



# Triggering factors and threshold analysis of baishuihe landslide based on the data mining methods

Fasheng Miao<sup>1</sup> · Yiping Wu<sup>1</sup> · Linwei Li<sup>1</sup> · Kang Liao<sup>1</sup> · Yang Xue<sup>1</sup>

Received: 3 July 2020 / Accepted: 5 November 2020 / Published online: 24 November 2020  
© Springer Nature B.V. 2020

## Abstract

The analysis of landslide monitoring data is important to the study and prediction of landslide deformation but is very challenging. In this research, a data mining method combining two-step clustering, Apriori algorithm and decision tree C5.0 model are proposed, and the Baishuihe Landslide in the Three Gorges Reservoir area is taken as the study case. 6 hydrologic factors related to rainfall and reservoir water level are chosen to carry out the data mining analysis. First, 6 hydrologic triggering factors and the deformation rate of the landslide are clustered by the two-step clustering. Then, the Apriori algorithm is used to mine the association rules between triggering factors and deformation rate. A total of 173 association rules are generated based on the data mining, and 20 rules are selected to be analyzed. At last, the decision tree C5.0 model is built to carry out threshold analysis of hydrologic triggering factors. The results show that monthly cumulative rainfall plays an important role in controlling landslide deformation, and 73.9 mm can be regarded as its threshold. Monthly average water level is the second factor to control landslide deformation. While the monthly maximum daily rainfall has no direct control over the acceleration stage of landslide deformation. The data mining method proposed in this paper has a high accuracy in the study of Baishuihe landslide, which could provide a significant basis for the data analysis and prediction of the accumulative landslide in the Three Gorges Reservoir area.

**Keywords** Baishuihe landslide · Three gorges reservoir · Data mining · Two-step clustering · Apriori algorithm · Decision tree C5.0

---

✉ Yiping Wu  
ypwu@cug.edu.cn

Fasheng Miao  
fsmiao@cug.edu.cn

Linwei Li  
linweili@cug.edu.cn

Kang Liao  
liaokang@cug.edu.cn

Yang Xue  
xueyang@cug.edu.cn

<sup>1</sup> Faculty of Engineering, China University of Geosciences, Wuhan 430074, China

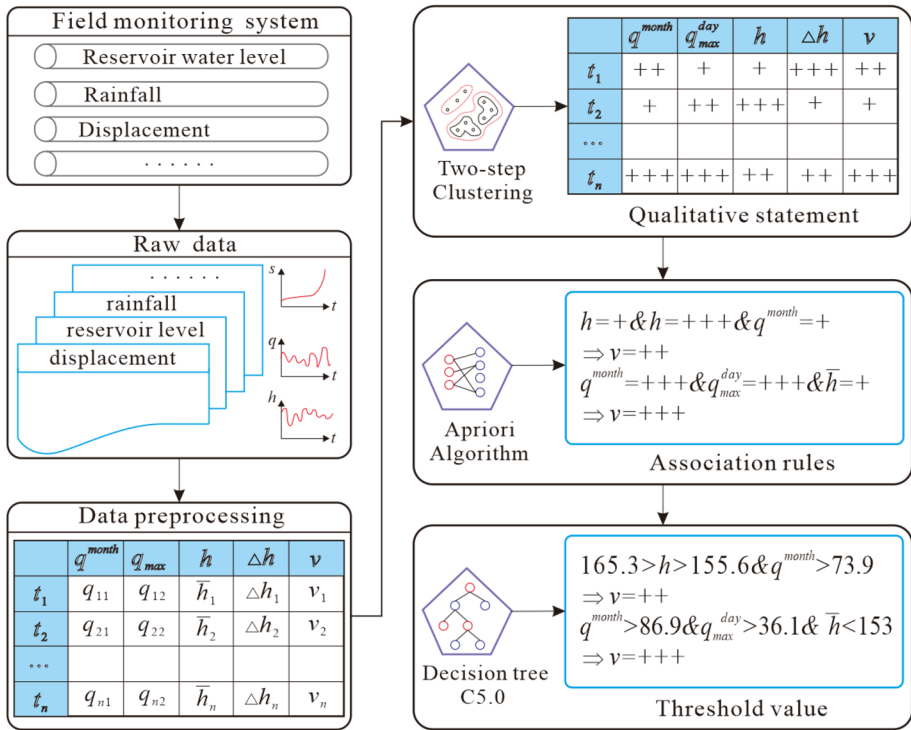
## 1 Introduction

Landslides are one of the worst types of natural disasters, which occur frequently around the world, particularly in mountainous regions (Sassa et al. 2010; Juang et al. 2019). The Three Gorges Reservoir area is in the middle and upper reaches of the Yangtze river. Since the impoundment of water in 2003, the reservoir bank has suffered periodic fluctuation of reservoir water level for a long time, which makes the rock and soil of the slope at the reservoir bank undergo the change of dynamic osmotic pressure repeatedly, thus causing great impact on the surrounding regional geological environment, resulting in the deformation and destruction of the original stable reservoir bank, and leading the reactivation and deformation of many ancient landslides (Tang et al. 2015, 2019). The political, economic, and social status of a large hydropower hub is significant (Wu et al. 2017; Li et al. 2019b, a). Therefore, it is of great significance to carry out the research on the riverbank landslides in the Three Gorges Reservoir area.

The deformation and failure of landslides are the manifestation of the coupling action of internal and external factors. Among them, the internal factors include landform, geological structure, rock and soil properties, etc. External factors include rainfall, reservoir water level, vibration, human activities and so on (Zhou et al. 2018a, b; Yao et al. 2019). For the landslide in the Three Gorges Reservoir area, the periodic change of reservoir water level fluctuated between 145 and 175 m every year, which results in the periodic change of dynamic seepage pressure of the rock and soil all the year round and has a negative impact on the stability of the landslide (Song et al. 2018; Huang et al. 2018). Rainfall is also an important external factor to cause deformation and failure of the landslide (Wang and Sassa 2001; Cao et al. 2020). On the one hand, rainfall infiltration increases the sliding force of slope. On the other hand, rainfall leads to the weakening of rock and soil mass, which reduces the stability of the landslide (Miao et al. 2019; Wang et al. 2020; Wu et al. 2020). Therefore, the reservoir water level and rainfall can be used as the hydrologic triggering factors of landslide deformation and failure in the Three Gorges Reservoir area (Xiong et al. 2019).

With the improvement of monitoring technology and accuracy, a large number of monitoring data are collected by real-time monitoring system of landslide (Zhang et al. 2018). For landslide monitoring data, most of the current research focuses on qualitative analysis or displacement prediction (Shihabudheen et al. 2017; Miao et al. 2018; Intrieri et al. 2019; Li et al. 2019b, a). In terms of data mining analysis, Tsai et al. (2013) used data mining technology to analyze terrain and vegetation factors to verify landslides induced by regional heavy rainfall in Taiwan. And then, decision tree and Bayesian network algorithms were used to extract information from landslides data. Huang et al. (2016) studied the correlation criteria between landslide displacement and reservoir water level and rainfall, and determined the triggering factors of landslide. Wu et al. (2016a, b), Ma et al. (2017a, 2018 and 2020) proposed a data mining method to investigate the hydrological causes of Majiagou landslide, and Apriori algorithm was used to mine association rules to determine the contribution of each hydrological parameter to the landslide movement. Ma et al. (2017b) proposed a hybrid method based on two-step clustering and decision tree C5.0 algorithm to establish a step type landslide deformation prediction model. Wang et al. (2019) proposed an improved parallel mining algorithm for cooperative frequent itemset in multiple data streams.

