

A HYBRID APPROACH FOR PREDICTION OF CHANGES IN LANDSLIDE RATES BASED ON CLUSTERING AND A DECISION TREE

Junwei Ma ^{1, *}, Huiming Tang ², Tao Wen ², Xiao Liu ¹

¹ Three Gorges Research Center for Geo-hazards of Ministry of Education, China University of Geosciences,
Wuhan, Hubei 430074, PR China – majw@cug.edu.cn

² Faculty of Engineering, China University of Geosciences, Wuhan, Hubei 430074, PR China

KEY WORDS: Landslides, Prediction, Two-step cluster, Classification and regression tree (CART), White box model

ABSTRACT:

Forecasting the complex deformation patterns (e.g., displacement, velocity, etc.) of landslides is required to prevent property damage and loss of human lives caused by landslide deformation and failure. In this study, a hybrid approach with clustering and a decision tree is proposed to predict changes in landslide rates. The performance of the hybrid approach is evaluated using multi-parameter monitoring data from the Majiagou landslide, Three Gorges Reservoir, China. A forecasting model consisting of a set of clear, interpretable decision rules was created, and the model achieved a satisfactory accuracy. The results indicate that the hybrid data mining approach can be used to build an explicit representation of the cause-effect relationships hidden in large, complex data sets and generate novel predictions of changes in landslide rates. It is believed that the approach employed in this study could be easily utilized by several categories of users, from beginner to expert, and provide support to improve studies of landslide deformation forecasting. Additionally, the proposed approach could be implemented in other domains characterized by large, complex data sets and cause-effect relationships.

1. INTRODUCTION

Landslides are important and common natural phenomena that occur annually worldwide (Salgueiro 1965; Sidle and Ochiai 2006). Landslide movement and failure can cause substantial damage and loss of life (Ma et al. 2017b, 2018). Forecasting the deformation patterns (e.g., displacement, velocity, etc.) of continuously deforming landslides is considered to an important and economical way of avoiding or reducing losses (Li et al. 2012, Bernardie et al. 2014) and remains a key challenge in natural hazard research. The prediction of landslide deformation patterns is complex because various triggers may influence this phenomenon (Bernardie et al. 2014; Ma et al. 2017b). To date, growth theories and models have been proposed to predict deformation patterns. In this section, we present a bibliometric literature review of landslide deformation forecasting.

The data for this bibliometric literature review were collected from the Science Citation Index Expanded (SCIE) using the “ISI Web of Knowledge” database of Thomson Reuters (version 5.13.1 – Web of Science) on 17 September 2016. Landslide (or slope) deformation (or displacement) forecasting (or prediction) were used as the keywords to search titles, abstracts, and keywords from 1995 to 2016. Carrot2 (<http://project.carrot2.org/>) was employed to analyze the search results for knowledge visualization. The literature analysis performed using Carrot2 yielded a foam tree visualization. Foam tree visualization has the potential to improve the representation quality of search results and provides a useful point of reference so that any substantial discrepancies can be considered (Chen and Chen 2003; Chen et al. 2014). The centralized topics in a foam tree are leading topics, and the size of each topic in a foam tree represents the frequency of the corresponding topic.

Based on 277 search results from the Web of Science, the leading topics regarding landslide deformation forecasting are visualized in Fig. 1. Stability prediction, landslide factors, and computer intelligence methods are among the leading topics in landslide deformation forecasting. These topics can be described in detail as follows.

(1) The objective of landslide deformation forecasting is to predict the deformation trend or stability of a continuously deformed landslide, and ultimately to deploy an early warning system.

(2) The factors that cause landslide deformation include rainfall (Hong et al. 2005; Guzzetti et al. 2007; Berti et al. 2012; Bernardie et al. 2014; Xu et al. 2016b), reservoir variations (Jia et al. 2009; Wang et al. 2004, 2008a, 2008b; Ma et al. 2017c), earthquakes (Khattak et al. 2010; Dai et al. 2011; Song et al. 2012; Xu et al. 2016a), and human activities (Li et al. 2012). Identifying the cause-effect relationships that can be used to accurately predict landslide deformation is of great interest to researchers (Ma et al. 2017b). Considerable efforts have been made to understand the effects of causal factors on landslide instability in natural systems. Several empirical models can describe these cause-effect relationships between causal factors and landslide instabilities, such as the intensity-duration (ID) model, the rainfall event-duration (ED) model, and the rainfall event-intensity (EI) model (Guzzetti et al. 2007). Most models, however, have only focused on a single factor that is not coincident with the fact that landslide movements are seldom linked to a single cause (Aleotti and Chowdhury 1999). Consequently, there are no available advisories or thresholds offered to residents when multiple causal factors are examined. Moreover, no quantitative model has been available to describe the relationships between causal factors and measured deformations. Namely, the aforementioned models do not

* Corresponding author

explicitly consider measured quantities such as displacement rates, as they are based on binary classification (e.g., occurrence or non-occurrence of a landslide).

