured fully by single AI models; (2) the input-output variables of the AI methods are determined empirically without appropriate determination law for input-outputs; and (3) the nonlinear fitting and predicting abilities of these AI models should be improved effectively[4,11,19,23,24]. In additional, related literature indicates that AI models have some disadvantages of local optimum, slow training and testing rate and over-fitting problem, which also hinder the prediction accuracy for the nonlinear time series[2,4]. To address these three limitations existed in AI models, this study proposes a chaos theory based wavelet analysis-Volterra filter model (chaotic WA-Volterra model) for displacements prediction.

For the first limitation, wavelet analysis (WA) is used to address the serious nonlinear and non-stationary problems in the landslide displacement. WA is an effective signal processing tool that can solve the nonlinear and non-stationary signal and offer the time-scale localization for signals[25]. The main features of WA method are exploring the time series from both time domain and frequency domain, which provide meaningful information for the physical structure of time series[15]. Hence, WA based AI models firstly decompose a time series into several multi-resolution frequency components and then these components are respectively predicted in the AI models with higher prediction performance[13,14]. Recently, WA-AI models have been effectively introduced to the areas of reservoir fluid contacts prediction[26], stream flow data series prediction[14] and seasonal variation of landslide displacement[27]. This study also introduces WA into AI methods for non-linear and non-stationary cumulative landslide displacements prediction.

In addition, the prediction performance of AI method is essentially depend on the input-output variables. In this study, the chaos theory, which can track the evaluation of nonlinear time series and rebuild its original evaluation system, is used to determine the input-output variables of the AI models. Chaos theory is an important method for input-output variables selection which is widely used in many areas[6,28]. In theory, based on the determination of the chaos evidence of landslide displacements, embedding theory and phase space reconstruction (*PSR*) methods of chaos theory can be used to build a chaos theory based model for displacements prediction. Based on embedding theory[29], the landslide displacements might to be predicted using a single variable displacement time series. Meanwhile, on the basis of *PSR* method[30], a single variable displacement time series can be constructed into a multi-dimensional phase-space to obtain the input-output variables of prediction models.

Furthermore, after the conductions of chaos theory and WA method on landslide displacements, the Volterra filter model is innovatively introduced into this study[31]. The Volterra filter model can predict the chaotic time series using a few time series data in the training and testing processes, and can automatically track the motion trace of the chaotic time series. Hence, the Volterra filter model can overcome the disadvantages of traditional machine learning models, and it has been successfully used in many areas with excellent generalization performance, such as air/fuel ratio prediction[32], multi-scale stream flow forecasting[33] and traffic flow prediction[34]. However, no attention has been attracted to use the Volterra filter model to predict landslide displacements, although the Volterra filter model has efficient prediction performance.

To summarize, a chaotic WA-Volterra model is innovatively proposed in this study to overcome the drawbacks of traditional machine learning models. Meanwhile, the chaotic WA-SVM model and the single chaotic Volterra filter model without WA method are used for comparisons. The GPS monitoring monthly displacement time series of the Shuping landslide and Baijiabao landslide, distance measuring equipment monitoring displacements on the Xintan landslide in Three Gorges Reservoir area (TGRA) are used as case study.

Study Area and Materials

The locations of the Shuping, Baijiabao and Xintan landslides are shown in Fig. 1; Shuping landslide locates on the south side of the Yangtze river in ZiGui County. The Baijiabao landslide is located on the west side of Xiangxi river, a tributary of the Yangtze river. Xintan landslide locates on the north side of the Yangtze river in ZiGui County.

Geological conditions and deformation characteristics of the shuping landslide. The topographic map of the research site and its GPS monitoring network are shown in Fig. 2. Figure 2 shows that the upper boundary of the chair-like shaped Shuping landslide is defined by the bedrock-soil interface. The maximum elevation of the upper boundary is 415 m. Its toe elevation is approximately 144 m. The longitudinal dimension and width of the landslide are 800 m and 700 m, respectively. In addition, the mean depth of the sliding surface is approximately 55 m.

The materials of the landslide are composed of medium permeable quaternary deposits. The slope mass structure is loose and it varies from 10° to 35°. The dip direction and dip angle are approximately 143° and 15°, respectively. The slip zone is composed of silty clay and fragmented rubble. The bedrock lithology is characterized by triassic formation of middle Badong formation.

Shuping landslide was activated in June 2003 and it exhibited a large amount of local deformation and failure. Some reservoir bank collapses occurred in the frontal part of this landslide from October 2003 to January 2004. Surface investigations showed that shearing and crush-pressing cracks had occurred since January 2004. The deformation of the study area had rapidly progressed since April 2007. Figure 3(a) illustrates a small debris flow that occurred. Figure 3(b) shows a large amount of significant road deformation and cracking that occurred in the road at the upper part of the landslide. The deformation and failure characteristics of Shuping landslide are serious.

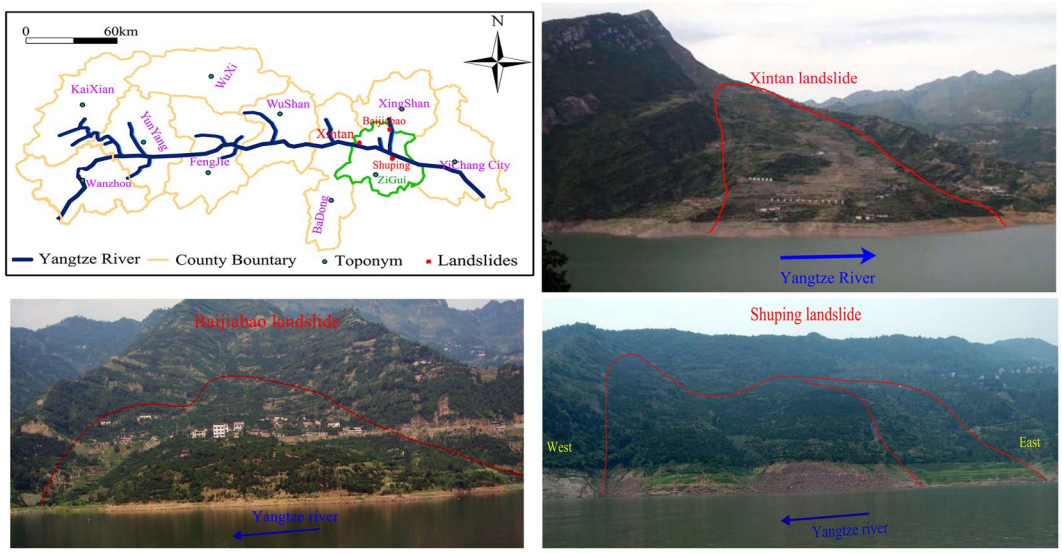

Figure 1. Geographic location information of Shuping and Baijiabao landslides (Adobe Photoshop CS3, http://www.adobe.com/cn/products/photoshop.html, drawn by Yuanyao Li; the photograph is taken from Yuanyao Li).

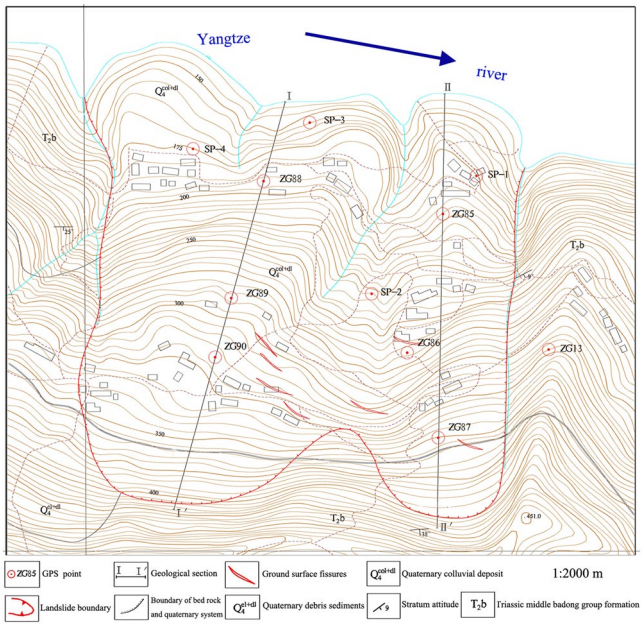

Figure 2. Topographical map of Shuping landslide, with locations of GPS points (AUTOCAD 2014, https://www.autodesk.com.cn/, drawn by Yuanyao Li).

In general, the deformation and failure of reservoir landslides are related to the rapid changes in the groundwater seepage field[35–37]. The mechanical effects of hydraulic uplift pressure and reverse seepage pressure on the landslide will increase when the groundwater level rises[38]. Meanwhile, the physical effects of groundwater seepage on landslides are mainly reflected in the reduction of soil shear strength and in the softening of sliding surface[39,40]. However, it is not easy to obtain sufficient groundwater level time series for building prediction models. Hence, this study does not put groundwater levels into the chaotic WA-Volterra model for landslide displacements prediction.

Geology and deformation characteristics of baijiabao landslide. Baijiabao landslide (Fig. 4), which is a fan shape and destructive landslide, is composed by the quaternary deposits (silty clay interspersed with fragmented stone). The structure of these quaternary deposits is loose and porous. The bedrock under this sliding mass is composed by Jurassic mudstone and sandstone. In addition, the Baijiabao landslide is a large landslide

Figure 3. Deformation characteristics of Shuping landslide (photographs taken by Yuanyao Li).

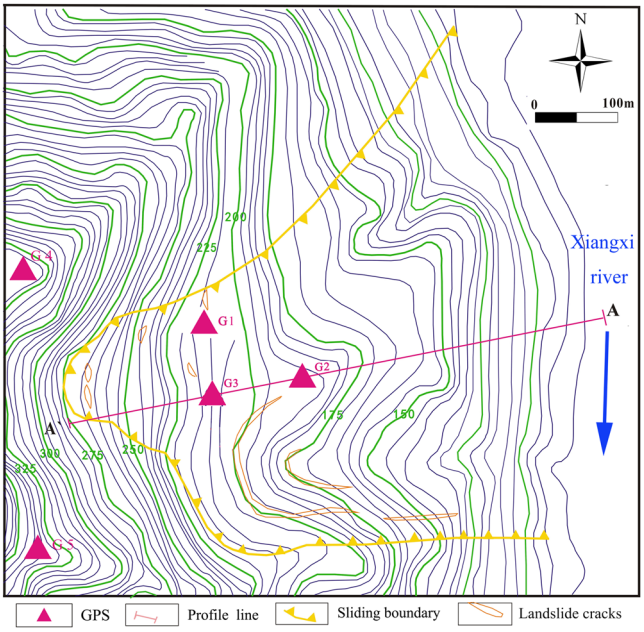

Figure 4. Terrain map of the Baijiabao landslide with GPS locations (AUTOCAD 2014, https://www.autodesk.com.cn/, drawn by Yuanyao Li).

with an approximately volume of $9.9 \times 10^6\,m^3$, an approximately area of $2.2 \times 10^5\,m^2$ and an approximately sliding depth of 45 m. The average slope of Baijiabao landslide is about 12°, the elevation of the frontal part of the landslide is about 125 m (extending to the bed of Xiangxi River) and the elevation of the upper part is about 275 m. Figure 4 shows that the boundaries of Baijiabao landslide are identified as gully and bed rock.