In this paper, a data mining method combining two-step clustering, Apriori algorithm and decision tree C5.0 were proposed, as shown in Fig. 1. The Baishuihe Landslide in the



**Fig. 1** Flow chart of the data mining process (+, ++, and +++ represent low, medium, and high performance)

Three Gorges Reservoir area was taken as the research object. First, 6 hydrologic induced factors were chosen to carry out the data mining analysis, including monthly cumulative rainfall ( $q^{month}$ ), monthly maximum daily rainfall ( $q_{max}^{day}$ ), monthly maximum continuous rainfall ( $q_{continuous}$ ), monthly average water level ( $\bar{h}$ ), monthly variation of water level ( $\Delta h$ ), and monthly maximum daily variation of water level ( $\Delta h_{max}^{daily}$ ). Then, the two-step clustering was used to cluster six triggering factors and deformation rate of the landslide, and the Apriori algorithm was used to mine the association rules between triggering factors and deformation rate. A total of 173 association rules are generated, and 20 rules are selected to be analyzed. At last, the decision tree C5.0 model was built to carry out threshold analysis of landslide triggering factors. The data mining method proposed in this paper has a high accuracy in the study of Baishuihe landslide monitoring data, which could provide a significant basis for the data analysis and prediction of the accumulative landslide in the Three Gorges Reservoir area.

## 2 Methodology

### 2.1 Two-step clustering

The two-step clustering method realizes the data clustering process by pre-clustering and clustering (Ding et al. 2012; Wu et al. 2016a, b), as shown in Fig. 2. The main

characteristics of the two-step clustering algorithm are: (1) it can deal with both numerical and categorical variables; (2) it can determine the number of clusters according to certain criteria automatically; (3) it can diagnose outliers and noise data in samples. Pre-clustering uses “sequential” method to roughly divide samples into several sub categories. At the beginning, all the data samples are regarded as a large class. After reading a sample data, it is determined whether this sample should be merged into the existing subclass, or a new class will be derived according to the degree of familiarity. This step is repeated, and the final sample data are divided into  $L$  classes. The number of clusters in the pre-clustering process is increasing. Based on pre-clustering, the clustering process also judges whether the sub-clusters generated in the pre-clustering can be merged according to the “degree of affinity” of the samples, and finally the sample data are divided into  $L$  categories. In the process of clustering, the number of clusters is decreasing. For numerical variables, Euclidean distance is usually used in two-step clustering. If the sample data contain both numerical variables and subtype variables, logarithmic likelihood distance should be used.

### 2.2 Apriori algorithm

Apriori algorithm was first proposed by Agrawal and Srikant and has become the core algorithm of association rule mining (Agrawal et al. 1993). This algorithm can only deal with categorical variables and cannot deal with numeric variables (Perego et al. 2001; Guo et al. 2019).

Frequent item set is the item set  $T$  containing item  $a$ . If its support is greater than or equal to the support threshold specified by the user, that is:

$$\frac{|T(a)|}{|T|} \geq \text{min supp} \tag{1}$$

then  $a$  is called frequent item set. The set including 1 item (length 1) is called frequent 1 item set, which is recorded as  $L_1$ . As shown in Fig. 3,  $a, b, c, d$  at the bottom layer can be called frequent 1-term set when the minimum support degree is met. The frequent item set

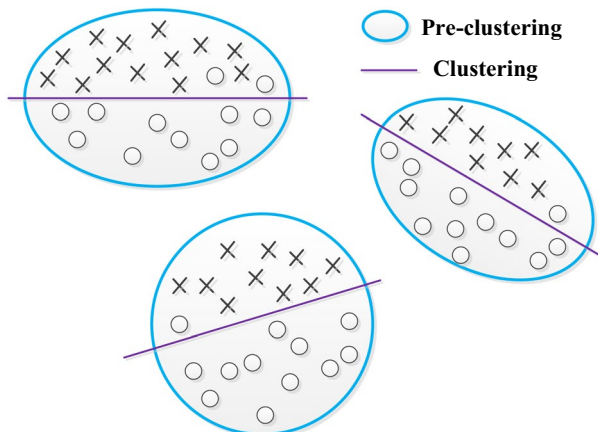


Fig. 2 The algorithm implementation process of two-step clustering

with  $k$  items is called frequent  $k$  item set, which is recorded as  $L_k$ . The upper level project sets  $ab$ ,  $abc$ , and  $abcd$  are frequent  $k$ -item sets when they meet the minimum support.

Apriori algorithm uses the iterative method of layer by layer search to generate frequent item sets. Frequent  $k$ -item sets are used to explore and generate  $(k + 1)$  item sets. The algorithm implementation process is shown in Fig. 3. The frequent item set with a length of 1 is searched out.  $L_1$  is used to generate frequent item set  $L_2$  with a length of 2, and  $L_2$  is used to generate frequent item set  $L_3$  with a length of 3. In this way, all frequent item sets are searched.

### 2.3 Decision tree C5.0

Decision tree model has advantages in estimating process and interpreting parameters (Pandya et al. 2015). Different from other statistical methods, decision tree model does not make statistical assumptions, and can process data representing different scales.

#### 2.3.1 Growth of decision tree C5.0

Decision tree C5.0 is based on the development of ID3 algorithm. The calculation method is derived from the concept of entropy, that is, the average uncertainty of the information source before it is sent out. A node  $n$ , assuming  $n$  is the whole sample set,  $C$  is a set of target variables,  $t$  is the number of  $C$  categories. Then entropy is defined as:

$$Ent(N) = - \sum_i^t p(C_i|N) \log_2 p(C_i|N) \tag{2}$$

where  $p(C_i|N)$  is the relative probability of  $C_i(i = 1, 2, \dots, t)$ . If a variable  $t$  with attribute is divided into  $k$  classes, the conditional entropy after the variable is introduced is defined as:

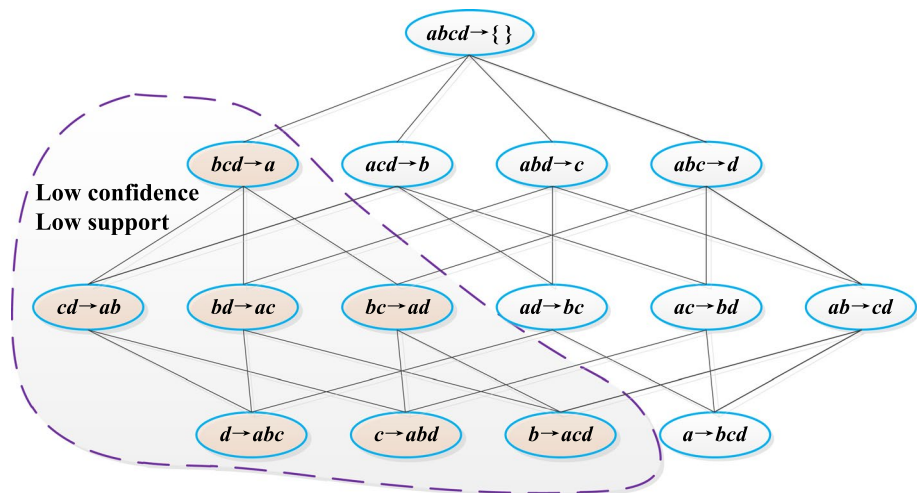


Fig. 3 The algorithm implementation process of Apriori

$$\text{Ent}(N|T) = \sum_j^k \frac{|T_j|}{|N|} \times \text{Ent}(N) \quad (3)$$

The entropy difference between the newly split node and the original node is the information gain, which can be expressed as:

$$\text{Gains}(N, T) = \text{Ent}(N) - \text{Ent}(N|T) \quad (4)$$

Normally,  $\text{Ent}(N) > \text{Ent}(N|T)$ . The degree of random uncertainty of information elimination is represented by information gain. Therefore, the growth of decision tree is determined by selecting the best grouping variable with the maximum information gain rate. Its definition is:

$$\text{GainRatio} = \frac{\text{Gains}(N, T)}{\text{Ent}(T)} \quad (5)$$

### 2.3.2 Pruning rules of decision tree C5.0

Decision tree C5.0 uses statistical confidence interval estimation method to evaluate the error of the training set. If node  $n$  contains samples of  $E_n$  prediction errors, the error rate of this node is:

$$f_n = \frac{E_n}{N} \quad (6)$$

In addition, the estimation error of node  $n$  is defined as:

$$e_n = f_n + z \sqrt{\frac{f_n(1-f_n)}{N}} \quad (7)$$

where  $z$  represents the threshold, which is generally equal to 1.15. On this basis, when the weighted error of the leaf node of the subtree to be pruned is greater than the estimation error of the parent node, the leaf node can be clipped, which is expressed as:

$$\sum_{n=1}^r P_n e_n > e, n = 1, 2, \dots, r \quad (8)$$

where  $r$  is the number of unmodified leaf nodes,  $P_n$  is the ratio of the sample size of leaf nodes to the sample size of subtree, and  $e$  is the estimated error value of the parent node.

