



Torrential rainfall-triggered shallow landslide characteristics and susceptibility assessment using ensemble data-driven models in the Dongjiang Reservoir Watershed, China

Jie Dou, et al. *[full author details at the end of the article]*

Received: 31 May 2018 / Accepted: 2 July 2019 / Published online: 10 July 2019
© Springer Nature B.V. 2019

Abstract

This study investigated the characteristics of rainfall-triggered landslides during the Typhoon Bilis in the Dongjiang Reservoir Watershed, China. The comparative shallow landslide susceptibility mappings (LSMs) were produced by the ensemble data-driven statistical models in a GIS environment. At first, the landslide inventory for the study area was prepared from the high-resolution QuickBird images, and China–Brazil Earth Resources Satellite images, and field survey. Other necessary data for landslide susceptibility analysis such as the amount of rainfall, geology, and topography were also collected from the respective agencies. Twelve predisposing factors were then prepared using this available dataset. To reduce the subjectivity of models and caution in the selection of predisposing factors, and to avoid the spatial autocorrelation redundancy, certainty factor approach was attempted to optimize these twelve set of parameters. For validating the accuracy of the model, the original landslide data were randomly divided into two parts: 70% (1545 landslides) for training the model and the remaining 30% (662 landslides) for validation. The verified results showed that using the optimized predisposing factors has a higher performance than using all the original twelve factors. The results of ensemble models also showed that LSM maps prepared using binary logistic regression (accuracy is 0.848) model are more accurate than those prepared using bivariate statistical analysis (accuracy is 0.837) model. Additionally, our analysis concludes that the short duration and high-intensity rainfall, drainage density, lithology, and curvature are the major influencing factors for landslide occurrences in this case study area. This research provides an improved understanding of the mechanism of landslides caused by the typhoons for the adjoining watersheds nearby the reservoir. The preliminary understandings and approach could also be applied in similar geological and rainfall-triggered case study sites in the other parts of the world for risk mitigation.

Keywords Shallow landslide · Certainty factor · Binary logistic regression · Torrential rainfall · Typhoon Bilis

1 Introduction

Very severe cyclonic storm Bilis struck the southeast coast of mainland China on July 14, 2006. The torrential rainfall accompanied by the typhoon triggered widespread floods, landslides, and debris flows, which significantly damaged the village of Zixing, in the

Hunan province as shown in Fig. 1 (Xu et al. 2011). More than two thousand debris flows and shallow landslides were induced by this heavy rainfall event. This catastrophic event damaged over 31,000 houses and led to over 345 fatalities, and about 89 people missing cases were reported in Hunan province. The unprecedented flood was estimated to have a 100-year return period. In all, this typhoon was responsible for 654 deaths and 208 missing and over USD 2.5 billion in damage to southeastern China (Xinhua, July 17, 2006).

In the wake of increasing landslide activities and associated hazards following the changes in the global climatic system, it is necessary to investigate the landslide characteristics and assess the landslide-prone area to mitigate damages associated with them.

Landslides are typical of mountainous terrains and are hazardous for people's life and habitat. Over the last few decades, it has been observed that the frequency of landslide occurrence is increasing worldwide (Petley 2012; Dou et al. 2015a, c; Zhu et al. 2017). Many mass movements have been induced by the rainfall accompanied by the typhoons that caused substantial loss of life and damage around the worldwide (Jebur et al. 2014; Wang et al. 2015; Dou et al. 2017). For instance, super Typhoon Haiyan in the year 2013 devastated the Leyte region in the Philippines with damage amounting to more than USD 2 billion (Rabonza et al. 2016). Heavy rainfall struck during the Typhoon Wipha on October 16, 2013, in the Izu Oshima Island in Japan, located about 100 km south of the Tokyo triggered many landslides, caused at the least 35 deaths and nearly 50 people missing (Ministry of Land Infrastructure and Transport and Japan-MLIT, 2013). According to MLIT, the torrential rainfall hit on August 2014 in Hiroshima city triggered 166 slope failures, caused 74 deaths, and damaged 429 houses (MLIT 2014; Wang et al. 2015). During June

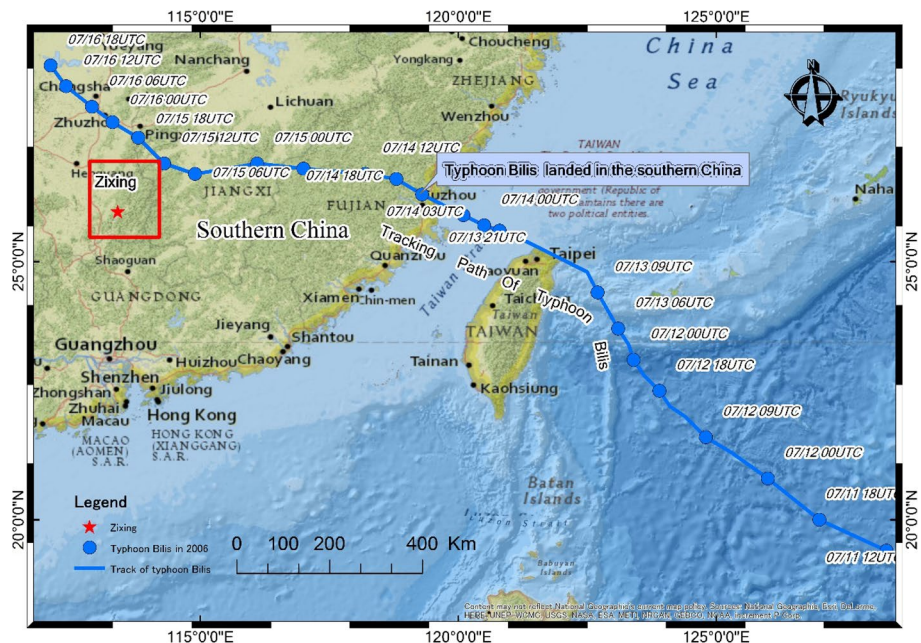


Fig. 1 Trajectory of the Typhoon Bilis in 2006: a red rectangle represents a severely damaged location in the study area (created using ArcGIS 10.4 software—data source from Japan National Institute of Informatics)

15–17, 2013, the cloudburst in Uttarakhand state in India triggered numerous landslides and caused the death of 6074 people and widespread damages to cultural properties (Martha et al. 2014). United State Geological Survey reports that an average of 25–50 people is killed by landslides each year in the USA (USGS 2019). In Italy, more than 7500 square miles of land areas are identified as high-risk zones for landslides (Parsons and Lister 2019). China, one of the largest countries with diverse topography, has no exemption to landslides. In fact, China has suffered from the most serious landslides in the past century that caused many human lives and economic destruction (Petley 2012). Historical records show that more than ninety thousand hazards associated with landslides have been recorded in several regions of China (Huang 2007). Southwestern part of the country is close to China Sea which is at high risk to landslides owing to the increased typhoon activities. Li et al. (2017) pointed out the number of rainfall-induced landslides in China has risen over to 90% of the total number of landslide events compared with last decades. Understanding these hazardous landslides and debris flows induced by heavy rainfall events has become an important and urgent issue in the view of emergency activities (Dou et al. 2015c; Wang et al. 2015).

Landslide spatial distribution in any region is influenced by physical rules that can be analyzed with the empirical, statistical, or deterministic approach (Reichenbach et al. 2018). Numerous models have been successfully applied for landslide susceptibility mapping worldwide (Youssef and Pradhan 2014; Chen et al. 2016; Camilo et al. 2017; Chen et al. 2018; Dou et al. 2018; Pham et al. 2018). In the early days, landslide susceptibility mapping is carried out using qualitative approaches (knowledge-driven methods). The pioneering works on data-driven methods and physically based models are dated to the late 1970s and early 1980s (Neuland 1976; Carrara 1983). In comparison with knowledge-driven methods, the latter one minimizes the subjectivity and attains reproducibility (Bui et al. 2011; Zêzere et al. 2017). The extensively used data-driven methods in susceptibility mapping are bivariate statistical analysis (BSA), binary logistic regression (BLR), artificial neural network (ANN), and support vector machines (Bui et al. 2012; Arnone et al. 2014; Dou et al. 2015b; Arnone et al. 2016; Pham et al. 2019). All of these techniques rely on a few assumptions (Rabonza et al. 2016; Dou et al. 2019a, b). One of the basic assumptions is “past is the key to future.” Therefore, bivariate statistical methods estimate landslide probabilities based on relationship analysis between historical landslide events and geo-environmental conditions inferred from heuristic investigations. The accuracy of such statistical techniques depends on the completeness of landslide inventory used to prepare the model. Landslide inventories can be either points (centroids of the landslide area or rupture zone) or polygons (Dou et al. 2014; Pham et al. 2018). Nowadays with the high-resolution imageries, polygon type is the most preferred. For achieving the likelihood ratio, landslide density analysis over the studied portion has to be established.

As per the literature, BSA and BLR are considered to be the most frequently used methods for the assessment of the likelihood of landslide occurrence at regional scales (Shahabi et al. 2014). Reichenbach et al. (2018) reviewed eighteen different landslide susceptibility models published over the last three decades and reported that logistic regression topped the chart accounting 18.5% of all the occurrences. The merit of BLR over other multivariate analysis methods is that it is independent of data distribution and can handle a variety of datasets such as continuous, categorical, and binary data (Bui et al. 2011). However, the BLR model has little to no predictive value, if a set of irrelevant independent parameters are involved. Because of such constraints, predicting landslide susceptibility needs a distributed model that ascertains all the relevant independent aspects of the method used. Effective landslide susceptibility mapping, therefore, requires optimal predisposing

factors as input to the LSM models. In LSM studies, selecting landslide-predisposing factors and their classes are key points. However, most scholars arbitrarily and subjectively selected the predisposing factors including geological, anthropogenic, geomorphological, and hydrological factors. There is no standard law to select predisposing factors. Hence, we address this issue by presenting the certainty factor (CF) model that has been applied to landslide factors. CF is a method using rule-based expert systems to handle certain problem classes.

The understanding of landslide mechanism of the rainfall-triggered event over a reservoir watershed is useful for geological disasters and warning systems. Several researchers have studied the impacts of tropical cyclones from the hydrological process in reservoir watersheds (Xu et al. 2011; Zou et al. 2013); however, to our knowledge, few studies have paid attention to the characteristics of rainfall-triggered landslides by tropical typhoons and assessment of landslide susceptibility in this study region. This study, therefore, focused on addressing: (1) characteristics of the landslides triggered by the extremely heavy rainfall even for the Dongjiang Reservoir Watershed, Hunan province, China; (2) constructing the event-based landslide inventory map using multi-high-resolution satellite images; (3) optimization of the best predisposing landslide factors using the CF model; (4) comparison with the LSM maps implemented by ensemble models and validation of the models.

2 Study area

The study area, Dongjiang Reservoir, which is situated in the southeast of Hunan Province, China, is an area vulnerable to heavy rainfall during the tropical cyclone seasons (Fig. 2). The elevation of the study area varies between 78 m a.s.l. and 1868 m a.s.l. with an average of 540 m. Three distinct geomorphological units represent the entire study reach: hills and valleys, hilly plains, and the Luoxiao Mountains near the eastern and southern borders. Geologically, the area is composed mostly of Paleozoic sedimentary and metamorphic rocks (sandstone, limestone, and slate) which were invaded by granitic rocks in places. The granitic rocks are severely weathered and thus are subjective to failure. The weathered soils are mostly composed of highly oxidized laterite, prone to erosion. Land use/cover in the study area is characterized by small-scale agro-industrial activities like a plantation, and paddy farming, and settlements. The case study area falls within the humid subtropical monsoon climate region. The mean annual precipitation is about 1932 mm (1953–2004), 80% of which occurred during the rainy months of March to August, typically influenced by cyclones. Each year numerous cyclones hit the province and cause severe damage to life and property in the region. The most recent one is tropical cyclone Mangkhut, which killed 2 people on September 16, 2018. Months before Mangkhut landfall, another cyclone Typhoon Ewiniar has brought torrential downpours recording over 250 mm of rain in 24 h, June 8 to 9, 2018. Wang et al. (2008) studied the extreme precipitation patterns in the Dongjiang River Basin using statistical parameters and noticed significant changes in several annual extreme flood flow and monthly precipitation processes in the region.