Geology and deformation characteristics of baijiabao landslide. The Xintan landslide (Fig. 1) is an accumulate landslide composed by silty clay and sandstone fragments. Its elevation ranges between 70 m and 500 m from frontal part to the upper part of the landslide with thickness of accumulative layer about 25.5 m. The Xintan landslide is generally inclined towards to the Yangtze river with average slope of 25° and a multilevel terraced terrain. The right and left landslide boundaries are characterized by steep cliffs. The bedrock lithology of Xintan landslide is mainly featured as Silurian sandstone and shale. Its groundwater is replenished by rainfall. There is almost no groundwater in this landslide because the slide mass is loose and the discharge of the groundwater is very good.

There are several reasons for the landslide instability. One reason is that, the surface slope of this landslide is steep, the accumulative layer of this landslide is loose and there is weak intercalated layer in this landslide. Another reason is that, the seasonal heavy rainfall decreases the shear strength of this landslide, as a result, the landslide is reactivated. The deformation and instability processes of Xintan landslide began in January 1977, and completely destroyed in May 1985. The local people's life and properties are seriously threatened by the deformation of Xintan landslide.

GPS system building on shuping and baijiabao landslides. To monitor the displacement time series of the Shuping and Baijiabao landslides, the GPS system was respectively built on the Shuping landslide in June

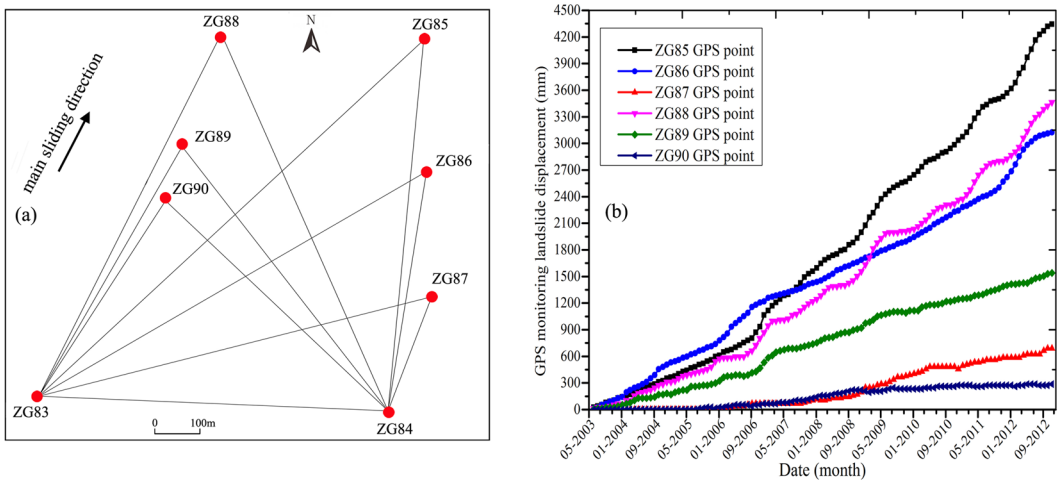

Figure 5. GPS system on Shuping landslide and obtained cumulative displacements with ZG83 ~ ZG90 GPS points (Drawn by Yuanyao Li).

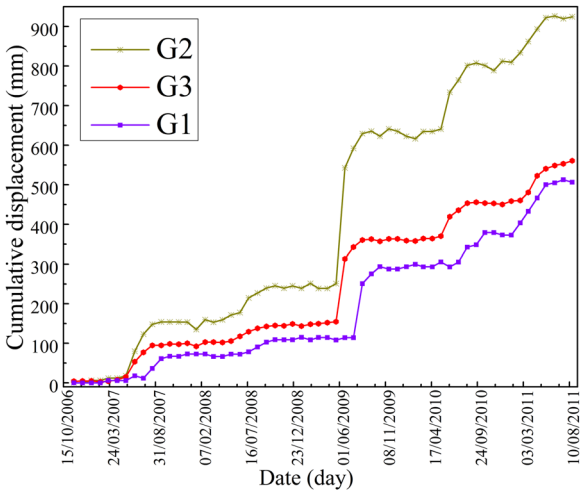

Figure 6. GPS monitoring cumulative displacements of Baijiabao landslide with G1 ~ G3 GPS points (Drawn by Yuanyao Li).

2003 and built on the Baijiabao landslide in September 2006. Figure 2 shows that the GPS points ZG85 ~ ZG90 were set as observation points, while the points ZG83 and ZG84 were set as reference points. Two GPS reference points were placed in the stable zones outside the Shuping landslide to ensure that there were no deformations in the reference points. Six GPS observation points were built in the deformation zone of the landslide. The same as Shuping landslide, Baijiabao landslide was also monitored by the GPS system with G4 ~ G5 as reference points and G1 ~ G3 as observation points (Fig. 4).

The reference stations respectively constructed a landslide monitoring control network with each GPS observation points ZG85 ~ ZG90 (Fig. 5(a)) and G1 ~ G3 in turn every month. Then, the displacements of GPS observation points are calculated comparing to the reference points based on the constructed landslide monitoring control network. The GPS receiving signals were processed by baseline processing and network adjustment in the GAMIT/GLOBK software[41]. Then, the monitored cumulative displacements can be obtained and displayed. In this study, two series of GPS monitoring landslide displacement time series were obtained from June 2003 to November 2012 on Shuping landslide (Fig. 5(b)) and from October 2006 to August 2011 on Baijiabao landslide (Fig. 6).

GPS monitoring displacements of shuping landslide. Figure 5(b) indicates that the cumulative displacements of the Shuping landslide are of non-linear and non-stationary characteristics. Figure 5(b) also shows that the cumulative displacement values of points ZG85, ZG86 and ZG88 on the frontal part of the landslide are greater than those of the points on the upper part, which suggests that the Shuping landslide is a translational landslide with earth sliding[42]. In this study, the cumulative displacements of ZG85 ~ ZG88 points with non-linear and non-stationary characteristics are predicted using the proposed models.

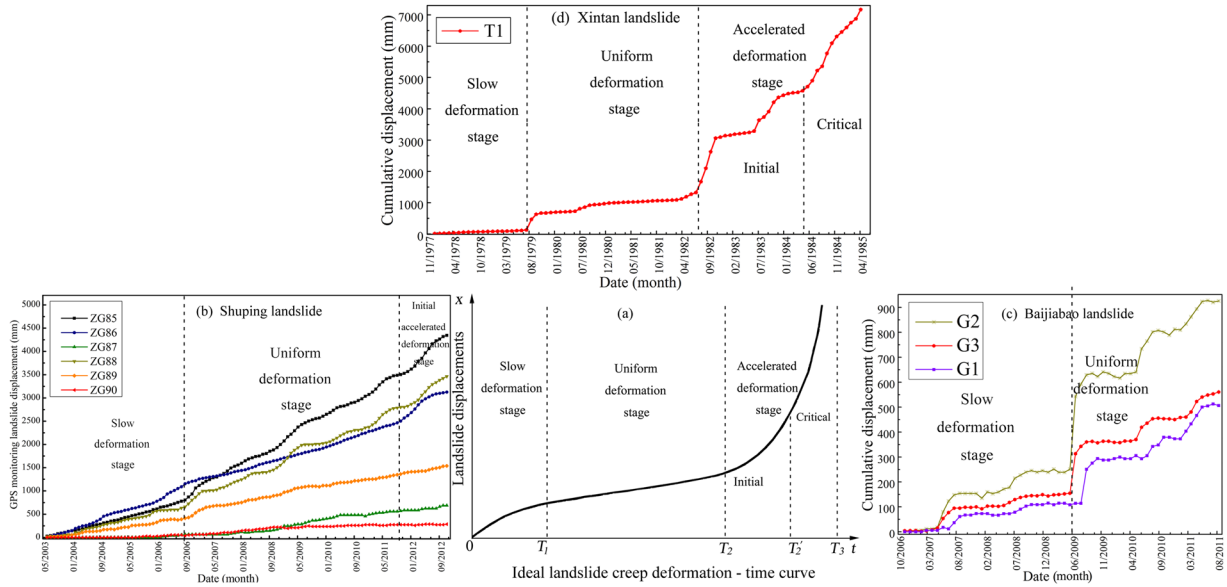

Figure 7. landslide ideal creep deformation-time curve (**a**), and deformation characteristics analysis of Shuping landslide (**b**) and Baijiabao landslide (**c**).

GPS monitoring displacements of baijiabao landslide. The Baijiabao landslide has been activated by the water level fluctuation of Three Gorges Reservoir and the seasonal rainfall since September, 2006. As a result, some shear and tensile fissures occurred in the middle and upper parts of this landslide (Fig. 4). Three GPS sensors have been placed on the landslide for landslide displacements monitoring since September, 2006 (Fig. 6). The nonlinear and non-stationary deformation features of Baijiabao landslide are reflected by these three displacement curves. Furthermore, in order to verify the prediction performance of WA-Volterra model for landslide displacement prediction, the G3 GPS point is used in this study as study case.

Creep deformation characteristics of the three landslides. Related literature shows that, the nature of soil landslide failure can be regarded as the creep deformation process of sliding mass[43,44]. In general, according to the landslide deformation rate and cumulative displacements, the ideal landslide creep deformation-time curve can be divided into three stages[44]: slow deformation stage with small deformation rate and low slope of cumulative displacements ($0 \sim T_1$), uniform deformation stage with almost uniform deformation rate and gradually increasing cumulative displacements ($T_1 \sim T_2$), and accelerated deformation stage with high deformation rate and rapidly increasing cumulative displacements until landslide failure occurs ($T_2 \sim T_3$) (Fig. 7(a)). The last stage can be further divided into initial accelerated deformation and critical failure stages.

The landslide deformation falling into the accelerated deformation stage is a very important warning indicator of landslide failure[45]. Therefore, it is significant to analysis the deformation features according to the creep deformation-time curves. Specially, The cumulative displacements of Shuping and Baijiabao landslides belong to non-ideal creep deformation-time curves. Instead, the cumulative displacements of Xintan landslide belongs to ideal creep deformation-time curve. The deformation of Shuping landslide is a type of continuous and almost uniform creep deformation feature, with slow deformation stage from June 2003 to September 2006, uniform deformation stage from October 2006 to October 2011 and initial accelerated deformation stage from November 2011 to January 2013 (Fig. 7(b)). Meanwhile, the deformation of Baijiabao landslide has a type of seasonal step-like creep deformation feature, with slow deformation stage from October 2006 to June 2009 and uniform deformation stage from July 2009 to August 2011(Fig. 7(c)). In addition, the Xintan landslide has also a seasonal step-like creep deformation feature, with slow deformation stage from January 1978 to August 1979, with uniform deformation stage from September 1979 to July 1982, and then with critical accelerated deformation stage from May 1984 until this landslide was completely destroyed(Fig. 7(d)).

