Regional landslide hazard assessment through integrating susceptibility index and rainfall process

Zhiheng Wang^{1,2} · Dongchuan Wang² · Qiaozhen Guo² · Daikun Wang³

Received: 3 May 2019 / Accepted: 24 August 2020 / Published online: 5 September 2020
© The Author(s) 2020

Abstract

Due to the difference of the spatial and temporal distribution of rainfall and the complex diversity of the disaster-prone environment (topography, geological, fault, and lithology), it is difficult to assess the hazard of landslides at the regional scale quantitatively only considering rainfall condition. Based on detailed landslide inventory and rainfall data in the hilly area in Sichuan province, this study analyzed the effects of both rainfall process and environmental factors on the occurrence of landslides. Through analyzing environmental factors, a landslide susceptibility index (LSI) was calculated using multiple layer perceptron (MLP) model to reflect the regional landslide susceptibility. Further, the characteristics of rainfall process and landslides were examined quantitatively with statistical analysis. Finally, a probability model integrating LSI and rainfall process was constructed using logistical regression analysis to assess the landslide hazard. Validation showed satisfactory results, and the inclusion of LSI effectively improved the accuracy of the landslide hazard assessment: Compared with only considering the rainfall process factors, the accuracy of the landslide prediction model both considering the rainfall process and landslide susceptibility is improved by 3%. These results indicate that an integration of susceptibility index and rainfall process is essential in improving the timeliness and accuracy of regional landslide early warning.

Keywords Landslides · Hazard assessment · Susceptibility influence · Logistic regression

✉ Zhiheng Wang
wangzhiheng@tcu.edu.cn

✉ Daikun Wang
derekwang@link.cuhk.edu.hk

¹ Tianjin Key Laboratory of Soft Soil Characteristics and Engineering Environment, Tianjin Chengjian University, Tianjin 300384, China

² School of Geology and Geomatics, Tianjin Chengjian University, Tianjin 300384, China

³ Department of Geography and Resource Management, The Chinese University of Hong Kong, Hong Kong 999077, China

1 Introduction

Landslides, defined as the downslope movement of rock, debris, or earth due to the influence of gravity, have caused significant property damages, personal injuries, or death worldwide (Abdallah and Faour 2017; Bui et al. 2017; Canli et al. 2018; Dowling and Santi 2014; Petley 2012; Segoni et al. 2018a, b). It is estimated that the annual economic loss caused by the landslide disaster in China is about \$ 50 million (Sang 2013). Due to the significant impact of landslides on properties and human lives, the assessment of causal factors of landslide hazard is essential. Landslides can be triggered by various external events such as intense rainfall, earthquakes, water-level changes, storm waves, or human activities, of which rainfall is the most common and important triggering factor (Brunetti et al. 2018; Olyazadeh et al. 2017; Palenzuela et al. 2016; Rossi et al. 2017). Most rainfall-induced landslides occur in the middle and late stages of the rainfall process or with a lag of a few days (Hong et al. 2018; Lee et al. 2014; Mathew et al. 2014). Moreover, landslides are also related to the pre-precipitation conditions, and the impact to rainfall-induced landslides varies differently (Bogaard and Greco 2018; Cardinali et al. 2006; Gabet et al. 2004; Segoni et al. 2018a, b; Yin 2004). In order to understand the effects of rainfall on landslides, two main perspectives have been carried out. The first one examines the thresholds of rainfall (e.g., length, intensity) causing landslides. The second explores the relationship between landslides and rainfall process by constructing statistical models such as multivariate regression, neural network, and support vector machine (De Luca and Versace 2016; Guzzetti et al. 2007; Jiusheng et al. 2004; Marra 2019; Naidu et al. 2018). In the first perspective, two major thresholds of rainfalls have been identified as (1) intensity–duration (ID) relationships and (2) antecedent precipitation (AP) schemes. ID relationships provide critical values of rainfall intensity, which can cause slope failure when the critical value is reached or exceeded (Brunetti et al. 2010; Caine 1980; Guzzetti et al. 2008; Rosi et al. 2016); AP schemes define the critical value of rainfall events that accumulate in a short period of time (equivalent to hours or days before the occurrence of landslide), depending on the long-term accumulated antecedent precipitation (Aleotti 2004; Chleborad 2003; Crozier 1999; Glade et al. 2000; Kanungo and Sharma 2014; Lagomarsino et al. 2015). With the second perspective, probability models have been constructed to assess regional landslides hazards. The relationships between landslides and rainfall process have been examined using logistic regression, artificial neural networks, multivariate linear regression, etc. (Gorsevski et al. 2001; Li et al. 2008; Yin 2004).

