

Application of statistical index and index of entropy methods to landslide susceptibility assessment in Gongliu (Xinjiang, China)

Qiqing Wang¹ · Wenping Li¹ · Yanli Wu¹ · Yabing Pei¹ · Peng Xie¹

Received: 3 August 2015 / Accepted: 25 January 2016 / Published online: 30 March 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract The purpose of this study is to evaluate and compare the results applying the statistical index and the index of entropy methods for estimating landslide susceptibility in Gongliu County, China. In order to do this, first, a landslide inventory map was constructed mainly based on earlier reports and aerial photographs as well as by carrying out field surveys. Then the landslide inventory was randomly divided into two datasets 70 % (163 landslides) for training the models and the remaining 30 % (70 landslides) was used for validation purpose. The landslide conditioning factors consist of slope angle, slope aspect, altitude, general curvature, plan curvature, profile curvature, distance to rivers, distance to roads, normalized difference vegetation index, sediment transport index, rainfall, and lithology. The relationships between landslide distributions and these parameters were analyzed using the two models, and the results of both the models were then used to calculate the landslide susceptibility of the entire study area. Finally, the accuracy of the landslide susceptibility maps was evaluated based on the area under the curve (AUC) method. The validation results showed that the statistical index model (AUC = 82.51 %) is slightly lower than the index of entropy model (AUC = 82.80 %) for success rate. Nevertheless, for the prediction rate, it was found that the statistical index model (AUC = 77.90 %) is slightly lower than the index of entropy model (AUC = 77.41 %). The landslide susceptibility maps produced from this study were successful and can be useful for preliminary general land use planning and hazard mitigation purpose.

Keywords Landslide · Susceptibility mapping · Statistical index (SI) · Index of entropy (IOE)

Introduction

Landslides are one of the most damaging hazards causing extensive damage of roads, bridges, human dwellings, agricultural lands, forests, communication network, etc., thereby resulting in injuries, loss of life, or damages to the property and environment (Kannan et al. 2015). Globally, landslides cause approximately 1000 deaths per year and property damage of about 4 billion (Lee and Pradhan 2007). In China, landslides happen rather frequently due to the topographical, geological, and environmental factors. It is reported that more than 30,737 hazards associated with landslides occurred in 2012, 2013 and 2014, which caused a total of 1256 people dead or missing, and a direct economic loss of 15.41 billion CNY (<http://www.cigem.gov.cn>). In order to mitigate or control such damage caused by landslides, it is necessary to study landslide phenomena including susceptibility mapping, hazard mapping, and risk assessment systematically (Bijukchhen et al. 2013).

Different researchers have employed different geographic information system (GIS) methods for assessing landslide susceptibility and hazard throughout the world (Lee and Pradhan 2007; Bijukchhen et al. 2013; Kannan et al. 2015; Wang et al. 2015). Many of the recent studies have applied probabilistic models (Lee and Dan 2005; Lee and Sambath 2006; Yilmaz 2009; Akgun et al. 2008). The statistical models, such as the logistic regression, bivariate and multivariate models, have also been applied to landslide susceptibility mapping (Ayalew and Yamagishi 2005; Lee 2005a; Yesilnacar and Topal 2005; Lee and Sambath 2006; Van Den Eeckhaut et al. 2006; Pradhan 2010).

✉ Qiqing Wang
qiqw89@qq.com

¹ School of Resources and Geosciences, China University of Mining and Technology, Xuzhou 221116, China

Varieties of classifications techniques such as fuzzy systems (Ercanoglu and Gokceoglu 2002, 2004; Kanungo et al. 2006; Pourghasemi et al. 2012a; Pradhan et al. 2010) and neural networks (Lee et al. 2003, 2004; Ermini et al. 2005; Pradhan and Lee 2010) etc., have also been mentioned in the various literature. These methods all belong to quantitative methods which are based on a comparison of a set of conditioning factors with spatial distribution of registered landslides. At present, the quantitative methods become more and more popular with the advances in computer and GIS technology. In addition, the qualitative methods such as analytic hierarchy process (Yalcin 2008; Komac 2006; Pourghasemi et al. 2012a; Yalcin et al. 2011; Ayalew et al. 2004) and weighted linear combination models (Ayalew et al. 2004; Akgun et al. 2008) etc., have also been widely used to landslide susceptibility mapping. The qualitative methods are based on heuristic approaches, which depend mainly on the experience of experts (Erener et al. 2015).

The Gongliu County of northwestern China, which is one of serious area in Xinjiang that happened landslide disaster, was selected as suitable for this study. This area had a great deal of landslide damage after heavy rain. These landslides were mainly shallow soil slips and debris flows and tend to occur in low and middle mountain steep area, which often cause enormous property damage and occasionally result in loss of life in this study area. For instance the Mohuer landslide, located in the middle of county, occurred on March 23, 1990; eight people lost their lives; a lot of livestock were buried. A landslide, taking place in the shadow of a freeway of area, killed eight people in April 9, 2004. In this study area, a total of 233 landslides were identified and mapped in the study area (Fig. 1). Up to date, few studies have been carried out on landslide susceptibility analysis in Gongliu. Therefore, it is necessary to assess and manage the study area that is susceptible to landslides to decrease landslide damage carrying out suitable mitigation measures.

In this paper, landslide susceptibility analysis was determined using two techniques including statistical index (SI) and index of entropy (IOE) models to prepare landslide susceptibility map and assess the most suitable results for the study area. For this purpose, the SI and IOE models to acquire the landslide susceptibility map using the ArcGIS 10.0 software (ESRI Inc., Redlands, CA, USA) were developed, applied, verified, and compared in the study area.

Study area

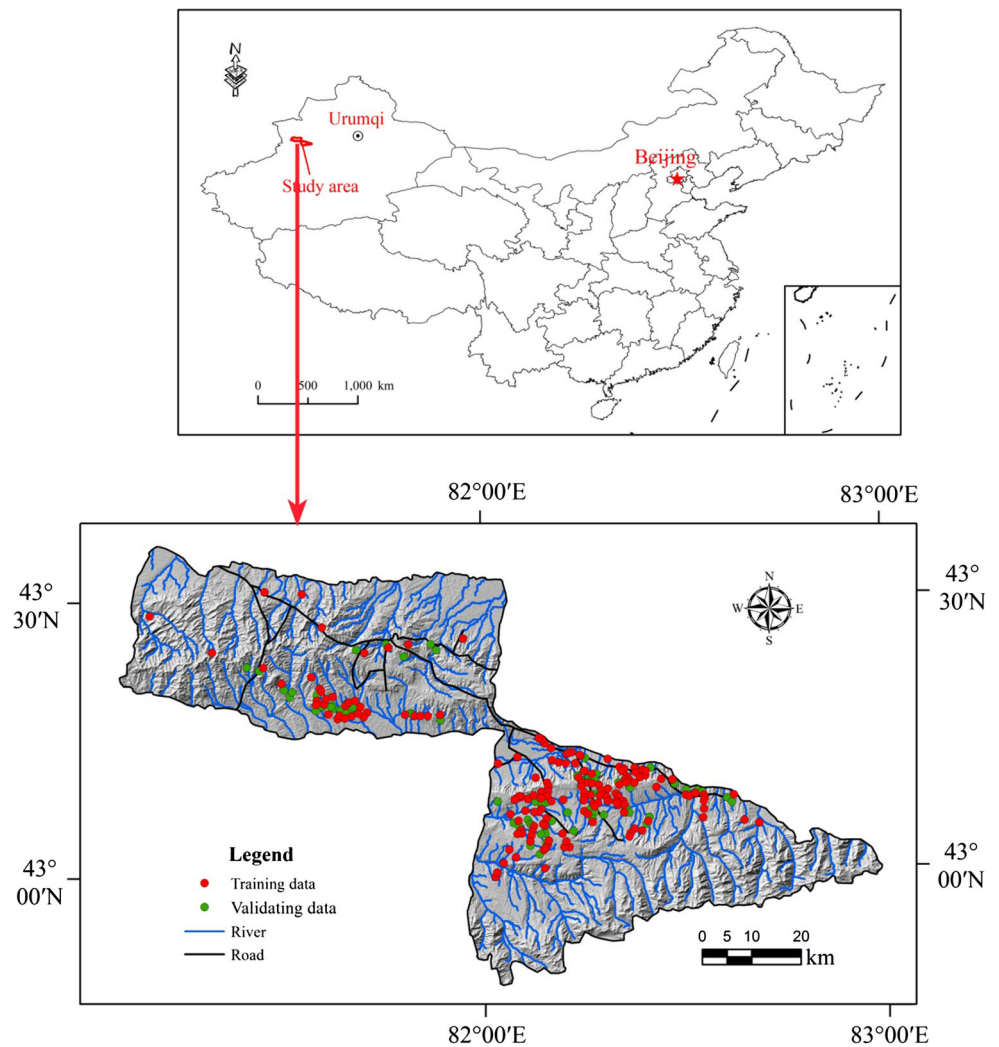
The study area lies in Xinjiang Uygur Autonomous Region of northwestern China and is bounded by the latitudes 42°54' and 43°38'N and the longitudes 81°34' and 83°35'E (Fig. 1).

The altitude of the area ranges from 767 m a.s.l. to as high as 4217 m a.s.l. and the total study area is about 4124 km². The terrain of the study area consists predominantly of the plain region, the mountain region, and the hill region, out of which mountainous and hilly region accounted for about 72.6 % of the total area. The local climate of the area is influenced by altitude. The area is characterized by the typical continental semi-arid climate, the winter is dry and cold, but the summer is hot and rainy. The maximum temperature reaches up to 37–39 °C during the summer, whereas the temperature falls down to –37 °C during the winter. The mean annual rainfall according to local weather station in a period of 40 years is around 200–700 mm, and the most rainfall appears in April–July [C.H. of China Meteorological Administration (CMA) 2014]. The main streams in the region are Jil essential Langhe and Nanshan Rivers, and these rivers and their tributaries form dendritic drainage system due to topographical and geological features of the area. The population of the county was about 196,400 in 2011 year. Major settlements are mainly distributed in the middle of area. Landslides are very common phenomenon in this area due to the coupling effects of special geological and climatic conditions and the influences of human engineering activities.

Data collection and database construction

The landslide inventory map, providing information for the assessment of the influence of different causative factors on landslide occurrence, is essential for the landslide susceptibility analysis (Ozdemir and Altural 2013; Jaafari et al. 2014; Shahabi et al. 2015). It can be prepared by using different techniques, such as field survey, aerial photograph, and satellite image interpretation, and literature search for historical landslide records (Lee and Talib 2005; Pradhan and Kim 2014; Solaimani et al. 2013). The reliability and accuracy of the collected data also affect the success of the applied method (Akgun et al. 2008). For landslide susceptibility assessment, the landslide inventory map of the region was prepared by using 1:50,000 scale aerial-photo (acquired in 2008) interpretation and extensive field surveys. Besides, the historical landslides records obtained from the internet and published literature were also used (Qin 2007). A total of 233 landslides were identified (Fig. 1). The landslide data set was randomly divided in two parts: 163 (70 %) cases are used for assessment and 70 (30 %) cases are kept for validating the landslide susceptibility map. In this study, the 12 factors contributing to landslide occurrence are slope angle, slope aspect, altitude, general curvature, plan curvature, profile curvature, distance to rivers, distance to roads, NDVI, STI, rainfall, and lithology. At first, a 30 m × 30 m digital elevation model (DEM) was collected from the advanced

Fig. 1 The location map of the study area



space-borne thermal emission and reflection radiometer (ASTER) acquired in 2010. Based on DEM, slope angle, slope aspect, altitude, general curvature, plan curvature, profile curvature, and STI maps were prepared. The distance to rivers and distance to roads maps were calculated by Euclidean distance tool of ArcGIS 10.0 based on the drainage and road maps (1:50,000-scale), respectively. The NDVI map was extracted from Landsat 7 ETM+ satellite image with 30-m resolution acquired on 12 November 2010. The lithology map was obtained from a 1:50,000-scale geologic map. All data layers were transformed in raster format with pixel size of 30 × 30 meters, hence the area grid was 3356 rows by 5830 columns with a total of 6430,157 pixels.

Slope angle is one of the parameters controlling the formation of mass movements and movement distance of mobilizing material (Demir et al. 2013). Because the slope angle is directly related to the landslides, it is frequently used in landslide susceptibility studies (Dai et al. 2001;

Dragičević et al. 2015) In this study, the slope angle was divided into seven classes (Fig. 2a) considering the steepness of the terrain (Liu et al. 2014; Kayastha et al. 2013). Slope aspect is also considered as an important factor in landslide studies since aspect affects parameters such as rainfall, discontinuities and exposure to sunlight (Süzen and Doyuran 2004; He et al. 2012). Generally, the slope facing towards the sunlight and rainfall zone have a greater tendency of landslide hazard in comparison with the slope in the rain shadow zone (Bijukchhen et al. 2013). The slope aspect of the present study area was grouped into nine classes (Fig. 2b): flat (−1), north (337.5°–360°, 0°–22.5°), north-east (22.5°–67.5°), east (67.5°–112.5°), southeast (112.5°–157.5°), south (157.5°–202.5°), southwest (202.5°–247.5°), west (247.5°–292.5°), and northwest (292.5°–337.5°).

Generally speaking, the elevation or altitude affects temperature, vegetation, rainfall, and gravitational energy of landslides. In turn, these conditions have the potential impact on slope stability (Meng et al. 2015; Kavzoglu et al.

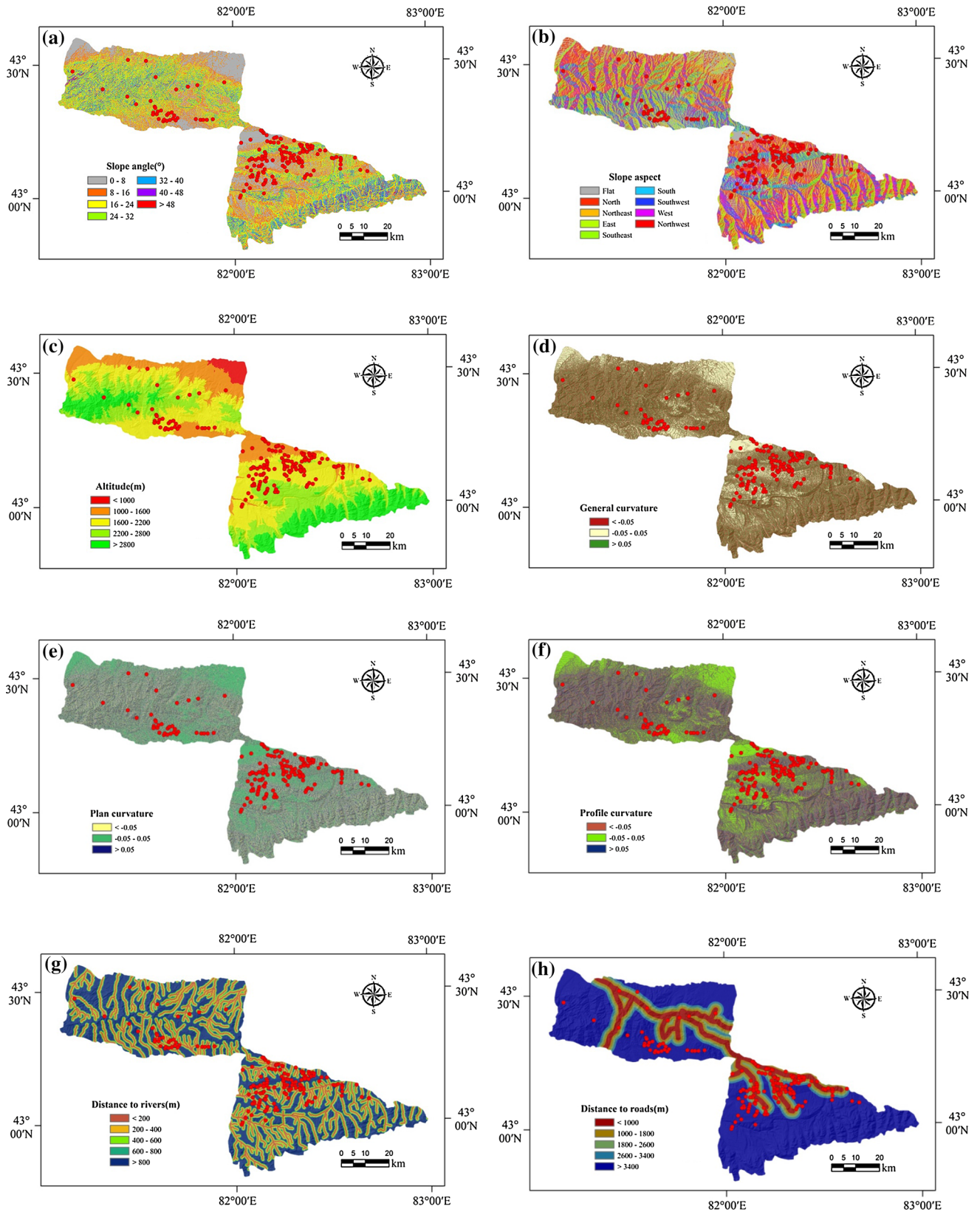


Fig. 2 **a** The slope angle map of the study area. **b** The slope aspect map of the study area. **c** The altitude map of the study area. **d** The general curvature map of the study area. **e** The plan curvature map of the study area. **f** The profile curvature map of the study area. **g** The distance to rivers map of the study area. **h** The distance to roads map of the study area. **i** The NDVI map of the study area. **j** The STI map of the study area. **k** The rainfall map of the study area. **l** The lithology map of the study area

2014). The altitude is one of the frequently used significant landslides conditioning parameter (Youssef et al. 2015; Ercanoglu et al. 2004; Rozos et al. 2011). In this study area, the elevation of ranges from 767 to 4217 m and was divided into five categories using an interval of 600 m (Fig. 2c). The slope curvature is an important variable that controls the superficial and subsurface hydrological regime of the slope. The characterization of slope morphology and flow can be analyzed with the help of the general curvature map (Nefeslioglu et al. 2008). The plan curvature is described as the curvature of a contour line formed by intersection of a horizontal plane with the surface. The profile curvature is the vertical plane parallel to the slope direction (Kannan et al. 2013). In the present study, the general curvature, plan curvature, and profile curvature were calculated in ArcGIS 10.0 based on DEM data and were divided into three classes, respectively (Fig. 2d–f).

Two proximity parameters including distance to rivers and distance to roads were taken into account in the study.

The distance to rivers is also considered as one of the most important factor for the landslide susceptibility analysis. Drainages adversely affect the stability of slope by saturating it and by eroding the toe of slopes (Bijukchhen et al. 2013; Dragičević et al. 2015). Similarly, the distance to roads is one of the causal factors for landslides since the load in the toe of slope can be reduced by road-cuts (Yalcin et al. 2011). In this study, both the proximity parameters were divided into five different buffer zones, respectively (Fig. 2g, h).

The normalized difference vegetation index (NDVI) is often considered as a controlling factor in landslide susceptibility mapping. In this study, the NDVI value was calculated using the formula $NDVI = (IR - R)/(IR + R)$, where IR is the infrared portion of the electromagnetic spectrum, and R is the red portion of the electromagnetic spectrum (Youssef et al. 2015). The NDVI map of this area was grouped into four classes (Fig. 2i). The sediment transport index (STI) characterizes the process of erosion and deposition (Devkota et al. 2013). STI is defined as in Eq. (1):

$$STI = \left(\frac{A_s}{22.13} \right)^{0.6} \left(\frac{\sin \beta}{0.0896} \right)^{1.3} \quad (1)$$

where, A_s is the specific catchment's area (m^2/m), and β the slope gradient. In the present study, STI was calculated as shown in Fig. 2j.

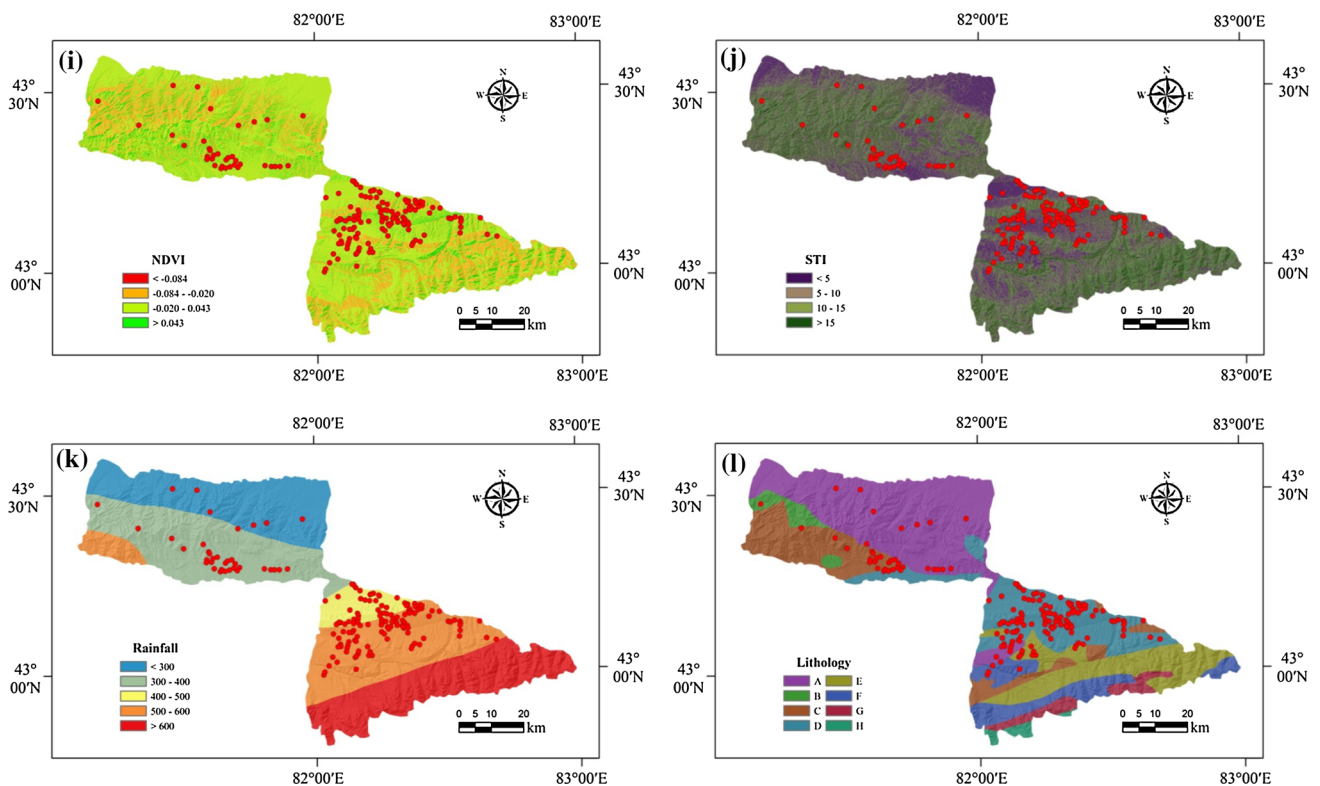


Fig. 2 continued

Rainfall, associated with landslide initiation by means of its influence on runoff and pore water pressure, is also a very important controlling factor (Shahabi et al. 2014; Van Westen et al. 2006; Yang et al. 2015). In the present study, the annual rainfall was reclassified into seven classes: <300, 300–400, 400–500, 500–600, >600 mm/year (Fig. 2k). Lithology plays an important role in landslide susceptibility studies since landslides are controlled by the rock properties of the land surface and different lithological units have different susceptibility values (Yesilnacar and Topal 2005; Yalcin et al. 2011). In this study, the lithological map was extracted from the geology database of the area. The study area is covered with various types of lithological units (Fig. 2l). Their names, lithologic characteristics, and ages of the geological units are provided in Table 1.

Methods

Statistical index model

The statistical index approach is considered as the simplest and quantitatively suitable method in landslide susceptibility mapping. It was introduced by van Westen et al. (1997) for landslide susceptibility analyses and has been adopted by various researchers (Yesilnacar 2005; Long 2008; Regmi et al. 2014). In this method, the weighting value for each categorical unit is defined as the natural logarithm of the landslide density in the categorical unit divided by the landslide density in the whole studied area (Kavzoglu et al. 2015; Regmi et al. 2014). This method is based upon the formula given by Van Westen (1997) as follows:

$$W_{ij} = \ln\left(\frac{E_{ij}}{E}\right) = \ln\left(\frac{N_{ij}/S_{ij}}{N/S}\right) \quad (2)$$

where, W_{ij} is the weight given to a certain class i of parameter j ; E_{ij} is the landslide density within class i of

parameter j ; E is the total landslide density within the entire map; N_{ij} is the number of landslides in a certain class i of parameter j ; S_{ij} is the number of pixels in a certain class i of parameter j ; N is the total number of landslides in the entire map; S is the total pixels of the entire map. In this study, the resultant weights for each thematic map for the statistical index model were calculated in ArcGIS 10.0 and Microsoft Excel, and the results are shown in Table 1. The higher resultant weight, the higher is the possibility that a mass movement occurs within the area covered by the considered class.

Index of entropy model

The second model used for landslide susceptibility analysis in the current study is index of entropy model (Devkota et al. 2013). The entropy indicates the extent of the instability, disorder, imbalance, and uncertainty of a system (Youssef et al. 2015). The entropy of a landslide refers to the extent that various factors influence the development of a landslide (Youssef et al. 2015; Jaafari et al. 2014). Several important factors provide additional entropy into the index system. As a result, the entropy value can be used to calculate objective weights of the index system (Yang et al. 2010; Pourghasemi et al. 2012b; Youssef et al. 2015; Bednarik et al. 2010). The equations used to calculate the information coefficient W_j representing the weight value for the parameter as a whole (Bednarik et al. 2010; Devkota et al. 2013) are given as follows:

$$P_{ij} = \frac{b}{a} \quad (3)$$

$$(P_{ij}) = \frac{P_{ij}}{\sum_{j=1}^{S_j} P_{ij}} \quad (4)$$

where a and b are the domain and landslide percentages, respectively; S_j is the number of classes; (P_{ij}) is the probability density. Here, H_j and $H_{j\max}$ represent entropy values (Eqs. 5, 6).

Table 1 Description of geological units of the study area (Qin 2007)

No.	Formation	Code	Lithology
A	–	Q _h , Q _p	Cohesive soil, sand gravel, pebbles, loess and gravel
B	Wulang, Shuixi Xiaouangou	P ₁ , J ₁₋₂ , T ₂₋₃	Argillaceous siltstone, sandstone, quartz sandstone, quartzite, mudstone, carbonaceous mudstone
C	Tuokuzidaban, Ishrick	C ₁ , T, C ₂	Conglomerate, limestone, siltstone, volcanic clastic limestone, sandy shale, clay shale
D	Akshak, Dahalajunshan	C ₁ , Q _p	Carbonate, clastic rocks, glutenite, limestone, rhyolitic porphyry, basaltic porphyrite
E	Shawan	D	Granite, granodiorite, plagioclase granite, granite porphyry, diorite
F	The Permian granite, Tekes	P	Monzonitic granite, intermediate-acid igneous rocks, neutral volcanic clastic rock
G	Xingditag, Tekes	C _h , P _t	Quartzite, quartz-schist, siltstone, phyllite, limestone and marble rock
H	Keketiekedaban	S	Limestone, pyroclastic rocks, monzonitic granite

$$H_j = - \sum_{i=1}^{S_j} (P_{ij}) \log_2(P_{ij}), \quad j = 1, 2, \dots, n \quad (5)$$

$$H_{jmax} = \log_2 S_j \quad (6)$$

I_j is the information coefficient (Eq. 7) and W_j represents the resultant weight value for the parameter as a whole (Eq. 8).

$$I_j = \frac{H_{jmax} - H_j}{H_{jmax}} I = (0, 1), \quad j = 1, 2, \dots, n \quad (7)$$

$$W_j = I_j P_{ij} \quad (8)$$

The complete calculation of weight determination for individual parameters is presented in Table 2. The final landslide susceptibility map is prepared based on the Eq. (9) using the ArcGIS 10.0 software.

$$Y_{IOE} = \sum_i^n \frac{z}{m_i} \times C \times W_j \quad (9)$$

where, Y_{IOE} is the sum of all the classes; i is the number of particular parametric map (1, 2, ..., n); z is the number of classes within parametric map with the greatest number of classes; m_i is the number of classes within particular parametric map; C is the value of the class after secondary classification and W_j is the weight of a parameter. The result of this summation represents the various levels of the landslide susceptibility (Bednarik et al. 2010; Devkota et al. 2013; Jaafari et al. 2014).

Results and discussion

Application of statistical index (SI) model in landslide susceptibility mapping

The correlation between the location of landslides and the landslide conditioning factors performed by statistical index model is shown in Table 2. From Table 2, it is seen that the slope angle in ranges of 0°–8° and 8°–16° has the high values of SI with the positive values (0.39 and 0.21), and other classes have negative value. The reason for this may be that the resistant lithologic units exist in the steep slopes and they are not covered by highly and completely weathered lithologic units which are more susceptible to landslide occurrence (Akgun et al. 2008; Yalcin et al. 2011). In the case of slope aspect, the SI value is highest for south facing with the positive value (0.39) The north-facing slopes are less prone to landslides as it has lower SI value (−0.34). The SI values of altitude show that they are positive for the ranges of 1000–1600, 1600–2200 m, with the highest value (0.57) for the altitude ranging between 1600 and 2200 m. However, it is clear that the landslide

susceptibility increases by the increase in altitude up to a certain extent (1600–2200 m) and then it decreases. In the case of general curvature, the SI values for each class were similar, which indicates that these classes have no obvious effect on the occurrence of landslides. The relation between plan curvature and landslide probabilities showed that −0.05 to 0.05 class has the highest value of SI (0.26), and for profile curvature, the class of <−0.05 shows a high SI value (0.11). In the case of distance to rivers, the distance of 200–400 m of rivers has highest correlation with landslide occurrence. For distance to roads, the SI values show that when the distance increases, the probability of landslides decreases. The highest probability for landslide occurrence is within distances less than 1000 m. The SI value for NDVI clearly showed that ranges of >0.043 have the most effect on landslide occurrence. The STI factor shows that the range of <5 is relatively conducive (high susceptible) to landslide occurrence. With regard to rainfall, the classes of 400–500, 500–600 mm/year seem to have the highest impact on landslide occurrence. Their SI values are 1.49 and 0.74, respectively. For the lithology, the class of D has highest SI value with the positive value (1.19). This indicates that lithological unit of the carbonate, clastic rocks, glutenite, limestone, rhyolitic porphyry, and basaltic porphyrite has the highest influence in triggering landslides.

In this study, the landslide susceptibility map (Fig. 3) was constructed using the landslide susceptibility index (LSI) value by the software of ArcGIS 10.0. The calculated LSI values for SI model of the study area range from about −7.98 to 6.27. Obviously, larger LSI values indicate a higher susceptibility for landsliding. The index values were reclassified as very low, low, moderate, high, and very high susceptibility classes using natural break approach in ArcGIS software. The distribution of observed landslides falling into various susceptibility classes of different landslide susceptibility zonation maps was counted in Fig. 5a. It can be observed from Fig. 5a that 16.95 % of the study area was placed in the group with very low susceptibility. Low, moderate and high susceptibility landslide classes comprised of 31.05, 23.61, and 16.38 % of the area, respectively. In all, 12.01 % of the region was placed in the class with very high landslide susceptibility. Meanwhile, the results show that the percentages of the total landslides in very low, low, moderate, high, and very high susceptibility classes are 1.29, 3.43, 22.32, 24.89, and 48.07 %, respectively.

Application of index of entropy (IOE) model in landslide susceptibility mapping

The density (P_{ij}) of the landslide occurrence in each class was calculated. The resultant weights for each thematic

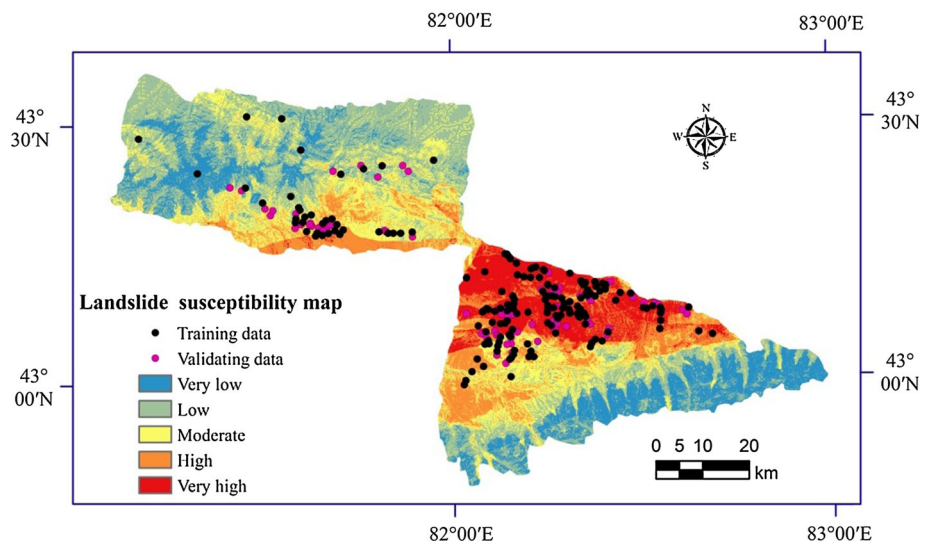
Table 2 Spatial relationship between each landslide conditioning factor and landslide by SI and IOE models

Factors	Classes	Percentage of landslide (%)	Percentage of domain (%)	SI	(P_{ij})	H_j	H_{jmax}	I_j	W_j
Slope angle (°)	0 to 8	36.81	24.90	0.39	0.30	2.40	2.81	0.15	0.11
	8 to 16	26.99	21.93	0.21	0.25				
	16 to 24	20.25	21.34	-0.05	0.19				
	24 to 32	10.43	16.82	-0.48	0.12				
	32 to 40	4.29	10.23	-0.87	0.08				
	40 to 48	1.23	3.85	-1.14	0.06				
	>48	0.00	0.94	0.00	0.00				
Slope aspect	Flat	10.43	9.18	0.13	0.12	3.14	3.17	0.01	0.01
	North	9.20	12.86	-0.34	0.08				
	Northeast	13.50	14.16	-0.05	0.10				
	East	11.66	12.25	-0.05	0.10				
	Southeast	9.20	9.21	0.00	0.11				
	South	11.04	7.49	0.39	0.16				
	Southwest	11.66	9.91	0.16	0.13				
	West	12.88	11.96	0.07	0.12				
	Northwest	10.43	12.98	-0.22	0.09				
Altitude (m)	<1000	0.00	2.12	0.00	0.00	1.55	2.32	0.34	0.23
	1000 to 1600	18.41	16.77	0.09	0.32				
	1600 to 2200	66.87	37.71	0.57	0.51				
	2200 to 2800	13.50	25.84	-0.65	0.15				
	>2800	1.23	17.56	-2.66	0.02				
General curvature	<-0.05	40.49	40.16	0.01	0.34	1.59	1.59	0.00	0.00
	-0.05 to 0.05	19.63	20.25	-0.03	0.33				
	>0.05	39.88	39.59	0.01	0.34				
Plan curvature	<-0.05	32.52	35.30	-0.08	0.30	1.56	1.59	0.02	0.02
	-0.05 to 0.05	37.42	28.89	0.26	0.42				
	>0.05	30.06	35.82	-0.18	0.28				
Profile curvature	<-0.05	39.26	35.20	0.11	0.37	1.58	1.59	0.01	0.01
	-0.05 to 0.05	30.06	28.32	0.06	0.35				
	>0.05	30.68	36.48	-0.17	0.28				
Distance to rivers (m)	0 to 200	18.41	17.96	0.03	0.19	2.28	2.32	0.02	0.02
	200 to 400	24.54	16.21	0.42	0.28				
	400 to 600	15.34	14.35	0.07	0.20				
	600 to 800	15.34	12.97	0.17	0.22				
	>800	26.38	38.52	-0.38	0.13				
Distance to roads (m)	0 to 1000	31.29	11.52	1.00	0.36	2.14	2.32	0.08	0.12
	1000 to 1800	9.20	7.77	0.17	0.16				
	1800 to 2600	12.88	6.46	0.69	0.27				
	2600 to 3400	5.52	5.62	-0.02	0.13				
	>3400	41.10	68.63	-0.51	0.08				
NDVI	<-0.084	0.00	0.00	0.00	0.00	1.56	2.00	0.22	0.16
	-0.084 to -0.020	19.02	25.40	-0.29	0.25				
	-0.020 to 0.043	69.33	64.61	0.07	0.36				
	>0.043	11.66	9.99	0.15	0.39				
STI	<5	42.95	32.01	0.29	0.33	1.96	2.00	0.02	0.02
	5 to 10	20.25	19.74	0.03	0.25				
	10 to 15	13.50	12.45	0.08	0.26				
	>15	23.31	35.80	-0.43	0.16				

Table 2 continued

Factors	Classes	Percentage of landslide (%)	Percentage of domain (%)	SI	(P_{ij})	H_j	H_{jmax}	I_j	W_j
Rainfall (mm/year)	<300	3.68	22.57	-1.81	0.02	1.41	2.32	0.40	0.59
	300 to 400	18.41	25.62	-0.33	0.10				
	400 to 500	24.54	5.51	1.49	0.60				
	500 to 600	53.37	25.36	0.74	0.28				
	>600	0.00	20.94	0.00	0.00				
Lithology	A	12.88	34.41	-0.98	0.07	1.84	3.00	0.39	0.26
	B	0.61	2.83	-1.53	0.04				
	C	9.82	15.21	-0.44	0.12				
	D	66.26	20.23	1.19	0.61				
	E	7.36	14.35	-0.67	0.10				
	F	3.07	8.57	-1.03	0.07				
	G	0.00	3.32	0.00	0.00				
	H	0.00	1.09	0.00	0.00				

Fig. 3 Landslide susceptibility map derived from the SI model



map for the IOE model are given in Table 2. The calculated weight values for different classes of the causative factors show the importance of the respective classes in the slope instability. From the W_j value, it is seen that the rainfall has the highest impact in the landslide susceptibility, followed by lithology, altitude, NDVI, distance to roads, and slope angle, while others are less significant in the landslide susceptibility of the region. From the result (P_{ij}) , it is seen that the slope angle of classes $0^\circ-8^\circ$ and $8^\circ-16^\circ$ has the high values of (P_{ij}) . In the case of aspect, south-facing slope is susceptible to landsliding. In the case of altitude, landslide density is highest at the elevation ranging from 1600 to 2200 m followed by from 1000 to 1600 m. In the case of general curvature, the (P_{ij}) values for each class were similar. This indicates that these classes have no obvious effect on the occurrence of landslides. In

terms of plan curvature and profile curvature, most of the existing landslides are distributed across the -0.05 to 0.05 class and <-0.05 class, respectively. In the case of the relationship between landslide occurrence and distance to drainage, the (P_{ij}) value was highest at the distance between 200 and 400 m indicating a high probability of landslide susceptibility. The distance to roads shows that the (P_{ij}) value decreases as the distance to roads increases. The (P_{ij}) value for NDVI showed that the class of >0.043 has the most effect on the occurrence of landslides with high value of 0.39. Relation between STI and landslide probability showed that <5 class has highest value of (P_{ij}) . In rainfall, the highest (P_{ij}) value (0.60) was located in the rainfall classes of 400–500 mm/year. In the case of lithology, it can be seen that most of the existing landslides are distributed the lithological unit of the carbonate, clastic

Fig. 4 Landslide susceptibility map derived from the IOE model

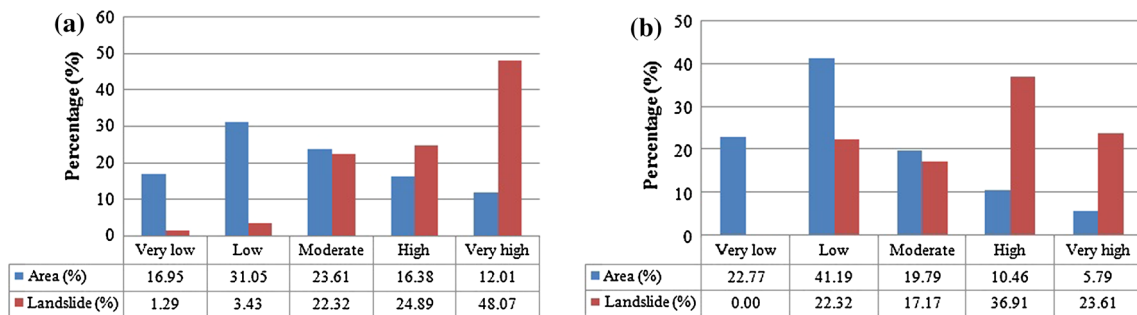
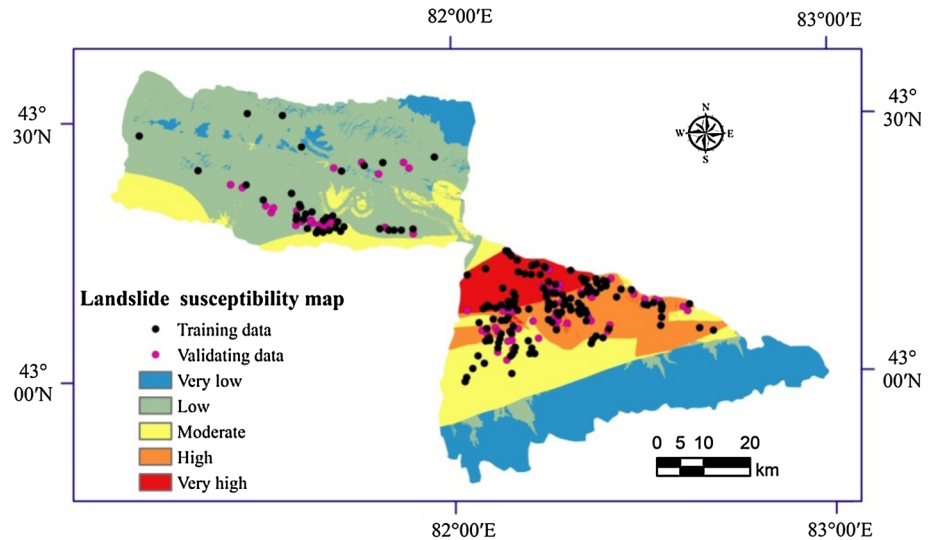


Fig. 5 Histograms representing the distribution of observed landslides falling into various susceptibility classes of different landslide susceptibility zonation maps: **a** statistical index model; **b** index of entropy model

rocks, glutenite, limestone, rhyolitic porphyry, and basaltic porphyrite. The findings based on entropy approach show that rainfall, lithology, and altitude, are the most important factors which explain better the landslide occurrence and distribution in study area. However, it should be noted that the landslide contributing factors may vary from region to region, such that the rating scheme followed in this study area may be not suitable anywhere else (Bijukchhen et al. 2013).

In this study, the final result of index of entropy (IOE) model is a LSI map, in which the LSI values vary from 0.22 to 4.64. Based on the natural breaks method in ArcGIS, the landslide susceptibility map was reclassified into five classes: very low, low, moderate, high, and very high (Fig. 4). Among the five susceptibility zones, 22.77 % of the study area was designated as a very low susceptible zone, and low, moderate, high and very high susceptible zones represent 41.19, 19.79, 10.46 and 5.79 %, respectively. In addition, 23.61 and 36.91 % of the total landslides falls in the very high and high susceptibility zones, respectively. Moderate, low, and very low susceptible

zones represent 17.17, 22.32, and 0.00 % of the landslides, respectively (Fig. 5b).

Validation of the landslide susceptibility maps

In order to determine the reliability of the evaluation results, it is important to perform validation of spatial results in a structured manner (Ahmed et al. 2013). To do this validation, the area under curve (AUC) method was used in this research. This method works by creating specific rate curve which explains percentage of known landslides that fall into each defined level of susceptibility rank and displays as the cumulative frequency diagram (Intarawichian and Dasananda 2011). The specific rate curves can be divided into success-rate curve and prediction-rate curve (Chung and Fabbri 2003). The success rate curve, considered as a degree of fit measure, is based on a comparison of the susceptibility map with the landslides used in modeling (Pradhan and Kim 2014; Chung and Fabbri 2003). In addition, the prediction rate curve is a good indicator of the predictive power of the susceptibility

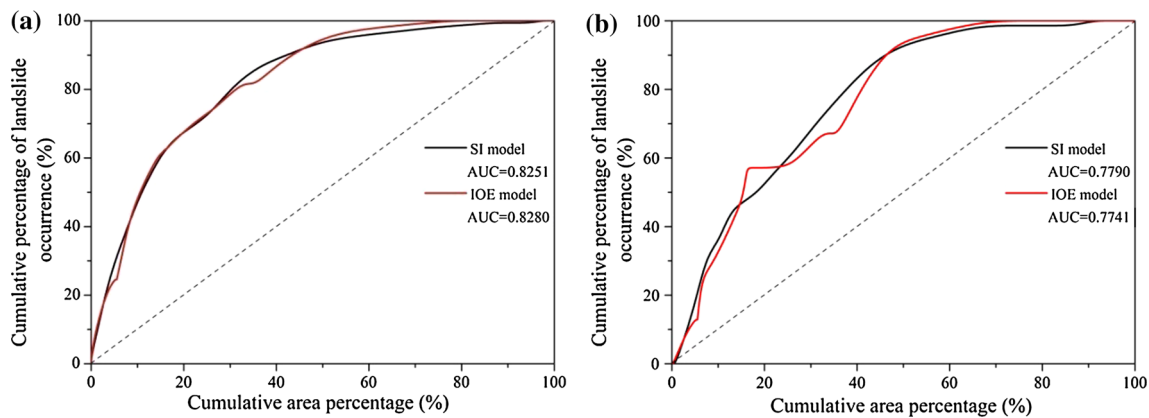


Fig. 6 AUC representing quality model **a** success rate curve and **b** prediction rate curve

map. The prediction rate curve can be created by the validation landslide inventory (Pradhan and Kim 2014; Chung and Fabbri 2003). The rate curves show the cumulative percentage of the area of the susceptibility classes on the x axis and the cumulative percentage of observed landslide occurrences in different susceptibility classes on the y axis. The area under the curve (AUC) can be used to determine prediction accuracy of the susceptibility map qualitatively in which larger area means higher accuracy achieved (Intarawichian and Dasananda 2011; Mathew et al. 2009; Lee 2005b). In this study, the rate curves were obtained by comparing the landslide training data (163 landslides) and validation data (70 landslides) with the susceptibility maps and the areas under the curve were calculated.

The success rate and prediction rate curves were shown in Fig. 6a, b, respectively. From calculation of the AUC for SI and IOE models, it was found that AUC for the success rate curves are 0.8251 and 0.8280, respectively. Namely, the training accuracy of the susceptibility maps is 82.51 and 82.80 %, respectively. The areas of prediction rate curves are 0.7790 and 0.7741, which means that the overall prediction rates are 77.90 and 77.41 %. The results showed that the SI and IOE models exhibited similar performance. Meanwhile, both the models are successful estimators, and the two models employed in this study have reasonably good accuracy in predicting the landslide susceptibility of the study area.

Conclusions

In this study, the statistical index and index of entropy models were employed and compared for landslide susceptibility assessment at the Gongliu County of Xinjiang Uygur Autonomous Region, China. Both the methods were employed using slope angle, slope aspect, altitude, general curvature, plan curvature, profile curvature, distance to

rivers, distance to roads, NDVI, STI, rainfall, and lithology. The selection of the 12 conditioning landslide factors, based on consideration of relevance, availability, and scale of data that was available for the study area, is relative and subjective, and can be improved in future research. In this study, a total of 233 landslides were identified and mapped. Out of which, 163 (70 %) were randomly selected for generating a model and the remaining 70 (30 %) were used for validation proposes. The susceptibility maps produced by SI and IOE models were divided into five different susceptibility classes such as very low, low, moderate, high, and very high. The area under the rate curve show that the SI model has a success rate of 82.51 % and predictive accuracy of 77.90 %, and IOE model has success rate of 82.80 % and predictive accuracy of 77.41 %. It is clear that both models showed almost similar results. The output results of the present study showed that the choice of suitable predisposing factors together with the statistical index and index of entropy methods and the application of geographical information systems are able to successfully identify this area that are susceptible to landslides. However, it should be noted that both the models were developed on some basic assumptions such as topography, geology, and stream etc. If the data of factors causing the landslides, such as extreme rainfall, earthquake shaking can be considered, then a more accurate analysis could be done.

The findings and results of the present study can help the decision makers, managers, urban planners, engineers, and land-use developers can for preliminary slope management and land-use planning. Also, this methodology can be used to assess landslide susceptibility in other areas of Xinjiang, as well as in other similar regions where the same geological and topographical feature prevails.

Acknowledgments The authors would like to express their gratitude to everyone who provided assistance for the present study. The study is jointly supported by the National Program on Key Basic

Research Project (Grant No. 2015CB251601), the State Key Program of National Natural Science of China (Grant No. 41430643) and the Natural Science Foundation of China (Grant No. 41302248). The authors would also like to acknowledge two anonymous reviewers and editor for their helpful comments on the previous version of the manuscript.

References

- Ahmed B, Ahmed R, Zhu X (2013) Evaluation of model validation techniques in land cover dynamics. *ISPRS Int J Geo Inf* 2(3):577–597
- Akgun A, Dag S, Bulut F (2008) Landslide susceptibility mapping for a landslide-prone area (Findikli, NE of Turkey) by likelihood-frequency ratio and weighted linear combination models. *Environ Geol* 54(6):1127–1143
- 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):15–31
- Ayalew L, Yamagishi H, Ugawa N (2004) Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. *Landslides* 1(1):73–81
- Bednarik M, Magulová B, Matys M, Marschalko M (2010) Landslide susceptibility assessment of the Kral'ovany-Liptovský Mikuláš railway case study. *Phys Chem Earth Parts A/B/C* 35(3):162–171
- Bijukchhen SM, Kayastha P, Dhital MR (2013) A comparative evaluation of heuristic and bivariate statistical modelling for landslide susceptibility mappings in Ghurmi-Dhad Khola, east Nepal. *Arab J Geosci* 6(8):2727–2743
- Chung CJF, Fabbri AG (2003) Validation of spatial prediction models for landslide hazard mapping. *Nat Hazards* 30(3):451–472
- Dai FC, Lee CF, Li J, Xu ZW (2001) Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environ Geol* 40(3):381–391
- Demir G, Aytekin M, Akgün A, İkizler SB, Tatar O (2013) A comparison of landslide susceptibility mapping of the eastern part of the North Anatolian Fault Zone (Turkey) by likelihood-frequency ratio and analytic hierarchy process methods. *Nat Hazards* 65(3):1481–1506
- Devkota KC, Regmi AD, Pourghasemi HR, Yoshida K, Pradhan B, Ryu IC, Althuwaynee OF (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(1):135–165
- Dragičević S, Lai T, Balram S (2015) GIS-based multicriteria evaluation with multiscale analysis to characterize urban landslide susceptibility map of a landslide prone area (West Black Sea Region, Turkey). *Eng Geol* 75(3):229–250
- Ercanoglu M, Gokceoglu C, Van Asch TW (2004) Landslide susceptibility zoning north of Yenice (NW Turkey) by multivariate statistical techniques. *Nat Hazards* 32(1):1–23
- Erener A, Mutlu A, Düzgün HS (2015) A comparative study for landslide susceptibility mapping using GIS-based multi-criteria decision analysis (MCDA), logistic regression (LR) and association rule mining (ARM). *Eng Geol*. doi:10.1016/j.enggeo.2015.09.007
- Ermini L, Catani F, Casagli N (2005) Artificial neural networks applied to landslide susceptibility assessment. *Geomorphology* 66(1):327–343
- He S, Pan P, Dai L, Wang H, Liu J (2012) Application of kernel-based Fisher discriminant analysis to map landslide susceptibility in the Qinggan River delta, Three Gorges, China. *Geomorphology* 171:30–41
- Intarawichian N, Dasananda S (2011) Frequency ratio model based landslide susceptibility mapping in lower Mae Chaem watershed, Northern Thailand. *Environ Earth Sci* 64(8):2271–2285
- Jaafari A, Najafi A, Pourghasemi HR, Rezaeian J, Sattarian A (2014) GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran. *Int J Environ Sci Technol* 11(4):909–926
- Kannan M, Saranathan E, Anbalagan R (2013) Landslide vulnerability mapping using frequency ratio model: a geospatial approach in Bodi-Bodimettu Ghat section, Theni district, Tamil Nadu, India. *Arab J Geosci* 6(8):2901–2913
- Kannan M, Saranathan E, Anbalagan R (2015) Comparative analysis in GIS-based landslide hazard zonation—a case study in Bodi-Bodimettu Ghat section, Theni District, Tamil Nadu, India. *Arab J Geosci* 8(2):691–699
- Kanungo DP, Arora MK, Sarkar S, Gupta RP (2006) A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. *Eng Geol* 85(3):347–366
- Kavzoglu T, Sahin EK, Colkesen I (2014) Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* 11(3):425–439
- Kavzoglu T, Sahin EK, Colkesen I (2015) An assessment of multivariate and bivariate approaches in landslide susceptibility mapping: a case study of Duzkoy district. *Nat Hazards* 76(1):471–496
- Kayastha P, Dhital MR, De Smedt F (2013) Evaluation of the consistency of landslide susceptibility mapping: a case study from the Kankai watershed in east Nepal. *Landslides* 10(6):785–799
- Komac M (2006) A landslide susceptibility model using the Analytical Hierarchy Process method and multivariate statistics in perialpine Slovenia. *Geomorphology* 74(1):17–28
- Lee S (2005a) Application and cross-validation of spatial logistic multiple regression for landslide susceptibility analysis. *Geosci J* 9(1):63–71
- Lee S (2005b) Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. *Int J Remote Sens* 26(7):1477–1491
- Lee S, Dan NT (2005) Probabilistic landslide susceptibility mapping in the Lai Chau province of Vietnam: focus on the relationship between tectonic fractures and landslides. *Environ Geol* 48(6):778–787
- Lee S, Pradhan B (2007) Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* 4(1):33–41
- Lee S, Sambath T (2006) Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. *Environ Geol* 50(6):847–855
- Lee S, Talib JA (2005) Probabilistic landslide susceptibility and factor effect analysis. *Environ Geol* 47(7):982–990
- Lee S, Ryu JH, Min K, Won JS (2003) Landslide susceptibility analysis using GIS and artificial neural network. *Earth Surf Proc Land* 28(12):1361–1376
- Lee S, Ryu JH, Won JS, Park HJ (2004) Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. *Eng Geol* 71(3):289–302

- Liu M, Chen X, Yang S (2014) Collapse landslide and mudslides hazard zonation. *Landslide science for a safer geoenvironment*. Springer, Berlin, pp 457–462
- Long NT (2008) Landslide susceptibility mapping of the mountainous area in A Luoi district, Thua Thien Hue province, Vietnam. Faculty of Engineering, Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Belgium
- Mathew J, Jha VK, Rawat GS (2009) Landslide susceptibility zonation mapping and its validation in part of Garhwal Lesser Himalaya, India, using binary logistic regression analysis and receiver operating characteristic curve method. *Landslides* 6(1):17–26
- Meng Q, Miao F, Zhen J, Wang X, Wang A, Peng Y, Fan Q (2015) GIS-based landslide susceptibility mapping with logistic regression, analytical hierarchy process, and combined fuzzy and support vector machine methods: a case study from Wolong Giant Panda Natural Reserve, China. *Bull Eng Geol Environ*. doi:10.1007/s10064-015-0786-x
- Nefeslioglu HA, Duman TY, Durmaz S (2008) Landslide susceptibility mapping for a part of tectonic Kelkit Valley (Eastern Black Sea region of Turkey). *Geomorphology* 94(3):401–418
- Ozdemir A, Altural T (2013) A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. *J Asian Earth Sci* 64:180–197
- Pourghasemi HR, Mohammady M, Pradhan B (2012a) Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. *Catena* 97:71–84
- Pourghasemi HR, Pradhan B, Gokceoglu C (2012b) Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Nat Hazards* 63(2):965–996
- Pradhan B (2010) Landslide susceptibility mapping of a catchment area using frequency ratio, fuzzy logic and multivariate logistic regression approaches. *J Indian Soc Remote Sens* 38(2):301–320
- Pradhan AMS, Kim YT (2014) Relative effect method of landslide susceptibility zonation in weathered granite soil: a case study in Deokjeok-ri Creek, South Korea. *Nat Hazards* 72(2):1189–1217
- Pradhan B, Lee S (2010) Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environ Model Softw* 25(6):747–759
- Pradhan B, Sezer EA, Gokceoglu C, Buchroithner MF (2010) Landslide susceptibility mapping by neuro-fuzzy approach in a landslide-prone area (Cameron Highlands, Malaysia). *Geosci Remote Sens IEEE Trans* 48(12):4164–4177
- Qin XM (2007) Based on GIS landslide geological disaster hazard evaluation research-taking Gongliu county as example. Master Thesis, Xinjiang University, Ürümqi, p 45
- Regmi AD, Devkota KC, Yoshida K, Pradhan B, Pourghasemi HR, Kumamoto T, Akgun A (2014) Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arab J Geosci* 7(2):725–742
- Rozos D, Bathrellos GD, Skillodimou HD (2011) Comparison of the implementation of rock engineering system and analytic hierarchy process methods, upon landslide susceptibility mapping, using GIS: a case study from the Eastern Achaia County of Peloponnesus, Greece. *Environ Earth Sci* 63(1):49–63
- Shahabi H, Khezri S, Ahmad BB, Hashim M (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
- Shahabi H, Hashim M, Ahmad BB (2015) Remote sensing and GIS-based landslide susceptibility mapping using frequency ratio, logistic regression, and fuzzy logic methods at the central Zab basin, Iran. *Environ Earth Sci* 73(12):8647–8668
- Solaimani K, Mousavi SZ, Kaviani A (2013) Landslide susceptibility mapping based on frequency ratio and logistic regression models. *Arab J Geosci* 6(7):2557–2569
- Süzen ML, Doyuran V (2004) Data driven bivariate landslide susceptibility assessment using geographical information systems: a method and application to Asarsuyu catchment, Turkey. *Eng Geol* 71(3):303–321
- Van Den Eeckhaut M, Vanwalleghem T, Poesen J, Govers G, Verstraeten G, Vandekerckhove L (2006) Prediction of landslide susceptibility using rare events logistic regression: a case-study in the Flemish Ardennes (Belgium). *Geomorphology* 76(3):392–410
- van Westen C (1997) Statistical landslide hazard analysis ILWIS 2.1 for Windows application guide. ITC Publication, Enschede, pp 73–84
- Van Westen CJ, Van Asch TW, Soeters R (2006) Landslide hazard and risk zonation—why is it still so difficult? *Bull Eng Geol Environ* 65(2):167–184
- Wang Q, Li W, Chen W, Bai H (2015) GIS-based assessment of landslide susceptibility using certainty factor and index of entropy models for the Qianyang County of Baoji City, China. *J Earth Syst Sci* 124(7):1399–1415
- Yalcin A (2008) GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. *Catena* 72(1):1–12
- Yalcin A, Reis S, Aydinoglu AC, Yomralioglu T (2011) A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. *Catena* 85(3):274–287
- Yang Z, Qiao J, Zhang X (2010) Regional landslide zonation based on entropy method in Three Gorges area, China. *IEEE International conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, vol 3. pp 1336–1339
- Yang ZH, Lan HX, Gao X, Li LP, Meng YS, Wu YM (2015) Urgent landslide susceptibility assessment in the 2013 Lushan earthquake-impacted area, Sichuan Province, China. *Nat Hazards* 75(3):2467–2487
- Yesilnacar EK (2005) The application of computational intelligence to landslide susceptibility mapping in Turkey (Doctoral Dissertation, University of Melbourne, Department, 200.)
- Yesilnacar E, Topal T (2005) Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Eng Geol* 79(3):251–266
- Yilmaz I (2009) Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey). *Comput Geosci* 35(6):1125–1138
- Youssef AM, Al-Kathery M, Pradhan B (2015) Landslide susceptibility mapping at Al-Hasher Area, Jizan (Saudi Arabia) using GIS-based frequency ratio and index of entropy models. *Geosci J* 19(1):113–134