The Dongjiang Reservoir is the biggest reservoir in the south of Hunan Province, which covers a water area of 160 km² and has a capacity of 8.12×10^9 m³. Owing to the intense rainfall triggered by the Typhoon Bilis in 2006, thousands of sediment-related disasters, including numerous slope failures (shallow landslides) and debris flows occurred and were identified from the high-resolution 0.6 m QuickBird images, China–Brazil Earth Resources Satellite (CBERS) images (20 m), and field surveys (Fig. 3). The torrential rainfall event

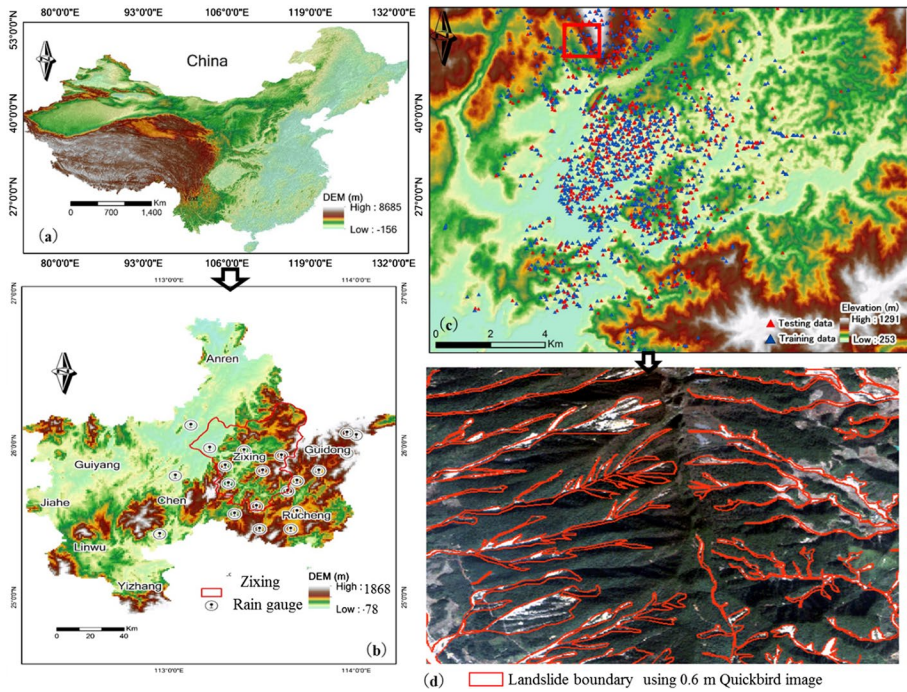


Fig. 2 Dongjiang Reservoir Watershed study area. **a** Location map of China. **b** Map of the study area with rain gauge distribution. **c** Distribution of shallow landslides on the elevation map derived from a 30-m DEM. **d** The lower map is the enlarged area of showing the landslide boundary

associated with Typhoon Bilis caused 246 deaths, 95 missing, and more than 300 million US dollars of economic loss just in and around Zixing City. Damages for destroyed or buried buildings by debris flows were serious. Flash floods also inundated the short and steep rivers in the hilly areas.

3 Data source

Rainfall data from the local records of 21 rain gauges in and around the Dongjiang Reservoir area were used to analyze the rainfall characteristics of the major rainstorm. Typhoon Bilis was a strong tropical storm with severe precipitation in a short duration, whose trail is shown in Fig. 1, and it landed on the coast of Fujian Province, China, on July 14 2006, with the maximum wind speed of 108 km/h. Then, it weakened into a tropical storm and moved westward and north-westward at the speed of 10–15 km/h until July 16 2006, when it disappeared in Hunan Province.

The rainfall observation data from the rain gauge networks around the reservoir on 14–15th July are displayed in Fig. 4. The Longxi rain gauge shows the maximum rainfall with a total 36-h rainfall of 507 mm and total monthly rainfall of 826 mm. One of the rain gauge data from Xingning was plotted as shown in Fig. 5. In 48 h, the cumulative rainfall in Xingning is more than 400 mm. The incremental rainfall of Xingning at 15–18 UTC was approximately 180 mm. More than 1600 landslides occurred when the accumulative

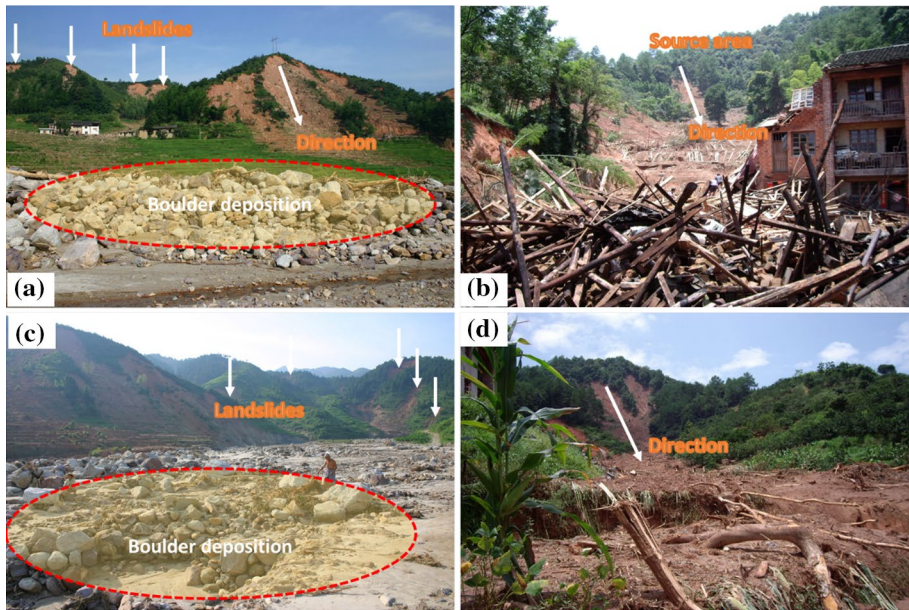


Fig. 3 Rainfall-induced landslides by the Typhoon Bilis. **a** Examples of shallow landslides (white arrows) and debris flow associated with boulder deposition; **b** seriously damaged houses; **c** rainfall-triggered debris flows caused substantial deposited materials, such as various sizes of boulder; **d** destroyed crops

rainfall reached 340 mm. Figure 6 shows the rainfall contour diagram of the Dongjiang Reservoir area in 36 h on July 14th–16th. The reservoir watershed area totally received a rainfall amount of around $6.6 \times 10^8 \text{ m}^3$, leading to a reservoir depth increase of 4.66 m. The reservoir was severely affected by the heavy rainfall in a short time.

The landslide inventory map is constructed through a combination of satellite image—interpretation of before and after the event (0.6-m QuickBird and 20-m CBERS) as listed in Table 1 and fieldworks. In order to identify the landslides triggered during the Bilis event, we firstly interpreted and mapped the landslides visible in the pre-event satellite CBERS imageries in a GIS environment. Following this, post-event satellite imageries from CBERS have interpreted for mapping all the landslides in the study area that have triggered before and after the event. Finally, high-resolution QuickBird images of December 2007 are interpreted and mapped for accurately delineating the boundary of landslide polygons. Then using analysis (erase function) toolbox in ArcGIS, landslide polygons of Bilis event are extracted from the entire database, assuming that no further landslides have occurred past the Bilis event till October 2006. This assumption is based on the fact that no major typhoons are reported in the study area during this time period. In this way, we built the entire database of landslide inventory from 2000 to 2009 as well as event-based landslide Atlas. A total of 2207 landslide polygons are mapped for the Bilis event from the interpretation of satellite imageries as shown in Fig. 7. The polygon data were then converted to landslide points. As more than 50% of the landslides in the study area are less than 10,000 m², the centroid technique was applied to deal with the transformation of landslide polygon to point. Although many studies have pointed out the lower accuracy in LSM while using point technique rather than landslide polygons (Simon et al. 2013), several other studies favor usage of centroid points for fast, easy to use, and automated

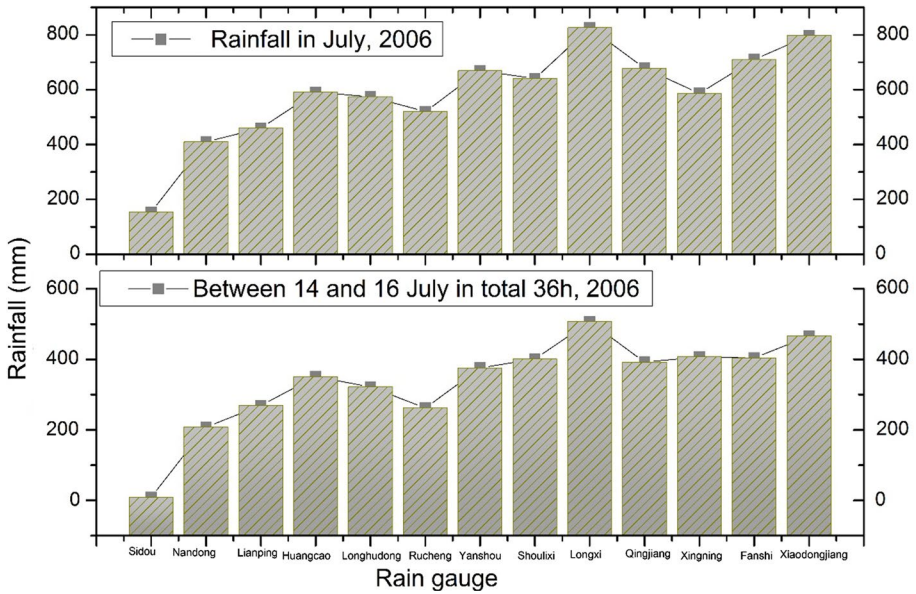


Fig. 4 Graphs showing rainfall around Dongjiang Reservoir in the month of July 2006 (top) and 36 h of rainfall between July 14 and 16, 2006 (bottom). The water level of the Dongjiang Reservoir increased to about 7.73 m during July 14–19, 2006

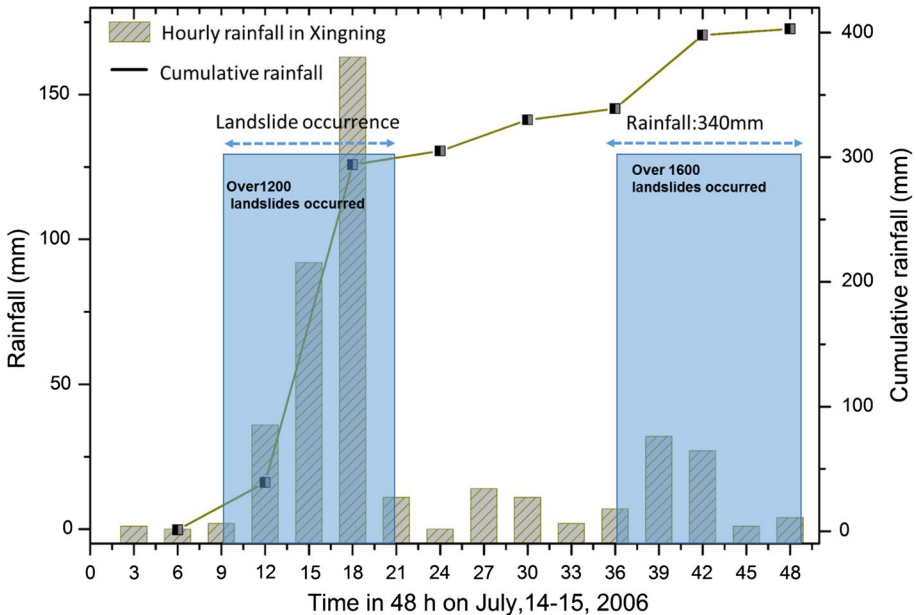


Fig. 5 Hourly and cumulative rainfall recorded by rain gauge in Xingning around the Dongjiang Reservoir