Results

Chaos evidences identification of monthly cumulative displacements. Related literature shows that the chaos theory based models have been widely studied in many areas of non-linear time series prediction, such as water level prediction[35] and electronic power prediction[46]. In this study, the phase spaces of GPS monitoring displacement time series are reconstructed firstly, then their chaos evidences are identified using the LLE and *CD* methods.

Reconstruct the phase spaces of cumulative displacements. The phase spaces of ZG85 (x_i^{ZG85}) ~ ZG88 (x_i^{ZG88})($i = 1, 2, \ldots\ldots, 114$) of Shuping landslide, G3 (x_i^{G3}) ($i = 1, 2, \ldots\ldots, 57$)of Baijiabao landslide, T1 (x_i^{T1}) ($i = 1, 2, \ldots\ldots, 89$) of Baijiabao landslide, cumulative displacement time series are reconstructed. The τ of ZG85 ~ ZG88, G3 and T1 displacements are set to 1 and the optimal m values of ZG85 ~ ZG88, G3 and T1 displacements are respectively

Displacements	Reconstructed phase spaces (input variables)	Output variables
ZG85	$X_i^{ZG85} = \left(x_i^{ZG85}, x_{(i-1)}^{ZG85}, x_{(i-2)}^{ZG85}\right), i = 3, 4, \cdots, 113$	$x_{(i+1)}^{ZG85}$
ZG86	$X_i^{ZG86} = \left(x_i^{ZG86}, x_{(i-1)}^{ZG86}\right), i = 2, 3, \cdots, 113$	$x_{(i+1)}^{ZG86}$
ZG87	$X_i^{ZG87} = \left(x_i^{ZG87}, x_{(i-1)}^{ZG87}, x_{(i-2)}^{ZG87}\right), i = 3, 4, \cdots, 113$	$x_{(i+1)}^{ZG87}$
ZG88	$X_i^{ZG88} = \left(x_i^{ZG88}, x_{(i-1)}^{ZG88}, x_{(i-2)}^{ZG88}\right), i = 3, 4, \cdots, 113$	$x_{(i+1)}^{ZG88}$
G3	$X_i^{G3} = \left(x_i^{G3}, x_{(i-1)}^{G3}, x_{(i-2)}^{G3}, x_{(i-3)}^{G3}, x_{(i-4)}^{G3}\right), i = 5, 6, \cdots, 56$	$x_{(i+1)}^{G3}$
T1	$X_i^{T1} = \left(x_i^{T1}, x_{(i-1)}^{T1}, x_{(i-2)}^{T1}\right), i = 3, 4, \cdots, 88$	$x_{(i+1)}^{T1}$

Table 1. Reconstructed phase spaces of ZG85 ~ ZG88, G3 and T1 GPS monitoring cumulative displacements.

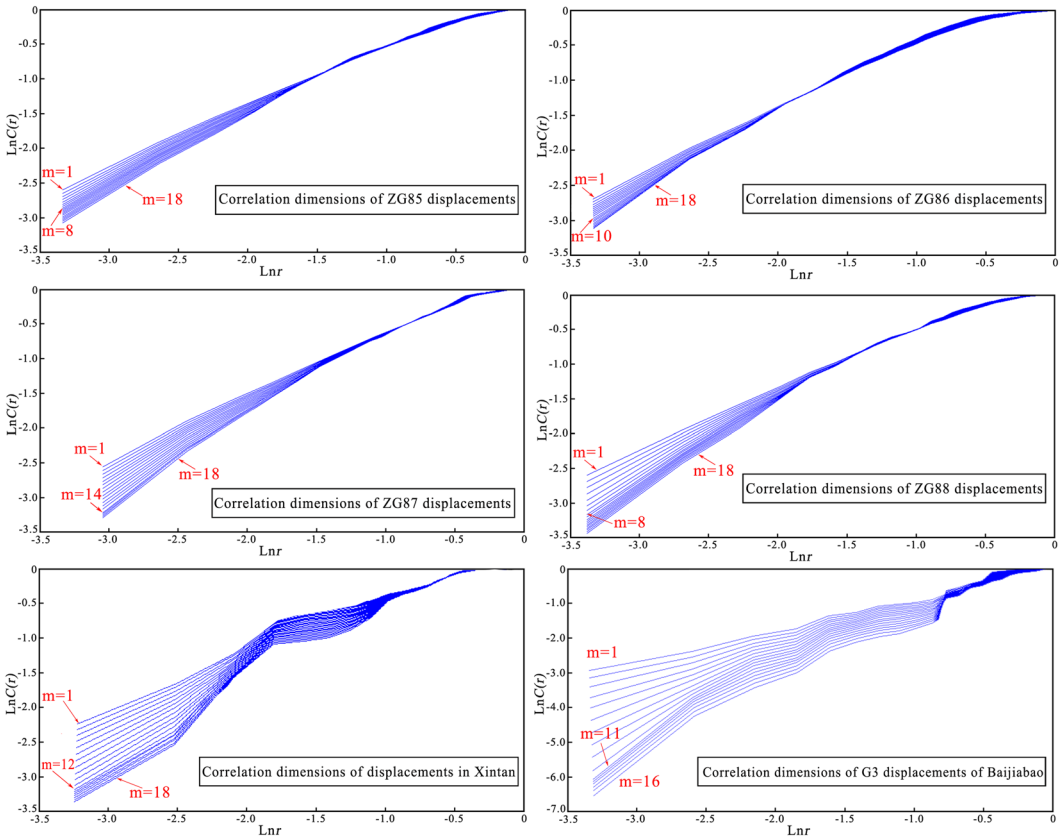

Figure 8. The correlation dimension curves of ZG85 ~ ZG88 of Shuping landslide, G3 of Baijiabao landslide and T1 of Xintan landslide cumulative displacements (drawn by Yuanyao Li).

3, 2, 3, 3, 5 and 3 based on the false nearest neighbor method. These reconstructed phase spaces are shown in Table 1.

Chaos evidence of monthly cumulative displacements. Final calculated *LLE* values of the ZG85 ~ ZG88, G3 and T1 displacement time series are respectively 0.0862, 0.0315, 0.0227, 0.0577, 0.1232 and 0.0526. The results show that there are chaos characteristics in all of the ZG85 ~ ZG88, G3 and T1 displacements because their *LLE* values are all greater than zero[47]. In addition, the non-linearity of the displacement time series can be measured by comparing the *LLE* values of the displacements. A greater *LLE* value means a higher non-linearity. The comparison results show that the G3 displacement has the highest non-linearity, while the ZG87 displacement has the lowest non-linearity.

CD method is also in this study, the corresponding $D(m)$ are all calculated with the m increases from 1 to 18. It can be seen from Fig. 8 the relationships between $C(r)$ and r of ZG85 ~ ZG88, G3 and T1 GPS monitoring displacements are respectively calculated. The calculation results indicate that the slopes of the lines in Fig. 8 respectively converge to a constant when m is increased to 8, 10, 14, 8, 11 and 12. Hence, we can draw conclusions that there are chaos characteristics in the ZG85 ~ ZG88, G3 and T1 GPS monitoring displacements, and the chaotic WA-Volterra and chaotic WA- SVM models can be adopted to predict the landslide displacements.

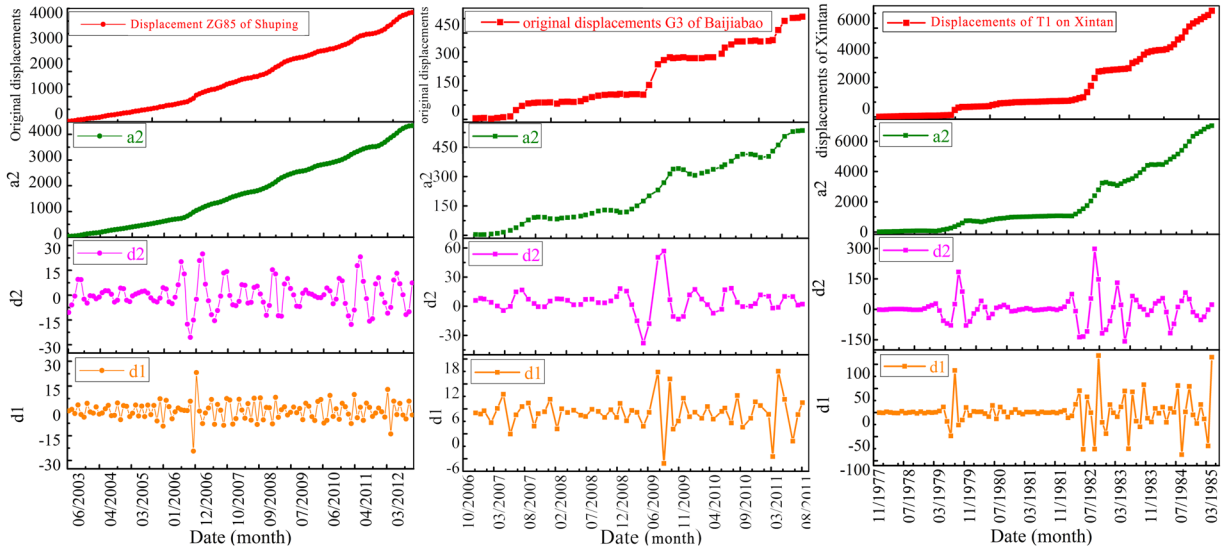

Figure 9. Displacement time series decomposition of ZG85 on Shuping landslide, G3 on Baijiabao landslide and T1 on Xintan landslide, with a2 reflecting low-frequency component, d1 reflecting the first high-frequency component and d2 reflecting the second high-frequency component of these cumulative displacements. (drawn by Yuanyao Li).

Wavelet analysis of GPS monitoring landslide cumulative displacements. The chaotic WA-Volterra and WA-SVM models predict cumulative displacements through forecasting all the low- and high-frequency components. These components are obtained from the wavelet analysis of GPS monitoring cumulative displacements. The decomposition levels of ZG85 ~ ZG88 of Shuping landslide, G3 of Baijiabao landslide and T1 point of Xintan landslide cumulative displacements are all two according to Eq. (11). Then we can decompose the ZG85 ~ ZG88, G3 and T1 landslide cumulative displacements using WA method as:

$$x_i^{ZG} = x_{a2,i}^{ZG} + x_{d1,i}^{ZG} + x_{d2,i}^{ZG} \tag{1}$$

$$x_i^{G} = x_{a2,i}^{G} + x_{d1,i}^{G} + x_{d2,i}^{G} \text{ and } x_i^{T} = x_{a2,i}^{T} + x_{d1,i}^{T} + x_{d2,i}^{T} \tag{2}$$

where x_i^{ZG}, x_i^{G} and x_i^{T} respectively presents the cumulative displacements of ZG85 ~ ZG88, G3 and T1; $x_{a2,i}$, $x_{d1,i}$ and $x_{d2,i}$ respectively donates the low-frequency, the first high-frequency and the second high-frequency components of ZG85 ~ ZG88, G3 and T1 landslide cumulative displacements. Taking the ZG85, G3 and T1 cumulative displacements as examples, the final decomposition results of ZG85, G3 and T1 cumulative displacements are respectively shown in Fig. 9.