(3) Computer intelligence methods, such as neural networks (Neaupane and Achet 2004; Lian et al. 2015, 2016; Yao et al. 2014, 2017), extreme learning machines (Lian et al. 2012, 2013, 2014a, 2014c), and support vector machines (Feng et al. 2004; Ren et al. 2015), have become increasingly popular approaches to drive forecasting models of landslide deformation. In building a model for landslide deformation forecasting, it is of fundamental importance to understand the complex landslide mechanisms (Yao et al. 2015). However, the fundamental mechanisms that control causal factors and landslide movements are not yet fully understood (Yao et al. 2015). This often occurs because landslides are complex, nonlinear and dynamic systems (Lian et al. 2014b), and the movements are seldom linked to a single cause (Aleotti and Chowdhury 1999). When deterministic and mechanistic models are lacking, researchers have attempted to use 'black box' models (e.g., network networks, extreme learning machines, support vector machines, etc.) to build landslide deformation models. These methods have allowed researchers to analyze large, incomplete, complex, and multi-parameter data sets.

Although most computer intelligence methods have been widely applied and proven useful for building forecasting models of landslide deformation, most have some limitations. The first limitation of the previously discussed methods is that these models with high prediction accuracy are too complex in terms of freedom. Consequently, it can be impossible to obtain the necessary input parameters at a sufficient forecasting accuracy for cases outside trained regions (Korup and Stolle 2014).

Second, forecasting models such as neural networks, super vector machines, and extreme learning machines are generally considered 'black box' methods (Cortez and Embrechts 2013; Kalteh 2013; Liu et al. 2013; Korup and Stolle 2014), as comprehensive explanations of the processes involved between the input and output stages are seldom provided; however, these methods yield satisfactory accuracy.

Unlike these complex 'black box' models, a decision tree algorithm in the data mining domain is a 'white box' model with an internal structure that can be viewed (Cortez and Embrechts 2013). The objective of a decision tree algorithm is to create a set of simple and clear rules that can be used to predict outcomes from a set of input variables (Ronowicz et al. 2015). Decision trees can be used in forecasting models and provide more insight into cause-effect relationships than do conventional black box models. However, few studies have applied decision tree methods to build forecasting models of landslide deformation.

The objective of this paper is to build a cause-effect model based on a hybrid data mining approach with clustering and a decision tree, which is expressed in explicit form for prediction of changes in landslide rates. The clustering procedure groups similar items based on landslide velocity, and the decision tree builds cause-effect functions between the clusters and multiple causal factors of landslide deformation. The performance of the hybrid approach is evaluated against the multi-parameter monitoring data acquired from the Majiagou landslide, one of the continuously deforming landslides in the Three Gorges Reservoir area, China.

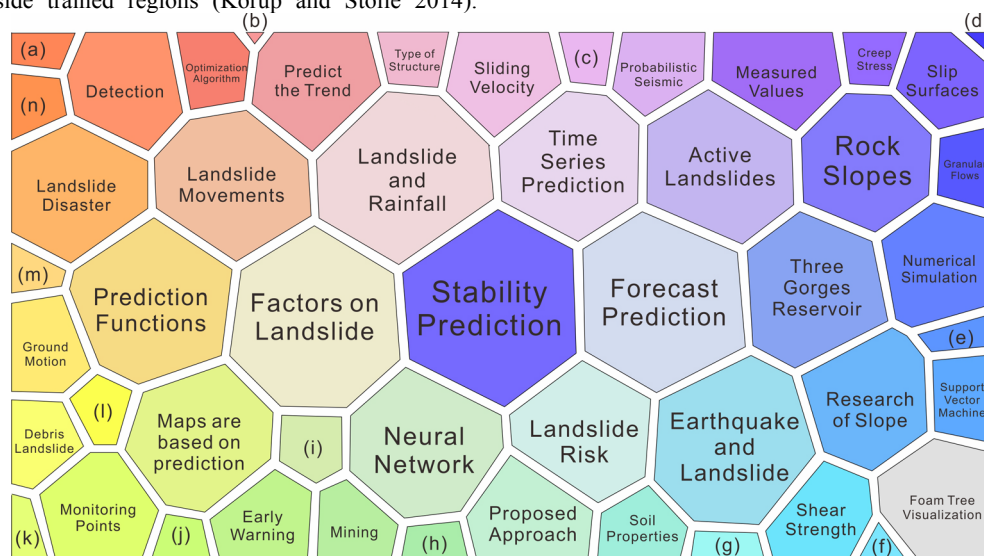


Fig. 1. Foam tree visualization of major topics in landslide deformation pattern forecasting. (a): Root Mean Square Error; (b) Other Topics; (c): Computer Intelligence; (d): Topographic Amplification; (e) Tunnel Construction; (f): Gray System; (g): Curve Fitting; (h): Tsunami Generation; (i): Transient Groundwater Regime; (j): Landslide Sensor; (k): Bed Sediment; (l): Wave Generated; (m): Spatial Scale and Data Accuracy; (n): Extreme Learning Machine.