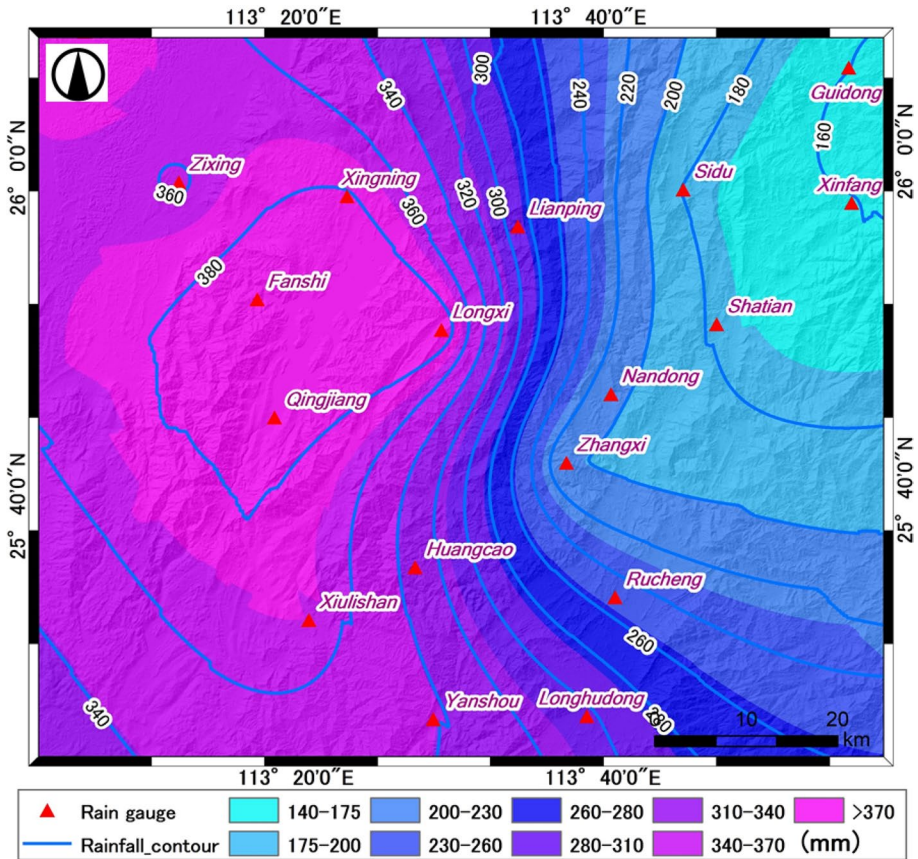


Fig. 6 Rainfall contour diagram of the Dongjiang Reservoir area in 36 h on July 14–16, 2006. The reservoir area suffered from the total rainfall of around 6.6×10^8 m³, leading to a reservoir depth increase of 4.66 m

LSM mapping (Bui et al. 2012; Chen et al. 2016; Pham et al. 2018). The landslides mostly located around the upper catchment of Dongjiang Reservoir corresponded with the zonal distribution. Field observations reveal the type of landslides as shallow landslides. The landslide density is approximately 8.2/km². Topographic data for analyses such as slope, aspect, and curvature are derived from the 30 m ASTER GDEM (version 2). In this case study, based on the analysis of landslide inventory map and availability of data, a total of 12 landslide-predisposing factors were prepared, namely elevation, slope angle, slope aspect, curvature, plan curvature, profile curvature, drainage density, distance to drainage network, stream power index (SPI), compound topographic index (CTI), 36-h cumulative rainfall, and lithology.

Table 1 Collected pre- and post-images in the study area

Serial number	Date	Path	Satellite and sensor
1	2000-03-26	373-71	CBERS01-CCD
2	2003-12-26	373-71	CBERS02-CCD
3	2005-11-27	373-71	CBERS02-CCD
4	2005-12-23	373-71	CBERS02-CCD
5	2006-10-05	373-71	CBERS02-CCD
6	2006-10-31	373-71	CBERS02-CCD
7	2006-12-22	373-71	CBERS02-CCD
8	2007-10-27	1-70	CBERS02-CCD
9	2008-11-11	373-71	CBERS02B-CCD
10	2008-05-13	373-71	CBERS02B-CCD
11	2009-01-02	373-71	CBERS02B-CCD
12	2009-10-24	372-71-A-2	CBERS02B-HR
13	2009-10-24	373-71-A-4	CBERS02B-HR
14	2009-01-02	373-71-B-1	CBERS02B-HR
15	2009-01-02	373-71-B-2	CBERS02B-HR
16	2009-01-02	373-71-B-3	CBERS02B-HR
17	2009-01-02	373-71-B-4	CBERS02B-HR
18	2009-01-02	373-71-B-5	CBERS02B-HR
19	2008-01-04	373-71-C-4	CBERS02B-HR
20	2008-01-04	373-71-C-5	CBERS02B-HR
21	2007-12-18		Quick bird

4 Methodology

4.1 CF model for selecting predisposing factors

The certainty factor (CF) model is a rule-based expert system developed by Shortliffe and Buchanan (1975) for managing uncertainty in computational fields. When comparing with other models, CF can provide probable favorability functions for incorporating heterogeneous data (Chung and Fabbri 1993). The CF weight can be computed by the subsequent functions:

$$CF = \begin{cases} \frac{P_a - P_s}{P_a(1 - P_s)} & \text{if } P_a \geq P_s \\ \frac{P_a - P_s}{P_s(1 - P_a)} & \text{if } P_a < P_s \end{cases} \tag{1}$$

Here P_a is the conditional likelihood of landslides in class a and P_s is the prior likelihood of a total number of landslides in the case study area. The CF values vary between -1 and 1 , and it indicates a measure of belief in the outcome (Lucas 2001). A positive CF value measures decreasing uncertainty, whereas negative values indicate an increasing uncertainty of landslide occurrence. If CF value is closed to 0 , no information on the certainty is indicated. Once the CF values for classes of the predisposing factors are

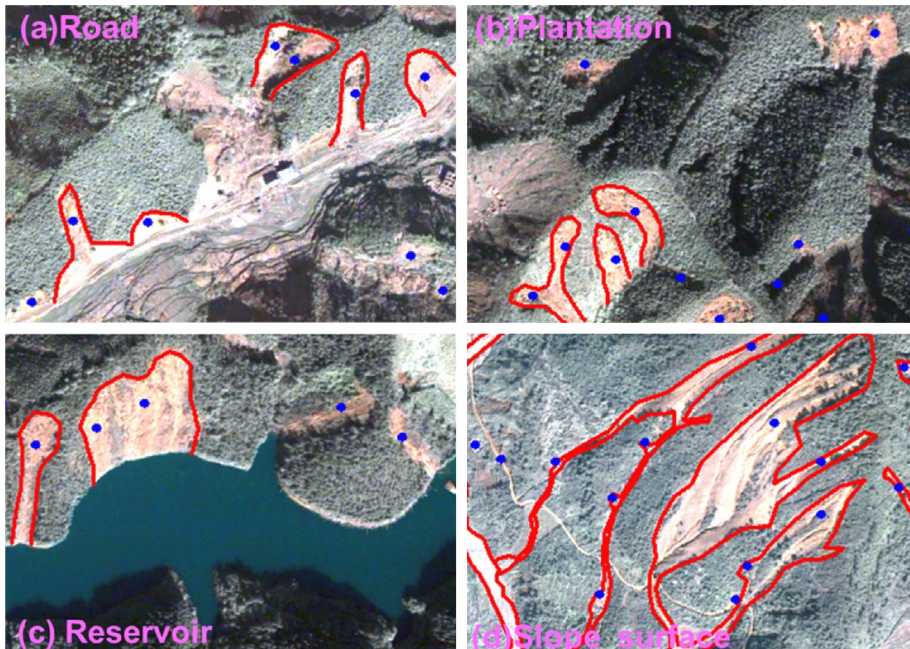


Fig. 7 Examples of the construction of landslide inventory-based satellite images and field survey, landslides mainly occurred in the four types of land use: **a** near the roads, **b** in the plantation area, **c** near the reservoir, and **d** in the slope surface in the hilly area

obtained, these factors are then integrated pairwise using the combination rule (Binaghi et al. 1998) as follows:

$$Z = \begin{cases} CF1 + CF2 - CF1CF2 & CF1, CF2 \geq 0 \\ CF1 + CF2 + CF1CF2 & CF1, CF2 < 0 \\ \frac{CF1+CF2}{1-\min(|CF1|,|CF2|)} & CF1, CF2, \text{ opposite signs} \end{cases} \quad (2)$$

where CF1 is a value in class 1, and CF2 is a value in class 2.

The pairwise combination is performed until all the CF layers are brought together, and the predisposing factors are optimized by computing the Z values. If the Z values are positive, we favor those factors have high correlations with landslide occurrence. Based on the range of CF values, predisposing factor weights were acquired. The weights are assessed as the sum of the ratio relative to those predisposing factors that provide a measurement of certainty in predicting landslides (Binaghi et al. 1998). According to the computed results, CF weights are then classified into six classes as shown in Table 2 (Binaghi et al. 1998).

4.2 Bivariate statistical analysis

Van Westen et al. (1997) proposed the bivariate statistical analysis (BSA) method, which is based on the assessment of the relationship of a landslide inventory map and predisposing factors. In the BSA method, the weight for each class of the landslide-predisposing factors was

Table 2 Weight classification based on the range of CF values

Level	Range	Explanation
1	– 1.0 to – 0.09	Extremely low certainty
2	– 0.09 to 0.09	Uncertainty
3	0.09–0.2	Low certainty
4	0.2–0.5	Medium certainty
5	0.5–0.8	High certainty
6	0.8–1.0	Extremely high certainty

initially determined. Landslide susceptibility indexes were then computed by summing up the weights. The weight (W_i) of each class i is defined as the natural logarithm of the landslide density in the class over the landslide density in the predisposing factor map as listed (van Westen et al. 1997):

$$W_i = \ln \left(\frac{\text{Density_landslide}}{\text{Density_area}} \right) = \ln \left(\frac{N_{ij}/A_{ij}}{N_i/A_T} \right) \tag{3}$$

where W_i is the weight given to an i th class of a certain thematic layer (e.g., limestone in the thematic layer—lithology); Density_landslide is the landslide density within the entire thematic layer; Density_area is the landslide density of the whole factor study area for all classes; N_{ij} is the number of landslide pixels in the class j of the predisposing factor i ; A_{ij} is the total area of the class j of the predisposing factor i ; N_i is the total number of landslides; and A_T is the pixels in the entire study area.

Finally, the LSM by BSA model was generated by the following equation:

$$\text{LSM}_{w_i} = (W_1) + (W_2) + (W_3) + \dots (W_i) \tag{4}$$

4.3 Binary logistic regression

Binary logistic regression (BLR) is one of the well-known multivariate analytical methods in the field of LSM assessment during the last decade (Chauhan et al. 2010; Dou et al. 2018). The BLR method is suitable for forecasting the presence or absence of a characteristic outcome from a set of parameters (Devkota et al. 2013). Here, we do not use the ordinary least squares regression (OLS) because of three problems: (1) the error terms are heteroskedastic; (2) the error terms are not normally distributed; (3) the predicted probabilities can be larger than 1 or less than 0. In this study, the purpose of BLR is thus to simulate the relationships between a dependent variable and multiple independent parameters (Bui et al. 2011). The advantage of BLR is that it does not compulsorily need normal distribution data. In addition, both continuous and discrete data can be used as an input for the BLR model.

The dependent parameter (Y) in the BLR method is a function of the possibility and can be calculated as follows (Lee and Pradhan 2006):

$$Y = \frac{1}{1 + e^{-z}} \tag{5}$$

where Y is the estimated likelihood of landslide occurrence and ranges [0 1]; z is the weighted linear combination of the independent parameters.

To linearize the stated model as well as eliminate the 0/1 boundaries for the dependent parameter, the estimated Y is transformed by the following equation:

$$Y' = \ln\left(\frac{Y}{1-Y}\right) \quad (6)$$

This modification is referred to as the logit transformation. Theoretically, the logit transformation of binary data can confirm that the dependent parameter is continuous and the logit transformation is limitless. Additionally, it can ensure that the likelihood surface can be continuous under $[0, 1]$. By means of the logit transformations, the standard linear regression models can be written by the following equation:

$$Y' = \ln\left(\frac{Y}{1-Y}\right) = \beta_{1*}x_1 + \beta_{2*}x_2 + \dots + \beta_{n*}x_n + \alpha \quad (7)$$

where α is the intercept of the equation, $\beta_1, \beta_2, \dots, \beta_n$ denotes the slope coefficients of the independent parameters. Landslide or non-landslide as the dependent determined the approximate equation that is meaningful at 0.01% error level.

5 Results

5.1 Characteristics of landslides triggered by the Typhoon Bilis

To investigate the landslide-predisposing factors contribution in the initiation of landslides, the landslides occurred in the case study area were interrelated with those factors contributing to landslide occurrence. These predisposing factors include elevation, slope angle, slope aspect, curvature, plan curvature, profile curvature, drainage density, distance to drainage network, SPI, CTI, cumulative rainfall, and lithology. Figure 8 shows the results of landslide frequency analysis that examines the relationships between landslide occurrence and the predisposing factors. The relationship of landslide frequency with elevation is shown in Fig. 8a. It can be seen that landslides (43.15%) mostly occurred at the intermediate elevation (320–400 m) taken a proportion of 29.11% total area. At the following elevation class (400–500 m), landslide frequency is around 21%. The results suggest that landslides are frequently in the middle elevations; this is because the area ratios in the middle elevations are greater than those in the higher elevations.