Phase space reconstructions of each frequency component. The reconstructed phase spaces of each frequency component provide input and output variables of Volterra filter and SVM models. In this study, we set the τ of all the frequency components as one, and we calculate the optimal m of all the frequency components using the FNN method. The results show that the optimal m of $x_{a2,i}^{ZG85}$, $x_{d1,i}^{ZG85}$, $x_{d2,i}^{ZG85}$ are respectively 2, 4 and 4; m of $x_{a2,i}^{ZG86}$, $x_{d1,i}^{ZG86}$, $x_{d2,i}^{ZG86}$ are respectively 2, 5 and 3; m of $x_{a2,i}^{ZG87}$, $x_{d1,i}^{ZG87}$, $x_{d2,i}^{ZG87}$ are respectively 2, 5 and 4; m of $x_{a2,i}^{ZG88}$, $x_{d1,i}^{ZG88}$, $x_{d2,i}^{ZG88}$ are respectively 2, 5 and 3; m of $x_{a2,i}^{G3}$, $x_{d1,i}^{G3}$, $x_{d2,i}^{G3}$ are respectively 2, 3 and 2 m of $x_{a2,i}^{T1}$, $x_{d1,i}^{T1}$, $x_{d2,i}^{T1}$ are respectively 2, 3 and 4;. The results of *PSR* are shown in Table 2. Then we can fit and predict landslide displacements using Voltterra filter and SVM models as:

$$x_{i+1} = Volterra(X_i) \text{ or } x_{i+1} = SVM(X_i) \tag{3}$$

Displacements prediction using chaotic WA-Volterra and chaotic WA-SVM models. In order to train and test the Volterra filter and SVM models, the ZG85 ~ ZG88 landslide displacements are divided into two subsets: the former 94 months of ZG85 ~ ZG88 GPS monitoring landslide displacements from June 2003 to March 2011 are adopted to train the model, while the remaining later 20 landslide displacements are adopted to test the models. The same as the Shuping landslide, there are a total of 57 months of monitoring data on the Baijiabao landslide, the former 42 monthly displacement time series of G3 from December 2006 to May 2010 are used to train the models, while the remaining 15 data points from June 2010 to August 2011 are used to test the models. Meanwhile, a total of 89 months of displacements on the Xintan landslide were monitored, the former 74 monthly displacements of T1 from January 1978 to February 1984 are used to train the models, while the remaining 15 displacements are used to test the models. The input-output variables of the chaotic WA-Volterra and chaotic WA-SVM models are obtained as shown in Table 2.

Displacements	Reconstructed phase spaces (input variables)	Output variables
ZG85	$X_{a2,i}^{ZG85} = \left(x_{a2,i}^{ZG85}, x_{(a2,i-1)}^{ZG85}\right), i = 2, 3, \cdots, 113$	$x_{a2,(i+1)}^{ZG85}$
	$X_{d1,i}^{ZG85} = \left(x_{d1,i}^{ZG85}, x_{d1,(i-1)}^{ZG85}, x_{d1,(i-2)}^{ZG85}, x_{d1,(i-3)}^{ZG85}\right), i = 4, 5, \cdots, 113$	$x_{d1,(i+1)}^{ZG85}$
	$X_{d2,i}^{ZG85} = \left(x_{d2,i}^{ZG85}, x_{d2,(i-1)}^{ZG85}, x_{d2,(i-2)}^{ZG85}, x_{d2,(i-3)}^{ZG85}\right), i = 4, 5, \cdots, 113$	$x_{d2,(i+1)}^{ZG85}$
ZG86	$X_{a2,i}^{ZG86} = \left(x_{a2,i}^{ZG86}, x_{(a2,i-1)}^{ZG86}\right), i = 2, 3, \cdots, 113$	$x_{a2,(i+1)}^{ZG86}$
	$X_{d1,i}^{ZG86} = \left(x_{d1,i}^{ZG86}, x_{d1,(i-1)}^{ZG86}, x_{d1,(i-2)}^{ZG86}, x_{d1,(i-3)}^{ZG86}, x_{d1,(i-4)}^{ZG86}\right), i = 5, 6, \cdots 113$	$x_{d1,(i+1)}^{ZG86}$
	$X_{d2,i}^{ZG86} = \left(x_{d2,i}^{ZG86}, x_{d2,(i-1)}^{ZG86}, x_{d2,(i-2)}^{ZG86}\right), i = 3, 4, \cdots, 113$	$x_{d2,(i+1)}^{ZG86}$
ZG87	$X_{a2,i}^{ZG87} = \left(x_{a2,i}^{ZG87}, x_{(a2,i-1)}^{ZG87}\right), i = 2, 3, \cdots, 113$	$x_{a2,(i+1)}^{ZG87}$
	$X_{d1,i}^{ZG87} = \left(x_{d1,i}^{ZG87}, x_{d1,(i-1)}^{ZG87}, x_{d1,(i-2)}^{ZG87}, x_{d1,(i-3)}^{ZG87}, x_{d1,(i-4)}^{ZG87}\right), i = 5, 6, \cdots 113$	$x_{d1,(i+1)}^{ZG87}$
	$X_{d2,i}^{ZG87} = \left(x_{d2,i}^{ZG87}, x_{d2,(i-1)}^{ZG87}, x_{d2,(i-2)}^{ZG87}, x_{d2,(i-3)}^{ZG87}\right), i = 4, 5, \cdots, 113$	$x_{d2,(i+1)}^{ZG87}$
ZG88	$X_{a2,i}^{ZG88} = \left(x_{a2,i}^{ZG88}, x_{(a2,i-1)}^{ZG88}\right), i = 2, 3, \cdots, 113$	$x_{a2,(i+1)}^{ZG88}$
	$X_{d1,i}^{ZG88} = \left(x_{d1,i}^{ZG88}, x_{d1,(i-1)}^{ZG88}, x_{d1,(i-2)}^{ZG88}, x_{d1,(i-3)}^{ZG88}, x_{d1,(i-4)}^{ZG88}\right), i = 5, 6, \cdots 113$	$x_{d1,(i+1)}^{ZG88}$
	$X_{d2,i}^{ZG88} = \left(x_{d2,i}^{ZG88}, x_{d2,(i-1)}^{ZG88}, x_{d2,(i-2)}^{ZG88}\right), i = 3, 4, \cdots, 113$	$x_{d2,(i+1)}^{ZG88}$
G3	$X_{a2,i}^{G3} = \left(x_{a2,i}^{G3}, x_{(a2,i-1)}^{G3}\right), i = 2, 3, \cdots, 56$	$x_{a2,(i+1)}^{G3}$
	$X_{d1,i}^{G3} = \left(x_{d1,i}^{G3}, x_{d1,(i-1)}^{G3}, x_{d1,(i-2)}^{G3}\right), i = 3, 4, \cdots, 56$	$x_{d1,(i+1)}^{G3}$
	$X_{d2,i}^{G3} = \left(x_{d2,i}^{G3}, x_{d2,(i-1)}^{G3}\right), i = 2, 3, \cdots, 56$	$x_{d2,(i+1)}^{G3}$
T1	$X_{a2,i}^{T1} = \left(x_{a2,i}^{T1}, x_{(a2,i-1)}^{T1}\right), i = 2, 3, \cdots, 88$	$x_{a2,(i+1)}^{T1}$
	$X_{d1,i}^{T1} = \left(x_{d1,i}^{T1}, x_{d1,(i-1)}^{T1}, x_{d1,(i-2)}^{T1}\right), i = 3, 5, \cdots, 88$	$x_{d1,(i+1)}^{T1}$
	$X_{d2,i}^{T1} = \left(x_{d2,i}^{T1}, x_{d2,(i-1)}^{T1}, x_{d2,(i-2)}^{T1}, x_{d2,(i-3)}^{T1}\right), i = 4, 5, \cdots, 88$	$x_{d2,(i+1)}^{T1}$

Table 2. *PSR* of each frequency component of ZG85 ~ ZG88, G3 and T1 cumulative displacements.

The second-order Volterra filter model is used to construct the models for predicting the landslide displacements on Shuping landslide and Baijiabao landslide. The predictive results of the chaotic WA-Volterra model are shown in Fig. 10. To compare Volterra filter model with SVM model, the same input-output variables that have been used in the chaotic WA-Volterra model are used again in the chaotic WA-SVM model. The optimal parameter combination (C, ε, γ) of SVM is shown in Table 3. In addition, single chaotic Volterra filter model is also adopted in this study for comparison. The final predictive values of chaotic WA-SVM and single chaotic Volterra models are also shown in Fig. 10, and their prediction performances are assessed in Table 4.

Discussion

Chaos characteristics identify. The calculated *LLE* and correlation dimensions values show that there are chaos characteristics in the GPS monitoring landslide displacements. Hence, we are sure that the ZG85 ~ ZG88 GPS of Shuping landslide, G3 of Baijibao landslide and T1 cumulative displacements of Xintan landslide can be forecasted using the chaos-based models. The predictive results shown in Fig. 10 and Table 4 suggest that, the chaotic WA-Volterra, chaotic WA-SVM and single chaotic Volterra models accurately predict the GPS monitoring cumulative displacements as a whole. We can presume that appropriate input and output variables for the three nonlinear models are acquired from the reconstructed phase spaces, and can further extrapolate that the original nonlinear evaluation process of landslide displacement time series is effectively rebuilt by the *PSR* method of chaos theory.

Comparison between volterra filter and svm models. The prediction performances of the three models are also shown in Fig. 10 and Table 4, which show that the chaotic WA-Volterra model is more accurate and credible than the chaotic WA-SVM model. We can conclude that the Volterra filter model is more appropriate for a finite nonlinear displacement time series than the SVM model. In actual, several nonlinear models are present to describe the nonlinear system, including Volterra filter, Taylor series and Wiener series, *et al.*[34]. Volterra filter model is one of the most popular nonlinear models for nonlinear dynamic time series prediction. There are several reasons, from a mathematical point of view, Voltera filter is essentially a functional series expansion of nonlinear time-invariant systems, and can also be regarded as the generalization of one-dimensional convolution in multidimensional convolution space[48]. From a physical point of view, the Volterra filter, which is similar to the linear impulse response function, can describe the substantive features of the nonlinear system with clear physical meaning. Moreover, the Volterra filter model can fit the nonlinear continuous functions with arbitrary precision because of its universal characteristic and ability of adaptive memory[49]. Similarly, the GPS monitoring landslide displacement time series is a nonlinear system with dynamic response and temporal memory features. Hence, the Volterra filter model can be used to accurately fit and predict the dynamic evaluation behavior of landslide displacements.