2. MATERIALS AND METHODS

2.1 Study site characterization

The Majiagou landslide located in the Three Gorges reservoir area in China was chosen as the experimental site (31°01' N, 110°42' E, see Fig. 2(a) and (b) for the location). The volume of the Majiagou landslide is estimated as 3.1 million cubic meters, and more than 132 people reside at the experimental site (Fig.

2(c)). Additionally, continuous deformation has occurred at the experimental site, and two noticeable cracks can be observed in the experimental surface. Therefore, there is an urgent need to build a forecasting model of landslide deformation for early warning.

This experimental area is characterized by narrow, steep valleys delimited by high, rugged mountains (Ma et al. 2017a, b). The

bedrock at the experimental site is mainly purple-red mudstone of the Late Jurassic. Purple-red mudstones are generally highly weathered and fractured. Many landslides have occurred in this rock formation in the Three Gorges Reservoir area due to its unfavorable properties and low strength. The experimental area receives heavy precipitation, with a mean annual value of 1029 mm, which is mainly concentrated between May and September (Ma et al. 2017b). The total rainfall from May to November

accounts for 66% of the annual total rainfall, while the total rainfall from November to March only accounts for 15% of the total. The reservoir level at the experimental site ranges from 145 to 175 m in elevation throughout the year. Rapid reservoir drawdown and prolonged intense rainfall occur annually from May to September. Heavy precipitation and rapid changes in reservoir water level are generally considered the two dominant causal factors of landslide instabilities at the experimental site.

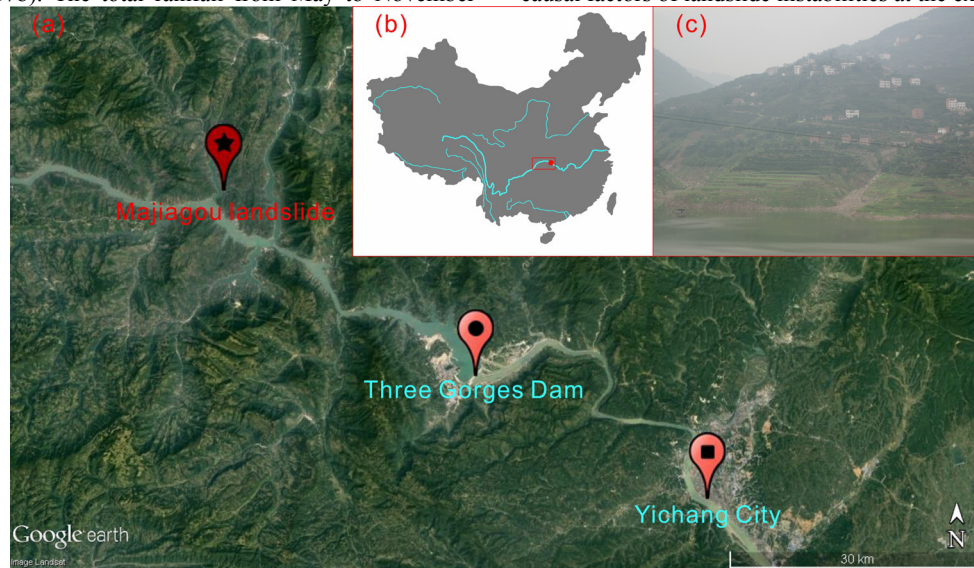


Fig. 2. (a) Location of the experimental site—the Majiagou landslide. (b) Location of the Three Gorges Reservoir area, China. (c) Photograph of the experimental site.

2.2 Landslide monitoring data

Approximately two years of multi-parameter time series were collected between May 2012 and May 2014 (Fig. 3). The collected raw data sets include rainfall intensity, reservoir level, reservoir variation, and landslide displacement at the experimental site. The rainfall, reservoir level, and reservoir variation data were monitored daily, while landslide displacement was surveyed every 5 days. Five-day average velocity was calculated from the raw displacement data. The last

5-day accumulated rainfall intensity was derived from the raw daily rainfall intensity. Raw daily reservoir values and raw daily reservoir variations were transformed into 5-day averages. The 5-day rainfall intensity, 5-day average reservoir level, 5-day average reservoir variation rate, and 5-day average velocity were adopted to build the forecasting model. The available data indicate that the Majiagou landslide is unstable and continuously active, and accelerations with velocities up to 0.75 mm per day can be observed.