## 3 Case study: baishuihe landslide

### 3.1 Geological conditions

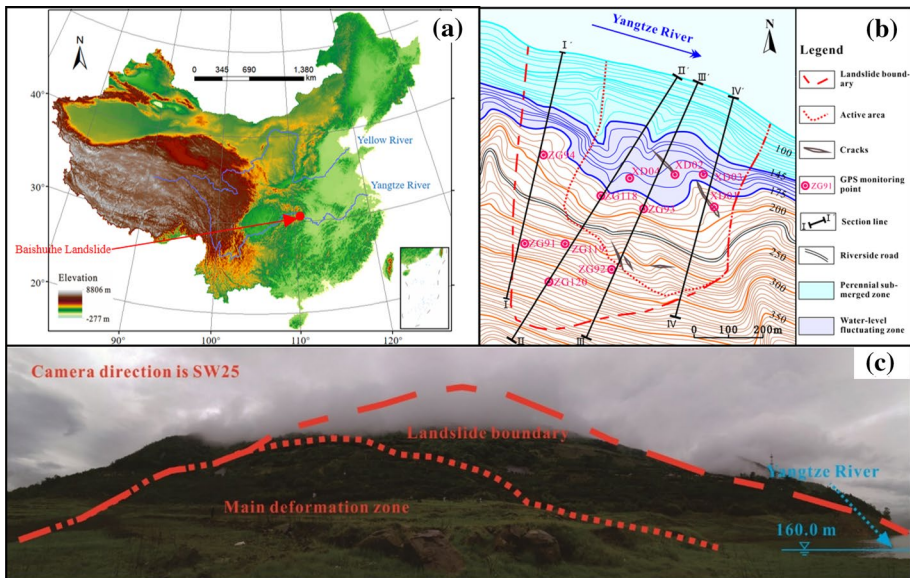
Baishuihe Landslide is located on the right bank of the Yangtze River, which belongs to the Shazhenxi town, Zigui County, Hubei Province (Fig. 4). It is approximately 56 km away from the Three Gorges Dam. Baishuihe landslide is a large-scale ancient cumulative landslide with the average slope inclination of  $30^\circ$  and the average thickness of 30 m.

The volume of the landslide is  $645 \times 10^4 \text{ m}^3$ , covering an area of  $21.5 \times 10^4 \text{ m}^2$ . The main sliding direction is NE15–NE20°. The north–south and east–west length of the landslide are about 600 m and 700 m, respectively. The Baishuihe Landslide formed in a nearly north–south gully with the south higher than the north and spread into the Yangtze River. The gradient of the toe and rear of the landslide is large, and the central portion is flat. In morphology, there is irregular flat concave terrain on both sides of the landslide, which is slightly higher than the middle of the landslide. The toe of the landslide extends to the bed of the Yangtze River, and the crown of the landslide is located at the boundary of the rock and soil with a height of 410 m. From the plane view, the boundary of the landslide is displayed in the shape of an irregular round-backed armchair.

Schematic geological profile of the Baishuihe Landslide is shown in Fig. 5. The materials of the landslide are Quaternary deposits, including silty clay and fragmented rubble with a loose and disorderly structure. The lithologies of the bedrock and strata that crop out around the landslide are mainly Jurassic siltstone, arenaceous shale, and quartz sandstone, with dip directions at 15° and dip angles of 36°. Physical and mechanics parameters of landslide materials were shown in Table 1.

Based on monitoring of surface displacement and the surface deformation characteristics, the Baishuihe Landslide was divided into two major areas in July 2004.

1. The active area (section A, namely main deformation zone or the warning area) is the front part of the landslide and has large deformation. Due to flooding by the reservoir water after the Three Gorges Dam was built, the landslide has obvious displacements, and multiple transverse tension cracks occur in the eastern part.
2. The relatively stable section B is the middle and rear of the landslide, where the accumulated deformation is small, and the deformation rate is slow, only 1.5–4.0 mm/a.



**Fig. 4** a Location of the Baishuihe Landslide; b Topographic map of the Baishuihe Landslide; c Overall view of the Baishuihe Landslide

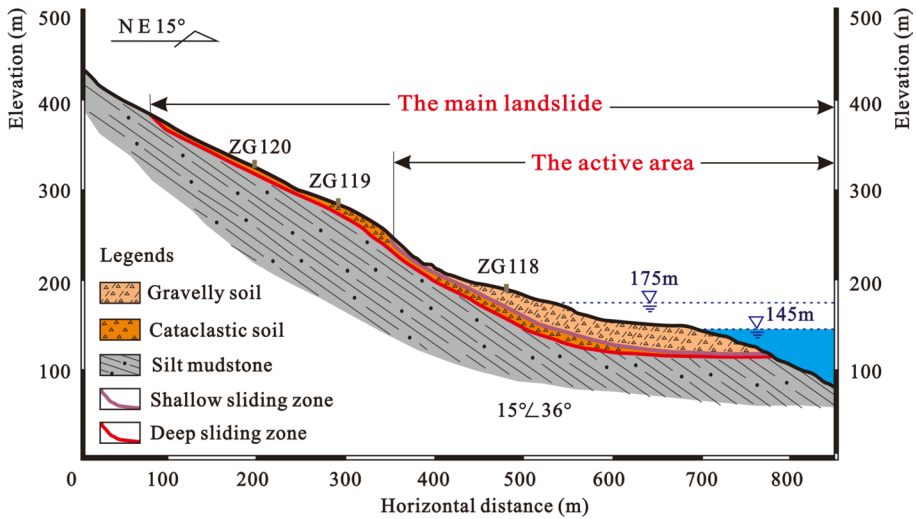


Fig. 5 Schematic geological profile of the Baishuihe Landslide (II–II')

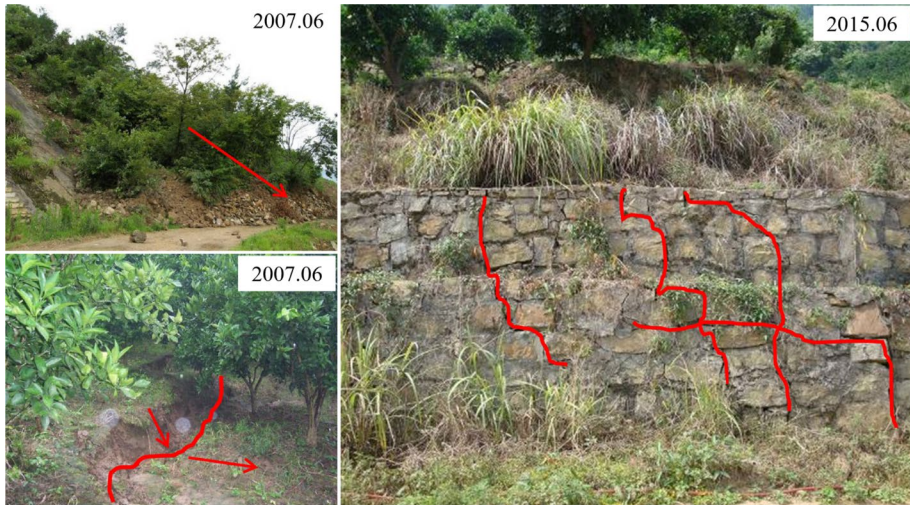
**Table 1** Physical and mechanics parameters of landslide materials