Comparisons between chaotic WA-Volterra and single chaotic Volterra models. Figure 10 and Table 4 show that the predictive displacements of both chaotic WA based models deviate slightly from the real monitoring displacements, indicating that both chaotic WA based models have great prediction performances. On the contrary, results in Fig. 10 also show that, cumulative displacements of ZG85 ~ ZG88, G3 and T1 are underestimated

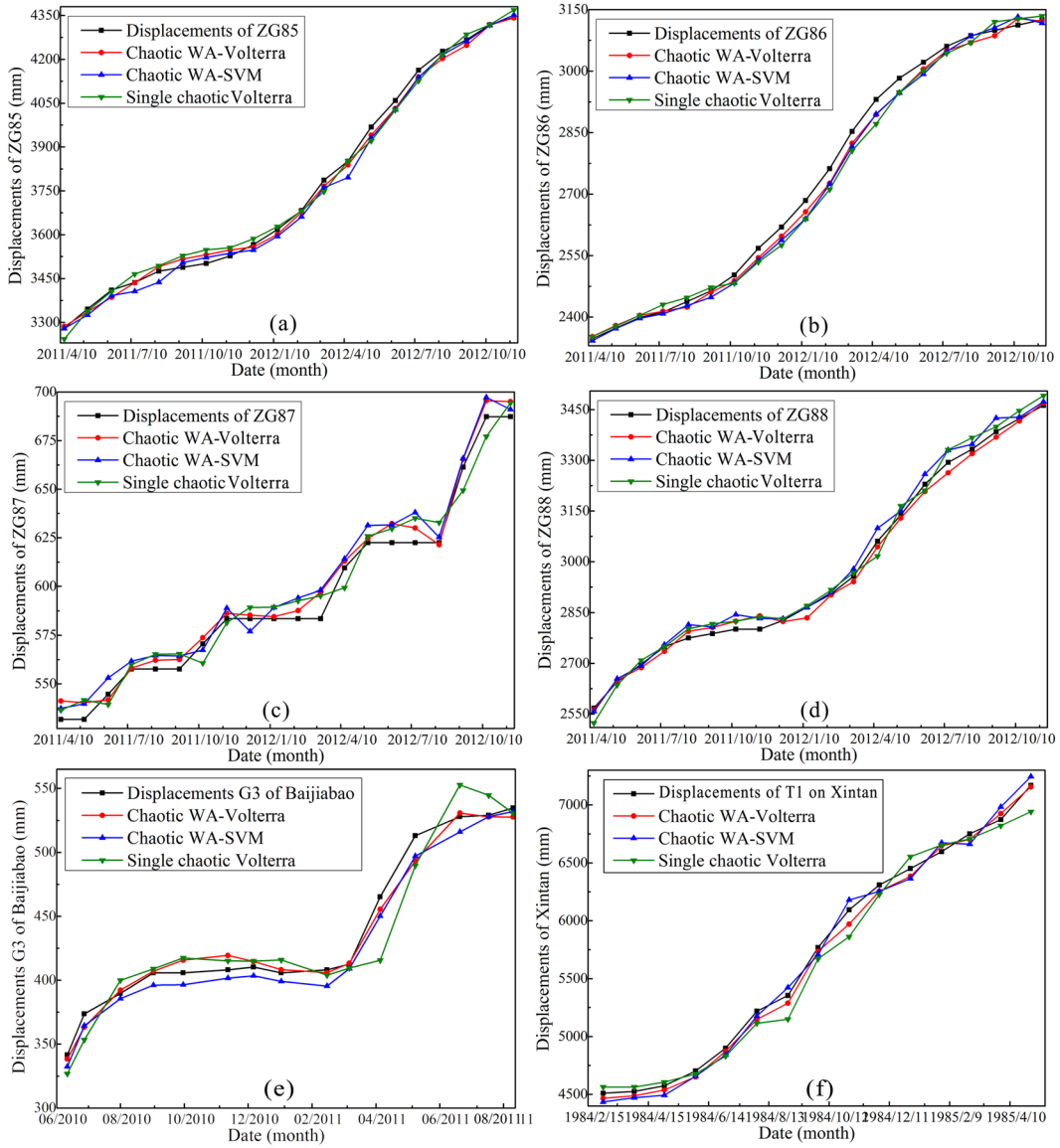

Figure 10. Comparison between predicting and monitoring the monthly ZG85 ~ ZG88 displacements of Shuping landslide, G3 displacements of Baijiabao landslide, and T1 displacements of Xintan landslide (drawn by Yuanyao Li).

overall by the single chaotic Volterra model. Specially, the chaotic Volterra model has difficulty in estimating the step-like features of these cumulative displacements, while chaotic WA-Volterra model overcomes this difficulty. In can be concluded that the chaotic WA based models have remarkably higher prediction performance than the single chaotic Volterra model. This is because that the nonlinear and non-stationary characteristics of GPS monitoring cumulative displacements are addressed reasonably well by the WA method before predict cumulative displacements using Volterra filter and SVM models. The "noises" existed in the original GPS monitoring landslide cumulative displacements are effectively removed by the WA method. As a result, the valuable information with less "noises" in the low- and high-frequency components can be fully extracted to train and test the Volterra filter and SVM models.

In addition to wavelet analysis, some other methods have also been proposed to implement time series decomposition and errors reduction. These methods include empirical mode decomposition[23], principal component decomposition[50,51], Fourier transformation[52], polynomial decomposition[10], *et al.* Comparing to the other time series decomposition methods, wavelet analysis is used more widely and is considered to be of higher decomposition efficiencies in both time and frequency domains[53,54]. Especially, Wickersham, Li and Lin[55] shows that wavelet analysis has advantages of great applicability to discrete signals, insensitivity to the choice of wavelet basis, and insensitivity to the target signal extracted from the raw measurements comparing to traditional Fourier and PCA methods. Hence, this study selects wavelet analysis to decompose GPS monitoring landslide displacement time series for building landslide displacements prediction model.

Unfortunately, the present chaotic WA-Volterra model is just an offline modeling procedure without practical application. Hence, it is necessary to software this proposed model for online real-time landslide displacement

	GPS points	Frequency components	(C, ε, γ)
Chaotic WA-Volterra, Chaotic WA-SVM	ZG85	$a2$	(97.64, 0.005, 0.48)
		$d1$	(18.12, 0.01, 0.18)
		$_{d2}$	(53.38, 0.005, 0.28)
	ZG86	$a2$	(486.52, 0.001, 0.31)
		$d1$	(45.42, 0.005, 0.19)
		$d2$	(38.22, 0.003, 0.32)
	ZG87	$a2$	(974.83, 0.005, 0.35)
		$d1$	(20.28, 0.008, 0.21)
		$d2$	(63.18, 0.031, 0.37)
	ZG88	$a2$	(52.47, 0.005, 0.39)
		$d1$	(42.36, 0.011, 0.21)
		$d2$	(73.62, 0.008, 0.22)
	G3	$a2$	(137, 0.02, 0.38)
		$d1$	(526, 0.06, 0.31)
		$d2$	(688, 0.005, 0.41)
	T1	$a2$	(73, 0.03, 0.45)
		$d1$	(81, 0.05, 0.33)
		$d2$	(125, 0.002, 0.32)

Table 3. Parameters of the SVM model.

GPS point	Prediction model	*RMSE*(*mm*)	*MAPE*(%)
ZG85	Chaotic WA-Volterra	19.13	0.45
	Chaotic WA-SVM	23.89	0.56
	Single chaotic Volterra	28.46	0.72
ZG86	Chaotic WA- Volterra	20.31	0.63
	Chaotic WA- SVM	24.22	0.76
	Single chaotic Volterra	29.67	0.83
ZG87	Chaotic WA- Volterra	6.32	0.86
	Chaotic WA- SVM	8.17	0.93
	Single chaotic Volterra	8.66	0.96
ZG88	Chaotic WA- Volterra	19.17	0.54
	Chaotic WA- SVM	23.64	0.66
	Single chaotic Volterra	26.48	0.75
G3	Chaotic WA- Volterra	7.88	1.59
	Chaotic WA- SVM	9.57	1.72
	Single chaotic Volterra	18.22	3.12
T1	Chaotic WA- Volterra	57.19	0.009
	Chaotic WA- SVM	72.93	0.012
	Single chaotic Volterra	116.92	0.016

Table 4. Displacements prediction results of chaotic WA-Volterra, chaotic WA-SVM and single chaotic Volterra models for ZG85 ~ ZG88 of Shuping, G3 of Baijiabao and T1 of Xintan landslides.

prediction. In addition, we can still improve these three models. For example, it is necessary to obtain a longer length displacement time series of the displacement rather than a limit length landslide displacement time series. As a result, the reconstructed strange attractors can be fully unfolded to reflect the original strange attractors.

Conclusions

In this study, chaos evidences of landslide displacements of Shuping landslide, Baijiabao landslide and Xintan landslide are determined based on LLE and *CD* methods, and a novel forecasting model is proposed for landslide displacements prediction. We can conclude that chaos characteristics existed in the nonlinear landslide displacements and these displacements are effectively predicted by the proposed chaotic WA-Volterra model. In addition, the chaotic WA-Volterra model obtains more accurate predictive displacements than the chaotic WA-SVM and single chaotic Volterra models. The main contributions of this study contain chaos characteristics identification of landslide displacements, landslide displacement time series decomposition using wavelet analysis, and Volterra filter model proposed for model construction.

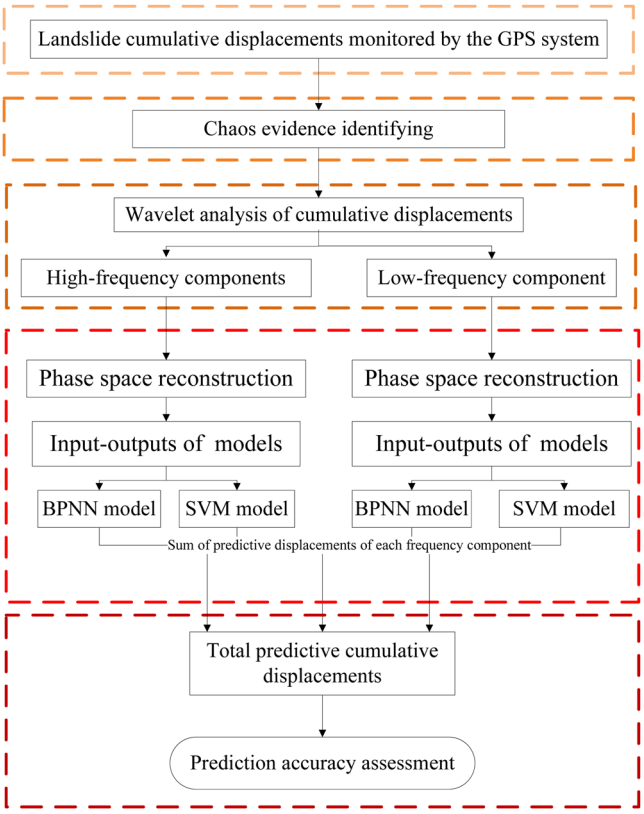

Figure 11. Flowchart of the chaotic WA-Volterra model (Drawn by Yuanyao Li).

Methods

The present chaotic WA-Volterra model has five steps as shown in Fig. 11: (1) the landslide cumulative displacements are obtained using the GPS system. (2) these cumulative displacement time series are normalized. (3) the chaos evidences of cumulative displacements are determined. (4) the cumulative displacements are decomposed into different low- and high-frequency components using WA method. (4) phase space of each frequency component is reconstructed using chaos theory, then Volterra filter model are trained and tested using these inputs-outputs obtained from reconstructed phase spaces. (5) through summing the predictive displacements of each frequency component, we can obtain the predictive cumulative displacements.