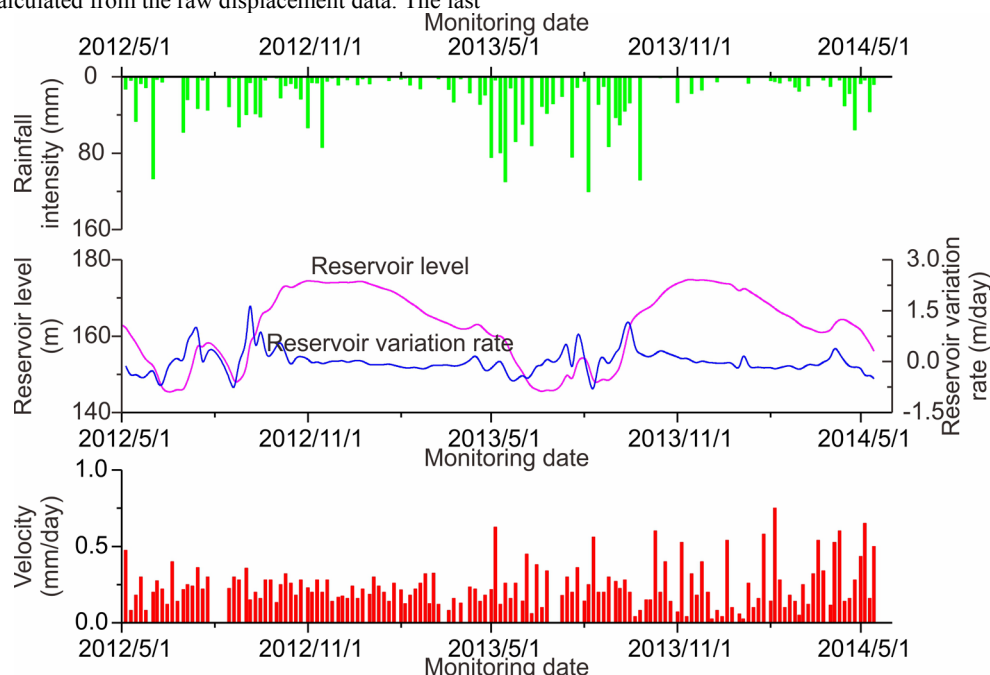


Fig. 3. Monitoring data from the experimental site—the Majiagou landslide

2.3 Prediction of changes in landslide rates using clustering and a decision tree

Artificial neural networks and decision trees are competitive methods that are considered effective prediction tools. However, decision trees outperform artificial neural networks in terms of conceptual understanding. The objective of a decision tree is to establish a set of decision rules that can be used to predict outcomes from a set of inputs. The prediction rules generated by a decision tree are always simple and accurate. Based on these rules, cause-effect relationships hidden in large, complex data sets can be identified (Ronowicz et al. 2015). Moreover, the decision tree method is fast and easy to execute. Therefore, decision trees have been successfully applied for prediction in a variety of fields (Pradhan 2013). Various algorithms have been proposed to build decision tree models, such as ID3, CART (classification and regression tree), CHAID (chi-square automatic interaction detector decision tree), and C5.0 (Pradhan 2013). CART, which is one of the most popular decision tree algorithms, was selected to model the decision tree in this study. A significant advantage of the CART algorithm is that it is a relatively automatic method in decision tree algorithms (Lewis 2000). Additionally, few input parameters are required for the CART algorithm. Because the decision tree method is generally more suitable for predicting categorical outcomes (Tso and Yau 2007), continuously monitored landslide velocity data were first discretized into categorical data via a clustering analysis.

The objective of a clustering analysis is to divide large data sets into groups. Data in a group are similar compared to data in other groups. The two-step cluster is simplistic but reasonably effective for clustering analysis, as this method can address both continuous and categorical data and automatically determine the optimum number of groups (Michailidou et al. 2009). Thus, the two-step cluster method was used in this study to divide the landslide velocity data into groups.

The hybrid approach used to build a forecasting model of changes in landslide rates in this study is illustrated in Fig. 4. First, the landslide velocities were divided into groups using a two-step cluster analysis. The clustering results showed a three-cluster solution based on the 142 inputs. The three clusters were named High (0.433–0.75 mm/day), Medium (0.229–0.40 mm/day), and Low (0.02–0.225 mm/day) according to a qualitative ranking of high, medium, and low performers. Second, a decision tree was built to establish easily understandable forecasting rules using the CART algorithm. A historical data matrix for decision tree modeling contained three input variables (5-day rainfall intensity, 5-day average reservoir level, and 5-day average reservoir variation rate containing continuous values) and one output variable (5-day average velocity containing categorical values). The data matrix was randomly split into training (60%) and testing sets (40%). The training subset was used for model construction, while the testing subset was used to test the predictive ability of the model. The forecasting of landslide deformation is based on decision rules applied to future events.

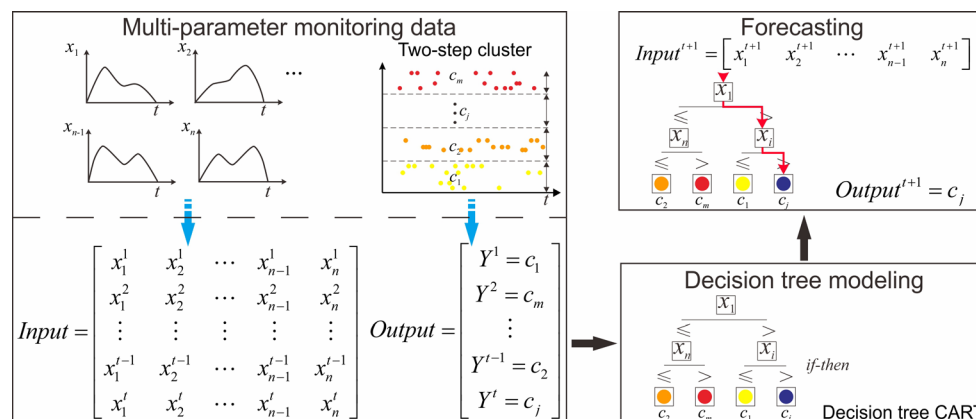


Fig. 4. A hybrid data mining approach for prediction of changes in landslide rates based on clustering and a decision tree