Materials	Unit weight (kN/m <sup>3</sup> )	<i>c</i> (kPa)	<i>φ</i> (°)	Saturated infiltration coefficient (m/s)
Landslide soils	20.7	32.4	18	1.27 E-6
Slip zone soils	22	24.5	16	6.48 E-7
Bedrock	24.8	228	34.2	5.79 E-8

### 3.2 Deformation of the landslide

As an active ancient landslide, Baishuihe landslide has slid for many times. On August 25, 1993, a landslide occurred at the back edge of the landslide, forcing 15 residents to move away. Since 2003, Baishuihe landslide has been warned for many times because of its strong deformation. In 2003, cracks with more than 300 m long were found in the eastern slip tongue, and 4 households were forced to leave. On the morning of June 30, 2007, approximately 100,000 m<sup>3</sup> of landslide piled on the road in the rear of the active area (Fig. 6). The Baishuihe landslide showed obvious deformation during the flood season from 2008 to 2012. By the end of August 2012, the maximum displacement reached 3148.3 mm. From May to August 2015, the retaining wall of Shahuang road at the back edge of the landslide cracked, with a crack width of 1–5 cm (Fig. 6). The deformation of the retaining wall is caused by the creep of the soil along the soil rock interface, which is a local deformation related to rainfall.





**Fig. 6** Macroscopic deformation of the Baishuihe Landslide

### 3.3 Analysis of the monitoring data

Since 2003, a total of 11 GPS displacement monitoring points are deployed on the Baishuihe landslide. Among them, ZG93, ZG118, and XD01 are in the active area, and the monitoring period of them is relatively long. Therefore, these 3 monitoring points can reflect the deformation characteristics of landslide accurately. Thus, in this study, the data of the monitoring points ZG93, ZG118, and XD01 from June 2006 to December 2016 were selected as the research object, as shown in Fig. 7. In June 2003, the water level of the reservoir was raised to 135 m for the first time. Until 2006, the reservoir level has been kept in the 135–140 m range. From 2006 to 2007, the water level of the reservoir has been raised to 155 m and then fluctuated in the 145–155 m. After 2008, the reservoir level was raised to over 170 m, and the normal operation mode of 145–175 m was started after 2010. According to the water level scheduling of the Three Gorges reservoir, the landslide deformation is divided into 3 stages:

1. Phase I (from June 2003 to June 2006): In this phase, the reservoir water level fluctuates between 135–140 m. Although the fluctuation range of the water level is small, the displacement of monitoring points begins to “step-like” increase steadily. The increase of displacement is mainly concentrated in the decline of reservoir water level and the subsequent period. The periodic decline of reservoir water level is the main factor causing the increase of displacement, and the displacement has a certain lag to reservoir level. In July 2005, there was a heavy rainfall process, and the displacement did not increase sharply.
2. Phase II (from July 2006 to June 2008): In this phase, the reservoir water level fluctuates between 145 and 155 m. Among them, during April June 2007, when the reservoir level dropped from 155 to 145 m for the first time, the large drop of the water level led to the increase of the hydrodynamic pressure in the landslide, and changed the seepage field of the landslide, which made the displacement of each monitoring point appear the first sudden increase. The increase of XD01 was more than 1000 mm.

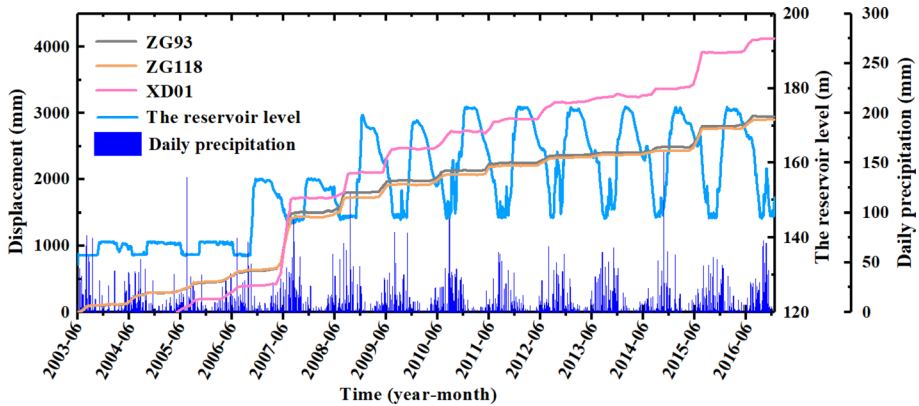


Fig. 7 Long-term monitoring data of the Baishuihe Landslide (displacement, reservoir level, precipitation)

- Phase III (from July 2008 to December 2016): In this phase, the reservoir water level fluctuates between 145 and 175 m. The landslide displacement increases in “step-like”, and the annual growth rate decreases before 2015. The displacement of XD01 monitoring point on the right side of the landslide is significantly larger than that of other monitoring points. The deformation on the right side of the landslide is relatively large, which is consistent with the field investigation. Cracks of Baishuihe landslide are mostly concentrated near the right edge of the landslide mass.

The above analysis shows that the fluctuation of reservoir water level and rainfall are the main factors affecting the deformation of landslide. Therefore, in this study, a total of 6 hydrologic factors were chosen to carry out the data mining analysis, including monthly cumulative rainfall ( $q^{\text{month}}$ ), monthly maximum daily rainfall ( $q_{\text{max}}^{\text{day}}$ ), monthly maximum continuous rainfall ( $q_{\text{continuous}}$ ), monthly average water level ( $\bar{h}$ ), monthly variation of water level ( $\Delta h$ ), monthly maximum daily variation of water level ( $\Delta h_{\text{max}}^{\text{daily}}$ ), as shown in Table 2.

## 4 Results

### 4.1 Clustering results

Based on the two-step clustering algorithm, the 6 triggering factors were clustered. The maximum and minimum categories of each triggering factors were set as 10 and 2,

Table 2 Division of the induced factors for the landslide

Rainfall	Reservoir water level
Monthly cumulative rainfall ( $q^{\text{month}}$ ) (mm)	Monthly average water level ( $\bar{h}$ ) (m)
Monthly maximum daily rainfall ( $q_{\text{max}}^{\text{day}}$ ) (mm)	Monthly variation of water level ( $\Delta h$ ) (m/month)
Monthly maximum continuous rainfall ( $q_{\text{continuous}}$ ) (mm)	Monthly maximum daily variation of water level ( $\Delta h_{\text{max}}^{\text{daily}}$ ) (m/day)

**Table 3** Clustering results of the rainfall factors

Induced factor	Clustering results	
Monthly cumulative rainfall ( $q^{\text{month}}$ ) (mm)	(183.5–517.6)	Heavy Rainfall
	(69.9–179.8)	Moderate-Rainfall
	(3.1–66.1)	Light-Rainfall
Monthly maximum daily rainfall ( $q_{\text{max}}^{\text{day}}$ ) (mm)	(55.9–160.7)	Heavy-daily-rainfall
	(26.5–55.2)	Moderate-Daily-Rainfall
	(1.3–25.6)	Light-Daily-Rainfall
Monthly maximum continuous rainfall ( $q_{\text{continuous}}$ ) (mm)	(110.5–239.4)	Heavy-Effective-Rainfall
	(36.6–109.8)	Moderate -Effective-Rainfall
	(1.5–36.1)	Light-Effective-Rainfall