Landslide displacement monitoring using the GPS system. *Introduction to the GPS system.* The proposed GPS system is composed of the signal receiver, signal processing and displacement presentation subsystems. The architecture of the GPS system is shown in Fig. 12. In the signal receiver subsystem, some GPS positioning methods used for monitoring the landslide displacements are proposed, such as the static relative positioning measurement[56] and carrier phase measuring method[57], *et al.* The static relative positioning measurement, which is of millimeter accuracy, is used in this study to construct the GPS monitoring system. The reason is that the static relative positioning measurement method can eliminate the orbit and atmosphere errors by a spatial correlation between the reference point and measuring points[25]. At the same time, the errors between the clocks in the GPS receivers and related satellites can also be removed by this method.

The production type of the GPS antenna is Trimble R8 GNSS with multi-channels and multi-frequencies. These GPS positioning messages contain a carrier phase, pseudo-range, the known coordinates and some other data[56]. The reference points and observation points receive one-day signals synchronously as the signal monitoring phase. The GPS antennas should receive at least five satellite signals at the same time, and then, the received signals will be transmitted to the signal processing subsystem.

The signal processing subsystem is mainly composed of GPS signal processing software (GAMIT/GLOBK software[41]) and a computer server, with display software. In the computer server, GAMIT/GLOBK software is used to obtain *mm* level positioning results of the measured points. First, the three-dimensional coordinate values of all of the points are measured. Then, the three-dimensional coordinate values of the measured points are compared to the reference point to determine the landslide displacements of the measured points.

Error analysis of the GPS system. It is important to analyze the errors of the GPS system. The displacement errors of the GPS system can be mainly divided into the system errors and random errors. Most of the system errors are removed when the received GPS signals are processed based on the principle of differential GPS; as a result, the random errors with a few system residual errors are the main type of errors of the GPS monitoring

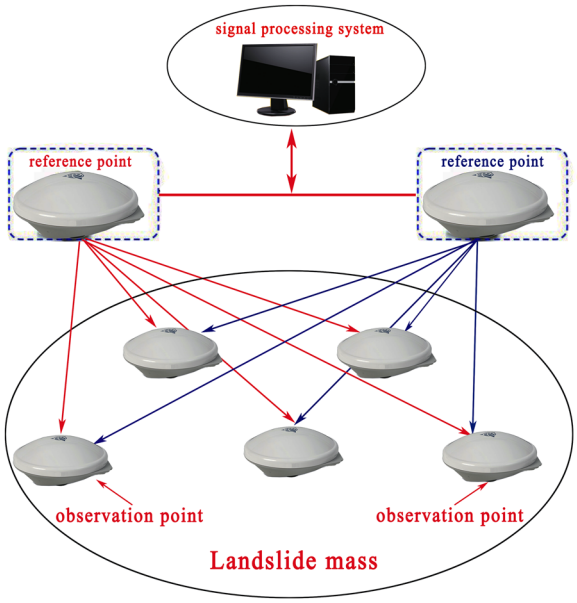

Figure 12. Structure of the GPS system (drawn by Yuanyao Li).

displacements[25]. The random errors reduce the accuracy of the fitting and prediction of the non-linear models. The errors of the GPS system on the Shuping and Baijiabao landslides have been calculated and shown that the horizontal errors and vertical errors in the monitoring landslide displacements are less than ±2.00 mm and ±5.00 mm, respectively[58]. Hence, the impact of the total errors on the prediction accuracy of the chaotic WA based model is limited[6].

Data normalization. It is necessary to perform data normalization to help the Volterra filter and SVM models converge to the optimal solution before model building. The original cumulative displacements are transformed into the desired range of [0, 1] as:

$$x_i = \frac{x_{old,i} - x_{old,\min}}{x_{old,\max} - x_{old,\min}} \tag{4}$$

where $x_{old,i}(i = 1, 2, \ldots\ldots, N)$ are the monitoring cumulative displacements, $x_{old,\min}$ and $x_{old,\max}$ are respectively their lower and upper bounds, N is the number of cumulative displacements. The predictive normalized displacement is back-transformed, so as to obtain final predictive displacement. The result is shown in Eq. (5), where y_i is the predictive normalized cumulative displacement.

$$\hat{y}_i = y_i \times (x_{old,\max} - x_{old,\min}) + x_{old,\min} \tag{5}$$

Wavelet analysis. Wavelet analysis (WA)[21] is a significant tool which deals with the time series based on the processes of dilation and translation. we define the mother wavelet function $\varphi(t)$:

$$\int_{-\infty}^{+\infty} \phi(t)dt = 0 \tag{6}$$

$$\phi_{a,b}(t) = |a|^{-\frac{1}{2}}\phi\left(\frac{t-b}{a}\right) \tag{7}$$

where a and b are set to real numbers, and $\varphi_{a,b}(t)$ donates the successive wavelet. This study use the discrete wavelet transform (DWT)[59], one wavelet analysis method, to decompose the landslide displacement time series. The parameters a and b of DWT can be determined as:

$$\phi_{u,v}\left(\frac{t-b}{a}\right) = a_0^{-\frac{u}{2}}\phi^*\left(\frac{t - nb_0a_0^u}{a_0^u}\right) \tag{8}$$

where a_0 is usually set to two, u and v are adopted to affect the processes of wavelet transform, b_0 is usually set to one. Then a finite discrete time series $f(t)$ can be defined by DWT as:

$$W_f(u, v) = 2^{-\frac{u}{2}} \sum_{t=0}^{K-1} f(t)\phi^*(2^{-u}i - v) \tag{9}$$

where $W_f(u, v)$ is regarded as the wavelet coefficient, $a = 2^u$, $b = 2^u v$ and $K = 2^U$; v is regarded as time translation parameter varying between 0 and $2^{U-u} - 1$, where $1 < u < U$. The DWT can effectively decompose the cumulative displacement time series into the trend, period and fluctuation characteristics using its high-pass filters and low-pass filters. Meanwhile, the Mallat algorithm[60] is used in the DWT method, then the cumulative displacement x_i is decomposed as:

$$x_i = x_{a,i} + x_{d1,i} + x_{d2,i} + \cdots\cdots + x_{dL,i} \tag{10}$$

where L donates the decomposition level; $x_{a,i}$ presents the low-frequency component suggesting the trend characteristics; $x_{d1,i}$, $x_{d2,i}$, ……, $x_{dL,i}$ respectively donates the first, second, ……, Lth level high-frequency components reflecting the periodic, fluctuation characteristics of the cumulative displacement time series. In addition, The *db*4 is used in this study as wavelet function according to Li, Huang, Jiang, Huang and Chang[25]. Further, Eq. (11) is used to determine the number of decomposition levels according to Nourani, Alami and Aminfar[61], where N presents the length of the GPS monitoring landslide displacement.

$$L = \text{int}[\log(N)] \tag{11}$$

Chaos theory and phase space reconstruction. *Introduction of chaos theory.* Chaos theory attempts to explain the fact that complex and unpredictable results can and will occur in nonlinear systems. A landslide displacement system has a temporal deterministic complexity. It has been proved by many studies that there is evidence of chaos in the monitoring displacement time series[6,28]. However, it is still necessary to determine the chaos evidence of the GPS monitoring landslide displacements of Shuping and Baijiabao landslides in this study. The reason is that the non-linearity of the landslide displacements can be established by determining the chaos evidence of the landslide displacements[6]. The chaos evidence of nonlinear time series is mainly identified through the Largest Lyapunov Exponent (LLE)[62] and Correlation Dimension (*CD*) methods[63]. This study uses both two methods to determine the chaos characteristics of GPS monitoring landslide displacements.

Phase space reconstruction. According to PSR, for example, a displacement time series x_i can be fully embedded into m-dimensional phase space. As a result, the inputs and output of nonlinear predictor can be obtained from the reconstructed m-dimensional phase spaces[35]. The inputs of nonlinear predictor are represented by vector X_i as:

$$X_i = (x_i, x_{i-\tau}, x_{i-2\tau}, \cdots\cdots, x_{i-(m-1)\tau}) \tag{12}$$

where $i = (m-1)\tau + 1, (m-1)\tau + 2$, ……, N is the number of vectors, N is the number of displacements, τ is delay time and m is embedding dimension ($m \geq d$, where d is the dimension of the attractor). It is important to select appropriate delay time and embedding dimension values to effectively reconstruct the original phase spaces of the GPS monitoring displacements. The delay time of displacement time series is generally set to one for landslide displacement predictions because the displacements are noisy and the length of displacement time series is very limit[24]. In addition, the false nearest neighbor (FNN) method[64] is used to calculate the appropriate m, this is because FNN is insensitive to finite and noisy displacements.

Identifying of chaos evidences of landslide cumulative displacements. *Chaos evidence identification using LLE method.* Small-data set method[65] is used to calculate the *LLE* value of landslide displacement. The LLE is set to L in this study then the average divergence at time t can be defined as

$$d(t) = ke^{Lt} \tag{13}$$

where k is a constant that normalizes the initial separation. The reconstructed trajectory X can be expressed as:

$$X = (X_1, X_2, \cdots\cdots, X_M)^T \tag{14}$$

where X_i is the ith data point of the dynamic system, M is the number of data points on the reconstructed attractor, and each X_i is given by Eq. (12). The nea rest neighbor X_j^{Near} is found by searching for the point that minimizes the distance to the particular reference point X_j:

$$d_j(0) = \min_{X_j^{Near}} \left\| X_j - X_j^{Near} \right\| \tag{15}$$

where $d_j(0)$ is the initial distance from the jth point to its nearest neighbor, and $\| \|$ denotes the Euclidean norm. Based on the definition of L given in Eq. (13), the jth pair of nearest neighbors diverges approximately at the rate given by

$$d_j(t_i) \approx k_j e^{L(i\Delta t)} \tag{16}$$

where $t_i = i\Delta t$, Δt is the sampling period of the time series, and k_j is the initial separation of the jth pair of nearest neighbors. Taking the logarithm of both sides of Eq. (16), we have

$$\ln d_j(t_i) \approx \ln k_j + L(i\Delta t) \tag{17}$$

Equation (18) represents a set of approximately parallel lines (for $j = 1, 2, \ldots\ldots, M$), each of which has a slope that is roughly proportional to L. The value of L can be calculated by

$$y(t_i) = \frac{1}{\Delta t}\left\langle \ln d_j(t_i) \right\rangle \tag{18}$$

where $\langle\cdot\rangle$ denotes the average over all j. This process of averaging is one very important step to determine an accurate L based on a finite and noisy time series. Landslide displacement is considered as a chaotic time series when LLE value satisfies the condition for chaos, i.e., $LLE > 0$.