Until now, most studies only consider the relationship between landslides and rainfall processes, such as daily rainfall and antecedent rainfall. Only a few studies have considered the coupling effect of rainfall processes and environmental factors (such as topography, landform, lithology, and fault) on the prediction of regional landslides (AUFLIĆ et al. 2016; Hong et al. 2007; Segoni et al. 2018a, b). However, these environmental factors may be crucial in affecting the occurrence of landslides. If these environmental factors are not taken into account, the accuracy of the developed model is questionable. Meanwhile, the structural relationship between rainfall processes and landslides is sensitive to different environments. In order to solve this problem, we analyze the impact on landslide occurrence with the coupling effect of landslide susceptibility and the rainfall process to improve the accuracy of regional landslides hazard assessment. In this paper, we first develop a landslide susceptibility index (LSI) using the multiple layer perceptron (MLP) model to

comprehensively analyze environmental factors. Then, the impact of rainfall on landslides is analyzed by calculating the effective rainfall based on the detailed landslide inventory and rainfall data. Finally, the integrated impact of LSI and rainfall processes on landslides is examined using a logistic regression model to evaluate the landslide hazard in hilly area of Sichuan province.

To examine the impact of environmental factors and rainfall processes on the occurrence of landslides, we have developed a procedure that includes (1) use multiple layer perceptron (MLP) to assess landslide susceptibility, (2) use statistical analysis to examine the relationship between landslides and rainfall processes, and (3) use logistic regression model to evaluate rainfall-induced landslide hazards. The workflow of this study is shown in Fig. 1.

2 Study area

In this study, the hilly area located in the eastern part of Sichuan province, China, is selected (Fig. 2). This site is located between the latitudes of 27.9°–32.4°N and longitudes of 102.9°~108.1°E, with the minimum elevation of 197 m and maximum elevation of 2234 m. The geographic area is 112,000 km², accounting for approximately 23% of Sichuan province. The study area belongs to the subtropical humid monsoon climate zone, with an annual precipitation exceeding 1000 mm, and rainfall (mainly rainstorm) mostly occurs in summer.

During the rainy season, landslides are very frequent. Meanwhile, this is also the area with high population density and economic development. The dense landslides lead to huge losses (Tang 1994). Affected by special topography, geology and meteorological

Fig. 1 Workflow of regional landslide hazard assessment

Fig. 2 The study area: hilly area of Sichuan province, China

conditions, the landslides in the study area are mainly rainfall-induced shallow landslides, most of which are small and medium scale (Zhang 2009).

3 Materials and methods

3.1 Data

3.1.1 Landslides data

The rainfall-induced landslide data in the study area are obtained from a field investigation conducted by China Geological Environmental Monitoring Institute in 2007 (Table 1). With the field investigation, a total of 648 landslides are identified, and the date of occurrence, inducing factors, and locations of the landslides are recorded. A portion of landslides data is shown in Table 1.

Through analysis from the information of 648 landslides, it is found the landslides in the study area are mainly shallow and medium-sized landslides, and the predominant factor is mainly continuous rainfall.

Table 1 Landslides in the hilly area of Sichuan province in 2007 (portion)

No.	Month	Day	Scale	Inducing factor	Location
1.	7	9	Small	Rainfall	Community 11 of Feiquan Village, Jine Town, Longchang County, Neijiang City
2.	7	10	Small	Heavy rainfall	Community 11 of Tiantang Village, Hulong Town, Anyue County, Ziyang City
3.	7	5	Small	Rainstorm	Community 4 of Tiantang Village, Hulong Town, Anyue County, Ziyang City
4.	7	17	Small	Rainstorm	Community 9 of Changxin Village, Qingliu Town, Anyue County, Ziyang City
5.	7	19	Small	Heavy rainfall	Communities 5 and 6 of Tianfu Village, Gaosheng Town, Anyue County, Ziyang City
6.	7	10	Small	Rainfall	Communities 1 and 2 of Yongjiang Village, Huguo Town, Naxi County, Luzhou City
7.	7	11	Small	Rainfall	Community 1 of Shuanggui Village, Dagu Town, Naxi County, Luzhou City
8.	7	11	Small	Rainfall	Community 4 of Yunhui Village, Guyun Town, Naxi County, Luzhou City
9.	7	19	Small	Rainfall	Group 6, Liangfeng Village, Anning Town, Matan County, Luzhou City
10.	7	10	Small	Rainfall	Dayan Village, Xianglin Town, Xuyong County, Luzhou City
.....

Fig. 3 Distribution map of meteorological stations

3.1.2 Rainfall data

The daily rainfall data used in this paper are provided by Sichuan Meteorological Data Service Center (<https://www.climate.sc.cn/gx/home.do>). In the study area, there exist 62 meteorological stations, and their spatial distributions are shown in Fig. 3. Based on the nearest distance principle, we associate each landslide with the nearest meteorological station and retrieve the precipitation data on the same day when the landslide occurred.