**Table 4** Clustering results of the reservoir water level factors

Induced factor	Clustering results	
Monthly average water level ( $\bar{h}$ ) (m)	(160.14–174.74)	High-Water-Level
	(144.21–158.47)	Medium-Water-Level
	(135.13–138.95)	Low-Water-Level
Monthly variation of water level ( $\Delta h$ ) (m/month)	(13.26–17.35)	Sharply-Rise
	(7.23–11.36)	Medium-Rise
	(1.57–5.89)	Slowly-Rise
	(– 1.56 to 1.31)	Smooth-Fluctuation
	(– 7.09 to – 3.41)	Medium-Drop
	(– 13.02 to – 8.59)	Sharply-Drop
Monthly maximum daily variation of water level ( $\Delta h_{\text{max}}^{\text{daily}}$ ) (m/day)	(1.66–3.223)	Sharply-Daily-rise
	(0.744–1.513)	Medium-Daily-rise
	(0.063–0.63)	Slowly-Daily-Rise
	(– 0.414 to 0)	Slowly-Daily-Drop
	(– 1.697 to – 0.49)	Medium-Daily-Drop

respectively. The distance measurement method used in the two-step clustering algorithm was Euclidean distance, and the cluster criterion was Bayesian Information Criterion (BIC). Clustering results of the hydrologic factors were shown in Table 3 and 4. Among them, monthly cumulative rainfall ( $q^{\text{month}}$ ) was clustered into Heavy Rainfall (183.5–517.6 mm), Moderate-Rainfall (69.9–179.8 mm), and Light Rainfall (3.1–66.1 mm). Monthly maximum daily rainfall ( $q_{\text{max}}^{\text{day}}$ ) was clustered into Heavy-Daily-Rainfall (55.9–160.7 mm), Moderate-Daily-Rainfall (26.5–55.2 mm), and Light-Daily-Rainfall (1.3–25.6 mm). Monthly maximum continuous rainfall ( $q_{\text{continuous}}$ ) was clustered into Heavy-Effective Rainfall (110.5–239.4 mm), Moderate-Effective Rainfall (36.6–109.8 mm), and Light-Effective Rainfall (1.5–36.1 mm).

Monthly average water level ( $\bar{h}$ ) was clustered into High-Water-Level (160.14–174.74 m), Medium-Water-Level (144.21–158.47 m), and Low-Water-Level (135.13–138.95 m). Monthly variation of water level ( $\Delta h$ ) was clustered into Sharply-Rise (13.26–17.35 m/month), Medium-Rise (7.23–11.36 m/month), Slowly-Rise (1.57–5.89 m/month), Smooth-Fluctuation (– 1.56 to 1.31 m/month), Medium-Drop (– 7.09 to – 3.41 m/

month), and Sharply-Drop ( $-13.02$  to  $-8.59$  m/month). Monthly maximum daily variation of water level ( $\Delta h_{\max}^{\text{daily}}$ ) was clustered into Sharply-Daily-Rise ( $1.66$ – $3.223$  m/day), Medium-Daily-Rise ( $0.744$ – $1.513$  m/day), Slowly-Daily-Rise ( $0.063$ – $0.63$  m/day), Slowly-Daily-Rrop ( $-0.414$  to  $0$  m/day), and Medium-Daily-Drop ( $-1.697$  to  $-0.49$  m/day), as shown in Table 3.

Monthly velocity ( $v$ ) was clustered in Table 5. The initial stage of deformation (Low I) indicates that the monitoring points deform at a rate of  $-0.195$  to  $0.078$  mm/month, which accounts for 42.3% of the total data set. The stable deformation stage (Medium II) indicates that the monitoring points deform at a rate of  $0.092$ – $0.939$  mm/month, which accounts for 40.5% of the total data set. The acceleration deformation (High III) indicates that the monitoring points deform at a rate of  $1.042$ – $10.669$  mm/month, which accounts for 17.2% of the total data set.

## 4.2 Data mining and analysis

In the data mining process, hydrologic factors of the landslide ( $q^{\text{month}}$ ,  $q_{\max}^{\text{day}}$ ,  $q_{\text{continuous}}$ ,  $\bar{h}$ ,  $\Delta h$ ,  $\Delta h_{\max}^{\text{daily}}$ ) are set as the former item of association rules, and the deformation rate (Monthly velocity  $v$ ) is set as the consequent item. The support and confidence threshold of Apriori algorithm are set at 1.5%, 80% to mine the association rules of the Baishuihe landslide. A total of 173 association rules are generated, most of which were I and II stages of the landslide. In these two stages, the deformation rate of the landslide is low, only from  $-0.195$  to  $0.939$  mm/day, and the displacement of monitoring points is almost stable. The III stage of landslide deformation should be paid more attention. Therefore, this paper only lists a few typical association rules about deformation stability stage. Selected association rules were shown in Table 6. The association rules of 1–6, 7–13 and 14–20 are the I, II, III stage, respectively.

Rules 1–6 are the association rules for landslide deformation with a low velocity of  $-0.195$  to  $0.078$  mm/month. Among them, rules 1–3 mean that if the water level is  $160.14$ – $174.74$  m (high-water-level), and one of the three rainfall induced factors is light rainfall ( $q^{\text{month}}$ ,  $q_{\max}^{\text{day}}$ ,  $q_{\text{continuous}}$ ), the landslide is likely to deform at a low rate (I stage). Rules 4–6 indicate that if the monthly variation of water level is between  $1.57$  and  $5.89$  m/month (slowly-rise), and the monthly maximum daily variation of water level is  $0.063$ – $0.63$  m/day (slowly-daily-rise), the landslide is likely to deform at a low rate (I stage).

Rules 7–13 are the association rules for landslide deformation with a medium velocity of  $0.092$ – $0.939$  mm/month. Hydrologic factors included in these rules are mainly the low-medium effective continuous rainfall (Moderate-Rainfall, Light-Daily-Rainfall, Moderate-Daily-Rainfall, Light-Effective-Rainfall, Moderate-Effective-Rainfall) and low-medium rate variation process of reservoir water level (Low-Water-Level, Slowly-Rise, Smooth-Fluctuation, Slowly-Daily-Rise, Medium-Daily-Drop). These factors cannot induce the large deformation of the landslide. The confidence level of each rule is as high as 100%.

**Table 5** Clustering results of the monthly velocity

Monthly velocity ( $v$ ) (mm/month)	Clustering results
(1.042–10.669)	High III
(0.092–0.939)	Medium II
( $-0.195$ to $0.078$ )	Low II