Slope angle plays an important role in the occurrence of landslides. On a relatively flat slope (0° – 5°), the force of gravity acts directly downward. Thus, the material remains on the flat slope and it will not move under the force of gravity, whereas on a steeper slope, the shear stress or tangential component of gravity increases, and the perpendicular component of gravity decreases (Dou et al. 2014). As observed in this study, the landslide frequency in the slope classes 10° – 15° , 15° – 20° , and 20° – 25° is 22.14%, 20.44%, and 16.82%, respectively, as shown in Fig. 8b. It could also be seen that gentle slope angles have a relatively lower frequency of landslide occurrence due to the lower shear stress at the slope angles 0° – 5° (Fig. 8b). The decrease in the frequency of landslides in steeper slope classes is attributed to the decrease in the percentage of an area ratio in that particular class.

Aspect that describes the orientation of slope is an important factor attributing the regions insolation, vegetative growth, soil moisture conditions and wind velocity (Aksoy and Ercanoglu 2012) and hence regarded as a highly important predisposing factor in LSM (Carrara 1983; Camilo et al. 2017). Also, when the hillsides suffer from the dense

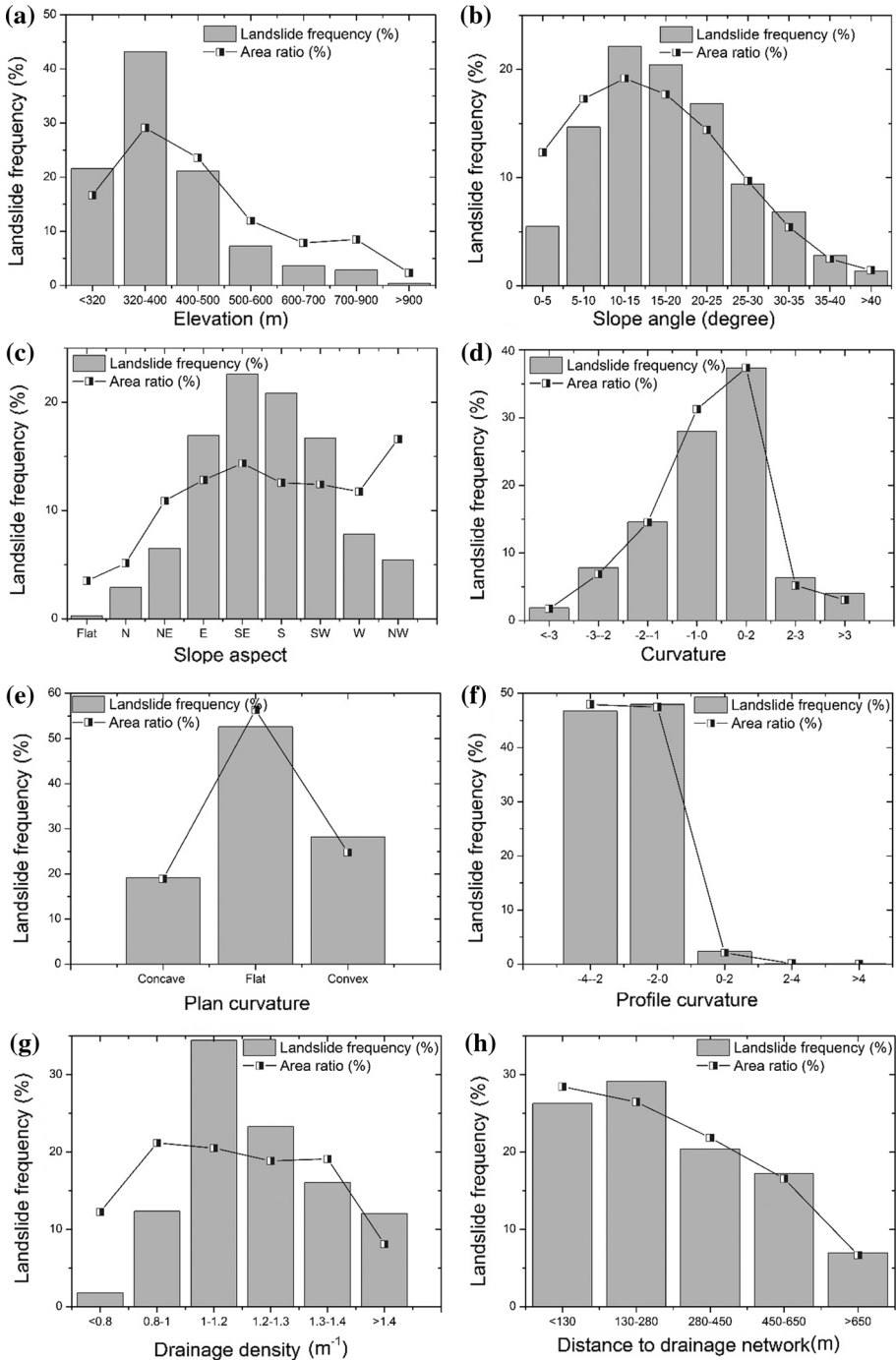


Fig. 8 Relationships between landslide frequency and the predisposing factors: **a** elevation; **b** slope angle; **c** slope aspect; **d** curvature; **e** plan curvature; **f** profile curvature; **g** drainage density; **h** distance to drainage network; **i** SPI; **j** CPI; **k** accumulative rainfall; **l** lithology

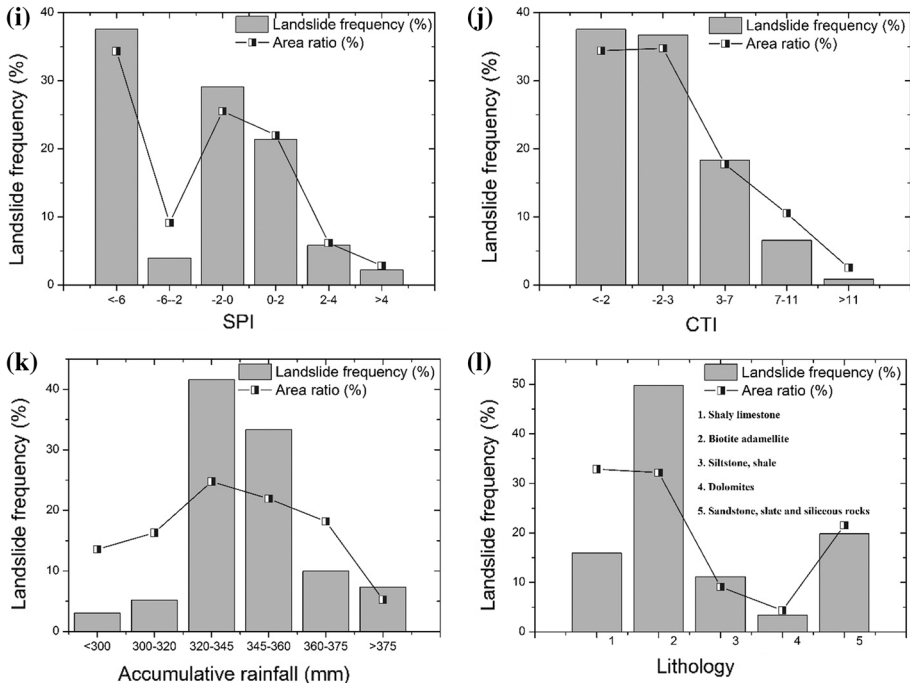


Fig. 8 (continued)

precipitation to reach saturation, it influences the infiltration capacity of the slope controlled by some parameters including the constitution of soil, permeability, and pore water pressure. With regard to the slope aspect, landslides mostly occurred among the east-, southeast-, south-, southwest-, and west-facing direction as shown in Fig. 8c. The results indicate that from east to west is greatly prone to landslide occurrence. The largest landslide frequency (22.59%) occurred along the southeastern slope direction, followed by south slope direction (20.84%). On north-facing slope direction, the landslide frequency is comparatively less. This is in agreement with many previous studies which states that north-facing slopes are favorable for the enhanced growth of vegetation (Olivero and Hix 1998; Ghimire et al. 2011; Måren et al. 2015). The higher solar radiation received in the south-facing slopes may dry out the vegetation cover faster and hence induces more landslides.

Figure 8d shows that landslides (37.37%) are mostly concentrated at the 0–2 class for the curvature, followed by the –1–0 class with a landslide frequency of 27.97%, while for the profile curvature, landslides mostly occurred at –2–0 class and –4–2 class (Fig. 8f). The curvature of the hillside in the horizontal plane is the plan curvature of that surface. Based on the hillsides, the plan curvatures are subdivided into concave (hollows), convex (noses), and flat (planar) regions. As for the plan curvature as shown in Fig. 8e, the landslides generally occur in the concave slope because it strengthens the soil moisture and causes the land sliding. However, in this study, the flat and convex slopes show higher landslide frequency than concave slopes. One reason is probably that hilly ridges in Dongjiang Watershed could be likely to collapse because of the impact of human activities

(building the reservoir) causing the higher ground acceleration. The other reason may be that the dropped intense rainfall flashed the surface of the hill slope; thus, the rainfall could not accumulate too much in short time.

Drainages undercut the hill slopes as the intensity of flow increases, thus resulting in increased landslide frequency with a higher drainage density. For example, drainage density and erosion rates in steep Japanese mountains are negatively correlated due to active landslides (Oguchi 1997). Several scholars have therefore studied the interrelationship of landslides and geomorphological characteristics of drainage networks (Hovius et al. 1998; Dou et al. 2015c). In this location, the Dongjiang River flows into the reservoir. It has been observed in our study area that the landslides mostly occurred at 1–1.4 m^{-1} and decreased further in proportion to the area ratio (Fig. 8g). For the distance to drainage network factor, the landslide highly occurred at 130–280 m followed by less than 130 m (Fig. 8h). With the increase in distance to the drainage network, the landslide frequently usually decreases because the topography change induced by erosion might influence the landslide initiation.

In the case of hydrological predisposing factors, SPI (the measures of the erosive power of overland flow) and CTI (soil wetness: topographic control on hydrological processes), landslides highly occurred at < -6 and at < -2 category, respectively, as shown in Fig. 8i, j. Rainfall increases the weight to the slope by seep into the bedrock beneath and replaces the pore space or fractures. This added weight force leads to an increase in stress and induces slope instability. Rainfall also induces a change in the angle of repose. In landslide studies, accumulated rainfall is considered as an important factor rather than simple rainfall statistics (Li et al. 2017). For the accumulative rainfall factor, landslides mostly occurred at the 320–345 mm, followed by 345–360 mm. The landslides also easily occurred at over 375 mm because it takes a relatively small percentage of the total study area in terms of this class.

Lithology is considered, landslides (around 50%) mostly occurred at the biotite adamellite type (one of the granite types), followed (about 20%) by the sandstone and slate type, and then by the limestone (about 16%). As mentioned previously in Sect. 2, the granitic rocks are highly weathered and are susceptible to failure. The sandstone type contains enough pore space to accumulate more rainfall that can saturate rock and increase its weight. Water also enters into the bedrock below through the bedding plane and ultimately reduces the cohesion. Similarly, the slate rocks which contain clay minerals generally tend to have a low shear strength and will be the most likely place for failure to occur, especially if the layer dips in a down-slope direction. Limestone units may have caverns and be leached in the rock due to chemical weathering by groundwater.

5.2 Predisposing factor selection for LSM maps

The results of the correlation analysis between the landslide occurrence and predisposing factors for the Dongjiang Reservoir area are shown in Table 3. The result of CF analysis shows that the Z value is positive for slope angle (0.25), curvature (0.82), plan curvature (0.21), drainage density (0.96), distance to drainage network (0.11), accumulative rainfall (0.97), and lithology (0.47) as shown in Fig. 9. The Z values are negative for the other factors. Hence, these seven factors are selected for producing LSM maps. This result also shows that the occurrence of landslides in the study area is mainly affected by some predisposing factors. Even Z values between those factors are different; they all contribute to a certain extent in the landslide occurrence. We conducted the objective method of CF analysis to avoid the “ghost effect” and get appropriate factors for modeling LSM maps.