3.2 Methods

As mentioned in Fig. 1, two critical methods are adopted in this study: multiple layer perceptron (MLP) and logistic regression model. The formation processes of the landslides are a complex physical process. It can be considered as a nonlinear system affected by both predisposing environments (landform, geologic structure, lithology, etc.) and triggering factors (rainfall, earthquake, etc.). The MLP model is based on the back-propagation learning algorithm and has the ability to model nonlinear systems, which meets the complicated characteristics of landslides disaster systems. Therefore, it is suitable for the regional landslides susceptibility assessment (Cong 1998; Moayedi et al. 2019; ThaiPham et al. 2017; Zare et al. 2012). Logistic regression model is used to build the probability model of landslides occurrences in the study area.

• MLP

In order to solve the problem of multi-category decomposition of nonlinear fractional data, multiple layer perceptron (MLP) model is proposed. MLP is a multilayer

feedforward network with one-way error propagation. As one of the most widely used and deeply studied network models, MLP has the prominent feature of high nonlinear mapping ability and has played an important role in image processing, pattern recognition, algorithm optimization, function approximation and self-adaptation (Ling 2004).

The structure of MLP is composed of nodes in different layers. The nodes in each layer are related to the next layer and output to the next layer. The output layer is amplified, attenuated, or suppressed by connecting values. In addition to the output layer, the excitation output value of each node is determined by the input node, excitation function, and bias value. The trained network can form a nonlinear mapping relationship between pattern space and classification space.

- **Logistic regression model**

Logistic regression model is a statistical analysis method that is often used on the regression analysis to binary categorical variables. Unlike a linear ordinary least square (OLS) regression, logistic regression is a nonlinear regression model whose parameters are estimated using the maximum likelihood method. It is proved that under the condition of random samples, the maximum likelihood estimation of the logistic regression model is consistent, progressive, and normal. Among the various statistical methods proposed in the landslide hazard assessment, logistic regression analysis has proved that it is a well-used and established technique (Ayalew and Yamagishi 2005; Chen and Wang 2007; Hemasinghe et al. 2018; Lee and Pradhan 2007).

When building a regional rainfall-induced landslide probability model, the dependent variable is often with a binary structure, that is, the landslide "occurs" or "does not occur," in accordance with the requirements of the logistic regression model. Meanwhile, the logistic regression model has a complete set of criteria for testing models structures and parameters and provides the probability of landslide occurrence in the form of regression values from 0 to 1.

Logistic regression analysis correlates the probability of landslide occurrence (values from 0 to 1) with logarithm (where indicates a higher probability of non-occurrence and indicates a higher probability of occurrence) (Hemasinghe et al. 2018). In logistic regression analysis, the " u " is assumed as a linear combination of independent variables, and formula (1) is given as follows:

$$P = e^u / (1 + e^u) \quad (1)$$

where output P is the probability of a landslide occurrence and u is the independent variable which is a linear combination of the influencing factors (slope, geology, rainfall, and land cover). In formula (2), $\beta_0, \beta_1, \dots, \beta_n$ are corresponding coefficients of each contribution factor indicating their contribution to the occurrences of landslides.

$$\mu = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots \dots + \beta_n x_n. \quad (2)$$

4 Landslide susceptibility assessment

Based on in-depth analysis on the formation mechanism of landslides in the study area, we have identified five types of variables, including topography (elevation, slope, and relief), rock and soil types (lithology), cut slope (river and road networks), geological structures

Table 2 Factors selected for landslide susceptibility assessment

Class	Subclass	Data format	Resolution (scale)	Meaning of factor
Topography	Elevation	Raster	30 m	Elevation has a certain effect on the distribution of rainfall, vegetation type, and vegetation coverage
	Slope	Raster	30 m	Slope plays an important role in controlling the distribution of landslides
	Relief	Raster	30 m	Relief is used to describe and reflect the macroscopic features of the terrain in a large
Rock and soil type	Lithology	Vector	1:500,000	Rock and soil type mainly affects the stable state of the landslide
Cut slope	River	Vector	1:250,000	River erosion will increase the free surface of the slope
	Road	Vector	1:250,000	Due to excessive excavation during the construction of the road, an effective free surface is formed
Geologic structure	Fault	Vector	1:500,000	Soft tectonic plane developed along the fault zone, and the rock is broken
Vegetation	NDVI	Raster	30 m	Vegetation coverage affects the shear strength of the soil, which in turn affects the stability of the slope