**Table 6** Association rule results for the Baishuihe Landslide

Rule ID	Rule	Support (%)	Confidence (%)	Lift
1	$\bar{h}$ = High-Water-Level & $q_{\text{month}} = \text{Light Rainfall}$ & $q_{\text{max}}^{\text{day}} = \text{Light-daily-rainfall} \Rightarrow \text{I}$	25.15	85.37	2.02
2	$\bar{h}$ = High-Water-Level & $q_{\text{continuous}}^{\text{day}} = \text{Low-Effective Rainfall} \Rightarrow \text{I}$	25.77	85.71	2.02
3	$\bar{h}$ = High-Water-Level & $q_{\text{max}}^{\text{day}} = \text{Light-Daily-Rainfall} \Rightarrow \text{I}$	28.22	86.96	2.05
4	$\Delta h = \text{Slowly-Rise} \ \& \ \Delta h_{\text{max}}^{\text{daily}} = \text{Slowly-Daily-Rise} \ \& \ q_{\text{month}} = \text{Moderate-Rainfall} \Rightarrow \text{I}$	3.07	100.00	2.36
5	$\Delta h = \text{Slowly-Rise} \ \& \ \Delta h_{\text{max}}^{\text{daily}} = \text{Slowly-Daily-Rise} \ \& \ \bar{h} = \text{High-Water-Level} \Rightarrow \text{I}$	2.45	100.00	2.36
6	$\Delta h = \text{Slowly-Rise} \ \& \ \Delta h_{\text{max}}^{\text{daily}} = \text{Slowly-Daily-Rise} \Rightarrow \text{I}$	4.29	100.00	2.36
7	$\Delta h = \text{Sharply-Drop} \ \& \ \Delta h_{\text{max}}^{\text{daily}} = \text{Medium-Daily-Drop} \ \& \ q_{\text{month}} = \text{Moderate-Rainfall} \ \& \ q_{\text{continuous}} = \text{Moderate-Effective Rainfall} \Rightarrow \text{II}$	2.45	100.00	2.47
8	$\Delta h_{\text{max}}^{\text{daily}} = \text{Slowly-Daily-Rise} \ \& \ \bar{h} = \text{Low-Water-Level} \ \& \ q_{\text{continuous}} = \text{Light-Effective Rainfall} \Rightarrow \text{II}$	3.07	100.00	2.47
9	$\Delta h_{\text{max}}^{\text{daily}} = \text{Slowly-Daily-Rise} \ \& \ \bar{h} = \text{Low-Water-Level} \ \& \ q_{\text{continuous}} = \text{Light-Effective Rainfall} \ \& \ q_{\text{max}}^{\text{day}} = \text{Light-Daily-Rain} \Rightarrow \text{II}$	3.07	100.00	2.47
10	$\bar{h} = \text{Low-Water-Level} \ \& \ q_{\text{month}} = \text{Moderate-Rainfall} \ \& \ q_{\text{max}}^{\text{day}} = \text{Light-Daily-Rainfall} \ \& \ \Delta h = \text{Slowly-Rise} \ \& \ \Delta h_{\text{max}}^{\text{daily}} = \text{Slowly-Daily-Rise} \Rightarrow \text{II}$	2.45	100.00	2.47
11	$q_{\text{month}} = \text{Moderate-Rainfall} \ \& \ q_{\text{continuous}} = \text{Moderate-Effective Rainfall} \ \& \ \Delta h = \text{Smooth-Fluctuation} \Rightarrow \text{II}$	2.45	100.00	2.47
12	$\Delta h = \text{Slowly-Rise} \ \& \ \Delta h_{\text{max}}^{\text{daily}} = \text{Medium-Daily-Rise} \ \& \ q_{\text{max}}^{\text{day}} = \text{Moderate-Daily-Rainfall} \ \& \ q_{\text{continuous}} = \text{Moderate-Effective Rainfall} \Rightarrow \text{II}$	1.84	100.00	2.47
13	$\Delta h = \text{Slowly-Rise} \ \& \ \Delta h_{\text{max}}^{\text{daily}} = \text{Medium-Daily-Rise} \ \& \ q_{\text{max}}^{\text{day}} = \text{Moderate-Daily-Rainfall} \ \& \ q_{\text{month}} = \text{Moderate-Rainfall} \Rightarrow \text{II}$	1.84	100.00	2.47
14	$q_{\text{continuous}} = \text{Heavy-Effective Rainfall} \ \& \ \bar{h} = \text{Medium-Water Level} \Rightarrow \text{III}$	6.13	90.00	5.24
15	$q_{\text{continuous}} = \text{Heavy-Effective Rainfall} \ \& \ q_{\text{max}}^{\text{day}} = \text{Heavy-Daily-Rainfall} \ \& \ \bar{h} = \text{Medium-Water-Level} \Rightarrow \text{III}$	5.52	88.89	5.17
16	$q_{\text{continuous}} = \text{Heavy-Effective Rainfall} \ \& \ q_{\text{month}} = \text{Heavy Rainfall} \ \& \ \bar{h} = \text{Medium-Water-Level} \Rightarrow \text{III}$	5.52	88.89	5.17
17	$q_{\text{continuous}} = \text{Heavy-Effective Rainfall} \ \& \ q_{\text{max}}^{\text{day}} = \text{Heavy-Daily-Rainfall} \ \& \ q_{\text{month}} = \text{Heavy Rainfall} \ \& \ \bar{h} = \text{Medium-Water Level} \Rightarrow \text{III}$	5.52	88.89	5.17
18	$q_{\text{continuous}} = \text{Heavy-Effective Rainfall} \ \& \ \bar{h} = \text{Medium-Water-Level} \ \& \ \Delta h = \text{Smooth-Fluctuation} \Rightarrow \text{III}$	3.07	80.00	4.66
19	$q_{\text{max}}^{\text{day}} = \text{Heavy-Daily-Rainfall} \ \& \ q_{\text{month}} = \text{Heavy Rainfall} \ \& \ \bar{h} = \text{Medium-Water-Level} \Rightarrow \text{III}$	6.13	80.00	4.66

Table 6 (continued)

Rule ID	Rule	Support (%)	Confidence (%)	Lift
20	$\Delta t = \text{Medium-Drop} \& q_{\text{max}}^{\text{day}} = \text{Moderate-Daily-Rainfall} \& q^{\text{month}} = \text{Moderate-Rainfall} \Rightarrow \text{III}$	3.07	80.00	4.66

This confidence index shows that when the front item of the rule occurs, the back item of the rule will be sure to occur.

Rules 14–20 are the association rules for landslide deformation with a high velocity of 1.042–10.669 mm/month. The hydrological induced factors included in the rules are mainly composed of heavy rainfall and high effective rainfall. These rules mean that when one of the three rainfall induced factors is reach to heavy rainfall( $q^{\text{month}}$ ,  $q_{\text{max}}^{\text{day}}$ ,  $q_{\text{continuous}}$ ), the landslide will deform at a high velocity (III stage), which indicates that rainfall controls the deformation rate of the Baishuihe Landslide.

### 4.3 Threshold values of the induced factors

In the Decision Tree C5.0 model, the hydrological factors of the landslide ( $q^{\text{month}}$ ,  $q_{\text{max}}^{\text{day}}$ ,  $q_{\text{continuous}}$ ,  $\bar{h}$ ,  $\Delta h, \Delta h_{\text{max}}^{\text{daily}}$ ) are set as the input parameters, and the deformation rate is set as the output parameter. The 80% of the total data is defined as training samples to build the decision tree model, and the rest 20% data are set as the testing samples to check the accuracy of the model. In order to improve the generalization ability of the model and prevent the model from over fitting, this paper adopts the method of combining cross validation and boosting technology in the construction of decision tree C5.0 model. The number of tests for boosting is set to 10, the number of cross validation folds is set to 10, and the expected noise is set to 10%. In this paper, a total of 8 decision tree models are built, and the model with the highest accuracy is selected for analysis. In this model, only 3 hydrological factors are contained, including  $q^{\text{month}}$ ,  $q_{\text{max}}^{\text{day}}$ ,  $\bar{h}$ . The importance degree of each factor is shown in Fig. 8. The monthly cumulative rainfall ( $q^{\text{month}}$ ) plays a significant role in controlling landslide deformation.

A total of 10 threshold criteria for deformation characteristics of Baishuihe landslide have been established in the decision tree model, as shown in Table 7. The threshold criteria of 1–4, 5–8, and 9–10 is the I, II, III stage, respectively.

The threshold criteria of 1–4 indicate that the deformation of the landslide is in the I stage. Criterion 1 can be interpreted as: when the average value of reservoir water level is less than 151.94 m, the landslide area suffers the accumulated rainfall with the intensity less than 23.6 mm, and the monthly maximum value of daily rainfall is less

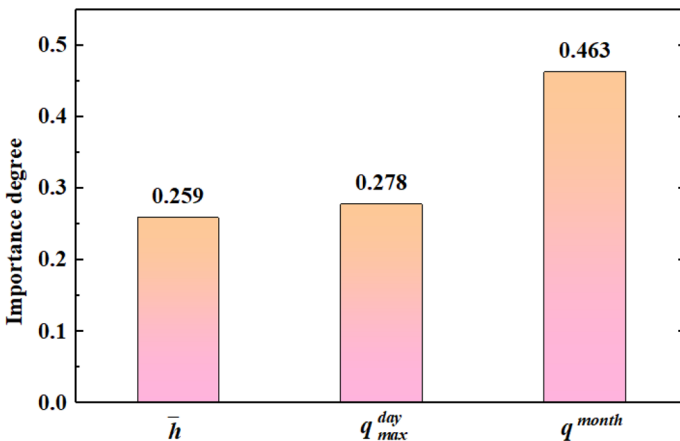


Fig. 8 The importance degree of each induced factor in decision tree C5.0 model

than 28.1 mm, the landslide enters the initial stage of deformation. Criterion 2 can be interpreted as: the slope monitoring point enters the initial stage of deformation when the intensity of accumulated rainfall in the landslide area is greater than 73.9 mm, the average value of reservoir water level is greater than 155.55 m, and the monthly maximum value of daily rainfall is less than 36.1 mm. The prediction accuracy of this criterion is as high as 92.8%. Criterion 3 can be interpreted as: the landslide enters the initial stage of deformation when the intensity of accumulated rainfall in the landslide area is less than 73.9 mm, the average value of reservoir water level is less than 151.94 m, and the monthly maximum value of daily rainfall is greater than 28.1 mm. The accuracy of this criterion is very high, reaching 100%. Criterion 4 can be interpreted as follows: the slope monitoring point enters the initial stage of deformation when the strength of the landslide area is less than 73.9 mm, and the average water level of the reservoir is more than 151.94 m. The number of samples of this criterion is the largest, and the accuracy is 83.3%.