Chaos evidence identification using CD method. *CD* method uses a non-integer fractal dimension to characterize the chaos characteristic of landslide displacement time series. In this study, the Grassberger-Procaccia (G-P) method[66] is proposed to calculate the *CD* values of ZG85 ~ ZG88, G3 and T1 landslide displacements. Suppose two points X_i and X_j in the reconstructed phase space, we can describe *CD* as:

$$C(r) = \lim_{N\to\infty} \frac{2}{M(M-1)} \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j\neq i}}^{M} S\left(r - \left\| X_i - X_j \right\|\right) \tag{19}$$

where S is the Heaviside step function, the $S(u)$ is set to 1 if $u \geq 0$ and is set to 0 if $u < 0$; r represents the radius of the sphere centered on X_i or X_j. Suppose the $D(m)$ as *CD* value, $C(r)$ can be related to radius r as Eq. (20) when the nonlinear landslide displacement is featured by an attractor.

$$C(r) \propto r^{D(m)} \tag{20}$$

Take the logarithm of Eq. (20), then we can rearrange it as:

$$D(m) = \lim_{r\to 0} \frac{\ln C(r)}{\ln r} \tag{21}$$

Based on Eq. (21), we can obtain a series of $D(m)$ through the increase of m value. When the nonlinear landslide displacements are of chaos characteristic, the $D(m)$ will continuously increases and then converges to a constant with the increase of m value[67].

Volterra filter model. The Volterra filter is a model which can be adopted to predict the non-linear time series[68]. Volterra filter model has the ability to obtain the memory information through the model training and testing processes. We can get a second-order Volterra filter model when setting the degree of Volterra filter as 2. Relational literature shows that a second-order Volterra filter model has efficient performance for nonlinear time series prediction:

$$\mathrm{y(t)} = \begin{cases} \sum_{\mathrm{n}=1}^{S} \sum_{\tau_1=1}^{p} h_1^{(n)}(\tau) x_n(t-\tau) \\ + \sum_{n=1}^{S} \sum_{\tau_1=1}^{p} \sum_{\tau_2=1}^{p} h_{2s}^{(n)}(\tau_1, \tau_2) x_n(t-\tau_1) x_n(t-\tau_2) \\ + \sum_{n_1=1}^{S} \sum_{n_2=1}^{n_1-1} \sum_{\tau_1=1}^{p} \sum_{\tau_2=1}^{p} h_{2x}^{(n_1,n_2)}(\tau_1, \tau_2) x_{n1}(t-\tau_1) x_{n2}(t-\tau_2) + \xi_{\mathrm{t}} \end{cases} \tag{22}$$

where $y(t)$ is single output prediction value, $x_1, x_2, x_3 \cdots x_S$ are time series and they are considered as input variable values, S represents the number of input variables, N donates length of time series about input variables, p represents the memory value which showing a significant lag relationship between different input variables, ξ_t represents the noises produced in model building.

Equation (22) shows that the second-order cross-kernels $h_{2x}^{(n_1,n_2)}$ represents a second-order nonlinear interactive relationships between each unique pair of input variables named x_{n1} and x_{n2}.We can simply Eq. (22) through combining the last two terms to yield Eq. (23) as:

$$y(t) = \left\{ \sum_{n=1}^{S} \sum_{\tau=1}^{p} h_1^{(n)}(\tau) x_n(t-\tau) + \sum_{n_1=1}^{S} \sum_{n_2=1}^{S} \sum_{\tau_1=1}^{p} \sum_{\tau_2=1}^{p} h_{2x}^{(n_1,n_2)}(\tau_1, \tau_2) x_{n1}(t-\tau_1) x_{n2}(t-\tau_2) + \xi_{\mathrm{t}} \right\} \tag{23}$$

The Orthogonal Least Squares-Error Reduction Ration[69] is used in this study to estimate the relative parameters of two-order Volterra filter model. Finally, the proposed Volterra filter model can be adopted to predict each frequency component of landslide displacement time series.

Support Vector machine. SVM[70] is a time series prediction technique with good performance. we can approximate the regression function as follows:

$$F(X) = \omega \cdot \varphi(X) + b \tag{24}$$

where b is a scalar threshold, ω is the weight vector, $\phi(X)$ is a high-dimensional feature space, and $\phi(X)$ is from the input space X by nonlinear mapping. The SVM model performs linear regression in the high-dimensional feature space by ε-insensitive loss. Then, we can estimate the coefficients ω and b by minimizing the regularized risk function Eq. (25)[71]:

$$\begin{aligned} &Min \ \frac{\|\omega\|^2}{2} \\ &s.\ t. \ \begin{cases} y_i - \varphi(\omega, x_i) - b \le \varepsilon \\ \varphi(\omega, x_i) + b - y_i \le \varepsilon \end{cases} \end{aligned} \tag{25}$$

Then, the regression problem can be transformed into the constrained formation:

$$\begin{aligned} &Min \ \frac{\|\omega\|^2}{2} + C\sum_{i=1}^{n}(\zeta_i + \zeta_i^*) \\ &s.\ t. \ \begin{cases} y_i - \langle \omega, x_i \rangle - b \le \varepsilon + \zeta_i \cdot \zeta_i^* \ge 0 \\ \langle \omega, x_i \rangle + b - y_i \le \varepsilon + \zeta_i \cdot \zeta_i^* \ge 0 \end{cases} \end{aligned} \tag{26}$$

where the constant C stands for the penalty degree of the sample with an error that exceeds e. We can use the optimization method to maximize the function to deal with the dual problem as:

$$\begin{aligned} &Max \sum_{i=1}^{n} y_i(a_i + a_i^*) - \varepsilon\sum_{i=1}^{n}(a_i + a_i^*) \\ &-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i - a_i^*)\left(a_j + a_j^*\right)K\left(X_j + X_j\right) \\ &s.\ t. \ \sum_{i=1}^{n}(a_i - a_i^*) = 0 \, and \, 0 \le a_i,\ a_i^* \le C \end{aligned} \tag{27}$$

where a_i and a_i^* are the Lagrange multiplier. The SVM predictor for the function fitting obtained by using the above-mentioned maximization function is then given as follows:

$$F(X) = \sum_{i=1}^{n}(a_i - a_i^*)K(x_i, x) + b \tag{28}$$

In Eq. (28), those sample points are called as support vectors. The radial basis kernel function $K(x_i, x) = \varphi(X_i, X_j)$ is used as the kernel function of SVM[35]. Then particle swarm optimization (PSO) method is introduced to determine the appropriate parameters (C, ε and r) of SVM[35].

Accuracy assessment. In this study, we use two assessment methods, Root Mean Square Error (*RMSE*) and Mean Absolute Percentage Error (*MAPE*), to evaluate the prediction accuracies of each model[72]. The *RMSE* is calculated as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N_0}(x_{old,i} - \hat{y}_i)^2}{N_0}} \tag{29}$$

where $x_{old,i}$ is the original cumulative displacement time sereis, $\hat{y}_i$ is the final predicted values, and N_0 is the length of the predicted data. *MAPE* can be described as:

$$MAPE = \frac{1}{N_0}\sum_{i=1}^{N_0}\frac{|x_{old,i} - \hat{y}_i|}{x_{old,i}} \times 100\% \tag{30}$$

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 28 January 2019; Accepted: 5 December 2019;
Published online: 27 December 2019