3. RESULTS AND DISCUSSION

The forecasting model of changes in landslide rates is illustrated in the form of a tree graph (Fig. 5). The obtained decision tree includes 19 root nodes (ID=1), 8 internal nodes (blue numbers), and 10 leaves (red numbers). A forecasting rule is an individual path from a root node to a leaf. The generated forecasting rules in Fig. 5 had an overall accuracy of 90%. Clearly, these rules can be used to effectively forecast potential landslide deformation. The rules can also be transformed as sequences of “if *a* & *b* & ... & *i* then *j*” statements. This means that if condition *a*, condition *b*, ..., and condition *i* occur, then outcome *j* occurs. Here, we present some specific rules of the tree.

if $x_1 > 1.1$ & $x_2 > 159.3$ then Low

if $x_1 < 0.6$ & $8.2 < x_2 \leq 109.1$ & $174.5 < x_3$ then Medium

if $x_1 \leq -0.23$ & $x_3 \leq 8.2$ then High

As shown, x_1 , which reflects the reservoir variation rate, is the most powerful variable. This finding is in agreement with the results described by Ma et al. (2017b). The mechanisms of reservoir variations that caused landslide movements can be described as follows. As a reservoir fills, water gradually drains into the landslide body and increases the pore water pressure. This increase in the pore water pressure and the associated softening of materials on the slip surface cause the shear resistance and effective normal stress along the slip surface to decrease, thereby reducing the landslide stability. During periods when the reservoir level drops, support provided by the water at the toe of the landslide decreases, while the piezometric level within the landslide mass may remain constant. In addition, high seepage pressures develop toward the toe of the landslide and reduce the landslide stability. Additionally, nodes 16 and 17 show that when the reservoir dropped at a high rate of 0.23 m per day, the landslide likely moved at a high rate. This result indicates that the threshold of reservoir variation for the landslide mass undergoing rapid deformation is -0.23 m per

day, where ‘-’ indicates a decrease in the reservoir level. This finding corresponds to those of He et al. (He et al. 2008), who found that landslides in the Three Gorges Reservoir area can be reactivated if the reservoir level rapidly decreases at a rate higher than 0.2 m per day. Conversely, landslide movements are relatively small when the reservoir level decreases at a rate of less than 0.2 m per day. These results confirm the usefulness of the decision tree model.

Internal nodes 6 and 7 show that moderate deformation occurred frequently when the reservoir level remained below 159.3 m. Additionally, internal nodes 14 and 15 reflect the trend in nodes 6 and 7, and the landslide mass moves at a moderate velocity more frequently when the reservoir level is low. Nodes 8 and 9 show that the landslide movements tend to be more severe when the experimental site experiences prolonged heavy rainfall, specifically, for a 5-day rainfall intensity exceeding 8.2 mm. This result indicates that the 5-day rainfall threshold required for rapid movement of the landslide mass is 8.2 mm. The mechanisms of rainfall-triggered landslide movements can be briefly described as follows. Landslide deformation triggered by rainfall is caused by increasing the water pressure in the ground. Specifically, rain infiltrates into the landslide body and increases the water pressure during

rainfall events. An increase in water pressure decreases the effective stresses on the slip surface, thereby reducing the shearing resistance and causing the landslide to deform.

Unlike ‘black box’ models such as artificial neural networks, support vector machines, extreme learning machines, etc., the proposed model captures some of the basic relationships hidden in the large, complex data sets and generates some novel predictions of changes in landslide rates. The important advantage of this forecasting model is its simplicity, as it can be easily understood, even by those with limited knowledge regarding such subjects.

The proposed approach works in a semi-automatic manner, as very few input parameters are required to drive the forecasting model, and a satisfactory performance is achieved. Thus, this approach could be utilized by several categories of users, from beginner to expert. Such a hybrid data mining approach benefits a beginner’s forecasting skills, as they can train forecasting models with a sufficient level of prediction accuracy for other practical cases using this approach. Experts can use such a tool to save time, and they can assess the quality of the final model and propose alternative models.