Fig. 4 Distributions of factors selected for landslide susceptibility assessment

Table 3 Classification result of factors for landslide susceptibility assessment

Factors	Classification	Processing method
Elevation (m)	0–400, 400–600, 600–800, 800–1000, 1000–1200, > 1200	Reclassify
Slope (°)	0–3, 3–5, 5–15, 15–25, 25–30, 30–45, > 45	
Relief (m)	0–200, 200–600, 600–1, 000	
NDVI	0.1–0.2, 0.2–0.3, 0.3–0.4, 0.4–0.5, 0.5–0.6, 0.6–0.7, 0.7–0.8, 0.8–0.9, 0.9–1.0	Vector to raster Buffer, reclassify
Lithology	Mudstone, limestone, basalt, dolomite, sandstone, conglomerate, etc.	
Distance to river (km)	0–2, 2–4, 4–6, 6–8, 8–10, 10–12, 12–14, > 14	
Distance to road (km)	0–2, 2–4, 4–6, 6–8, 8–10, > 10	
Distance to fault (km)	0–2, 2–4, 4–6, 6–8, 8–10, 10–12, 12–14, 14–16, > 16

Fig. 5 Distribution of historical landslides in the study area

(fault), and vegetation conditions (NDVI), as the major factors for landslide susceptibility assessment (Table 2 and Fig. 4).

The distribution of factors selected for landslide susceptibility assessment is shown in Fig. 4.

In order to build the model, the factors should be preprocessed, and the results are shown in Table 3.

In this study, historical landslides (Fig. 5) used for landslide susceptibility assessment are obtained through field investigations and interpretations from high-resolution remote sensing images.

Based on the distribution of historical landslides and the factors selected for susceptibility assessment, we proposed an MLP model to assess landslide susceptibility in the study area (Wang et al. 2015). In order to obtain reliable results, on the basis of summarizing previous studies and combining the actual situation of the study area, this paper adopts the method of selecting positive samples and negative samples at the same time for sample selection.

Randomly select 80% of the 375 landslide data, that is, 300 landslides (positive samples), and an equivalent of 300 stable slopes (negative samples) as the training samples of the model. The model validation sample utilizes 20% of landslide data and the same number of stable slopes, namely 75 landslides (positive samples) and 75 non-landslides (negative samples). The rainfall-induced landslide susceptibility index is calculated as shown in Fig. 6.

Based on the landslide susceptibility index, the natural breaks method is used to divide the study area into three susceptibility zones: high, middle, and low, as shown in Fig. 7.

Fig. 6 Distribution of landslide susceptibility index in the study area

Fig. 7 Result of landslide susceptibility partitioning in the study area

Fig. 8 The relationship between landslides and susceptibility level

Through statistical analysis, it can be found that the distribution of landslides in the study area is closely related to the susceptibility partitioning. As shown in Fig. 8, 45.1% of the landslides are located in the high-susceptibility area, covering an area of 28,754 km², which is only 25.7% of the total area; 26.4% of the landslides are in the moderate-susceptibility area, which covers an area of 34,917 km², accounting for 31.2% of the total area; and only 28.6% of the landslides area are in the low-susceptibility area, with an area of 48,312 km², accounting for 43.1% of the total area.

In Fig. 8, it is also indicated that landslide susceptibility assessment is a comprehensive assessment of the landform, lithology, geological structure, etc., that are susceptible. It reflects the impact of disaster environment factors on landslides in the study area. Therefore, it is necessary to consider landslide susceptibility index when building the landslide probability model.

5 Examination of the relationship between landslides and rainfall process

In order to determine the relationship between the landslides and rainfall process (daily rainfall, antecedent rainfall, etc.) in the study area, a detailed statistical analysis based on the landslides inventory has been carried out. Through the analysis, it can be found that 72% of landslides occur on the day when the precipitation is greater than 10 mm, and 94% of landslides

Fig. 9 Relationship between rainfall and landslides occurrence

Fig. 10 Fitting curve of the relationship between rainfall and landslides occurrence

occur within 3 days after heavy rainfall. As time goes on, the number of landslides decreases rapidly.

It can be considered that the changes reflected in Fig. 9 represent the contribution of the antecedent rainfall to the occurrence of landslide. The contribution can also be called the weight coefficient of rainfall to landslides. The univariate negative exponential function is used to fit the relationship between the days before the occurrence of landslides and the probability of landslides (Eq. 3 and Fig. 10).

Based on the curve shown in Fig. 10, Eq. (3) can be obtained by fitting the function:

$$f(x) = 0.7194 \times \exp(-1.373 \times x) \quad (3)$$

where x represents the days before the landslides occurred in the study area and $f(x)$ represents the probability of landslides. The curve fitting accuracy of Eq. (3) is high with a R^2 of 0.997.

Equation (3) reflects the relationship between the landslides occurrence and time of rainfall accompanied by different intensities of rainfall infiltration to the slope. It helps to deepen the understanding of the formation mechanism of rainfall-induced landslides in the study area. From the statistical analysis, it is found that the rainfall-induced landslides in the study area are closely related to the daily rainfall and the rainfall of the first 5 days in the hilly area of Sichuan province. Therefore, the days after a major rainfall can be considered as the factors involved in the construction of probability model which is used to evaluate the landslides hazard in the study area. Based on Eq. (3), the corresponding attenuation coefficient is calculated as shown in Table 4.