Table 3 Spatial relationship between the predisposing factors and landslide occurrence based on the CF and BSA methods

Factors	Class	Percentage of domain	Percentage of land-slides	CF	Z	BSA
Elevation (m)	< 320	16.63	21.57	0.56		0.26
	320–400	29.11	43.15	0.80		0.39
	400–500	23.59	21.18	−0.22		−0.11
	500–600	11.95	7.25	−0.61	−0.99	−0.50
	600–700	7.88	3.62	−0.74		−0.78
	700–900	8.51	2.83	−0.83		−1.10
	> 900	2.33	0.40	−0.92		−1.77
Slope angle (°)	0–5	12.33	5.49	−0.75		−0.81
	5–10	17.28	14.67	−0.30		−0.16
	10–15	19.17	22.14	0.33		0.14
	15–20	17.71	20.44	0.33		0.14
	20–25	14.42	16.82	0.35	0.25	0.15
	25–30	9.70	9.40	−0.07		−0.03
	30–35	5.43	6.85	0.51		0.23
	35–40	2.50	2.83	0.29		0.12
	40–62	1.45	1.36	−0.14		−0.06
	Slope aspect	Flat	3.50	0.28	−0.97	
N		5.13	2.89	−0.66		−0.57
NE		10.89	6.51	−0.62		−0.51
E		12.82	16.93	0.60		0.28
SE		14.35	22.59	0.90	−0.01	0.45
S		12.55	20.84	0.98		0.51
SW		12.41	16.70	0.63		0.30
W		11.75	7.81	−0.55		−0.41
Curvature	NW	16.60	5.44	0.59		−1.12
	< −3	1.75	1.87	0.16		0.07
	−3–2	6.85	7.81	0.30		0.13
	−2–1	14.50	14.61	0.02		0.01
	−1–0	31.26	27.97	−0.22	0.82	−0.11
	0–2	37.36	37.37	0.00		0.00
	2–3	5.21	6.34	0.44		0.20
	> 3	3.08	4.02	0.58		0.27
Plan curvature	Concave	18.86	19.20	0.04		0.02
	Flat	56.41	52.60	−0.15	0.21	−0.07
	Convex	24.74	28.20	0.30		0.13
Profile curvature	−4–2	47.95	46.72	−0.06		−0.03
	−2–0	47.48	47.96	0.02		0.01
	0–2	2.10	2.32	0.24	−0.01	0.10
	2–4	0.12	0.11	−0.05		−0.02
	> 4	0.01	0.06	1.87		1.44

Table 3 (continued)

Factors	Class	Percentage of domain	Percentage of land-slides	CF	Z	BSA
Drainage density (m^{-1})	<0.8	12.27	1.81	-0.93		-1.91
	0.8–1	21.19	12.34	-0.64		-0.54
	1–1.2	20.51	34.43	0.99		0.52
	1.2–1.3	18.85	23.27	0.47	0.96	0.21
	1.3–1.4	19.11	16.08	-0.32		-0.17
	1.4–1.7	8.07	12.06	0.81		0.40
Distance to drainage network (m)	<130	28.45	26.27	-0.17		-0.08
	130–280	26.47	29.16	0.23		0.10
	280–450	21.85	20.39	-0.15	0.11	-0.07
	450–650	16.56	17.21	0.09		0.04
	>650	6.68	6.96	0.10		0.04
SPI	<-6	34.34	37.54	0.21		0.09
	-6 to 2	9.14	3.96	-0.76		-0.84
	-2 to 0	25.51	29.11	0.30		0.13
	0–2	22.02	21.35	-0.07	-0.79	-0.03
	2–4	6.16	5.83	-0.12		-0.05
	>4	2.84	2.21	-0.41		-0.25
CTI	<-2	34.41	37.54	0.21		0.09
	-2–3	34.75	36.69	0.13		0.05
	3–7	17.76	18.35	0.08	-0.89	0.03
	7–11	10.54	6.57	-0.60		-0.47
	>11	2.54	0.85	-0.83		-1.10
Accumulative rainfall (mm)	<300	13.58	3.06	-0.89		-1.49
	300–320	16.32	5.21	-0.84		-1.14
	320–345	24.76	41.62	1.00		0.52
	345–360	21.92	33.35	0.84	0.97	0.42
	360–375	18.17	9.97	-0.67		-0.60
	375–390	5.25	7.36	0.70		0.34
Lithology	Shaly limestone	32.87	15.91	-0.72		-0.73
	Biotite adamellite	32.18	49.72	0.87	0.47	0.44
	Siltstone, shale	9.08	11.16	0.46		0.21
	Dolomites	4.34	3.40	-0.41		-0.25
	Sandstone, slate and siliceous rocks	21.53	19.82	-0.17		-0.08

5.3 Mapping landslide susceptibility using BSA

The correlations between the landslide occurrence and predisposing factors using BSA are represented in Table 3. Two landslide susceptibility maps were generated: (1) using the seven selected factors ($CF > 0$) and (2) using all the original 12 factors (Fig. 10). Based on the natural breaks, the susceptibility level was divided into six classes, i.e., extremely low,

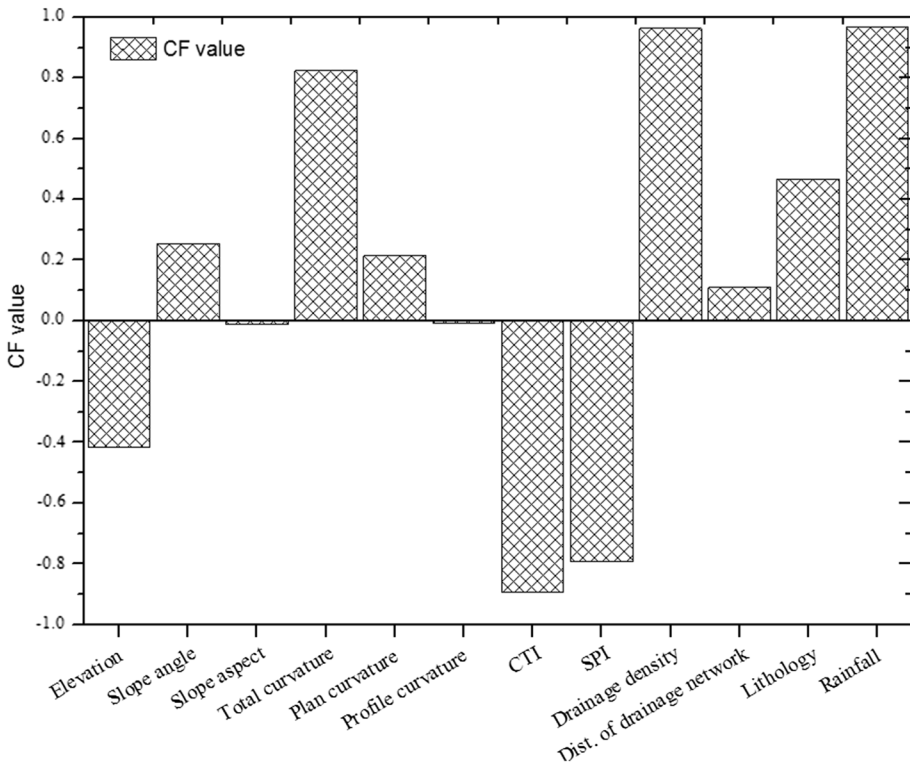


Fig. 9 Calculation of CF values in the Dongjiang Reservoir Watershed

low, moderate, high, very high, and extremely high. Visual interpretation reveals that there are much more red color areas (very high susceptible class) in Fig. 10b, whereas there are more dark blue areas (very low susceptible class) in Fig. 10a. Quantification of the same as shown in Fig. 11 and Table 4 reveals that 90.84% of the total landslides occurred in the 52.56% of the area which are classified as high, very high, and extremely high susceptibilities when the original factors were used, while 51.73% of the total landslides occurred in the 92.03% of the area which are classified as high, very high, and extremely high susceptibilities if the optimized seven factors were used (Fig. 12 and Table 5).

5.4 Mapping landslide susceptibility using BLR

The forward stepwise BLR approach was used to incorporate the predictor variables using the SPSS 20 software. The training dataset (1545 of total landslides) represented by points was assigned the value of 1. The same number of non-landslide points was randomly sampled from the landslide-free area and assigned the value of 0. The result based on all original factors is shown in Table 6. According to this table obtained by logistic regression, all the predisposing factors have a P value less than 0.05, indicating a statistical correlation between factors and the susceptibility of landslides at the 90% confidence level (Bui et al. 2011). Based on the equation, the occurrence of landslide probability (P) can be computed as mentioned before.

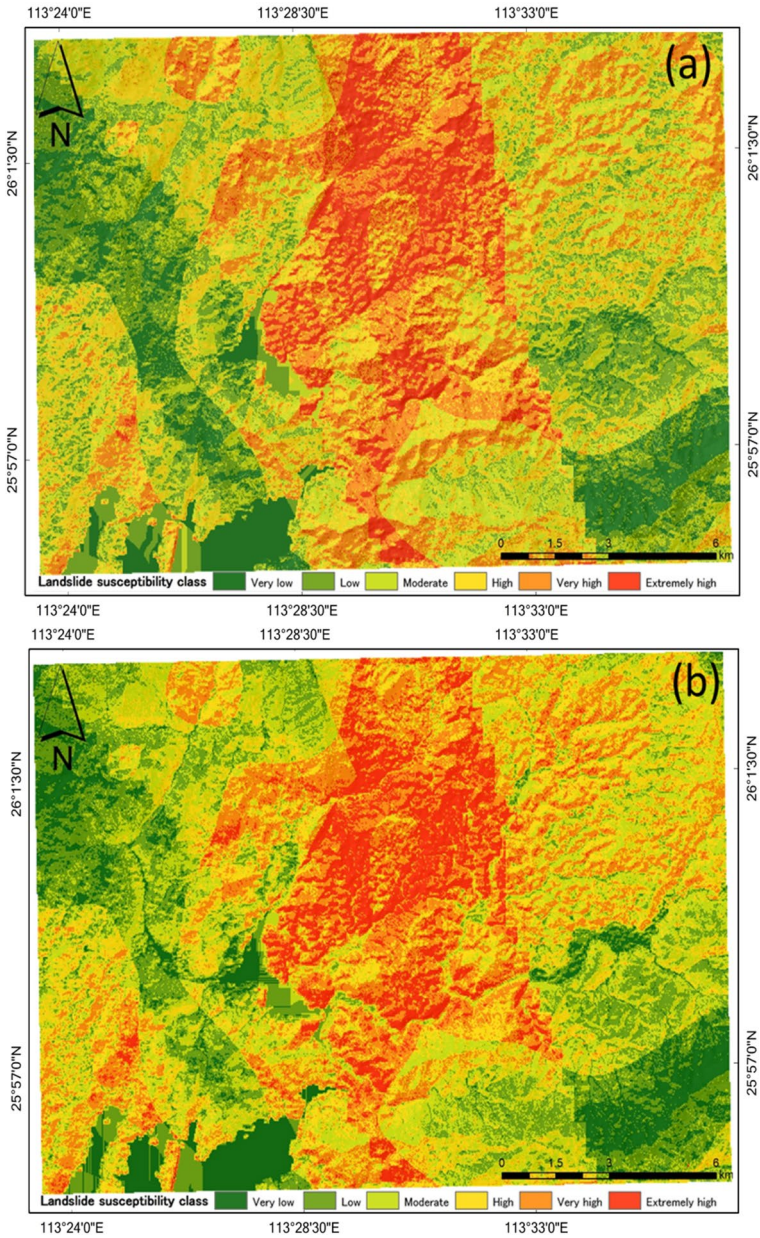


Fig. 10 LSM maps produced by the BSA method: **a** selected seven factors and **b** original 12 factors. Maps show the spatial probability of landslide occurrence in six classes. The upstream of the reservoir is located at the lower left corner of the map

Lastly, the regression coefficients of the predictors, GIS, and the natural break criterion were used to generate the landslide susceptibility maps (Fig. 13). In the maps, there are places where differences are subtle but also areas with obvious dissimilarities. There are

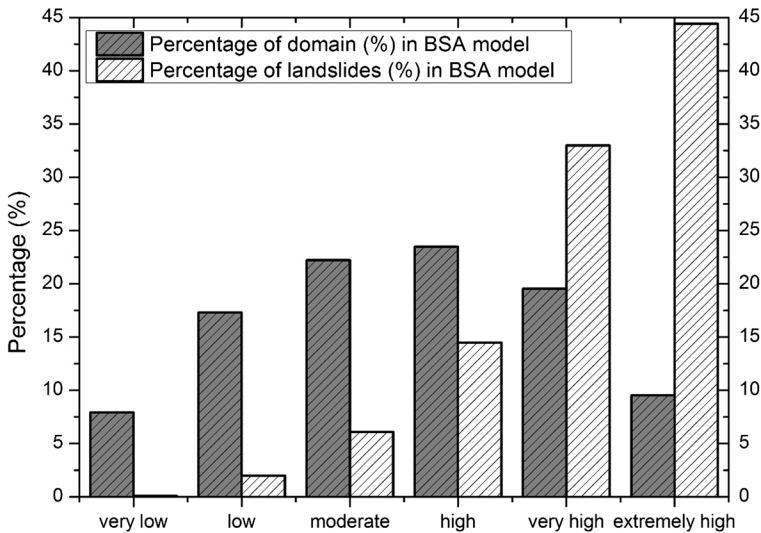


Fig. 11 Susceptibility class distribution within the study area and the occurrence of landslides according to the classification scheme for LSM using the BSA method with the original 12 factors

Table 4 Result of statistical analysis concerning landslide susceptibility from the BSA method with the original 12 factors

Class	Area of each class in pixels	Percentage of domain	No. of landslides	Percentage of landslides
Very low	23,530	7.90	2	0.09
Low	51,529	17.31	44	1.99
Moderate	66,159	22.22	134	6.07
High	69,919	23.49	319	14.45
Very high	58,197	19.55	728	32.99
Extremely high	28,374	9.53	980	44.40

more red colors in the map when using all factors, which segregate at the very high and extremely high ends of the color ramp than the seven-factor counterpart. The map from the seven factors is less heterogeneous. Figure 14 and Table 7 show that 95.51% of the total landslides occurred in the 66.73% of the area which are classified as high, very high, and extremely high susceptibilities if the all the original factors were used, while if the optimal seven factors were used 96.1% of the total landslides occurred in the 64.09% of the area which are classified as high, very high, and extremely high susceptibilities (Fig. 15 and Table 8).