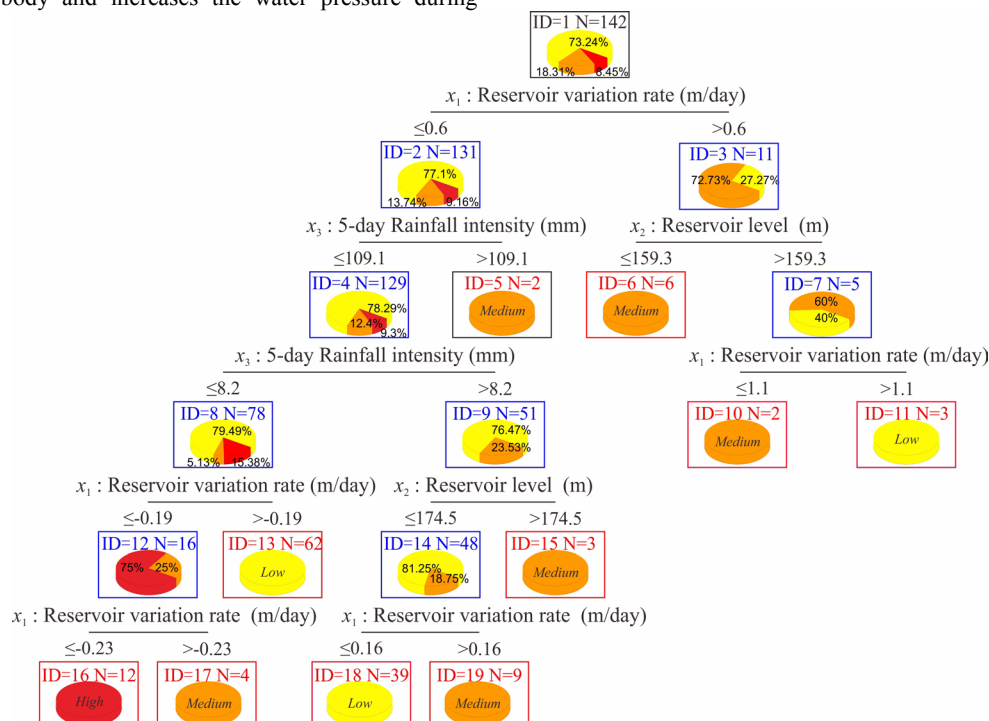


Fig. 5. Illustration of the forecasting model of changes in landslide rates for the Majiagou landslide. ID is the node number of the decision tree, and N is the number of cases in each node.

4. CONCLUSIONS

Predicting the complex deformation patterns of landslides is an important issue for early warning. A hybrid approach using clustering and a decision tree was proposed to interpret time series of multi-parameter landslide data. The hybrid approach proved useful and made it possible to capture the cause-effect relationships hidden in the complex multi-parameter time series and generate some novel predictions of changes in landslide rates. Given the satisfactory accuracy of the trained model, such a hybrid approach could be useful for building forecasting models of changes in landslide rates. In addition to improving accuracy, the trained model can be easily understood and

executed by a wide range of users, from beginner to expert. Notably, the significant advantage of the hybrid data mining approach is its ease of implementation. The authors of this study believe that the methodology described herein will provide support to improve landslide deformation forecasting.

ACKNOWLEDGEMENTS

This study was financially supported by the National Natural Science Foundation of China (41702328), and the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan) (CUGL170813). All support is gratefully acknowledged.