Criteria 5–8 indicate that the deformation of the landslide is in the II stage. Criterion 5 can be interpreted as follows: the landslide area suffers from the accumulated rainfall with the intensity greater than 73.9 mm, and the landslide maintains the II stage when the average reservoir water level is less than 149.91 m. The accuracy of this criterion is 73.2%, and many deformation examples are included. Criterion 6 can be interpreted as: when the average water level of the reservoir is less than 151.94 m, the landslide area suffers rainfall with the intensity of 23.6–73.9 mm, and the monthly maximum rainfall is less than 28.1 mm, the landslide maintains the II stage. The accuracy of the criterion is high, reaching 100%. Criterion 7 can be interpreted as: when the mean value of reservoir water level is greater than 155.55 m, the landslide area suffers rainfall with intensity greater than 73.9 mm, and the monthly maximum value of daily rainfall is greater than 36.1 mm, the landslide maintains the II stage. The accuracy of the criterion is high, reaching 100%. Criterion 8 can be interpreted as: when the average water level of the reservoir is 151.11–153.12 m, and the landslide area suffers rainfall with intensity greater than 73.9 mm, the landslide maintains the II stage.

**Table 7** Threshold values of the induced factors based on decision tree C5.0 model

Deformation velocity	Number	Criterion	Accuracy (%)	Accuracy
I	1	$q^{month} \leq 23.6 \ \& \ \bar{h} \leq 151.94 \ \& \ q_{max}^{day} \leq 28.1 \Rightarrow v=I$	66.7 (4/6)	Medium
	2	$q^{month} > 73.9 \ \& \ \bar{h} > 155.55 \ \& \ q_{max}^{day} \leq 36.1 \Rightarrow v=I$	92.8(13/14)	High
	3	$q^{month} \leq 73.9 \ \& \ \bar{h} \leq 151.94 \ \& \ q_{max}^{day} > 28.1 \Rightarrow v=I$	100(2/2)	Extreme High
	4	$q^{month} \leq 73.9 \ \& \ \bar{h} > 151.94 \Rightarrow v=I$	83.3(35/42)	High
II	5	$q^{month} > 73.9 \ \& \ \bar{h} \leq 149.91 \Rightarrow v=II$	73.2(30/41)	High
	6	$23.6 < q^{month} \leq 73.9 \ \& \ \bar{h} \leq 151.94 \ \& \ q_{max}^{day} \leq 28.1 \Rightarrow v=II$	100(2/2)	Extreme High
	7	$q^{month} > 73.9 \ \& \ \bar{h} > 155.55 \ \& \ q_{max}^{day} > 36.1 \Rightarrow v=II$	100(4/4)	Extreme High
	8	$q^{month} > 73.9 \ \& \ 151.11 < \bar{h} \leq 153.12 \Rightarrow v=II$	80(4/5)	High
III	9	$q^{month} > 73.9 \ \& \ 149.91 < \bar{h} \leq 151.11 \Rightarrow v=III$	71.4(5/7)	High
	10	$q^{month} > 73.9 \ \& \ 153.12 < \bar{h} \leq 155.55 \Rightarrow v=III$	100(7/7)	Extreme High



Criteria 9–10 are the criteria for the landslide to enter the deformation acceleration stage (III). Criterion 9 can be interpreted as: when the average reservoir water level is 149.91–151.11 m, and the landslide area suffers rainfall with intensity greater than 73.9 mm, the landslide enters the deformation acceleration stage. Criterion 10 can be interpreted as: when the average reservoir water level is 153.12–155.55 m, and the landslide area suffers rainfall with intensity greater than 73.9 mm, the landslide enters the deformation acceleration stage. The accuracy of the criterion is as high as 100%.

Accuracy rate of the training and testing samples based on decision tree C5.0 model is shown in Table 8. In the data mining process, 130 samples were selected for training, of which the correct number was 106 (81.5%). 33 samples were selected for testing, of which the correct number was 28 (84.9%). The accuracy of training samples and test samples is higher than 80%, which can be used as the basis of judgment. The deformation and failure of landslide are affected by many external factors. In addition to the reservoir water level and rainfall selected in this paper, the landslide is also affected by human engineering activities. This is also the main reason for the training and prediction errors in this paper.

## 5 Discussion

Baishuihe landslide is a large-scale consequent slope, which is classified as traction landslide in the stress form. The deformation and failure mode can be basically divided into the following stages: (1) the long-term effect of reservoir water level causes the toe slope to soften and the slide residence section to gradually fail; (2) after the loss of the original sliding residence force at the toe slope, the upper rock, and soil mass lose their original support. At the same time, the rainfall and seepage pressure of reservoir water level cause the deformation and weakening of soil mass in the fluctuation zone, and the unstable area is moving upward; (3) when the deformation develops to a certain stage, the state of soil reaches the limit, resulting in tension shear cracks and stress redistribution; (4) when the cracks extend to the rear slope, the whole landslide is in an extreme unstable state, and it is likely to slide along the sliding surface under the action of external factors.

In general, the acceleration stage of landslide deformation is the focus of researchers and engineers. According to the data mining results, monthly cumulative rainfall ( $q^{\text{month}}$ ) plays an important role in controlling landslide deformation. Monthly cumulative rainfall of 73.9 mm can be regarded as the threshold of rainfall (Table 6). When the rainfall in that month does not reach this threshold, the Baishuihe landslide will not enter the deformation acceleration stage. The second factor to control landslide deformation is the monthly average water level ( $\bar{h}$ ). According to monitoring data of Baishuihe landslide, generally, the rainfall is concentrated in June to September every year, and the reservoir will be controlled from high water level to low water level before this period. In other words, there is a certain negative correlation between monthly rainfall and monthly average water level elevation. This period is also the time when the landslide deformation is severe. Therefore, the monthly average water level of 155.55 m can be regarded as the threshold of  $\bar{h}$ . When the reservoir level in that month

**Table 8** Accuracy rates for the training and testing samples based on decision tree C5.0 model

	Training sample	Percent (%)	Testing sample	Percent (%)
Correct	106	81.5	28	84.9
Wrong	24	18.5	5	15.1

does not reach this threshold, the Baishuihe landslide will not enter the deformation acceleration stage. Monthly maximum daily rainfall ( $q_{\max}^{\text{day}}$ ) usually has an impact on the I and II states of landslide deformation, but has no direct control over the acceleration stage of landslide deformation.

## 6 Conclusion

In this research, a data mining method combining two-step clustering, Apriori algorithm and decision tree C5.0 was proposed. The Baishuihe Landslide in the Three Gorges Reservoir area was taken as the research object to analyze the relationship between induced factors and landslide deformation. The following conclusions can be reached:

1. The fluctuation of reservoir water level and rainfall were the main factors affecting the deformation of landslide, and 6 hydrologic factors were chosen to carry out the data mining analysis, including three factors related to rainfall, and three factors related to reservoir water level.
2. A total of 173 association rules were generated based on the data mining, and 20 rules were selected to be analyzed. The association rules showed that rainfall controls the deformation rate of the Baishuihe Landslide.
3. Monthly cumulative rainfall played an important role in controlling landslide deformation, and 73.9 mm can be regarded as its threshold. Monthly average water level was the second factor to control landslide deformation. The monthly maximum daily rainfall had no direct control over the acceleration stage of landslide deformation.
4. The data mining method proposed in this paper has a high accuracy in the study of Baishuihe landslide monitoring data. Therefore, it is of great significance for the data analysis and prediction of the accumulative landslide in the Three Gorges Reservoir area.