References

1. Huang, F., Yin, K., Huang, J., Gui, L. & Wang, P. Landslide susceptibility mapping based on self-organizing-map network and extreme learning machine. *Engineering Geology* **223**, 11–22 (2017).
2. Xu, S. & Niu, R. Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China. *Computers & Geosciences* **111**, 87–96 (2018).
3. Li, D. *et al*. Landslide Susceptibility Prediction Using Particle-Swarm-Optimized Multilayer Perceptron: Comparisons with Multilayer-Perceptron-Only, BP Neural Network, and Information Value Models. *Applied Sciences* **9**, 3664 (2019).
4. Cai, Z., Xu, W., Meng, Y., Shi, C. & Wang, R. Prediction of landslide displacement based on GA-LSSVM with multiple factors. *Bulletin of Engineering Geology and the Environment* **75**, 637–646 (2016).
5. Lian, C., Zeng, Z., Yao, W. & Tang, H. Ensemble of extreme learning machine for landslide displacement prediction based on time series analysis. *Neural Computing & Applications* **24**, 99–107 (2014).
6. Huang, F., Huang, J., Jiang, S. & Zhou, C. Landslide displacement prediction based on multivariate chaotic model and extreme learning machine. *Engineering Geology* **218**, 173–186 (2017).
7. Guo, Z. *et al*. Regional rainfall warning system for landslides with creep deformation in Three Gorges using a statistical black box model. *Scientific reports* **9**, 8962 (2019).
8. Peng, W. F., Wang, C. L., Chen, S. T. & Lee, S. T. Incorporating the effects of topographic amplification and sliding areas in the modeling of earthquake-induced landslide hazards, using the cumulative displacement method. *Computers & Geosciences* **35**, 946–966 (2009).
9. Deng, W. *et al*. GPS-PAIL: prediction of lysine acetyltransferase-specific modification sites from protein sequences. *Scientific reports* **6**, 39787 (2016).
10. Zhou, C. *et al*. Displacement prediction of step-like landslide by applying a novel kernel extreme learning machine method. *Landslides*, 1–15 (2018).
11. Huang, F., Yin, K., He, T., Zhou, C. & Zhang, J. Influencing factor analysis and displacement prediction in reservoir landslides- a case study of Three Gorges Reservoir (China). *Tehnički vjesnik* **23**, 617–626 (2016).
12. Jibson, R. W. Regression models for estimating coseismic landslide displacement. *Engineering Geology* **91**, 209–218 (2007).
13. Fahimi, F., Yaseen, Z. M. & El-Shafie, A. Application of soft computing based hybrid models in hydrological variables modeling: a comprehensive review. *Theoretical & Applied Climatology*, 1–29 (2016).
14. Yaseen, Z. M., El-Shafie, A., Jaafar, O., Afan, H. A. & Sayl, K. N. Artificial intelligence based models for stream-flow forecasting: 2000–2015. *Journal of Hydrology* **530**, 829–844 (2015).
15. Afan, H. A., El-Shafie, A., Wan, H. M. W. M. & Yaseen, Z. M. Past, present and prospect of an Artificial Intelligence (AI) based model for sediment transport prediction. *Journal of Hydrology* **541**, 902–913 (2016).
16. Huang, F. *et al*. A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction. *Landslides* (2019).
17. Yaseen, Z. M., Sulaiman, S. O., Deo, R. C. & Chau, K.-W. An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *Journal of Hydrology* (2018).
18. Gentili, P. L., Gotoda, H., Dolnik, M. & Epstein, I. R. Analysis and prediction of aperiodic hydrodynamic oscillatory time series by feed-forward neural networks, fuzzy logic, and a local nonlinear predictor. *Chaos* **25**, 1383–2693 (2015).
19. Miao, F., Wu, Y., Xie, Y. & Li, Y. Prediction of landslide displacement with step-like behavior based on multialgorithm optimization and a support vector regression model. *Landslides*, 1–14 (2017).
20. Wu, L. C. *et al*. Detection of American Football Head Impacts Using Biomechanical Features and Support Vector Machine Classification. *Scientific Reports* **8** (2018).
21. Deka, P. C. Support vector machine applications in the field of hydrology: a review. *Applied soft computing* **19**, 372–386 (2014).
22. Huang, F., Yao, C., Liu, W., Li, Y. & Liu, X. Landslide susceptibility assessment in the Nantian area of China: a comparison of frequency ratio model and support vector machine. *Geomatics, Natural Hazards and Risk* **9**, 919–938 (2018).
23. Lian, C., Zeng, Z., Yao, W. & Tang, H. Displacement prediction model of landslide based on a modified ensemble empirical mode decomposition and extreme learning machine. *Natural hazards* **66**, 759–771 (2013).
24. Huang, F. *et al*. Landslide displacement prediction using discrete wavelet transform and extreme learning machine based on chaos theory. *Environmental Earth Sciences* **75**, 1376 (2016).
25. Li, Y., Huang, J., Jiang, S.-H., Huang, F. & Chang, Z. A web-based GPS system for displacement monitoring and failure mechanism analysis of reservoir landslide. *Scientific reports* **7**, 17171 (2017).
26. Heidary, M. & Javaherian, A. Wavelet analysis in determination of reservoir fluid contacts. *Computers & Geosciences* **52**, 60–67 (2013).
27. Tomás, R., Li, Z., Lopez-Sanchez, J. M., Liu, P. & Singleton, A. Using wavelet tools to analyse seasonal variations from InSAR time-series data: a case study of the Huangtupo landslide. *Landslides* **13**, 1–14 (2016).
28. Huang, Z., Law, K. T., Liu, H. & Jiang, T. The chaotic characteristics of landslide evolution: a case study of Xintan landslide. *Environmental geology* **56**, 1585–1591 (2009).
29. Takens, F. *Detecting strange attractors in turbulence*. (Springer, 1981).
30. King, G. & Stewart, I. Phase space reconstruction for symmetric dynamical systems. *Physica D: Nonlinear Phenomena* **58**, 216–228 (1992).
31. Tahara, T. *et al*. Asymptotic stability of a modified Lotka-Volterra model with small immigrations. *Scientific Reports* **8**, 7029 (2018).
32. Shi, Y., Yu, D. L., Tian, Y. & Shi, Y. Air–fuel ratio prediction and NMPC for SI engines with modified Volterra model and RBF network. *Engineering Applications of Artificial Intelligence* **45**, 313–324 (2015).
33. Rathinasamy, M., Adamowski, J. & Khosa, R. Multiscale streamflow forecasting using a new Bayesian Model Average based ensemble multi-wavelet Volterra nonlinear method. *Journal of Hydrology* **507**, 186–200 (2013).
34. Zhang, Y. M. & Shi-Ru, Q. U. Adaptive Volterra prediction model for traffic flow based on chaos. *Application Research of Computers* (2010).
35. Huang, F., Huang, J., Jiang, S.-H. & Zhou, C. Prediction of groundwater levels using evidence of chaos and support vector machine. *Journal of Hydroinformatics* **19**, 586–606 (2017).
36. Huang, F., Luo, X. & Liu, W. Stability Analysis of Hydrodynamic Pressure Landslides with Different Permeability Coefficients Affected by Reservoir Water Level Fluctuations and Rainstorms. *Water* **9**, 450 (2017).
37. Liu, W., Xiaoyan, L., Faming, H. & Mingfu, F. Uncertainty of the Soil–Water Characteristic Curve and Its Effects on Slope Seepage and Stability Analysis under Conditions of Rainfall Using the Markov Chain Monte Carlo Method. *Water* **9**, 758-.
38. Huang, F., Chen, L., Yin, K., Huang, J. & Gui, L. Object-oriented change detection and damage assessment using high-resolution remote sensing images, Tangjiao Landslide, Three Gorges Reservoir, China. *Environmental Earth Sciences* **77**, 183 (2018).
39. Jiang, S.-H. *et al*. Modelling of spatial variability of soil undrained shear strength by conditional random fields for slope reliability analysis. *Applied Mathematical Modelling* **63**, 374–389 (2018).
40. Liu, W., Luo, X., Huang, F. & Fu, M. Prediction of soil water retention curve using Bayesian updating from limited measurement data. *Applied Mathematical Modelling* **76**, 380–395 (2019).
41. Li, X. *et al*. Estimation of the Precipitable Water Vapor from ground-based GPS with GAMIT/GLOBK. *IEEE* **1**, 210–214 (2010).
42. Varnes, D. J. Landslide Types and Processes. *Landslides-invesitgation and mitigation* **247**, 36–75 (1996).

43. Sun, M., Tang, H., Wang, M., Shan, Z. & Hu, X. Creep behavior of slip zone soil of the Majiagou landslide in the Three Gorges area. *Environmental Earth Sciences* **75**, 1199 (2016).
44. Wang, S. *et al*. Residual-state creep of clastic soil in a reactivated slow-moving landslide in the Three Gorges Reservoir Region, China. *Landslides* (2018).
45. Liu, Y. & Zhang, Y. X. Application of Optimized Parameters SVM in Deformation Prediction of Creep Landslide Tunnel. *Applied Mechanics &. Materials* **675–677**, 265–268 (2014).
46. Niu, D., Wang, Y. & Wu, D. D. Power load forecasting using support vector machine and ant colony optimization. *Expert Systems with Applications* **37**, 2531–2539 (2010).
47. An, X., Jiang, D., Liu, C. & Zhao, M. Wind farm power prediction based on wavelet decomposition and chaotic time series. *Expert Systems with Applications* **38**, 11280–11285 (2011).
48. Kashani, M. H., Ghorbani, M. A., Dinpasho, Y., Shahmorad, S. & Kundzewicz, Z. W. Comparative study of different wavelets for developing parsimonious Volterra model for rainfall-runoff simulation. *Water Resources* **44**, 568–578 (2017).
49. Maheswaran, R. & Khosa, R. Wavelet Volterra Coupled Models for forecasting of nonlinear and non-stationary time series. *Neurocomputing* **149**, 1074–1084 (2015).
50. Alinia, H. S., Tiampo, K. F. & James, T. S. GPS coordinate time series measurements in Ontario and Quebec, Canada. *Journal of Geodesy* **91**, 653–683 (2017).
51. Dong, D. *et al*. Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis. *Journal of Geophysical Research Solid Earth* **111** (2006).
52. Hassanieh, H., Adib, F., Katabi, D. & Indyk, P. In International Conference on Mobile Computing & Networking (2012).
53. Zhang, A. B., Chen, T. Y., Liu, X. X., Zhang, Y. J. & Yang, Y. T. Monitoring Data Filter and Deformation Information Extraction Based on Wavelet Filter and Empirical Mode Decomposition. *Applied Mechanics & Materials* **742**, 261–271 (2015).
54. Pavlov, A. & Makarov, V. A. Sorting of neural spikes: When wavelet based methods outperform principal component analysis. *Natural Computing An International Journal* **6**, 269–281 (2007).
55. Wickersham, A. J., Li, X. & Lin, M. Comparison of Fourier, principal component and wavelet analyses for high speed flame measurements. *Computer Physics Communications* **185**, 1237–1245 (2014).
56. Liu, C., Lin, P., Zhao, X. & Gao, J. Reducing GPS carrier phase errors in the measurement and position domains for short-distance static relative positioning. *Acta Geodaetica et Geophysica* **51**, 1–13 (2016).
57. Cellmer, S., Wielgosz, P. & Rzepecka, Z. Modified ambiguity function approach for GPS carrier phase positioning. *Journal of Geodesy* **84**, 267–275 (2010).
58. Yuan, Y., Xiaohan, M. A., Qingyue, L. I., Qi, D. U. & Jiahua, L. A study of landslide inducing factors and early warning criterion by analyzing the automatic monitoring curves of the Shuping Landslide. *Hydrogeology & Engineering Geology* (2015).
59. Qiu, J.-D., Huang, J.-H., Liang, R.-P. & Lu, X.-Q. Prediction of G-protein-coupled receptor classes based on the concept of Chou's pseudo amino acid composition: an approach from discrete wavelet transform. *Analytical Biochemistry* **390**, 68–73 (2009).
60. Tiwari, M. K. & Chatterjee, C. A new wavelet–bootstrap–ANN hybrid model for daily discharge forecasting. *Journal of Hydroinformatics* **13**, 500–519 (2011).
61. Nourani, V., Alami, M. T. & Aminfar, M. H. A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Engineering Applications of Artificial Intelligence* **22**, 466–472 (2009).
62. Wolf, A., Swift, J. B., Swinney, H. L. & Vastano, J. A. Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena* **16**, 285–317 (1985).
63. Broock, W., Scheinkman, J. A., Dechert, W. D. & LeBaron, B. A test for independence based on the correlation dimension. *Econometric reviews* **15**, 197–235 (1996).
64. Kennel, M. B. & Abarbanel, H. D. False neighbors and false strands: a reliable minimum embedding dimension algorithm. *Physical Review E* **66**, 026209 (2002).
65. Rosenstein, M. T., Collins, J. J. & De Luca, C. J. A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena* **65**, 117–134 (1993).
66. Grassberger, P. & Procaccia, I. Characterization of strange attractors. *Physical review letters* **50**, 346–349 (1983).
67. Lai, Y.-C. & Lerner, D. Effective scaling regime for computing the correlation dimension from chaotic time series. *Physica D: Nonlinear Phenomena* **115**, 1–18 (1998).
68. Tiao, W. C. Practical approach to investigate the statistics of nonlinear pressure on a high-speed ship by using the Volterra model. *Ocean Engineering* **37**, 847–857 (2010).
69. Chen, S., Billings, S. A. & Luo, W. Orthogonal least squares methods and their application to non-linear system identification. *International Journal of Control* **50**, 1873–1896 (1989).
70. Cortes, C. & Vapnik, V. Support-vector networks. *Machine learning* **20**, 273–297 (1995).
71. Yaseen, Z. M., Kisi, O. & Demir, V. Enhancing long-term streamflow forecasting and predicting using periodicity data component: application of artificial intelligence. *Water resources management* **30**, 4125–4151 (2016).
72. Chadalawada, J. & Babovic, V. Review and comparison of performance indices for automatic model induction. *Journal of Hydroinformatics* **21**, 13–31 (2019).

Acknowledgements

This work was supported by the Project of "Failure mechanism and deformation prediction of rainfall induced landslides in the Wulin mountain area, No. 2017-52", was also supported by the National Natural Science Foundation of China (No. 41572256 and No. 41572292).

Author contributions

Yuanyao Li designed the experiments, analyzed the data and wrote the paper. Ronglin Sun performed the experiments. Kunlong Yin and Yong Xu help Yuanyao Li to perform the experiments and write the paper. Bo Chai contributed reagents/materials/analysis tools. Lili Xiao help modify this paper.

Competing interests

Dr Li's work has been funded by the project of "Failure mechanism and deformation prediction of rainfall induced landslides in the Wulin mountain area, No. 2017-52". He has received compensation as a member of Geological Survey, China University of Geosciences, Wuhan, China. Ronglin Sun, Kunlong Yin, Yong Xu, Bo Chai and Lili Xiao declare no potential conflict of interests.

Additional information

Correspondence and requests for materials should be addressed to Y.L.