REFERENCES

- Aleotti P, Chowdhury R, 1999. Landslide hazard assessment: summary review and new perspectives. *B ENG GEOL ENVIRON* 58, pp. 21-44. doi: 10.1007/s100640050066
- Bernardie S, Desramaut N, Malet JP, Gourlay M, Grandjean G, 2014. Prediction of changes in landslide rates induced by rainfall. *LANDSLIDES* 12, pp. 481-494. doi: 10.1007/s10346-014-0495-8
- Berti M, Martina MLV, Franceschini S, Pignone S, Simoni A, Pizzolo M, 2012. Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. *J Geophys Res Earth Surf* 117, pp. F04006. doi: 10.1029/2012JF002367
- Chen C, Chen C, 2003. Mapping scientific frontiers-The Quest for Knowledge Visualization (Second Edition). Springer, Berlin
- Chen C, Dubin R, Kim MC, 2014. Emerging trends and new developments in regenerative medicine: a scientometric update (2000-2014). *EXPERT OPIN BIOL TH*, 14, pp. 1295-1317. doi: 10.1517/14712598.2014.920813
- Cortez P, Embrechts MJ, 2013. Using sensitivity analysis and visualization techniques to open black box data mining models. *INFORM SCIENCES*, 225, pp. 1-17. doi: 10.1016/j.ins.2012.10.039
- Dai FC, Xu C, Yao X, Xu L, Tu XB, Gong QM, 2011. Spatial distribution of landslides triggered by the 2008 Ms 8.0 Wenchuan earthquake, China. *J ASIAN EARTH SCI*, 40, pp. 883-895. doi: 10.1016/j.jseas.2010.04.010
- Feng XT, Zhao HB, Li SJ, 2004. Modeling non-linear displacement time series of geo-materials using evolutionary support vector machines. *INT J ROCK MECH MIN*, 41, pp. 1087-1107. doi: 10.1016/j.ijrmms.2004.04.003
- Guzzetti F, Peruccacci S, Rossi M, Stark CP, 2007. Rainfall thresholds for the initiation of landslides in central and southern Europe. *METEOROL ATMOS PHYS*, 98, pp. 239-267. doi: 10.1007/s00703-007-0262-7
- He KQ, Li XR, Yan XQ, Guo D, 2008. The landslides in the Three Gorges Reservoir Region, China and the effects of water storage and rain on their stability. *Environ Geol*, 55, pp. 55-63. doi: 10.1007/s00254-007-0964-7
- Hong Y, Hiura H, Shino K, Sassa K, Suemine A, Fukuoka H, Wang G, 2005. The influence of intense rainfall on the activity of large-scale crystalline schist landslides in Shikoku Island, Japan. *LANDSLIDES*, 2, pp. 97-105. doi: 10.1007/s10346-004-0043-z
- Jia GW, Zhan TLT, Chen YM, Fredlund DG, 2009. Performance of a large-scale slope model subjected to rising and lowering water levels. *ENG GEOL*, 106, pp. 92-103. doi: 10.1016/j.enggeo.2009.03.003
- Kalteh AM, 2013. Monthly river flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform. *COMPUT GEOSCI-UK*, 54, pp. 1-8. doi: 10.1016/j.cageo.2012.11.015
- Khattak GA, Owen LA, Kamp U, Harp EL, 2010. Evolution of earthquake-triggered landslides in the Kashmir Himalaya, northern Pakistan. *GEOMORPHOLOGY*, 115, pp. 102-108. doi: 10.1016/j.geomorph.2009.09.035
- Korup O, Stolle A, 2014. Landslide prediction from machine learning. *Geol Today*, 30, pp. 26-33. doi: 10.1111/gto.12034
- Lewis RJ, 2000. An Introduction to classification and regression tree (CART) analysis. Annual Meeting of the Society for Academic Emergency Medicine, San Francisco, California, pp 1-14
- Li XB, Dong LJ, Zhao GY, Huang M, Liu AH, Zeng LF, Dong L, Chen GH, 2012. Stability analysis and comprehensive treatment methods of landslides under complex mining environment-A case study of Dahu landslide from Linbao Henan in China. *SAFETY SCI*, 50, pp. 695-704. doi: 10.1016/j.ssci.2011.08.049
- Lian C, Zeng ZG, Yao W, Tang HM, 2012. Displacement prediction model of landslide based on ensemble of extreme learning machine. In: Huang T, Zeng Z, Li C and Leung CS (eds) Neural Information Processing: 19th International Conference, ICONIP 2012, Doha, Qatar, November 12-15, 2012, Proceedings, Part IV, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 240-247
- Lian C, Zeng ZG, Yao W, Tang HM, 2013. Displacement prediction model of landslide based on a modified ensemble empirical mode decomposition and extreme learning machine. *NAT HAZARDS*, 66, pp. 759-771. doi: 10.1007/s11069-012-0517-6
- Lian C, Zeng ZG, Yao W, Tang HM, 2014a. Extreme learning machine for the displacement prediction of landslide under rainfall and reservoir level. *STOCH ENVIRON RES RISK ASSESS*, 28, pp. 1957-1972. doi: 10.1007/s00477-014-0875-6
- Lian C, Zeng ZG, Yao W, Tang HM, 2014b. Ensemble of extreme learning machine for landslide displacement prediction based on time series analysis. *Neural Computing and Applications*, 24, pp. 99-107. doi: 10.1007/s00521-013-1446-3
- Lian C, Zeng ZG, Yao W, Tang HM, 2014c. Extreme learning machine for the displacement prediction of landslide under rainfall and reservoir level. *STOCH ENVIRON RES RISK ASSESS*, 28, pp. 1957-1972. doi: 10.1007/s00477-014-0875-6
- Lian C, Zeng ZG, Yao W, Tang HM, 2015. Multiple neural networks switched prediction for landslide displacement. *ENG GEOL*, 186, pp. 91-99. doi: 10.1016/j.enggeo.2014.11.014
- Lian C, Zeng ZG, Yao W, Tang HM, Chen CLP, 2016. Landslide displacement prediction with uncertainty based on neural networks with random hidden weights. *IEEE T NEUR NET LEAR*, pp. 1-13. doi: 10.1109/TNNLS.2015.2512283
- Liu XY, Li P, Gao CH, 2013. Symmetric extreme learning machine. *Neural Computing and Applications*, 22, pp. 551-558. doi: 10.1007/s00521-012-0859-8
- Ma JW, Tang HM, Hu XL, Bobet A, Yong R, Ez Eldin MAM, 2017a. Model testing of the spatial-temporal evolution of a landslide failure. *B ENG GEOL ENVIRON*, 76, pp. 323-339. doi: 10.1007/s10064-016-0884-4