**Acknowledgements** This research was supported by the National Natural Science Foundation of China (42007267, 41977244), the Fundamental Research Funds for the Central University, China University of Geosciences (Wuhan), and the National Key R&D Program of China (2017YFC-1501301). The authors thank the colleagues in our laboratory for their constructive comments and assistance.

## Compliance with ethical standards

**Conflict of interest** We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work. There is no professional or other personal interest of any nature or kind in any product, service or company that could be construed as influencing the position presented in, or the review of the manuscript entitled “Triggering Factors and Threshold Analysis of Baishuihe Landslide based on the Data Mining Methods”.

## References

- Agrawal R, Imieliński T, Swami A (1993). Mining association rules between sets of items in large databases. In: Proceedings of the 1993 ACM SIGMOD international conference on management of data (p 207–216)
- Cao Y, Yin K, Zhou C, Ahmed B (2020) Establishment of landslide groundwater level prediction model based on GA-SVM and influencing factor analysis. *Sensors* 20(3):845

- Ding L, Gonzalez-Longatt FM et al (2012) Two-step spectral clustering controlled islanding algorithm. *IEEE Trans Power Syst* 28(1):75–84
- Guo W, Zuo X, Yu J et al (2019) Method for mid-long-term prediction of landslides movements based on optimized apriori algorithm. *Appl Sci* 9(18):3819
- Huang H, Yi W, Lu S et al (2016) Use of monitoring data to interpret active landslide movements and hydrological triggers in three gorges reservoir. *J Perform Constr Facil* 30(1):C4014005
- Huang D, Gu DM, Song YX et al (2018) Towards a complete understanding of the triggering mechanism of a large reactivated landslide in the three gorges reservoir. *Eng Geol* 238:36–51
- Intrieri E, Carlà T, Gigli G (2019) Forecasting the time of failure of landslides at slope-scale: a literature review. *Earth-science Rev* 193:333–349
- Juang CH, Dijkstra T, Wasowski J et al (2019) Loess geohazards research in china: advances and challenges for mega engineering projects. *Eng Geol* 251:1–10
- Li H, Xu Q, He Y et al (2019a) Modeling and predicting reservoir landslide displacement with deep belief network and EWMA control charts: a case study in Three Gorges Reservoir. *Landslides* 17(3):1–15
- Li S, Xu Q, Tang M et al (2019b) Characterizing the spatial distribution and fundamental controls of landslides in the three gorges reservoir area, China. *Bull Eng Geol Environ* 78(6):4275–4290
- Ma, J., Niu, X., Tang, H., Wang, Y., Wen, T., Zhang, J. (2020). Displacement prediction of a complex landslide in the Three Gorges Reservoir Area (China) using a hybrid computational intelligence approach. *Complexity* 2020
- Ma J, Tang H, Hu X et al (2017) Identification of causal factors for the Majiagou landslide using modern data mining methods. *Landslides* 14(1):311–322
- Ma J, Tang H, Liu X et al (2017) 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(3):1275–1281
- Ma J, Tang H, Liu X, Wen T, Zhang J, Tan Q, Fan Z (2018) Probabilistic forecasting of landslide displacement accounting for epistemic uncertainty: a case study in the Three Gorges Reservoir area China. *Landslides* 15(6):1145–1153
- Miao F, Wu Y, Li L et al (2019) Weakening laws of slip zone soils during wetting–drying cycles based on fractal theory: a case study in the Three Gorges Reservoir (China). *Acta Geotech*. <https://doi.org/10.1007/s11440-019-00894-8>
- Miao F, Wu Y, Xie Y et al (2018) Prediction of landslide displacement with step-like behavior based on multialgorithm optimization and a support vector regression model. *Landslides* 15(3):475–488
- Pandya R, Pandya J (2015) C5. 0 algorithm to improved decision tree with feature selection and reduced error pruning. *Int J Comput Appl* 117(16):18–21
- Perego R, Orlando S, Palmerini P (2001) Enhancing the apriori algorithm for frequent set counting. In: international conference on data warehousing and knowledge discovery (p. 71–82). Springer, Berlin, Heidelberg
- Sassa K, Nagai O, Solidum R et al (2010) An integrated model simulating the initiation and motion of earthquake and rain induced rapid landslides and its application to the 2006 Leyte landslide. *Landslides* 7(3):219–236
- Shihabudheen KV, Pillai GN, Peethambaran B (2017) Prediction of landslide displacement with controlling factors using extreme learning adaptive neuro-fuzzy inference system (ELANFIS). *Appl Soft Comput* 61:892–904
- Song K, Wang F, Yi Q et al (2018) Landslide deformation behavior influenced by water level fluctuations of the Three Gorges Reservoir (China). *Eng Geol* 247:58–68
- Tang H, Li C, Hu X et al (2015) Evolution characteristics of the Huangtupo landslide based on in situ tunneling and monitoring. *Landslides* 12(3):511–521
- Tang H, Wasowski J, Juang CH (2019) Geohazards in the three Gorges Reservoir area, China-lessons learned from decades of research. *Eng Geol* 261:105267
- Tsai F, Lai JS, Chen WW et al (2013) Analysis of topographic and vegetative factors with data mining for landslide verification. *Ecol Eng* 61:669–677
- Wang G, Sassa K (2001) Factors affecting rainfall-induced flowslides in laboratory flume tests. *Geotechnique* 51(7):587–599
- Wang J, Xiao L, Zhang J et al (2020) Deformation characteristics and failure mechanisms of a rainfall-induced complex landslide in Wanzhou county, Three Gorges Reservoir. *China Landslides* 17(2):419–431
- Wang Q, Wang X (2019) Improved algorithm for parallel mining collaborative frequent itemsets in multiple data streams. *Clust Comput* 22(3):6133–6141
- Wu LZ, Zhu SR, Peng J (2020) Application of the Chebyshev spectral method to the simulation of groundwater flow and rainfall-induced landslides. *Appl Math Model* 80:408–425

- Wu X, Zhan FB, Zhang K, Deng Q (2016) Application of a two-step cluster analysis and the apriori algorithm to classify the deformation states of two typical colluvial landslides in the Three Gorges. *China Environ Earth Sci* 75(2):146
- Wu X, Zhan FB, Zhang K et al (2016) Application of a two-step cluster analysis and the apriori algorithm to classify the deformation states of two typical colluvial landslides in the Three Gorges. *China Environ Earth Sci* 75(2):146
- Wu Y, Miao F, Li L et al (2017) Time-varying reliability analysis of huangtupo riverside No. 2 landslide in the Three Gorges Reservoir based on water-soil coupling. *Eng Geol* 226:267–276
- Xiong X, Shi Z, Xiong Y et al (2019) Unsaturated slope stability around the Three Gorges Reservoir under various combinations of rainfall and water level fluctuation. *Eng Geol* 261:105231
- Yao W, Li C, Zuo Q et al (2019) Spatiotemporal deformation characteristics and triggering factors of Baijiabao landslide in Three Gorges Reservoir region, China. *Geomorphology* 343:34–47
- Zhang Y, Hu X, Tannant DD et al (2018) Field monitoring and deformation characteristics of a landslide with piles in the Three Gorges Reservoir area. *Landslides* 15(3):581–592
- Zhou C, Yin K, Cao Y, Ahmed B, Fu X (2018) A novel method for landslide displacement prediction by integrating advanced computational intelligence algorithms. *Sci Rep* 8(1):1–12
- Zhou C, Yin K, Cao Y, Intriери E, Ahmed B, Catani F (2018) Displacement prediction of step-like landslide by applying a novel kernel extreme learning machine method. *Landslides* 15(11):2211–2225

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.