5.5 Accuracy estimation

For the verification, the total landslides were randomly divided into two groups, training data and validation data. The evaluation of the prediction skills of susceptibility models was made using receiver operating characteristics (ROC) curves and

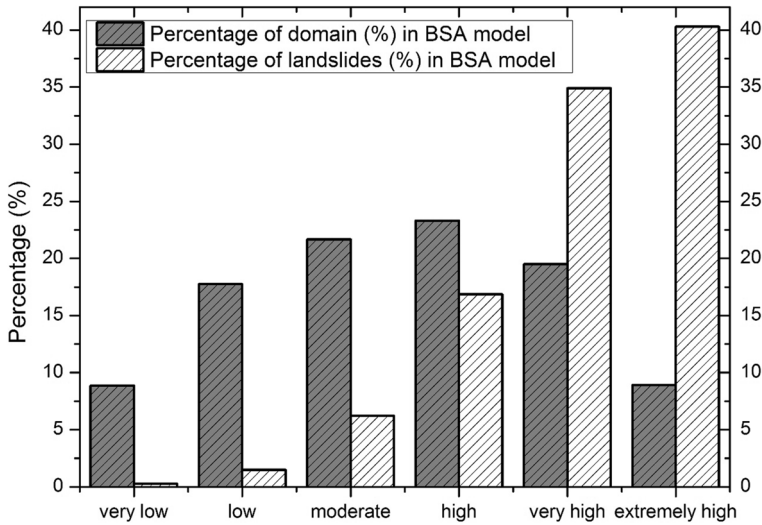


Fig. 12 Susceptibility class distribution within the study area and the occurrence of landslides according to the classification scheme for LSM using the BSA method with the selected seven factors

Table 5 Result of statistics analysis concerning landslide susceptibility from the BSA method with the selected seven factors

Class	Area of each class in pixels	Percentage of domain	No. of landslides	Percentage of landslides
Very low	26,388	8.86	6	0.27
Low	52,840	17.75	33	1.50
Moderate	64,488	21.66	137	6.21
High	69,376	23.30	372	16.86
Very high	58,052	19.50	770	34.89
Extremely high	26,564	8.92	889	40.28

computing the receiver operating characteristic (ROC) plot of sensitivity (% of terrain units containing landslides that are correctly classified) and 1-specificity (% of terrain units containing landslides that are correctly classified). The ROC area under the curve (AUC) evaluates the overall performance of the landslide susceptibility models (Bui et al. 2011). As a rule, the closer the ROC AUC value to 1, the better is the landslide model performance (Shahabi et al. 2014). For the BSA method, AUC value (0.837) is higher when the optimal seven factors were used than 0.794 from all the original factors (Fig. 16a). For the BLR model, the AUC value of the prediction rate curve (84.8%) from the seven factors is higher than that from all factors (80.8%) as shown in Fig. 16b. Consequently, using the seven factors gives a higher accuracy than using all the original factors. In addition, BLR has a slightly higher accuracy than BSA.

Table 6 Coefficients, statistics of the factors with all factors used in the BLR equation

Factors	<i>B</i>	SE	Wald	<i>df</i>	Sig.	Exp(<i>B</i>)	95% C.I. for EXP(<i>B</i>)	
							Lower	Upper
Elevation	−0.50	0.04	187.79	1	0.00	0.61	0.56	0.65
Slope angle	0.07	0.02	9.50	1	0.00	1.08	1.03	1.13
Slope aspect	−0.06	0.02	10.75	1	0.00	0.95	0.91	0.98
Plan curvature	0.13	0.08	2.52	1	0.01	1.14	0.97	1.34
Profile curvature	−0.25	0.07	12.86	1	0.00	1.28	1.12	1.47
Curvature	0.08	0.06	1.88	1	0.02	1.08	0.97	1.20
Drainage density	0.03	0.03	0.92	1	0.03	1.03	0.97	1.09
Dist. drainage network	0.21	0.03	42.40	1	0.00	1.24	1.16	1.32
SPI	0.22	0.05	21.43	1	0.00	1.25	1.14	1.37
CTI	−0.45	0.06	55.03	1	0.00	0.64	0.57	0.72
Accumulative rainfall	0.01	0.03	0.17	1	0.05	1.01	0.95	1.08
Lithology	0.01	0.02	0.06	1	0.01	1.01	0.97	1.05
Constant	−0.18	0.15	1.48	1	0.02	0.84		

6 Discussions

Devastating landslides as a result of intense rainfall are common in many places around the world every year. Predicting the exact locations of the instabilities and therefore landslide susceptibility assessment is rather difficult due to the uncertainty of the spatial and temporal distribution of rainfall. We investigated the landslide characteristics triggered during the torrential rainfall caused by Typhoon Bilis in the Dongjiang Reservoir Watershed region. In the study area, intense rainfall caused slope failures associated with severely weathered granite, resulting in numerous shallow landslides. While there are many factors that lead to landslides such as rainfall, slope, aspect, curvature, bedrock, drainage density, elevation, SPI, CPI are the important ones. Though the selection of factors is a fundamental step for landslide susceptibility evaluation, universal standard or rule to select the predisposing factors is absent (Dou et al. 2019a, b). These issues are commonly addressed by GIS-based landslide susceptibility studies.

To address this problem, we proposed the CF method to select the principal factors. Different scholars use various landslide-predisposing factors for LSM. Using this method, we selected the predisposing factors highly related to landslide occurrence. Our study of rainfall-induced landslides in Dongjiang Reservoir Watershed can be applicable in many similar cases. The resultant improvement in the values of AUC validates our approach. The use of the optimized factors led to a higher accuracy than when all possible factors were simultaneously used. Spatial autocorrelation and data redundancy among the predisposing factors before optimization are the possible causes for this observation.

Analysis of CF suggests that drainage density and total curvature are important in the case study area besides the other common factors such as lithology and rainfall. Total curvature represents the morphological measurement of the topography (Lee and Pradhan 2006). A more upwardly concave or convex slope holds more water and keeps it longer, and these hydrological controls of topography are more expressed in mountainous areas and lower in the flat areas. Furthermore, another important factor that does not represent in

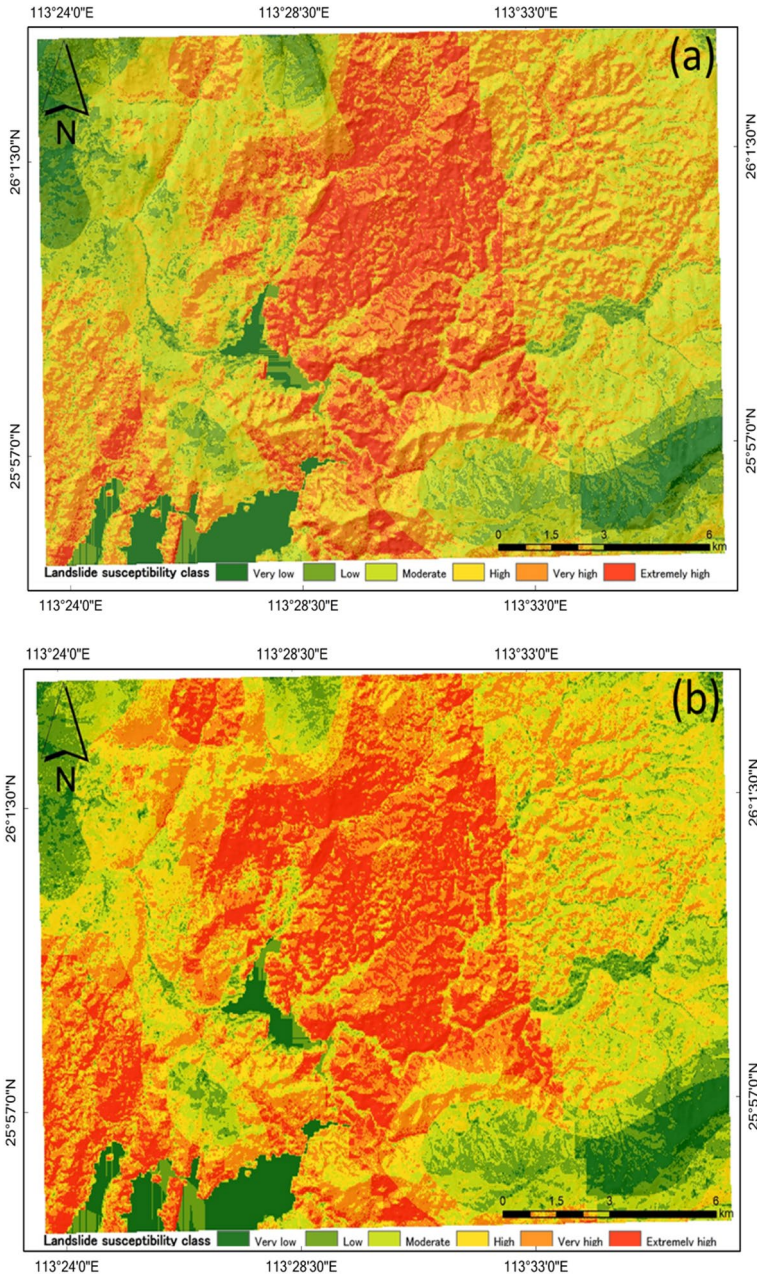


Fig. 13 LSM maps produced by the BLR method: **a** selected seven factors and **b** original 12 factors. Maps show the spatial probability of landslide occurrence in six classes. The upstream of the reservoir is located at the lower left corner of the map

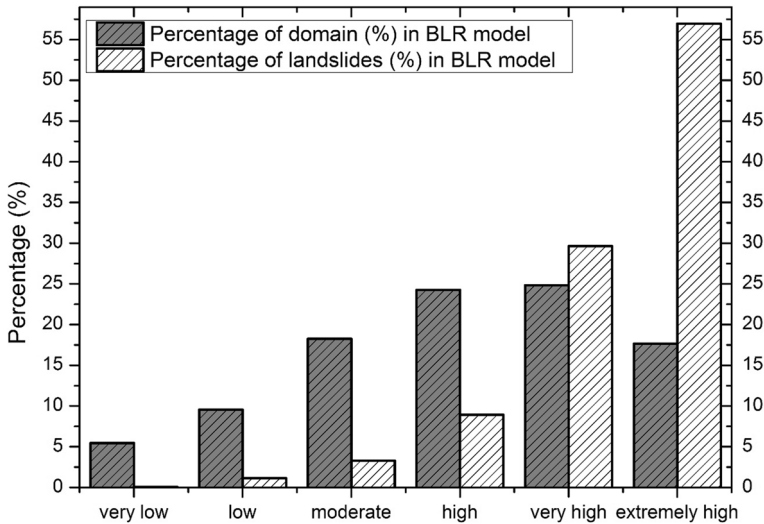


Fig. 14 Susceptibility class distribution within the study area and the occurrence of landslides according to the classification scheme for LSM using the BLR method with the original 12 factors

Table 7 Result of statistical analysis concerning landslide susceptibility from the BLR method with the original 12 factors