The calculation method of effective rainfall in the study area is shown in Eq. (4) based on the attenuation coefficient shown in Table 4.

$$f(x) = 0.719 \times R_0 + 0.182 \times R_1 + 0.046 \times R_2 + 0.012 \times R_3 + 0.003 \times R_4 + 0.001 \times R_5 \quad (4)$$

Table 4 Attenuation coefficient of the time before the landslide happened

Days before landslides occurred	Attenuation coefficient
0	0.719
1	0.182
2	0.046
3	0.012
4	0.003
5	0.001

where $f(x)$ is the attenuation function; R_0 is the daily rainfall; and R_n is the previous daily precipitation ($1 \leq n \leq 5$).

Based on Eq. (4), the effective rainfall of daily rainfall and first 1, 2, 3, 4, and 5 days is calculated as shown in Table 5.

The effective rainfall, as the representative factor of the rainfall process affecting landslides, is used to build the probability model of landslides occurrence.

6 Construction of the rainfall-induced landslide probability model

6.1 Model construction

Based on the rainfall process and landslides susceptibility index, a logistic regression algorithm is adopted to establish the rainfall-induced probability model. Table 6 shows the process of stepwise regression.

According to the standard significance level of 0.05, we have found the effective rainfall before 2 and 4 days (Day_2 and Day_4) has no significant relationship with the landslides occurrence, while Day_0 , Day_3 , Day_5 , and LSI have significant relationship with the landslides occurrence. The insignificance of Day_2 and Day_4 may be due to the correlation with other variables.

Table 5 Partial data of the effective rainfall (portion)

No.	Effective rainfall of different days to landslides (mm)					
	0	1	2	3	4	5
1	378.55	145.37	0.20	1.07	4.65	18.41
2	662.11	20.91	0.19	0.84	3.54	14.02
3	784.26	437.64	378.55	145.37	0.20	1.07
4	731.07	662.11	20.91	0.19	0.84	3.54
5	474.46	731.07	662.11	20.91	0.19	0.84
6	137.54	145.30	258.25	769.28	378.55	145.37
7	103.16	137.54	145.30	258.25	769.28	378.55
8	1.61	6.42	26.00	103.16	137.53	145.30
9	33.84	0.14	0.58	2.32	9.36	11.52
10	8.57	33.84	0.14	0.58	2.32	9.36
11						

As shown in Table 6, Day₀, Day₃, Day₅, and LSI are selected to construct a logistic regression model. The rainfall-induced landslides probability model in the hilly area of Sichuan province is shown in Eq. (5).

$$P = \frac{e^z}{1 + e^z} \quad (5)$$

where

$$z = -4.965 + 11.57 \times r_0 + 11.59 \times r_3 + 8.26 \times r_5 + 10.73 \times \text{LSI}$$

P represents the probability of the landslides occurrence, LSI represents the landslide susceptibility index, r_0 represents the effective rainfall in the day, and r_3 and r_5 represent the effective rainfall before 3 and 5 days respectively.

Table 6 Influential factors in the rainfall-induced landslide probability model

Steps	Factors	Coefficient	Standard error	Significance	Exp (B)
1	Day ₀	9.774	2.641	.000	99.310
	Day ₁	2.737	2.396	.253	.141
	Day ₂	1.336	2.212	.546	.050
	Day ₃	10.301	3.546	.004	28.526
	Day ₄	-.564	2.175	.795	.008
	Day ₅	9.174	3.318	.006	14.436
	LSI	11.341	3.872	.003	42.626
	Constant	-5.285	1.340	.000	
2	Day ₀	9.840	2.630	.000	108.313
	Day ₁	2.715	2.384	.255	.141
	Day ₂	1.307	2.199	.552	.050
	Day ₃	9.958	3.263	.002	35.292
	Day ₅	9.029	3.255	.006	14.150
	LSI	11.642	3.714	.002	78.506
	Constant	-5.387	1.289	.000	
3	Day ₀	9.705	2.595	.000	101.423
	Day ₁	3.358	2.170	.122	.409
	Day ₃	10.686	3.112	.001	98.090
	Day ₅	9.119	3.268	.005	15.097
	LSI	11.843	3.706	.001	97.522
	Constant	-5.441	1.289	.000	
4	Day ₀	11.570	2.505	.000	780.814
	Day ₃	11.590	3.095	.000	250.399
	Day ₅	8.265	3.169	.009	7.802
	LSI	10.730	3.555	.003	43.041
	Constant	-4.965	1.220	.000	

(Day₀: the effective rainfall in the day; Day_n: the effective rainfall before n day, (1 ≤ n ≤ 5))

6.2 Model validation

To analyze the accuracy of the probability model, the classification table is used to quantitatively analyze the simulation results of the model based on the qualitative analysis shown in Table 7. From Table 7, we can also find that the prediction results are strongly dependent on the threshold value of the LSI that is used to consider a pixel as stable or unstable.

7 Conclusions

7.1 Results

Based on the above analysis, this article has achieved the following results.

- (1) In the study area, the landslide susceptibility index comprehensively reflects the impact of various disaster environment factors (landform, lithology, geological structure, etc.) on the distribution of landslides. The density of the landslides is significantly raised with the increase in the susceptibility level.
- (2) The occurrence of landslides in the study area is closely related to the rainfall on the day and the first five days. The relationship can be fitted by negative exponential function of one variable, and this equation contributes to deepening the understanding of the mechanism of landslides induced by rainfall in the study area, which approximately reflects the relationship between the time and the landslides occurrence by infiltration into slope under different intensities of rainfall.
- (3) Based on the effective rainfall factor and landslide susceptibility index, the logistic regression model is utilized to build the probability model of rainfall-induced landslides. Through the conditional likelihood ratio backward elimination, it is found the effective rainfall on the day (Day_0), before 3 days (Day_3), before 5 days (Day_5), and LSI had significant relationship with landslides. The AUC value is 0.86; therefore, the prediction accuracy of the model is good.

7.2 Discussion

In order to analyze the influence of landslide susceptibility index on the accuracy of the landslide probability model, Cox and Snell R^2 and Nagelkerke R^2 are used to test the goodness of fit of logistic regression equations. From Table 8, the Cox and Snell R^2 considering the susceptibility index is 0.351 and the Nagelkerke R^2 is 0.467, which are significantly

Table 7 Classification of the model fitting result

Observed		Predicted		
		Landslide		Percentage correct (%)
		0	1	
Landslide	0	65	11	85.5
	1	20	58	74.4
Overall percentage				80.0

Table 8 Result of the model fit test

	Cox and Snell R^2	Nagelkerke R^2
Considering LSI	.351	.467
Not considering LSI	.303	.404

Fig. 11 ROC curve of the model verification without considering landslide susceptibility

higher than the corresponding value without considering the susceptibility index factor. It can be explained that the substitution of landslide susceptibility index factor effectively improves the model's goodness of fit.

Meanwhile, the ROC curve method is utilized to compare the accuracy changes before and after considering the landslides susceptibility index in the landslide probability model.

As shown in Fig. 11, the rainfall-induced landslides probability model constructed has a high value of “diagnosis” value, because the ROC curve area of the model considering the landslide susceptibility index is 0.86; in contrast, the accuracy of the rainfall-induced landslides probability model without considering the landslide susceptibility index is verified by the ROC curve, and the AUC is 0.83. It can be found that the accuracy of the rainfall-induced landslides probability model is improved by 3% by considering the landslide susceptibility index.

In addition, the mechanism of causative factors (rainfall, earthquake, human activities, etc.) affecting the stability of landslides with quantitative method has been the difficulty in regional landslide temporal prediction and will retain an important direction in future

research. The innovation of this paper is that through the statistical analysis of relationship between the landslides occurrences and rainfall processes in this study, the weight coefficient function reflecting the rainfall contribution to landslide was obtained. This function approximately reflects the occurrences of landslides and rainfall with different intensities infiltrating into the slope body, which is helpful to deepen the understanding of the rainfall-induced landslide's formation mechanism in the hilly area of Sichuan province. Based on the function, a more accurate probability model can be built to evaluate the landslides hazard in the study area.

Disaster risk monitoring plays an increasingly important role in the emergency management and is an important measure to strengthen the sustainable development of society, economy, and environment. This paper establishes a probability model of the landslides occurrences in the hilly area of Sichuan province and uses this model to dynamically assess regional landslide hazards in the coming days based on the rainfall forecast data and rainfall real-time monitoring data with map algebra. Combined with the vulnerability assessment results of the disaster bodies in the study area, it is possible to assess the risk of regional landslides in the next few days. The research results can be directly applied to the landslide risk monitoring work in the study area, and the research ideas and methods have important reference value for the landslide hazard assessment and risk monitoring in other geology disaster-prone areas.

However, there are some limitations in this study. Detailed recording of landslides occurrence data (location and time) and the corresponding rainfall data are essential for regional landslide hazard assessment. Based on the statistical analysis, this paper obtains the rainfall weight coefficient function for the hilly area in Sichuan province with the 648 landslide records in the study area and the corresponding rainfall data in 2007. Due to the relatively small amount of the available sample data, the universality of the weight coefficient function obtained above needs to be further verified.

Acknowledgements The authors acknowledge China Institute of Geo-Environment Monitoring for providing historical landslides data. This research was supported by the Natural Science Foundation of Tianjin, China (Grant Nos. 16JCYBJC21400, 18JCYBJC90900), the National Key R&D Program of China (Grant No. 2017YFC1502901), and the Scientific Research Project of Tianjin Municipal Education Commission, China (Grant No. 2018KJ164). The authors also sincerely thank the academic editors and anonymous referees for thorough and constructive reviews.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdallah C, Faour G (2017) Landslide hazard mapping of Ibrahim River Basin, Lebanon. *Nat Hazards* 85(1):237–266
- Aleotti P (2004) A warning system for rainfall-induced shallow failures. *Eng Geol* 73(3–4):247–265
- Auflüch MJ, Šinigoj J, Krivic M, Podboj M, Peternel T, Komac M (2016) Landslide prediction system for rainfall induced landslides in Slovenia (Masprem). *Geologija* 59(2):259–271
- Ayalew L, Yamagishi H (2005) The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* 65(1–2):15–31

- Bogaard T, Greco R (2018) Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: proposing hydro-meteorological thresholds. *Nat Hazards Earth Syst Sci* 18(1):31–39
- Brunetti MT, Melillo M, Peruccacci S, Ciabatta L, Brocca L (2018) How far are we from the use of satellite rainfall products in landslide forecasting? *Remote Sens Environ* 210:65–75
- Brunetti MT, Peruccacci S, Rossi M, Luciani S, Valigi D, Guzzetti F (2010) Rainfall thresholds for the possible occurrence of landslides in Italy. *Nat Hazards Earth Syst Sci* 10(3):447–458
- Bui DT, Tuan TA, Hoang N-D, Thanh NQ, Nguyen DB, Liem NV, Pradhan B (2017) Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization. *Landslides* 14(2):447–458
- Caine N (1980) The rainfall intensity—duration control of shallow landslides and debris flows. *Geografiska Ann Ser A Phys Geogr* 62(541):23–27
- Canli E, Loigge B, Glade T (2018) Spatially distributed rainfall information and its potential for regional landslide early warning systems. *Nat Hazards* 91(1–2):1–25
- Cardinali M, Galli M, Guzzetti F, Ardizzone F, Reichenbach P, Bartocchini P (2006) Rainfall induced landslides in December 2004 in south-western Umbria, central Italy: types, extent, damage and risk assessment. *Nat Hazards Earth Syst Sci* 6(2):237–260
- Chen Z, Wang J (2007) Landslide hazard mapping using logistic regression model in Mackenzie Valley, Canada. *Nat Hazards* 42(1):75–89
- Chleborad AF (2003) Preliminary evaluation of a precipitation threshold for anticipating the occurrence of landslides in the Seattle, Washington, Area. *US Geol Surv Open-File Rep* 3(463):1–17
- Cong S (1998) Neural network theory and applications with MATLAB toolboxes. University of Science and Technology of China Press, Hefei
- Crozier MJ (1999) Prediction of rainfall-triggered landslides: a test of the antecedent water status model. *Earth Surf Process Landforms J Br Geomorphol Res Group* 24(9):825–833
- De Luca DL, Versace P (2016) A general formulation to describe empirical rainfall thresholds for landslides. *Proc Earth Planet Sci* 16:98–107
- Dowling CA, Santi PM (2014) Debris flows and their toll on human life: a global analysis of debris-flow fatalities from 1950 to 2011. *Nat Hazards* 71(1):203–227
- Gabet EJ, Burbank DW, Putkonen JK, Pratt-Sitaula BA, Ojha T (2004) Rainfall thresholds for landsliding in the Himalayas of Nepal. *Geomorphology* 63(3–4):131–143
- Glade T, Crozier M, Smith P (2000) Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical “antecedent daily rainfall model”. *Pure Appl Geophys* 157(6–8):1059–1079
- Gorsevski PV, Foltz RB, Gessler PE, Cundy TW (2001) Statistical modeling of landslide hazard using GIS. In: *Proceedings of the Seventh Federal Interagency Sedimentation Conference*. Reno, Nevada. Washington, DC: US, pp 103–109
- Guzzetti F, Peruccacci S, Rossi M, Stark CP (2007) Rainfall thresholds for the initiation of landslides in central and southern Europe. *Meteorol Atmos Phys* 98(3–4):239–267
- Guzzetti F, Peruccacci S, Rossi M, Stark CP (2008) The rainfall intensity–duration control of shallow landslides and debris flows: an update. *Landslides* 5(1):3–17
- Hemasinghe H, Rangali RSS, Deshapriya NL, Samarakoon L (2018) Landslide susceptibility mapping using logistic regression model (a case study in Badulla District, Sri Lanka). *Proc Eng* 212:1046–1053
- Hong M, Kim J, Jeong S (2018) Rainfall intensity-duration thresholds for landslide prediction in South Korea by considering the effects of antecedent rainfall. *Landslides* 15(3):523–534
- Hong Y, Adler RF, Huffman G (2007) An experimental global prediction system for rainfall-triggered landslides using satellite remote sensing and geospatial datasets. *IEEE Trans Geosci Remote Sens* 45(6):1671–1680
- Jiusheng S, Xiufen L, Li W, Xingqiu Z (2004) Rainfall characteristics analysis on landslide in Jiangxi Province. *Meteorol Month* 1:13–15+21
- Kanungo DP, Sharma S (2014) Rainfall thresholds for prediction of shallow landslides around Chamoli-Joshimath region, Garhwal Himalayas, India. *Landslides* 11(4):629–638
- Lagomarsino D, Segoni S, Rosi A, Rossi G, Battistini A, Catani F, Casagli N (2015) Quantitative comparison between two different methodologies to define rainfall thresholds for landslide forecasting. *Nat Hazards Earth Syst Sci Discuss* 3(10):2413–2423
- Lee ML, Ng KY, Huang YF, Li WC (2014) Rainfall-induced landslides in Hulu Kelang area, Malaysia. *Nat Hazards* 70(1):353–375
- Lee S, Pradhan B (2007) Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* 4(1):33–41

- Li C, Ma S, Zhu X (2008) Forecasting of landslide triggered by rainfall: theory, methods and applications. Geological Publishing House, Beijing, China
- Ling H (2004) The classification model of RS images based on artificial neural network-MLP. *Bull Surv Mapp* 9:29–42
- Marra F (2019) Rainfall thresholds for landslide occurrence: systematic underestimation using coarse temporal resolution data. *Nat Hazards* 95(3):883–890
- Mathew J, Babu DG, Kundu S, Kumar KV, Pant CC (2014) Integrating intensity–duration-based rainfall threshold and antecedent rainfall-based probability estimate towards generating early warning for rainfall-induced landslides in parts of the Garhwal Himalaya, India. *Landslides* 11(4):575–588
- Moayedi H, Osouli A, Bui DT, Foong LK (2019) Spatial landslide susceptibility assessment based on novel neural-metaheuristic geographic information system based ensembles. *Sensors* 19(21):4698
- Naidu S, Sajinkumar KS, Oommen T, Anuja VJ, Samuel RA, Muraleedharan C (2018) Early warning system for shallow landslides using rainfall threshold and slope stability analysis. *Geosci Front* 9(6):1871–1882
- Olyazadeh R, Sudmeier-Rieux K, Jaboyedoff M, Derron M-H, Devkota S (2017) An offline–online Web-GIS Android application for fast data acquisition of landslide hazard and risk. *Nat Hazards Earth Syst Sci* 17(4):549–561
- Palenzuela JA, Jiménez-Perálvarez JD, Chacón J, Irigaray C (2016) Assessing critical rainfall thresholds for landslide triggering by generating additional information from a reduced database: an approach with examples from the Betic Cordillera (Spain). *Nat Hazards* 84(1):185–212
- Petley D (2012) Global patterns of loss of life from landslides. *Geology* 40(10):927–930
- Rosi A, Peternel T, Jemec-Auflič M, Komac M, Segoni S, Casagli N (2016) Rainfall thresholds for rainfall-induced landslides in Slovenia. *Landslides* 13(6):1571–1577
- Rossi M, Kirschbaum D, Valigi D, Mondini AC, Guzzetti F (2017) Comparison of satellite rainfall estimates and rain gauge measurements in Italy, and impact on landslide modeling. *Climate* 5(4):90
- Sang K (2013) Statistics and analysis of landslide disaster data in China in recent 60 years. *Public Commun Sci Technol* 5(10):129+124
- Segoni S, Rosi A, Lagomarsino D, Fanti R, Casagli N (2018) Brief communication: using averaged soil moisture estimates to improve the performances of a regional-scale landslide early warning system. *Nat Hazards Earth Syst Sci* 18(3):807–812
- Segoni S, Tofani V, Rosi A, Catani F, Casagli N (2018) Combination of rainfall thresholds and susceptibility maps for dynamic landslide hazard assessment at regional scale. *Front Earth Sci* 6:1–11
- Tang X (1994) The exploration on the causes of geotectonic to form the regularity of distribution of the mountain cal amity landforms surrounding Sichuan Basin. *J Soil Eros Soil Water Conserv* 2(8):76–84
- ThaiPham B, Bui D, IndraPrakash MBD (2017) Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. *CATENA* 149:52–63
- Wang Z, Hu Z, Zhao W, Guo Q, Wan S (2015) Research on regional landslide susceptibility assessment based on multiple layer perceptron-take the hilly area in Sichuan as example. *J Disaster Prev Mitig Eng* 35(5):691–698
- Yin K (2004) Landslide hazard prediction and evaluation. China University of Geoscience Press, Wuhan, China
- Zare M, Pourghasemi HR, Vafakhah M, Pradhan B (2012) Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: a comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. *Arab J Geosci* 6(8):1–16
- Zhang S (2009) Basics of geological disaster meteorological forecast. China Meteorological Press, Beijing