- Ma JW, Tang HM, Hu XL, Bobet A, Zhang M, Zhu TW, Song YJ, Ez Eldin MAM, 2017b. Identification of causal factors for the Majiagou landslide using modern data mining methods. *LANDSLIDES*, 14, pp. 311-322. doi: 10.1007/s10346-016-0693-7
- Ma JW, Tang HM, Liu X, Hu XL, Sun MJ, Song YJ, 2017c. Establishment of a deformation forecasting model for a step-like landslide based on decision tree C5.0 and two-step cluster algorithms: a case study in the Three Gorges Reservoir area, China. *LANDSLIDES*, 14, pp. 1275–1281. doi: 10.1007/s10346-017-0804-0
- Ma JW, Tang HM, Liu X, Wen T, Zhang JR, Tan QW, Fan ZQ, 2018. Probabilistic forecasting of landslide displacement accounting for epistemic uncertainty: a case study in the Three Gorges Reservoir area, China. *LANDSLIDES*, pp. 1-9. doi: 10.1007/s10346-017-0941-5
- Michailidou C, Maheras P, Arseni-Papadimitriou A, Kolyva-Machera F, Anagnostopoulou C, 2009. A study of weather types at Athens and Thessaloniki and their relationship to circulation types for the cold-wet period, part I: two-step cluster analysis. *THEOR APPL CLIMATOL*, 97, pp. 163-177. doi: 10.1007/s00704-008-0057-x
- Neaupane KM, Achet SH, 2004. Use of backpropagation neural network for landslide monitoring: a case study in the higher Himalaya. *ENG GEOL*, 74, pp. 213-226. doi: 10.1016/j.enggeo.2004.03.010
- Pradhan B, 2013. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *COMPUT GEOSCI-UK*, 51, pp. 350-365. doi: 10.1016/j.cageo.2012.08.023
- Ren F, Wu X, Zhang K, Niu R, 2015. Application of wavelet analysis and a particle swarm-optimized support vector machine to predict the displacement of the Shuping landslide in the Three Gorges, China. *ENVIRON EARTH SCI*, 73, pp. 4791-4804. doi: 10.1007/s12665-014-3764-x
- Ronowicz J, Thommes M, Kleinebudde P, Krysiński J, 2015. A data mining approach to optimize pellets manufacturing process based on a decision tree algorithm. *Eur J Pharm Sci*, 73, pp. 44-48. doi: 10.1016/j.ejps.2015.03.013
- Salgueiro PR, 1965. Landslide investigation by means of photogrammetry. *Photogrammetria*, 20, pp. 107-114. doi: 10.1016/0031-8663(65)90008-6
- Sidle RC, Ochiai H, 2006. Landslides: processes, prediction, and land use. American Geophysical Union, Washington D.C., USA
- Song Y, Gong J, Gao S, Wang D, Cui T, Li Y, Wei B, 2012. Susceptibility assessment of earthquake-induced landslides using Bayesian network: A case study in Beichuan, China. *COMPUT GEOSCI-UK*, 42, pp. 189-199. doi: http://dx.doi.org/10.1016/j.cageo.2011.09.011
- Tso GKF, Yau KKW, 2007. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *ENERGY*, 32, pp. 1761-1768. doi: 10.1016/j.energy.2006.11.010
- Wang FW, Zhang YM, Huo ZT, Matsumoto T, Huang BL, 2004. The July 14, 2003 Qianjiangping landslide, Three Gorges Reservoir, China. *LANDSLIDES*, 1, pp. 157-162. doi: 10.1007/s10346-004-0020-6
- Wang FW, Zhang YM, Huo ZT, Peng XM, Araiba K, Wang GH, 2008a. Movement of the Shuping landslide in the first four years after the initial impoundment of the Three Gorges Dam Reservoir, China. *LANDSLIDES*, 5, pp. 321-329. doi: 10.1007/s10346-008-0128-1
- Wang FW, Zhang YM, Huo ZT, Peng XM, Wang SM, Yamasaki S, 2008b. Mechanism for the rapid motion of the Qianjiangping landslide during reactivation by the first impoundment of the Three Gorges Dam reservoir, China. *LANDSLIDES*, 5, pp. 379-386. doi: 10.1007/s10346-008-0130-7
- Xu C, Xu X, Shen L, Yao Q, Tan X, Kang W, Ma S, Wu X, Cai J, Gao M, Li K, 2016a. Optimized volume models of earthquake-triggered landslides. *SCI REP-UK* 6: 29797. doi: 10.1038/srep29797
- Xu Q, Liu HX, Ran JX, Li WH, Sun X, 2016b. Field monitoring of groundwater responses to heavy rainfalls and the early warning of the Kualiangzi landslide in Sichuan Basin, southwestern China. *LANDSLIDES*, pp. 1-16. doi: 10.1007/s10346-016-0717-3
- Yao W, Zeng ZG, Lian C, 2017. Generating probabilistic predictions using mean-variance estimation and echo state network. *NEUROCOMPUTING*, 219, pp. 536-547. doi: 10.1016/j.neucom.2016.09.064
- Yao W, Zeng ZG, Lian C, Tang HM, 2014. A hybrid approach based on reservoir computing for landslide displacement prediction. Control Conference (CCC), 2014 33rd Chinese, IEEE, pp 5026-5030
- Yao W, Zeng ZG, Lian C, Tang HM, 2015. Training enhanced reservoir computing predictor for landslide displacement. *ENG GEOL*, 188, pp. 101-109. doi: 10.1016/j.enggeo.2014.11.008