Class	Area of each class	Percentage of domain	No. of landslides	Percentage of landslides
Very low	16,245	5.46	1	0.05
Low	28,468	9.56	25	1.13
Moderate	54,336	18.25	73	3.31
High	72,201	24.25	197	8.93
Very high	73,890	24.82	654	29.63
Extremely high	52,568	17.66	1257	56.96

this study is the location of the reservoir and its implications. During the heavy rains that drenched the area, the fluctuation of groundwater might have played a very important role in triggering landslides around the reservoir. The slopes tend to lose their stability due to the loss of suction under this circumstance. Previous studies have indicated precipitation, subsequent infiltration, groundwater circulation patterns, and the resultant increase in the hydrostatic pressures that have cumulated over long periods in triggering the landslides (de Montety et al. 2007; Ronchetti et al. 2009). Debieche et al. (2012) in their study pointed out that the influence of flow path and aquifer complexity in the hydrogeology of a landslide. Susceptibility assessments may also be influenced by other important factors such as lithology as noticed in the CF analysis. The weathering of granite bedrock provided a source for forming into the residual soil. Under the unsaturated conditions, residual soil depositions are probably the frequent prone to induce landslides associated with long-duration rainfall (Regmi et al. 2013; Yamagishi et al. 2004). The permeability and drainage characteristics of the area also affected the large-scale movement of boulders and sediments. Based on

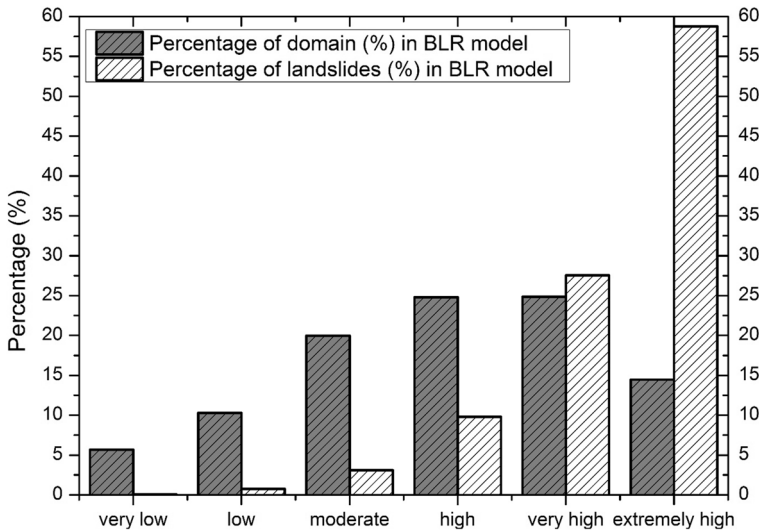


Fig. 15 Susceptibility class distribution within the study area and the occurrence of landslides according to the classification scheme for LSM using the BLR method with selected seven factors

Table 8 Result of statistical analysis concerning landslide susceptibility from the BLR method with the selected seven factors

Class	Area of each class	Percentage of domain	No. of landslides	Percentage of landslides
Very low	16,889	5.67	1	0.05
Low	30,600	10.28	17	0.77
Moderate	59,412	19.96	68	3.08
High	73,816	24.79	216	9.79
Very high	74,014	24.86	608	27.55
Extremely high	42,977	14.44	1297	58.77

the degree of fracturing and weathering, the underlying rock could have acted as a sink or as a source for groundwater in the overlying landslide and should be very crucial for slope stability analyses. Studies in the Japanese archipelago by various researchers in granitic terrains of central Japan found that groundwater flowing in permeable weak and fractured rocks seeps into the overlying unconsolidated sediments (Asano et al. 2003; Katsura et al. 2008), resulting in landslides.

Additionally, previous research by the authors in Sado Island, Japan (Dou et al. 2015c), has found that the drainage density, lithology, and slope angle are the typical factors. These findings also agree with the other studies around the world (Jebur et al. 2014; Dou et al. 2015c). For instance, the drainage density can provide an indirect measure of groundwater conditions that play an important role in landslide activity (Dou et al. 2015c). Thus, these landslide factors may be common to various areas in the world. We believe that our research findings differ from the others in a way that we provide a method to select and qualify the landslide-predisposing factors. The comparison of BSA and BLR with the support of

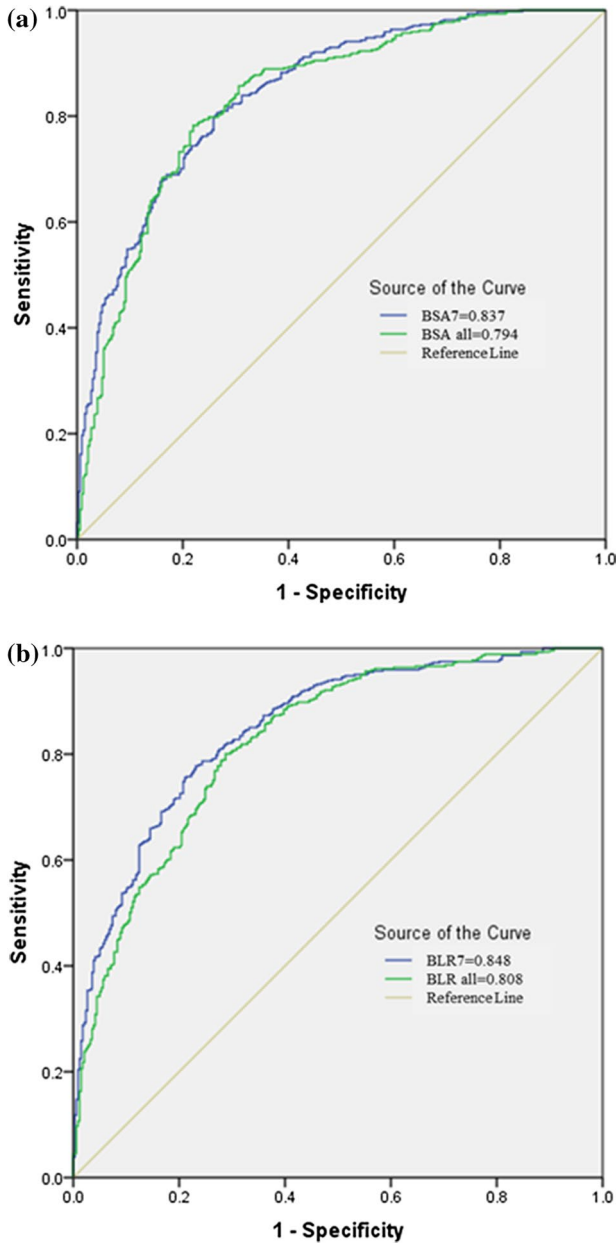


Fig. 16 **a** ROC curves for landslide susceptibility maps produced using BSA with the selected seven and original 12 factors; **b** ROC curves for landslide susceptibility maps produced using BLR with the selected seven and original 12 factors

respective AUC values suggests that logistic regression has a better performance than BSA. This conclusion is also in a good agreement with the other researchers around the world (Chen and Wang 2007; Devkota et al. 2013). However, both the BLR- and BSA-derived

LSM maps ought to leave some stripping, called here as ghost effect (Fig. 16). These ghost effects can be largely attributed from the buffer zone reproduction of drainage density and distance to drainage networks. Saha et al. (2005) also reported ghost effects in their LSM because of structural discontinuity buffering while producing landslide nominal susceptibility factor for Himalayas. Nevertheless, the resultant prediction maps from data-driven models are very much helpful in emergency response and management of the Dongjiang region.

7 Conclusions

This study explores characteristics of landslides induced by the Typhoon Bilis. Due to the orographic effects, around the reservoir areas are likely to have received extremely high rainfall totals. Two main reasons are responsible for landslide event: (1) torrential rainfall at the high intensity and rainfall duration and (2) serious weathering rock formed into considerable sediment, thus combined with the water formation into mudslides downstream. Additionally, this research determines the usefulness of the CF model in identifying the fitted predisposing factors for LSM mapping. Based on the CF model, seven influencing factors with the high correlations to landslide occurrence were selected from a set of original factors. The LSM maps were then produced by the BSA and BLR methods for the CF-identified predisposing factors and the original set of factors. Both the success rate and prediction rate indicated for both the BSA and BLR methods that the seven factors obtain better results than that of all factors. In addition, we noticed that the maps prepared by using seven predisposing factors have much more homogeneous classes than the original factors. The proposed certainty factor method provides a useful way to select the predisposing factors of landslides in particular where data redundancy or scarcity is critical. The findings acknowledge that in the mountainous regions suffering from data scarcity, it is possible to select key factors related to landslide occurrence based on the CF models in a GIS platform. Moreover, in this research, BLR has slightly outperformed the others such as frequency ratio, BSA, which agrees with results from some other researchers in the world.

We believe that the results of our studies provide helpful information for disaster management, urban planning, risk mitigation, and related decision making in landslide-prone areas. For example, in the study areas, the resultant landslide susceptibility maps can be conducive to select appropriate locations for urban development to increase economic benefits and decrease future damages and loss of lives.

Acknowledgements The authors would like to thank Professor Dr. Li Tiefeng of China Geological Survey (CGS) for providing the satellite image data. Dou also expresses his great gratitude to Dr. Uttam, and Dr. Zou Yi for their constructive comments and support. This research work has been supported by the National Key R&D Program of China (ID: 2018YFC1504803) and the National Nature Science Foundation of China (Grant Nos. 51679127 and 51439003).

References


- Aksoy B, Ercanoglu M (2012) Landslide identification and classification by object-based image analysis and fuzzy logic: an example from the Azdavay region (Kastamonu, Turkey). *Comput Geosci* 38:87–98. <https://doi.org/10.1016/j.cageo.2011.05.010>

- Arnone E, Francipane A, Noto LV et al (2014) Strategies investigation in using artificial neural network for landslide susceptibility mapping: application to a Sicilian catchment. *J Hydroinform*. <https://doi.org/10.2166/hydro.2013.191>
- Arnone E, Francipane A, Scarbaci A et al (2016) Effect of raster resolution and polygon-conversion algorithm on landslide susceptibility mapping. *Environ Model Softw*. <https://doi.org/10.1016/j.envsoft.2016.07.016>
- Asano Y, Uchida T, Ohte N (2003) Hydrologic and geochemical influences on the dissolved silica concentration in natural water in a steep headwater catchment. *Geochim Cosmochim Acta* . [https://doi.org/10.1016/S0016-7037\(02\)01342-X](https://doi.org/10.1016/S0016-7037(02)01342-X)
- Binagli E, Luzzi L, Madella P et al (1998) Slope instability zonation: a comparison between certainty factor and fuzzy Dempster–Shafer approaches. *Nat Hazards* 17:77–97. <https://doi.org/10.1023/A:1008001724538>
- Bui DT, Lofman O, Revhaug I, Dick O (2011) Landslide susceptibility analysis in the Hoa Binh province of Vietnam using statistical index and logistic regression. *Nat Hazards* 59:1413–1444. <https://doi.org/10.1007/s11069-011-9844-2>
- Bui DT, Pradhan B, Lofman O, Revhaug I (2012) Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and naive Bayes models. *Math Probl Eng*. <https://doi.org/10.1155/2012/974638>
- Camilo DC, Lombardo L, Mai PM et al (2017) Handling high predictor dimensionality in slope-unit-based landslide susceptibility models through LASSO-penalized generalized linear model. *Environ Model Softw*. <https://doi.org/10.1016/j.envsoft.2017.08.003>
- Carrara A (1983) Multivariate models for landslide hazard evaluation. *J Int Assoc Math Geol* 15:403–426. <https://doi.org/10.1007/BF01031290>
- Chauhan S, Sharma M, Arora MK (2010) Landslide susceptibility zonation of the Chamoli region, Garhwal Himalayas, using logistic regression model. *Landslides* 7:411–423. <https://doi.org/10.1007/s10346-010-0202-3>
- Chen Z, Wang J (2007) Landslide hazard mapping using logistic regression model in Mackenzie Valley, Canada. *Nat Hazards* 42:75–89. <https://doi.org/10.1007/s11069-006-9061-6>
- Chen W, Li W, Hou E, Li X (2016) GIS-based landslide susceptibility mapping using analytical hierarchy process (AHP) and certainty factor (CF) models for the Baozhong region of Baoji City, China. *Environ Earth Sci* 75:3951. <https://doi.org/10.1007/s12665-015-4795-7>
- Chen W, Li H, Hou E et al (2018) GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. *Sci Total Environ*. <https://doi.org/10.1016/j.scitotenv.2018.04.055>
- Chung C-J, Fabbri AG (1993) The representation of geoscience information for data integration. *Nonrenew Resour* 2:122–139. <https://doi.org/10.1007/BF02272809>
- de Montety V, Marc V, Emblanch C et al (2007) Identifying the origin of groundwater and flow processes in complex landslides affecting black marls: insights from a hydrochemical survey. *Earth Surf Proc Land*. <https://doi.org/10.1002/esp.1370>
- Debieche TH, Bogaard TA, Marc V et al (2012) Hydrological and hydrochemical processes observed during a large-scale infiltration experiment at the Super-Sauze mudslide (France). *Hydrol Process*. <https://doi.org/10.1002/hyp.7843>
- Devkota KC, Regmi AD, Pourghasemi HR et al (2013) Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. *Nat Hazards* 65:135–165. <https://doi.org/10.1007/s11069-012-0347-6>
- Dou J, Oguchi T, S. Hayakawa Y et al (2014) GIS-based landslide susceptibility mapping using a certainty factor model and its validation in the Chuetsu area, Central Japan. In: *Landslide science for a safer geoenvironment*. Springer, Cham, pp 419–424
- Dou J, Chang KT, Chen S et al (2015a) Automatic case-based reasoning approach for landslide detection: integration of object-oriented image analysis and a genetic algorithm. *Remote Sens* 7:4318–4342. <https://doi.org/10.3390/rs70404318>
- Dou J, Paudel U, Oguchi T et al (2015b) Shallow and deep-seated landslide differentiation using support vector machines: a case study of the Chuetsu Area, Japan. *Terr Atmos Ocean Sci* 26:227. [https://doi.org/10.3319/TAO.2014.12.02.07\(EOSI\)](https://doi.org/10.3319/TAO.2014.12.02.07(EOSI))
- Dou J, Yamagishi H, Pourghasemi HR et al (2015c) An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan. *Nat Hazards* 78:1749–1776. <https://doi.org/10.1007/s11069-015-1799-2>

- Dou J, Yamagishi H, Xu Y et al (2017) Characteristics of the torrential rainfall-induced shallow landslides by Typhoon Biliis, in July 2006, using remote sensing and GIS. In: Yamagishi H, Bhandary NP (eds) GIS landslide. Springer Japan, Tokyo, pp 221–230
- Dou J, Yamagishi H, Zhu Z et al (2018) TXT-tool 1.081-6.1 A comparative study of the binary logistic regression (BLR) and artificial neural network (ANN) models for gis-based spatial predicting landslides at a regional scale. In: Landslide dynamics: ISDR-ICL landslide interactive teaching tools. Springer, Cham, pp 139–151
- Dou J, Yunus AP, Tien Bui D et al (2019a) Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. *Sci Total Environ* 662:332–346. <https://doi.org/10.1016/j.scitotenv.2019.01.221>
- Dou J, Yunus AP, Tien Bui D et al (2019b) Evaluating GIS-based multiple statistical models and data mining for earthquake and rainfall-induced landslide susceptibility using the LiDAR DEM. *Remote Sens* 11:638. <https://doi.org/10.3390/rs11060638>
- Ghimire B, Mainali KP, Lekhak HD et al (2011) Regeneration of *Pinus wallichiana* AB Jackson in a trans-Himalayan dry valley of north-central Nepal. *Himal J Sci* 6:19–26. <https://doi.org/10.3126/hj.s.v6i8.1798>
- Hovius N, Stark CP, Tutton MA, Abbott LD (1998) Landslide-driven drainage network evolution in a pre-steady-state mountain belt: Finisterre Mountains, Papua New Guinea. *Geology* 26:1071–1074
- Huang R (2007) Large-scale landslides and their sliding mechanisms in China since the 20th century. *Chin J Rock Mech Eng* 26:433–454
- Jebur MN, Pradhan B, Tehrany MS (2014) Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. *Remote Sens Environ* 152:150–165. <https://doi.org/10.1016/j.rse.2014.05.013>
- Katsura S, Kosugi K, Mizutani T et al (2008) Effects of bedrock groundwater on spatial and temporal variations in soil mantle groundwater in a steep granitic headwater catchment. *Water Resour Res*. <https://doi.org/10.1029/2007WR006610>
- Lee S, Pradhan B (2006) Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* 4:33–41. <https://doi.org/10.1007/s10346-006-0047-y>
- Li WY, Liu C, Scaioni M et al (2017) Spatio-temporal analysis and simulation on shallow rainfall-induced landslides in China using landslide susceptibility dynamics and rainfall I-D thresholds. *Sci China Earth Sci* 60:720–732. <https://doi.org/10.1007/s11430-016-9008-4>
- Lucas PJF (2001) Certainty-factor-like structures in Bayesian belief networks. *Knowl Based Syst* 14:327–335. <https://doi.org/10.1007/3-540-46238-4-3>
- Måren IE, Karki S, Prajapati C et al (2015) Facing north or south: does slope aspect impact forest stand characteristics and soil properties in a semiarid trans-Himalayan valley? *J Arid Environ* 121:112–123. <https://doi.org/10.1016/j.jaridenv.2015.06.004>
- Martha TR, Roy P, Govindharaj KB et al (2014) Landslides triggered by the June 2013 extreme rainfall event in parts of Uttarakhand state, India. *Landslides*. <https://doi.org/10.1007/s10346-014-0540-7>
- Neuland H (1976) A prediction model of landslides. CATENA. [https://doi.org/10.1016/0341-8162\(76\)90011-4](https://doi.org/10.1016/0341-8162(76)90011-4)
- Oguchi T (1997) Drainage density and relative relief in humid steep mountains with frequent slope failure. *Earth Surf Proc Land* 22:107–120. [https://doi.org/10.1002/\(SICI\)1096-9837\(199702\)22:2%3c107:AID-ESP680%3e3.0.CO;2-U](https://doi.org/10.1002/(SICI)1096-9837(199702)22:2%3c107:AID-ESP680%3e3.0.CO;2-U)
- Olivero AM, Hix DM (1998) Influence of aspect and stand age on ground flora of southeastern Ohio forest ecosystems. *Plant Ecol* 139:177–187. <https://doi.org/10.1023/A:1009758501201>
- Parsons S (2019) Italy, Austria and China top the list of countries at high risk of landslides right now. <https://resourcewatch.org/blog/2018/08/27/italy-austria-and-china-top-the-list-of-countries-at-high-risk-of-landslides-right-now/>. Accessed 9 May 2018
- Petley D (2012) Global patterns of loss of life from landslides. *Geology* 40:927–930. <https://doi.org/10.1130/G33217.1>
- Pham BT, Prakash I, Dou J et al (2018) A novel hybrid approach of landslide susceptibility modeling using rotation forest ensemble and different base classifiers. *Geocarto Int*. <https://doi.org/10.1080/10106049.2018.1559885>
- Pham BT, Nguyen MD, Bui KTT et al (2019) A novel artificial intelligence approach based on multi-layer perceptron neural network and biogeography-based optimization for predicting coefficient of consolidation of soil. CATENA. <https://doi.org/10.1016/j.catena.2018.10.004>
- Rabonza ML, Felix RP, Lagmay AMFA et al (2016) Shallow landslide susceptibility mapping using high-resolution topography for areas devastated by super typhoon Haiyan. *Landslides* 13:201–210. <https://doi.org/10.1007/s10346-015-0626-x>

- Regmi AD, Yoshida K, Nagata H et al (2013) The relationship between geology and rock weathering on the rock instability along Mugling–Narayanghat road corridor, Central Nepal Himalaya. *Nat Hazards* 66:501–532. <https://doi.org/10.1007/s11069-012-0497-6>
- Reichenbach P, Rossi M, Malamud BD et al (2018) A review of statistically-based landslide susceptibility models. *Earth Sci Rev* 180:60–91
- Ronchetti F, Borgatti L, Cervi F et al (2009) Groundwater processes in a complex landslide, northern Apennines, Italy. *Nat Hazards Earth Syst Sci*. <https://doi.org/10.5194/nhess-9-895-2009>
- Saha AK, Gupta RP, Sarkar I et al (2005) An approach for GIS-based statistical landslide susceptibility-zonation-with a case study in the Himalayas. *Landslides* 2:61–69. <https://doi.org/10.1007/s10346-004-0039-8>
- Shahabi H, Khezri S, Bin B et al (2014) Landslide susceptibility mapping at central Zab basin, Iran: a comparison between analytical hierarchy process, frequency ratio and logistic regression models. *CATENA* 115:55–70. <https://doi.org/10.1016/j.catena.2013.11.014>
- Shortliffe EH, Buchanan BG (1975) A model of inexact reasoning in medicine. *Math Biosci* 23:351–379
- Simon N, Crozier M, de Roiste M, Rafek AG (2013) Point based assessment: Selecting the best way to represent landslide polygon as point frequency in landslide investigation. *Electron J Geotech Eng* 18:775–784
- USGS (2019) How many deaths result from landslides each year? https://www.usgs.gov/faqs/how-many-deaths-result-landslides-each-year?qt-news_science_products=0#qt-news_science_products. Accessed 9 May 2018
- van Westen CJ, Rengers N, Terlien MTJ, Soeters R (1997) Prediction of the occurrence of slope instability phenomenal through GIS-based hazard zonation. *Geol Rundsch* 86:404–414. <https://doi.org/10.1007/s005310050149>
- Wang W, Chen X, Shi P, van Gelder PHAJM (2008) Detecting changes in extreme precipitation and extreme streamflow in the Dongjiang River Basin in southern China. *Hydrol Earth Syst Sci* 12:207–221. <https://doi.org/10.5194/hess-12-207-2008>
- Wang F, Wu Y-H, Yang H et al (2015) Preliminary investigation of the 20 August 2014 debris flows triggered by a severe rainstorm in Hiroshima City, Japan. *Geoenviron Disasters* 2:17. <https://doi.org/10.1186/s40677-015-0025-6>
- Xu Y, He H, Shen X (2011) Case study on dynamic survey of group geological disasters in dongjiang reservoir region, Zixing City using CBERS images. *Acta Scientiarum Naturalium Universitatis Pekinensis* 47:689–697
- Yamagishi H, Marui H, Ayalew L et al (2004) Estimation of the sequence and size of the Tozawagawa landslide, Niigata, Japan, using aerial photographs. *Landslides* 1:299–303. <https://doi.org/10.1007/s10346-004-0032-2>
- Youssef AM, Pradhan B (2014) Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia. *Environ Earth Sci*. <https://doi.org/10.1007/s12665-014-3661-3>
- Zêzere JL, Pereira S, Melo R et al (2017) Mapping landslide susceptibility using data-driven methods. *Sci Total Environ* 589:250–267. <https://doi.org/10.1016/j.scitotenv.2017.02.188>
- Zhu Z, Peng D, Dou J (2017) Changes in the two-dimensional and perimeter-based fractal dimensions of kaolinite flocs during flocculation: a simple experimental study. *Water Sci Technol*. <https://doi.org/10.2166/wst.2017.603>
- Zou Y, Qiu S, Kuang Y, Huang N (2013) Analysis of a major storm over the Dongjiang reservoir basin associated with Typhoon Bilis (2006). *Nat Hazards* 69:201–218. <https://doi.org/10.1007/s11069-013-0696-9>

Affiliations

Jie Dou^{1,2}  · Ali P. Yunus^{3,4} · Yueren Xu⁵ · Zhongfan Zhu⁶ · Chi-Wen Chen⁷ · Meheub Sahana⁸ · Khabat Khosravi⁹ · Yong Yang¹⁰ · Binh Thai Pham¹¹

✉ Jie Dou
douj888@gmail.com

✉ Binh Thai Pham
binhpt@utt.edu.vn

Ali P. Yunus
yunusp@csis.u-tokyo.ac.jp

Yueren Xu
yuefei189@gmail.com

Zhongfan Zhu
zhuzhongfan1985@bnu.edu.cn

Chi-Wen Chen
kevin4919@gmail.com

Meheub Sahana
mehebubsahana@gmail.com

Khabat Khosravi
Khabat.khosravi@gmail.com

Yong Yang
yyhefei888@163.com

¹ Key Laboratory of Geological Hazards on Three Gorges Reservoir Area (China Three Gorges University), Ministry of Education, Yichang, China

² Department of Civil and Environmental Engineering, Nagaoka University of Technology, Nagaoka, Japan

³ Department of Remote Sensing and GIS Applications, Aligarh Muslim University, Aligarh 202002, India

⁴ State Key Laboratory of Geohazard Prevention and Geoenvironment Protection (SKLGP), Chengdu University of Technology, Chengdu, China 610059

⁵ Key Laboratory of Earthquake Prediction, Institute of Earthquake Forecasting, China Earthquake Administration, Beijing 100036, China

⁶ College of Water Sciences, Beijing Normal University, Xijiekouwai Street 19, Beijing 100875, China

⁷ National Science and Technology Center for Disaster Reduction, No. 200, Sec. 3, Beixin Road, Xindian District, New Taipei City, Taiwan

⁸ Department of Geography, Jamia Millia Islamia, New Delhi, India

⁹ Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

¹⁰ Department of Watershed Management, Sari Agricultural Science and Natural Resources University, Sari, Iran

¹¹ Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam