Geocarto International

ISSN: 1010-6049 (Print) 1752-0762 (Online) Journal homepage: http://www.tandfonline.com/loi/tgei20

Optimizing landslide susceptibility mapping in the Kongtong District, NW China: comparing the subdivision criteria of factors

Ge Yan, Shouyun Liang, Xiangang Gui, Yu Xie & Hongliang Zhao

To cite this article: Ge Yan, Shouyun Liang, Xiangang Gui, Yu Xie & Hongliang Zhao (2018): Optimizing landslide susceptibility mapping in the Kongtong District, NW China: comparing the subdivision criteria of factors, Geocarto International, DOI: 10.1080/10106049.2018.1499816

To link to this article: https://doi.org/10.1080/10106049.2018.1499816

Accepted author version posted online: 13 Jul 2018.
Published online: 18 Oct 2018.

GEOCARTO INTERNATIONAL
https://doi.org/10.1080/10106049.2018.1499816

Taylor & Francis
Taylor & Francis Group

Optimizing landslide susceptibility mapping in the Kongtong District, NW China: comparing the subdivision criteria of factors

Ge Yan[a,b], Shouyun Liang[b], Xiangang Gui[b], Yu Xie[c] and Hongliang Zhao[b]

[a]State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai, China; [b]Key Laboratory of Mechanics on Disaster and Environment in Western China, Ministry of Education, Lanzhou University, Lanzhou, China; [c]Department of Land and Resources of Gansu Province, Lanzhou, China

ABSTRACT

The standards applied to reclassify landslide-conditioning factors differ among studies and may change the accuracy of identifying landslide-prone areas. Therefore, we identified two standards per factor (elevation, aspect, slope, proximity to roads and proximity to streams) from the existing literature and set them as predisposing criteria in this paper. In addition to the five factors, lithology represented by types and a landslide inventory map produced from field surveys were also used in mapping. Thirty-two landslide susceptibility maps were generated based on weights-of-evidence and evaluated using the relative operative characteristic method. The results show that the subdivision criteria of factors change the accuracy, with the success rate varying from 84.34% to 87.51%. The map with the highest value captures more landslides in relatively higher susceptibility classes and is therefore considered the optimal one. Ultimately, a simplified mode of combining subdivision criteria is proposed to simplify comparison.

ARTICLE HISTORY
Received 8 December 2017
Accepted 3 July 2018

KEYWORDS
Landslide; susceptibility mapping; factor subdivision; optimization

1. Introduction

Landslides are among the most damaging geologic hazards (He et al. 2012; Hadji et al. 2013) and are responsible for extensive loss of life and property. One way to mitigate the damage caused by landslides is hazard zoning (Pourghasemi et al. 2012).

Researchers have developed many models and methods, such as frequency ratios (Lee and Lee 2006; Meten et al. 2015), logistic regression (Bai et al. 2010; Timilsina et al. 2014), weight of evidence (Piacentini et al. 2012), three-dimensional deterministic approach (Xie et al. 2007; Jia et al. 2012), analytical hierarchy process (Hasekioğulları and Ercanoglu 2012; Kayastha et al. 2013), artificial neural network (Conforti et al. 2014) and support vector machine (Yao et al. 2008; Xu et al. 2012). It is important to compare these different models and methods (Bui et al. 2012a) because the most suitable method remains undecided. In recent decades, a large number of studies have compared the different models (Bui et al. 2012a; Mohammady et al. 2012; Shahabi et al. 2014; Cui et al.

CONTACT Shouyun Liang shouyunl@163.com

2016; Ding et al. 2016; Hong et al. 2016a; Kornejady et al. 2017; Chen et al. 2017a, 2017b) to cross check the results from different models and determine which is optimal.

In addition to the models and methods, the quality of the database also influences the accuracy of susceptibility assessment. To accurately prepare the database of conditioning factors, ASTER imagery and LiDAR data have been used to produce altitude, slope, aspect, curvature, stream power index, topographic wetness index, topographic roughness index and sediment transport index measurements (Nefeslioglu et al. 2012; Song et al. 2012; Jebur et al. 2014). To improve the landslide inventory map, the random forest algorithm was employed to recognize landslides from very high resolution imagery (Stumpf and Kerle 2011a, 2011b; Chen et al. 2014). Landslides typically occur under complex geological conditions. It is important to determine the optimal combination of predisposing factors and to use them to produce an accurate landslide susceptibility map (Wang et al. 2015a). Comparison analysis is therefore also often applied in this process. Hasekioğulları and Ercanoglu (2012) successively increased factors in the order of their relative importance and found that a susceptibility map produced from more factors had better accuracy. Meten et al. (2015) applied mathematical combination theory on eight selected predisposing factors to analyse all combinations and optimized the susceptibility map using a combination of seven factors, without distance from water, according to prediction accuracy. Wang et al. (2015a) combined factors by adding new ones to four basic factors and showed that the optimum susceptibility map was obtained from the combination of slope, lithology, drains, annual rainfall, faults, roads and vegetation. These studies indicated that some landslide-conditioning factors may cause noise that reduces predictive capability and should be abandoned. For this reason, the random forest method (Chen et al. 2018) and the linear support vector machine method (Chen et al. 2017b) were adopted for the selection of landslide-conditioning factors. Furthermore, distinct spatial resolutions were used and compared to determine the suitable range of pixel sizes for landslide susceptibility assessment (Lee et al. 2004; Palamakumbure et al. 2015).

Comparisons have become a mainstream approach in landslide susceptibility assessment and provide a new perspective for improving mapping accuracy. Prior to the actual mapping, a factor implemented with a continuous variable must be reclassified into multi-ranges as input data. The criteria for reclassification vary among studies. For example, aspect is often classified into Flat, N, E, S and W (Yalcin 2008; Tangestani 2009; Conforti 2014) or into Flat, N, E, S, W, NE, SE, SW and NW (Regmi et al. 2010; Bui et al. 2012a; Kayastha et al. 2013; Ozdemir and Altural 2013; Hong et al. 2016a; Zhang et al. 2016), among many other factors (e.g. slope, elevation, distance from roads and distance from rivers). Theoretically, the distinction of factor subdivision can change the topological structure of the input and output data. However, no comparison has been carried out to assess the effects of the division standards on the accuracy of susceptibility mapping. This article therefore introduces a novel comparison between different criteria of factor subdivision. Using the Kongtong District as a case study, six factors were chosen and classified using the predisposing criteria collected from the existing literature. The resultant maps combined from different criteria were compared using the relative operative characteristic method. The optimal combination of criteria for factor subdivision was then determined based on comparison and found to enhance the susceptibility mapping results.

2. Study area

Kongtong is a district in the Loess Plateau region of China with frequent landslides. Malan loess is widely distributed, occupying approximately 80% of the area. The study

area is adjacent to the eastern Liupan Mountains, located at longitudes from 106°25' to 107°21′E and latitudes from 35°12' to 35°45'N, covering an area of approximately 1,800 km^2. The study area has an elevation range of 1,120–2,240 m with slope angles of 0–71° (Figures 1 and 2). According to statistics recorded at Pingliang Weather Station, the temperature ranges from −25.7 to 37.3 °C with an annual average temperature of 9.4 °C. The annual rainfall varies from 744.5 mm (in 1964) to 249.9 mm (in 1942) with an average of 503.7 mm. Rainfall is concentrated mainly from July to September.

Jinghe River, flowing through the middle of the district, is the only river in the region and has a total of 154 tributaries. The loess tablelands, or the residual tablelands, are mainly distributed north of Jinghe River, spreading NW–SE. The loess ridges and hills are more developed around the loess tablelands and the intermountain basin. There are also fairly frequent engineering activities, such as road construction, in the region. Cutting of the surface often occurs during the construction process, changing the original stress balance of the slope body, and providing favourable potential energy conditions for the occurrence of a landslide. According to records in the investigation report on geological disasters in Kongtong District accomplished by Geo-Environment Monitoring Institute of Gansu Province in 2007, landslides have occurred frequently and are widely distributed in this area (Figure 1). In July 1992, Caofeng Town experienced landslides with a direct economic loss of approximately 45,000 Yuan, and two people were killed by the Sishe landslide in Liuhu Township. And in August 1993, the Hongzhaobi landslide in Liuhu Township killed seven villagers and caused a direct economic loss of almost 120,000 Yuan.

3. Data

3.1. Collection

According to Ayalew et al. (2005), the factors selected for the GIS-based assessment of landslide susceptibility must be operational, non-redundant, non-uniform, measurable, and represented over the entire area. Glade et al. (2005) concluded that the selection of predisposing factors depended on, for example, the scale of analysis, the characteristics of the study area and the landslide type. Oh and Pradhan (2011) suggested that the input parameters should be reliable, representative and easily obtained. The most suitable mode is still under debate in terms of selecting and combining the independent variables for landslide hazard analysis (Hadji et al. 2013). We selected six parameters as landslide-conditioning factors: elevation, aspect, slope, distance from roads, proximity to streams and lithology.

Elevation is an important factor in landslide occurrence because of the large variability in weather conditions and climate at different elevations, which causes differences in soil and vegetation (Aniya 1985). For the study area, elevation can be generated from contour lines provided by a 1:50,000-scale topographic map. In contrast to elevation that was generated with acceptable accuracy, the data sets used for the rainfall and vegetation indices have poor precision. Elevation was therefore considered as an indicator for rainfall and the amount of vegetation.

The other topographic conditions involving aspect and slope also play significant roles in landslide events. Aspect often controls the amount of water in the slopes and hillsides (Shahabi et al. 2013). Slope is well known as the major factor controlling landslide formation (Shahabi et al. 2014). Aspect and slope were both calculated from the layer elevation using ArcMap v.10.3 software.

The river is an important factor as gully erosion may influence the initiation of landslides. In the mountains, roads often cut the slope, change the surface morphology

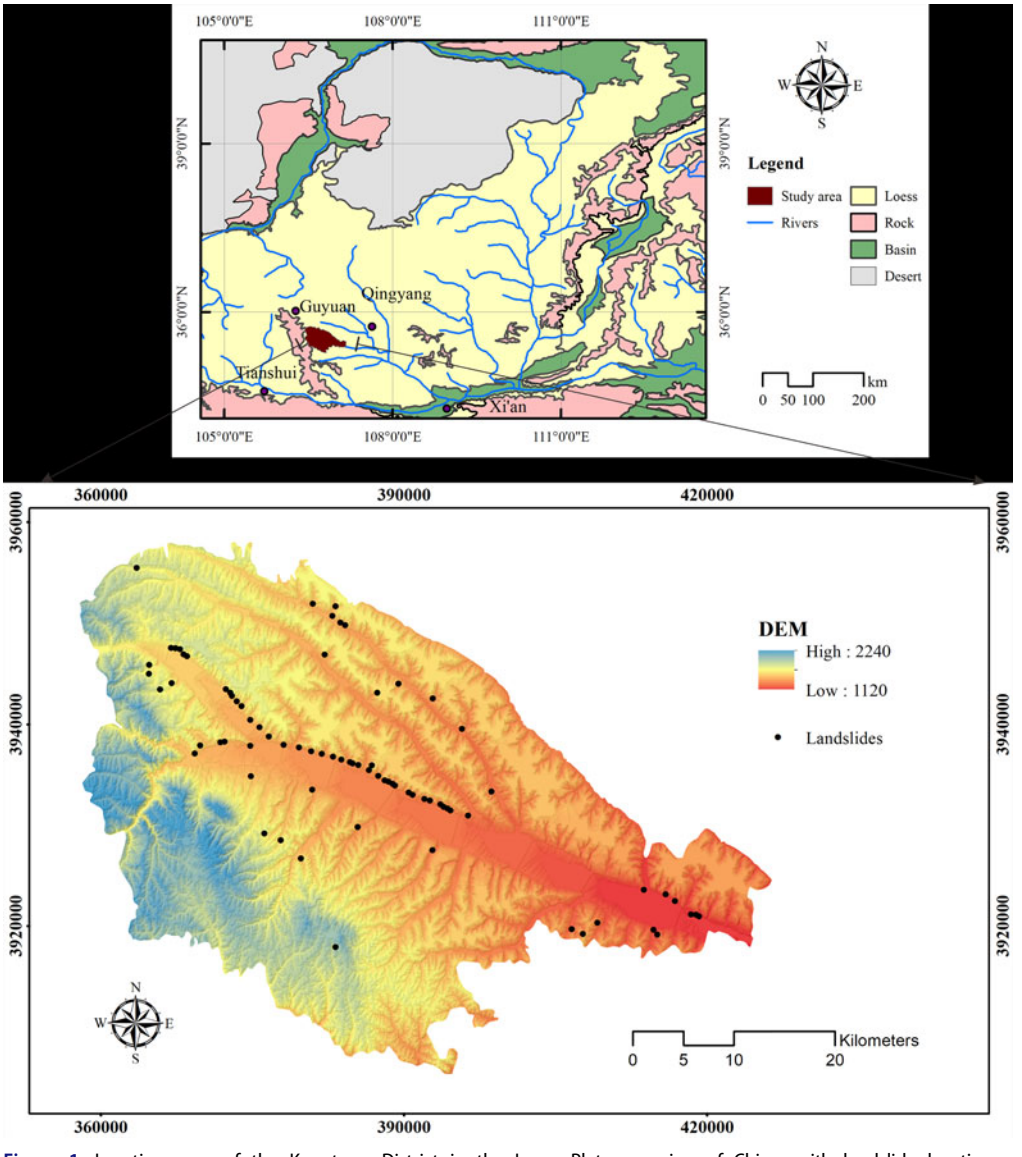

Figure 1. Location map of the Kongtong District in the Loess Plateau region of China, with landslide locations and elevation.

and affect slope stability. The distances from streams and roads were used to describe the level of influence of the rivers and roads, respectively. In ArcGIS, proximity to streams and distance from roads were generated using the Euclidean distance function along the streams and roads. The stream and road information is contained by the topographic map, and the vector data of streams and roads were extracted in advance from the map.

Lithology affects the occurrence of landslides because it determines different properties of weathering processes, infiltration rates (Kavzoglu et al. 2014) and shear strength. A 1:100,000 geological map was produced from 1:200,000 regional geological maps with appropriate supplementation during the field survey. The lithological map was produced from the 1:100,000-scale geological map.

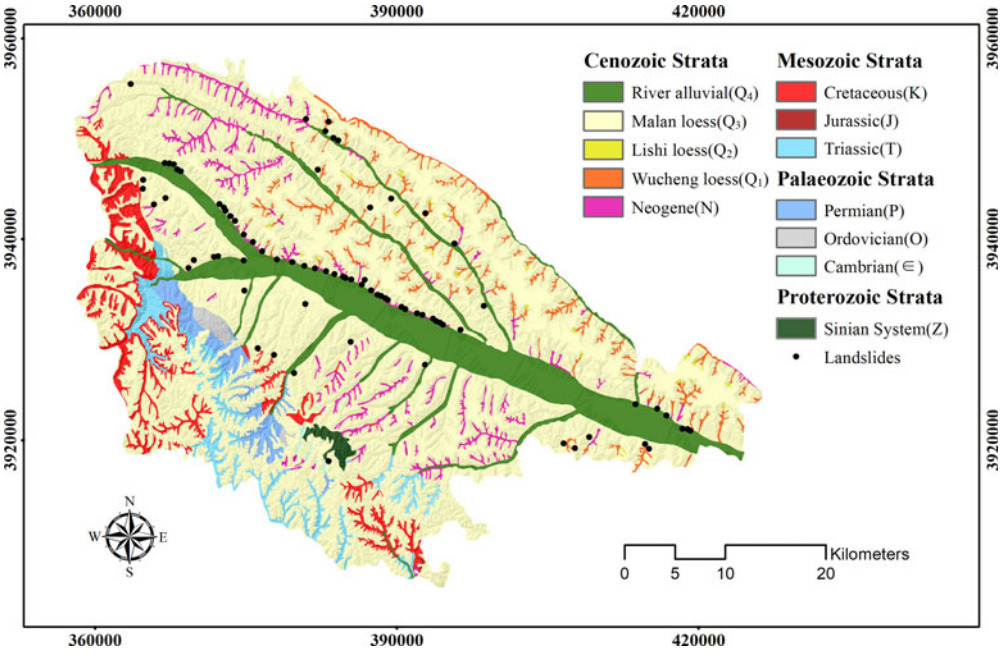

Figure 2. Spatial distribution map of lithology in the Kongtong District, modified from the Geological Environmental Monitoring Institute of Gansu Province (2007).

With the support of a special funded project for geological hazards, a comprehensive field investigation was completed and landslide statistics were collected in the study area according to 'the basic requirements of county (city) geological disaster investigation and zoning' issued by the Ministry of Land and Resources of China. A total of 78 landslide points were identified on the map by their GPS coordinates. All landslides were used for training and testing. The landslide areas range from 270 m^2 to .46 km^2 and cover a total of 3.4 km^2. Most of them occurred in the middle of the study area (Figure 2). According to records in the investigation report on geological disasters in Kongtong District accomplished by Geo-Environment Monitoring Institute of Gansu Province in 2007, landslides in the study area destroyed almost 43 buildings and caused 11 deaths, and still threaten the lives of approximately 13,331 local residents. The landslide inventory map passed the expert review organized by the Department of Land and Resources of Gansu Province.

3.2. Map resolution

The choice of resolution has an impact on the precision of mapping landslide susceptibility. Lee et al. (2004) conducted a comparative study on landslide susceptibility assessments using distinct spatial resolutions and found that a 30-m pixel was the recommended maximum to achieve acceptable predictive capacity. Palamakumbure et al. (2015) concluded that 10 m is the optimal pixel resolution for the Sydney Basin. Hengl (2006) recommended that guidelines must be followed in selecting a pixel size and calculated the average spaces between contour lines to extract a reasonable pixel size (p), which is expressed as follows:

$$p = A/\left(2 \cdot \sum L\right) \tag{1}$$

where A is the area of the entire study region, and ΣL is the total length of all contours. A recommended size of 58.6 m was obtained based on Equation (1). However, this region has many well-developed loess tablelands with relatively thin contour lines and valleys where the densities of the contours are greater than the average regional value. Therefore, the optimal pixel size of 20 m was adopted, which is less than 58.6 m. As a result, the maps of landslide-conditioning factors were converted to form a 20 m $\times$ 20 m grid with 2,636 rows and 3,423 columns, with a total of 4,815,503 pixels.

Slope and aspect were derived from the elevation and therefore shared the same resolution. The distances from water and roads were converted into a cell size of 20 m as the vector data of streams, and roads were also extracted from 1:50,000-scale topographic maps. Lithology was resampled at the same resolution as the elevation because the geological map is at a relatively smaller scale.

4. Method

4.1. Factor subdivision

Lithology distinguishes attributes by types, whereas other factors, such as elevation, aspect, slope, proximity to streams and distance from roads, describe properties based on continuous variables. Many studies have classified these variables into multi-ranges to express them as category. Because the reclassification criteria of the landslide-conditioning factors are the focus of our investigation, the following two sub-sections clarify the standards for the topographic conditions (elevation, slope and aspect) and linear features (proximity to roads and proximity to streams; Figure 3).

4.1.1. Topographic conditions

Two criteria were used to reclassify the aspect. Standard i divides the surface into N, E, S, W and Flat indicating upward facing (Yalcin 2008; Tangestani 2009; Conforti 2014). Standard ii divides the slope orientation into N, E, S, W, NE, SE, SW, NW and Flat (Regmi et al. 2010; Bui et al. 2012a; Kayastha et al. 2013; Ozdemir and Altural 2013; Hong et al. 2016a; Zhang et al. 2016). These two criteria cover almost all the subdivision modes of aspect used in the existing literature and were used as the predisposing division techniques for the landslide-susceptibility mapping of the Kongtong District.

The two main standards of slope subdivision were identified from previous studies. Criterion i regroups the slope into meaningful subclasses. They include flat to gentle slope (less than 15°), moderate slope (15–25°), fairly moderate slope (25–35°), steep slope (35°–45°) and very steep slope (greater than 45°) (Tangestani 2009; Kayastha et al. 2013; Pradhan and Kim 2016). Criterion ii divides the slope into six classes with 10° intervals: 0–10°, 10–20°, 20–30°, 30–40°, 40–50° and >50° (Yalcin 2008; Bui et al. 2012a; Zhang et al. 2016).

In contrast to aspect and slope, it is impossible to determine the range of elevations from existing studies similar to that of the Kongtong District, in this case 1,120–2,240 m. Therefore, the criterion for elevation based on existing research will not completely match with the study region. However, some guidelines are summarized as follows. For standard i, elevation is often reclassified in 200-m intervals (Regmi et al. 2010; Hong et al. 2016b; Pradhan and Kim 2016), whereas 100-m intervals are used in standard ii (Kawabata and Bandibas 2009; Ozdemir and Altural 2013; Conforti et al. 2014).

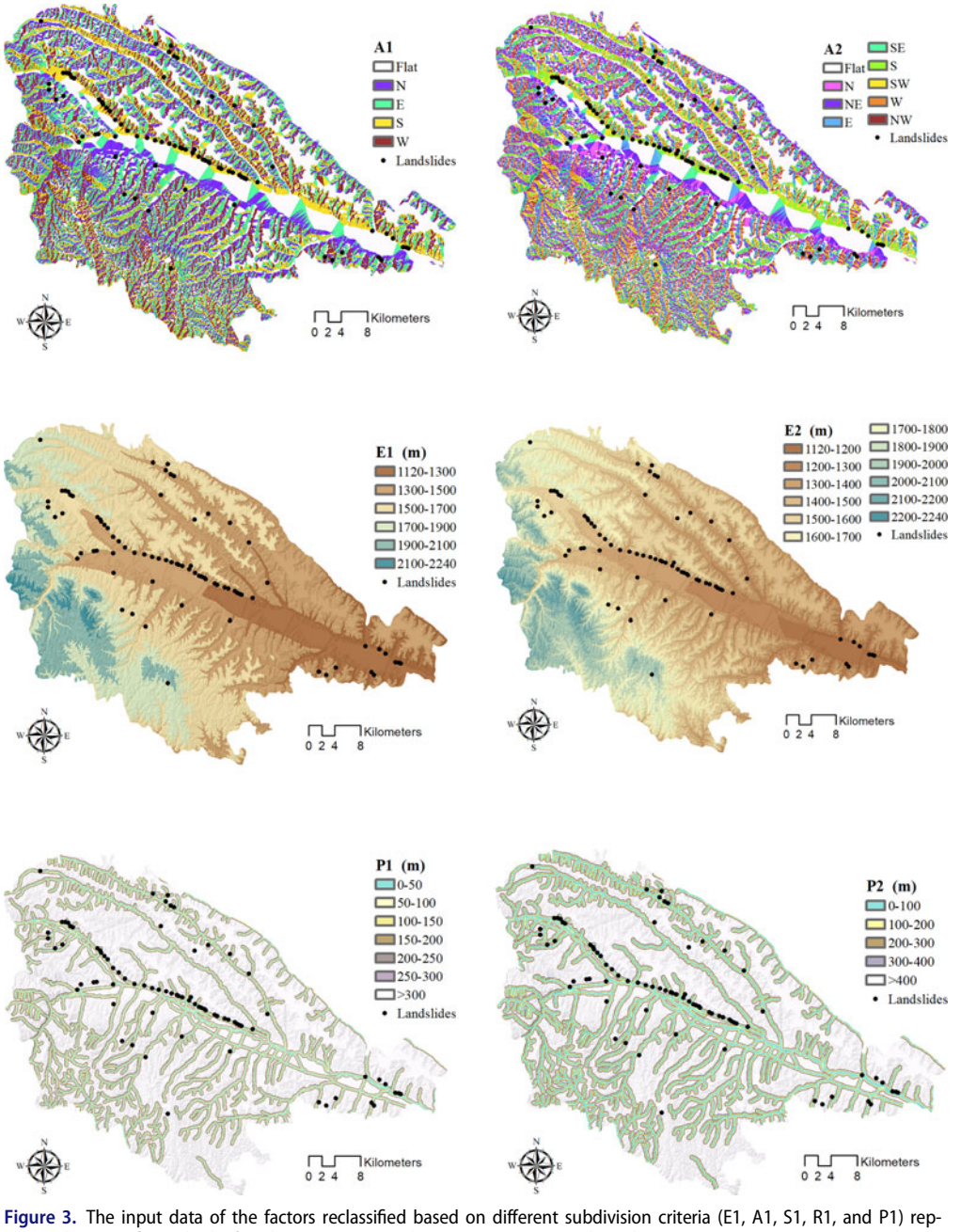

Figure 3. The input data of the factors reclassified based on different subdivision criteria (E1, A1, S1, R1, and P1) represent the use of standard i for elevation, aspect, slope, proximity to roads and proximity to streams, respectively; E2, A2, S2, R2 and P2 represent standard.

4.1.2. Linear features

The influence of the stream is assumed negligible beyond the longest impact distance, considered as the effect range. Areas farther than a specific distance from roads were considered to be simply influenced by instability related to roads. All pixels located farther than the longest impact distance from rivers or roads had little influence and were therefore grouped into a single class.

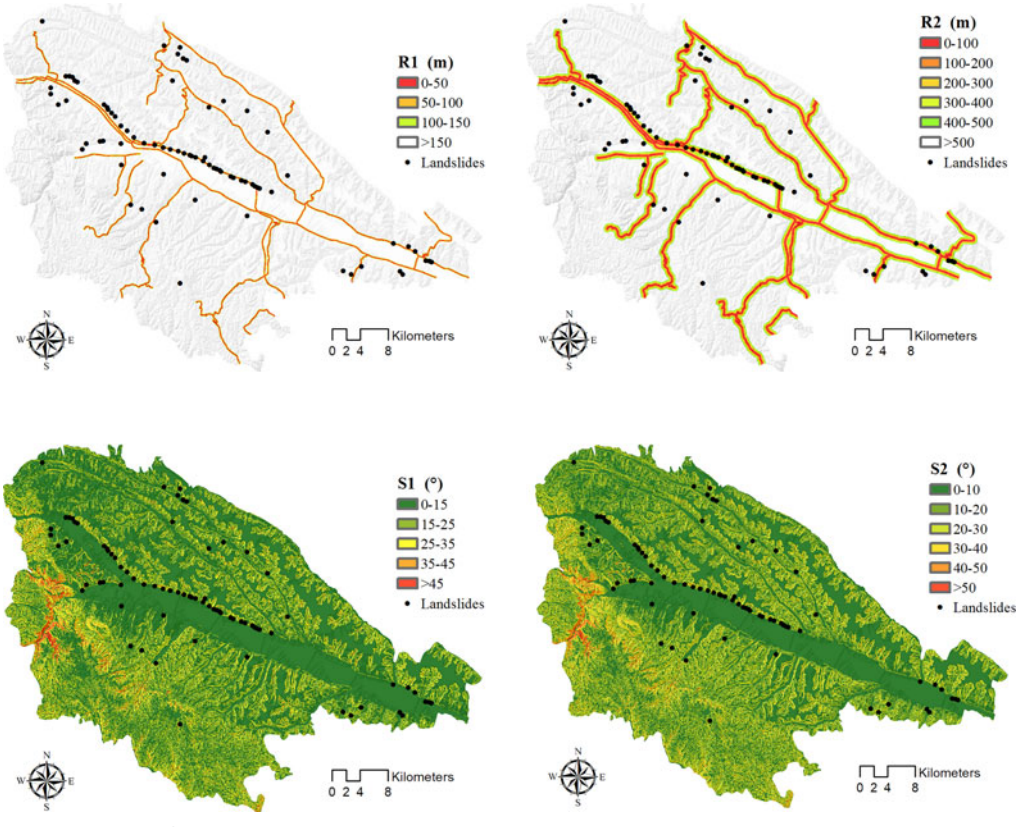

Figure 3. Continued

We collected two subdivision criteria from previous studies for the proximity to roads and the proximity to streams. Standard i for roads produced an effect range value of 150 m, and the distance was regrouped into 0–50, 50–100, 100–150 and >150 m (Che et al. 2012; Hong et al. 2015), whereas the effect range of standard ii was set as 500 m with classes of 0–100, 100–200, 200–300, 300–400, 400–500, and >500 m (Bui et al. 2012b; Mohammady et al. 2012).

The effect ranges for rivers for standards i and ii were set as 300 m and 400 m, respectively. The distance from streams was divided into 0–50, 50–100, 100–150, 150–200, 200–250, 250–300 and >300 m (Sezer et al. 2011; Che et al. 2012) for standard i, and 0–100, 100–200, 200–300, 300–400 and >400 m for standard ii (Lee and Lee 2006; Schicker and Moon 2012).

The first criterion for distance from roads or streams had a relatively small effect range on the occurrence of landslides, whereas the second had a larger effect range. Therefore, a reasonable setting for the effect range could be analysed by comparing susceptibility maps based on the two distinct criteria.

4.2. Evaluation methods

Weights-of-evidence was first used in mineral exploration in 1988 as a bivariate method (Bonham-Carter et al. 1988) and was then applied to assess landslide susceptibility after

Van Western (2002). This method depends on Bayes' rule:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{2}$$

where four symbols, B,$\bar{B}$, L, and $\bar{L}$, are used to represent the presence of a specific factor class, the absence of a specific factor class, the presence of landslides and the absence of landslides, respectively. Then, $p(B|L)$, $p(B|\bar{L})$, $p(\bar{B}|L)$ and $p(\bar{B}|\bar{L})$ can be calculated using Equation (2).

This method calculates weights (w^+ and w^-) to represent the relationship between landslides (L) and specific factor classes (B) (Bonham-Carter 1994) as follows:

$$w_i^+ = \ln \frac{p(B|L)}{p(B|\bar{L})} \tag{3}$$

$$w_i^- = \ln \frac{p(\bar{B}|L)}{p(\bar{B}|\bar{L})} \tag{4}$$

where w^+ indicates the presence of predictable variables (factor class) at the landslide positions, and the value of this weight suggests positive mutuality between the presence of predictable variables and landslides; w^- indicates the absence of predictable variables, representing the degree of negative mutuality.

Finally, the landslide susceptibility index (LSI) is calculated as the product of the probabilities relevant to the different constituent parts of the model, as represented in Equation (5):

$$\mathrm{LSI} = \sum w_i^+. \tag{5}$$

4.3. Validation

The mapping accuracy of the resultant map was evaluated by analysing the relative operative characteristics (ROC) and the frequency distributions in the susceptibility classes. Using the size under the ROC curve, abbreviated as AUC, both prediction and success rates were measured to check the reliability and efficiency of the landslide probability map (Mohammady et al. 2012; Umar et al. 2014; Wang et al. 2015b). In this paper, the landslide point used for training was also compared with the resulting maps for validation. The success rate curve provides information on the degree of fitness of the assessment method with the observed landslides (Chung and Fabbri 2003; Remondo et al. 2003). The magnitude of AUC varies from .5 to 1, where a large value indicates high accuracy, and a small one indicates inaccuracy (Fawcett 2006).

The frequency ratio was calculated for the landslide susceptibility maps generated under distinct combinations of subdivision criteria by comparing landslide locations with the susceptibility classes. All landslide points were superimposed over different landslide susceptibility zones to calculate the frequency ratio for each susceptibility zone. Theoretically, the frequency ratio values should increase from a low susceptibility zone to a very high one (Pourghasemi et al. 2012).

5. Results

Table 1 lists the results of the factor subdivision under 11 criteria involving two standards each for five factors (elevation, aspect, slope, proximity to roads and proximity to streams)

Table 1. Spatial relationships of landslides with different factors and their weights using the weights-of-evidence method with 11 subdivision criteria for the factors.

Factor	Subdivision criterion	Classes	Number Of grids	% of area	Number of landslide	% of landslide	W^+	W^-
Elevation (m)	E1	1,120–1,300	511,449	10.6	9	11.5	.083	.016
		1,300–1,500	1,579,500	32.8	44	56.4	.542	–.198
		1,500–1,700	1,569,301	32.6	24	30.8	–.057	.113
		1,700–1,900	674,968	14.0	1	1.3	–2.392	.141
		1,900–2,100	430,493	9.0	0	.0	none	.094
		2,100-2,240	49,792	1.0	0	.0	none	.010
	E2	1,120–1,200	165,744	3.4	6	7.7	.804	–.028
		1,200–1,300	345,705	7.2	3	3.8	–.624	.043
		1,300–1,400	721,847	15.0	19	24.4	.486	–.053
		1,400–1,500	857,653	17.8	25	32.0	.588	–.098
		1,500–1,600	984,791	20.5	16	20.5	.003	.051
		1,600–1,700	584,510	12.1	8	10.3	–.168	.044
		1,700–1,800	406,580	8.5	0	.0	none	.088
		1,800–1,900	268,388	5.6	1	1.3	–1.470	.047
		1,900–2,000	280,854	5.8	0	.0	none	.060
		2,000–2,100	149,639	3.1	0	.0	none	.032
		2,100–2,200	49,287	1.0	0	.0	none	.010
		2,200–2,240	505	.0	0	.0	none	.000
Aspect	A1	Flat	1,277,483	26.6	12	15.4	–.545	.178
		N	871,961	18.1	8	10.2	–.568	.115
		E	1,093,996	22.7	13	16.7	–.310	.115
		S	752,474	15.6	33	42.3	.996	–0.241
		W	819,589	17	12	15.4	–.101	.056
	A2	Flat	1,277,483	26.5	12	15.4	–.545	.178
		N	395,440	8.2	5	6.4	–.248	.033
		NE	569,766	11.8	4	5.1	–.836	.084
		E	553,177	11.5	2	2.6	–1.500	.101
		SE	466,159	9.7	11	14.1	.376	–.017
		S	347,088	7.2	12	15.4	.758	–.056
		SW	386,513	8.0	25	32.1	1.385	–.211
		W	402,616	8.4	3	3.8	–.776	.056
		NW	417,261	8.7	4	5.1	–.525	.049
Slope (°)	S1	0–15	2,817,022	58.5	40	51.3	–.132	.355
		15–25	123,5126	25.7	17	21.8	–.163	.106
		25–35	589,343	12.2	18	23.1	.634	–.072
		35–45	145,543	3.0	3	3.8	.241	.000
		>45	28,469	.6	0	.0	none	.006
	S2	0–10	2,248,727	46.7	29	37.2	–.228	.278
		10–20	1,225,992	25.4	20	25.6	.007	.066
		20–30	957,379	19.9	21	26.9	.303	–.020
		30–40	311,125	6.5	6	7.7	.174	.004
		40–50	60,658	1.3	2	2.6	.711	–.008
		>50	11,622	.2	0	.0	none	.002
Proximity to roads (m)	R1	0–50	111,816	2.3	13	16.7	1.971	–.119
		50–100	106,295	2.2	5	6.4	1.066	–.030
		100–150	85,699	1.8	5	6.4	1.282	–.034
		>150	4,511,693	93.7	55	70.5	–.284	1.940
	R2	0–100	218,111	4.5	18	23.1	1.628	–.157
		100–200	184,453	3.8	9	11.5	1.103	–.057
		200–300	171,063	3.6	2	2.6	–.326	.016
		300–400	164,094	3.4	5	6.4	.632	–.018
		400–500	159,592	3.3	3	3.8	.149	.003
		>500	3,918,190	81.4	41	52.6	–.437	1.138
Proximity to streams (m)	P1	0–50	361,431	7.5	2	2.6	–1.074	.057
		50–100	335,620	7.0	2	2.6	–1.000	.052
		100–150	268,507	5.6	6	7.7	.322	–.006
		150–200	307,614	6.4	9	11.5	.591	–.030
		200–250	255,802	5.3	16	20.5	1.351	–.124

(*continued*)

Table 1. Continued.

Factor	Subdivision criterion	Classes	Number Of grids	% of area	Number of landslide	% of landslide	W^+	W^-
		250–300	280,768	5.8	17	21.8	1.319	−.130
		>300	300,5761	62.4	26	33.3	−.627	.670
	P2	0–100	697,051	14.5	4	5.1	−1.038	.115
		100–200	576,121	12.0	15	19.2	.475	−.039
		200–300	536,570	11.1	33	42.3	1.334	−.292
		300–400	486,664	10.1	8	10.3	.015	.021
		>400	2,519,097	52.3	18	23.1	−.818	.538
Lithology	–	Q_4	588,287	12.2	8	10.3	−.175	.045
		Q_3	345,4614	71.8	62	79.5	.103	.262
		Q_2	8,793	.2	0	.0	none	.002
		Q_1	105,835	2.2	0	.0	none	.022
		N	147,429	3.1	8	10.2	1.209	−.054
		K	230,897	4.8	0	.0	none	.049
		J	789	.0	0	.0	none	.000
		T	146,168	3.0	0	.0	none	.031
		P	86,661	1.8	0	.0	none	.018
		O	21,040	.4	0	.0	none	.004
		$\in$	4,772	.1	0	.0	none	.001
		Z	20,218	.4	0	.0	none	.004

and a single standard for lithology. The spatial relationship between the landslides and each factor was considered by calculating the percentages of area, landslides and weights of landslides for the classes. Only the positive weight (w^+) was used as a part of the landslide susceptibility index, as indicated in Equation (5). Table 1 demonstrates that if a specific class has no landslides, w^+ will have no value in the results. In this case, the LSI is assigned 'none' as the lowest value. This situation was detected, for example, for the 1700–1800 m elevation class, but ignored when 200 m was used as the interval to reclassify the elevation. This finding suggests that standard ii can increase the 'none' component of the LSI, or the lowest value, compared with standard i. For the aspect, the w^+ value varied from −.568 to .996 under criterion i, but between −1.500 and 1.385 under criterion ii. Standard ii broadened the range of the w^+ value, which is expected to improve the range of the LSI and the mapping result because the greater the scope of the LSI, the better the susceptibility map, as suggested by Meten et al. (2015). Positive values of w^+, .632 and .149, were found for classes 300–400 m and 400–500 m, respectively, for the proximity to roads, but negative values from 300 to 500 m were found when 150 m was used as the effect range. For proximity to streams, a positive w^+ of .015 occurred for 300–400 m distance in contrast to a negative value of −.627 for distances greater than 300 m because of the difference in the effect range. This result suggests that the selection of criteria to reclassify the distance from roads and streams likely affects the outputs, considering that the LSI is based on w^+.

Table 2 lists the success rates (represented by AUC) of 32 susceptibility maps corresponding to the combination of the subdivision criteria (Figure 4). The subdivision standard of lithology was not included in the combination because this factor was only classified by a single mode. All the factors (except lithology) combined in Group 1 were subdivided by criterion i, producing a map with the lowest value of .8434. The highest value of .8751 was obtained by Group 29, where elevation, aspect, proximity to roads and proximity to streams were reclassified using standard ii. To save space, only the maps generated by groups 1 and 29, indicating the largest difference, were selected for the division of the susceptibility classes, which helps make the map easier to read. Most researchers develop class boundaries in accordance with their own expert opinion, as there are no common rules to automatically divide such continuous data (Ayalew et al. 2004). In this

Table 2. Area under the curves (success rate curve) obtained using different factor classification standards.

Group ID	Combination of subdivision criteria	AUC value	Group ID	Combination of subdivision criteria	AUC value
1	E1 + A1 + S1 + R1 + P1	.8434	17	E2 + A1 + S1 + R1 + P1	.8493
2	E1 + A1 + S1 + R1 + P2	.8480	18	E2 + A1 + S1 + R1 + P2	.8539
3	E1 + A1 + S1 + R2 + P1	.8501	19	E2 + A1 + S1 + R2 + P1	.8524
4	E1 + A1 + S2 + R1 + P1	.8443	20	E2 + A1 + S2 + R1 + P1	.8487
5	E1 + A1 + S1 + R2 + P2	.8535	21	E2 + A1 + S1 + R2 + P2	.8564
6	E1 + A1 + S2 + R2 + P1	.8513	22	E2 + A1 + S2 + R2 + P1	.8525
7	E1 + A1 + S2 + R1 + P2	.8486	23	E2 + A1 + S2 + R1 + P2	.8532
8	E1 + A1 + S2 + R2 + P2	.8550	24	E2 + A1 + S2 + R2 + P2	.8566
9	E1 + A2 + S1 + R1 + P1	.8664	25	E2 + A2 + S1 + R1 + P1	.8706
10	E1 + A2 + S1 + R1 + P2	.8694	26	E2 + A2 + S1 + R1 + P2	.8739
11	E1 + A2 + S1 + R2 + P1	.8706	27	E2 + A2 + S1 + R2 + P1	.8723
12	E1 + A2 + S2 + R1 + P1	.8643	28	E2 + A2 + S2 + R1 + P1	.8703
13	E1 + A2 + S1 + R2 + P2	.8727	29	E2 + A2 + S1 + R2 + P2	.8751
14	E1 + A2 + S2 + R2 + P1	.8690	30	E2 + A2 + S2 + R2 + P1	.8722
15	E1 + A2 + S2 + R1 + P2	.8675	31	E2 + A2 + S2 + R1 + P2	.8731
16	E1 + A2 + S2 + R2 + P2	.8713	32	E2 + A2 + S2 + R2 + P2	.8745

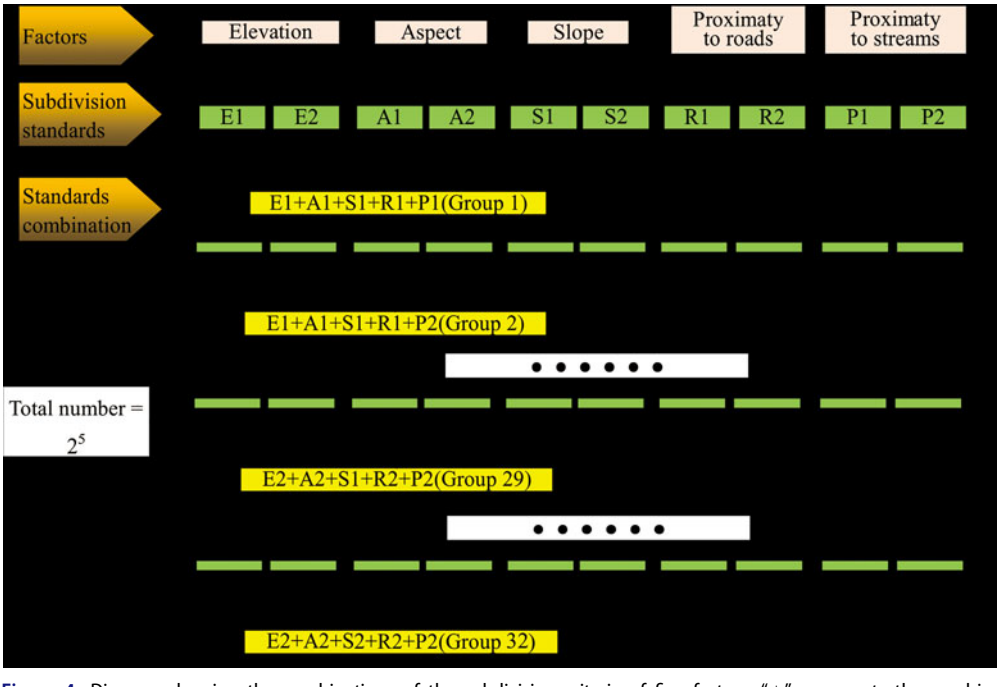

Figure 4. Diagram showing the combinations of the subdivision criteria of five factors; "+" represents the combination of subdivision criteria.

study, the two selected LSI maps were reclassified into low, moderate, high and very high susceptible zones (Figure 5), representing 10%, 20%, 30% and 40% of the study area, respectively (Kayastha et al. 2012, 2013). Figure 6 shows the percentage and the frequency ratio of landslides in the susceptibility classes calculated by comparing the occurrence of landslides with the results of the landslide susceptibility map. Over 60% of the landslides occurred in the very high susceptibility classes for both maps (Figure 6(a)). However, the high and very high susceptibility classes for the map with the highest AUC contain more landslides than the same classes for the map with the lowest AUC. This finding means

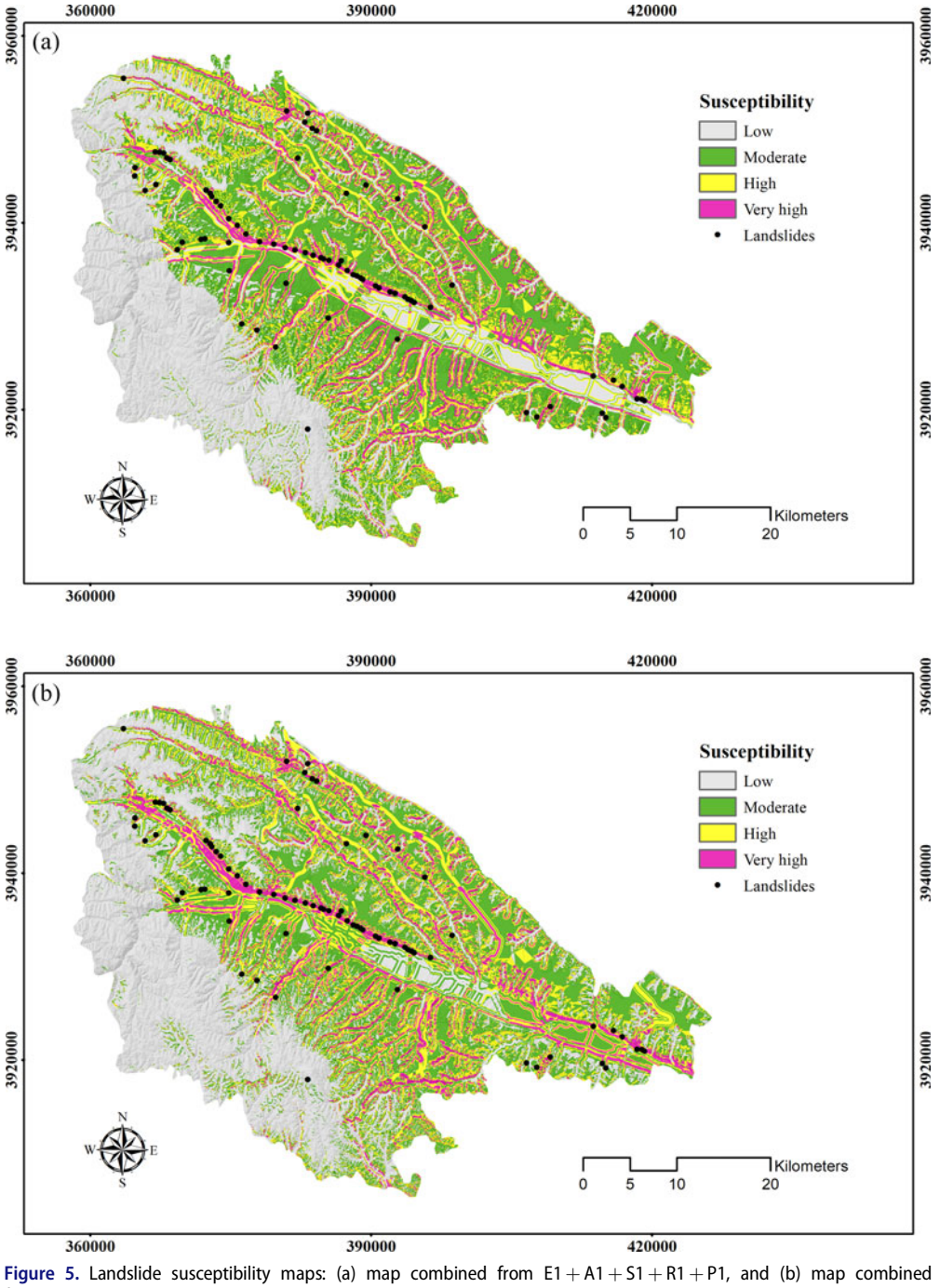

Figure 5. Landslide susceptibility maps: (a) map combined from E1 + A1 + S1 + R1 + P1, and (b) map combined from E2 + A2 + S1 + R2 + P2.

that altering the subdivision criteria helps identify more landslides located in higher susceptibility classes. Figure 6(b) presents a slight increase in the frequency ratio from low to high susceptibility classes and an extreme increase from high to very high susceptibility classes, which indicates that the two maps are both reliable.

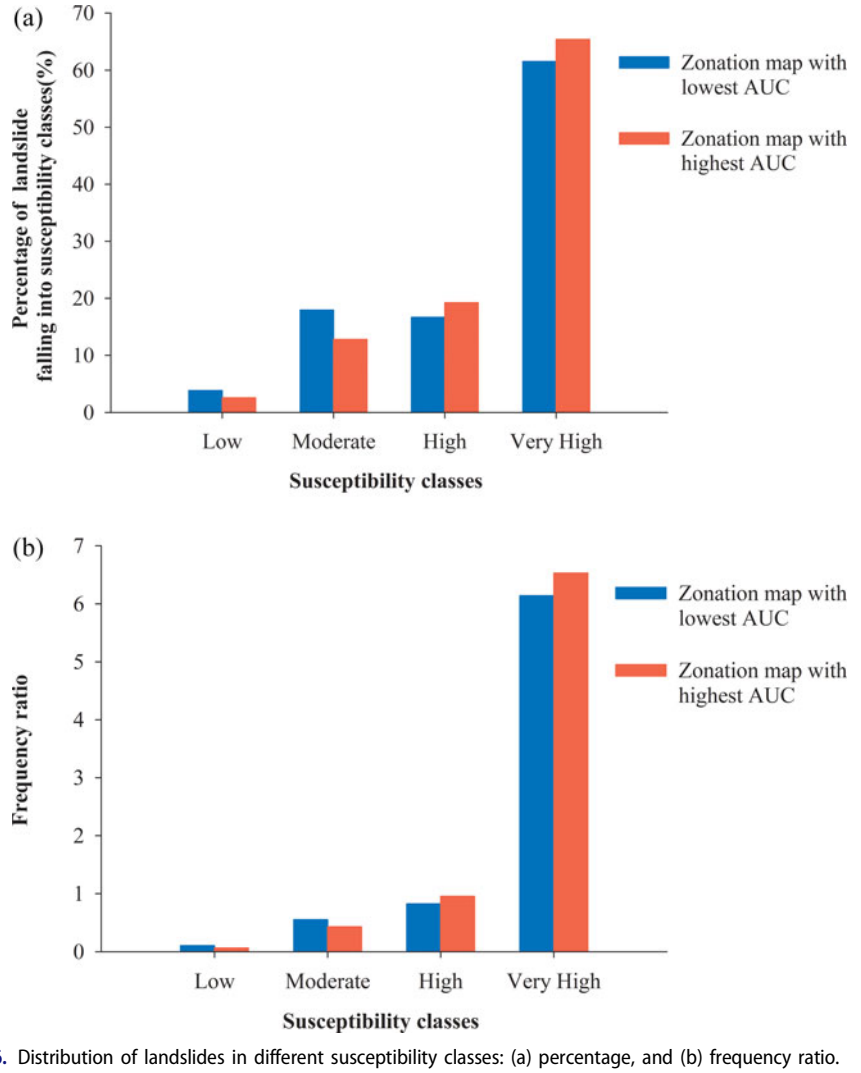

Figure 6. Distribution of landslides in different susceptibility classes: (a) percentage, and (b) frequency ratio.

6. Discussion

Landslide points that were used in this study were collected through comprehensive field investigation, but not all landslides were identified. In the middle portion of the study area, landslides pose serious threats to dense populations and are therefore the focus of field investigations. In addition, surveys are easy to conduct in the central region because of relatively easy access by roadway. Landslides occur in other areas, but they attract little attention because of poor access and low population densities, as is the case in the Southern and Northern Mountains, and are therefore not mapped and studied. Therefore, the landslides considered in this research are only those in the more populated areas. The concentrated distribution of landslides used in this study has a relatively small influence on the results of this assessment.

The value of AUC decreases in order of Group 9 (.8664), Group 3 (.8501), Group 17 (.8493), Group 2 (.8480), Group 4 (.8443) and Group 1 (.8434; Table 2). Comparing with

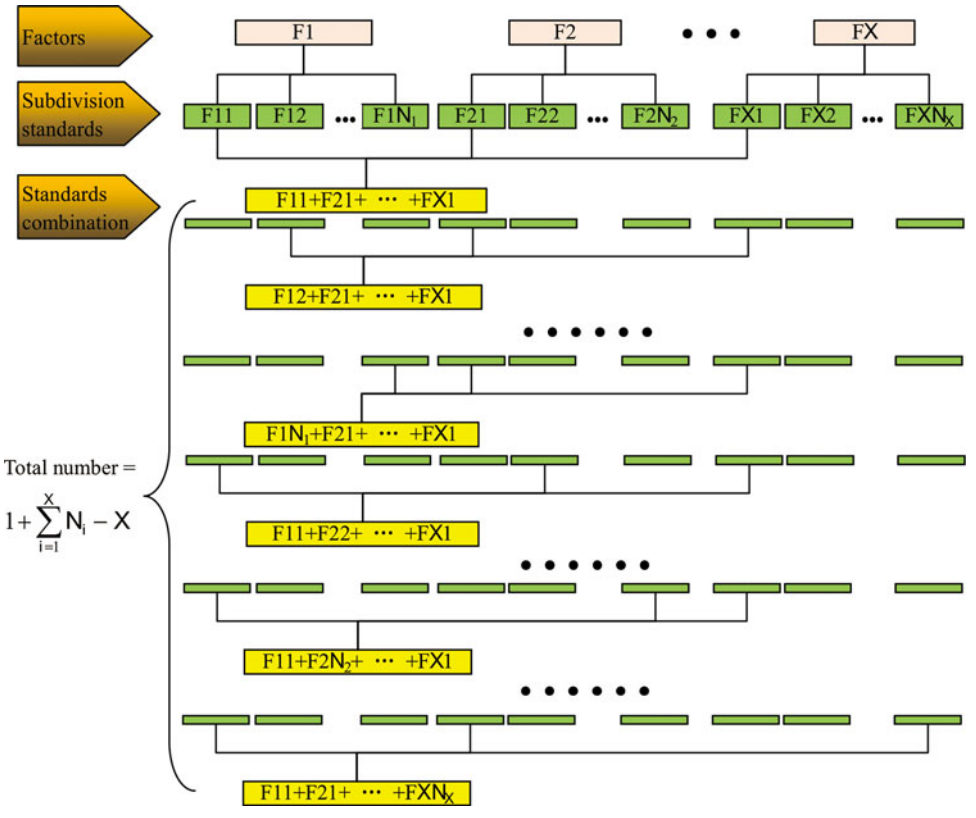

Figure 7. Simplified mode to combine different subdivision criteria of X factors to determine the best combination.

Group 1, where all the factors were reclassified using standard i, groups 9, 3, 17, 2 and 4 applied standard ii to subdivide aspect, proximity to roads, elevation, proximity to water and slope, respectively. The values indicate that dividing aspect by standard ii instead of standard i results in the clearest improvement in the accuracy of the susceptibility mapping, followed by proximity to roads, elevation and proximity to streams. However, the subdivision criteria of slope hardly affected the accuracy because the AUC of Group 4 is similar to that of Group 1. The selection of modes for aspect subdivision has a significant effect on the accuracy of susceptibility mapping and that of elevation. The more detailed the subdivision of the aspect and elevation, the higher the accuracy. Making the effect range of roads and rivers more suitable is meaningful, and we suggest that the longest distance influenced by roads and rivers should not be set too small, although an effect range that is too large may not be realistic.

Groups 25, 20, 19 and 18 used standard ii to regroup aspect, slope, proximity to roads and proximity to streams after elevation was subdivided by standard ii (Table 2). In this case, the largest increment of the AUC value was also produced when aspect was reclassified using standard ii, still followed by proximity to roads and streams, whereas the two standards for slope remained similar. This result means that the subdivision standard causes a relatively independent effect on the AUC value. Therefore, we propose a simplified and general mode to combine the different subdivision criteria of factors (Figure 7). In this mode, it is assumed that there are *X* landslide-conditioning factors, F1, F2, … , F*X*. For any factor F*x*, *Nx* of the subdivision criteria are considered. The combination of

'F11 + F21 + … + FX1' is set as the basic reference. For each factor Fi ($i = 1, 2, \ldots, X$), there are $N_{i\text{-}1}$ combinations compared with the reference to determine the best one, and a total of $1 + \sum_{i=1}^{X} N_i - X$ combinations are considered for comparison in contrast to $\prod_{i=1}^{X} N_i$ because the subdivision criteria influence accuracy independently. Taking the data of this article as an example, Group 1 is considered as the reference, and groups 17, 9, 4, 3 and 2 are compared with Group 1 to determine the best combination for elevation, aspect, slope, proximity to roads and proximity to streams, respectively. Standard ii is the best option for all the factors, and Group 32 is considered to be the best combination based on this mode. The AUC of Group 32 is .8745, very close to the highest value (.8751; Table 2). Using this mode, only some precision is lost, but a large amount of energy is saved during the process of comparing the subdivision criteria of the factors.

7. Conclusions

In the susceptibility mapping of landslides, the factors implemented based on continuous variables are often represented by multi-ranges obtained by the subdivision of these factors. The subdivision criteria vary between studies, even for the same factor. Using the Kongtong District as a case study, a range of criteria were used to reclassify factors and compared to enhance the accuracy of the assessment. Two standards for each parameter, including elevation, aspect, slope, proximity to roads and proximity to rivers, were collected from the literature and set as predisposing modes of subdivision, whereas lithology was implemented based on types without subdivision. In addition to the above six landslide-conditioning factors, a landslide inventory map was prepared based on field surveys to produce a spatial database, and all observed landslides were used for training and testing. The possible spatial distribution of the landslides was identified by comparing the occurrences of landslides with the layers of subdivided factors and a positive weight, w^{+}, determined using a weights-of-evidence method that worked as a component for calculating the susceptibility index of landslides.

A total of 32 maps were produced from the combinations of the criteria. The success rate was represented by the area under the ROC curve and calculated to analyse the accuracy of the resultant maps. Thus, when aspect, elevation, proximity to roads and proximity to streams were subdivided using standard ii, the optimal susceptibility map with the maximum AUC value of .8751 was produced, whereas the minimum value, approximately .03 less than the former, was calculated when all factors were reclassified using criterion i. To save space, only the maps with the highest and lowest AUC were reclassified into four susceptibility classes and compared with the landslide points to calculate the frequency ratio of landslides. The two maps are reliable because over 60% of the landslides occurred in very high susceptibility zones and the frequency ratio values increase from the class of low susceptibility to that of very high susceptibility.

In addition, the relationship between the value of the AUC and the combination of criteria shows that aspect produces the largest improvement to the accuracy using standard ii, which divides this parameter into Flat, N, E, S, W, NE, SE, SW and NW. This result also indicates that determining a suitable effect range for roads and rivers is significant. Finally, a simplified and general mode was proposed to combine the different subdivision criteria of the factors considering that the subdivision standard causes a relatively independent effect on the AUC value. The mode saves effort and time during the comparison of the subdivision criteria of the factors. This study improves the identification of optimized landslide-susceptibility mapping, and the results will serve as an important reference for future research.

Acknowledgements

We would like to thank two anonymous reviewers and editor of professor Lulla for their critical reviews and constructive comments which have improved the manuscript. The geological disaster data sets used in this analysis come from the report accomplished by Geo-Environment Monitoring Institute of Gansu Province.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project was supported by the State Key Development Program for Basic Research of China under Grant Number 2014CB744701 and by the National Natural Science Foundation of China under Grant Number 41272326.

ORCID

Ge Yan http://orcid.org/0000-0003-0499-085X
Shouyun Liang http://orcid.org/0000-0001-5728-7311

References

Aniya M. 1985. Landslide-susceptibility mapping in the Amahata river basin, Japan. Ann Assoc Am Geogr. 75:102–114.

Ayalew L, Yamagishi H, Marui H, Kanno T. 2005. Landslide in Sado Island Japan: GIS-based susceptibility mapping with comparison of results from two methods and verifications. Eng Geol. 81:432–445.

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. 1:73–81.

Bai SB, Jian W, Zhou PG, Hou SS, Xu SN. 2010. GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. Geomorphology. 115:23–31.

Bonham-Carter GF. 1994. Geographic information systems for geoscientists: modelling with GIS. Comput. Methamphetamine Geosci. 4:1–2.

Bonham-Carter GF, Agterberg FP, Wright DF. 1988. Integration of Geological datasets for gold exploration in Nova Scotia. Photogramm Eng Remote Sens. 54:1585–1592.

Bui DT, Pradhan B, Lofman O, Revhaug I, Dick OB. 2012a. Landslide susceptibility assessment in the HoaBinh province of Vietnam: a comparison of the Levenberg-Marquardt and Bayesian regularized neural networks. Geomorphology. 171–172:12–29.

Bui DT, Pradhan B, Lofman O, Revhaug I, Dick OB. 2012b. Landslide susceptibility mapping at HoaBinh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS. Comput Geosci. 45:199–211.

Che VB, Kervyn M, Suh CE, Fontijn K, Ernst GGJ. del Marmol MA, Trefois P, Jacobs P. 2012. Landslide susceptibility assessment in Limbe (SW Cameroon): a field calibrated seed cell and information value method. Catena. 92:83–98.

Chen W, Li X, Wang Y, Chen G, Liu S. 2014. Forested landslide detection using LiDAR data and the random forest algorithm: a case study of the Three Gorges, China. Remote Sens Environ. 152:291–301.

Chen W, Pourghasemi HR., Zhao Z. 2017a. A GIS-based comparative study of dempster-shafer, logistic regression, and artificial neural network models for landslide susceptibility mapping. Geocarto Int. 32:367–385.

Chen W, Xie X, Peng J, Shahabi H, Hong H, Bui DT, Duan Z, Li S, Zhu AX. 2018. GIS-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method. Catena. 164:135–149.

Chen W, Xie X, Wang J, Pradhan B, Hong HY, Bui DT, Duan Z, Ma JQ. 2017b. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena. 151:147–160.

Chung CF, Fabbri AG. 2003. Validation of spatial prediction models for landslide hazard mapping. Nat Hazards. 30:451–472.
Conforti M, Pascale S, Robustelli G, Sdao F. 2014. Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy). Catena. 113:236–250.
Cui K, Lu D, Li W. 2016. Comparison of landslide susceptibility mapping based on statistical index, certainty factors, weights of evidence and evidential belief function models. Geocarto Int. 32:935–955.
Ding Q, Chen W, Hong H. 2016. Application of frequency ratio, weights of evidence and evidential belief function models in landslide susceptibility mapping. Geocarto Int. 32:619–639.
Fawcett T. 2006. An introduction to ROC analysis. Pattern Recognit Lett. 27:861–874.
Glade T, Anderson M, Crozier MJ (Eds). 2005. Landslide hazard and risk. New York: Wiley. p. 824.
Hadji R, Boumazbeur AE, Limani Y, Baghem M, Chouabi AEM, Demdoum A. 2013. Geologic topographic and climatic controls in landslide hazard assessment using GIS modeling: a case study of Souk Ahras region, NE Algeria. Quat Int. 302:224–237.
Hasekioğulları GD, Ercanoglu M. 2012. A new approach to use AHP in landslide susceptibility mapping: a case study at Yenice (Karabuk, NW Turkey). Nat Hazards. 63:1157–1179.
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. 171–172:30–41.
Hengl T. 2006. Finding the right pixel size. Comput Geosci. 32:1283–1298.
Hong H, Chen W, Xu C, Youssef AM, Pradhan B, Bui DT. 2016a. Rainfall-induced landslide susceptibility assessment at the chongren area (China) using frequency ratio, certainty factor, and index of entropy. Geocarto Int. 32:139–154.
Hong H, Pourghasemi HR, Pourtaghi ZS. 2016b. Landslide susceptibility assessment in Lianhua County (China): a comparison between a random forest data mining technique and bivariate and multivariate statistical models. Geomorphology. 259:105–118.
Hong H, Pradhan B, Xu C, Bui DT. 2015. Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines. Catena. 133:266–281.
Jebur MN, Pradhan B, Tehrany MS. 2014. Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. Remote Sens Environ. 152:150–165.
Jia N, Mitani Y, Xie M, Djamaluddin I. 2012. Shallow landslide hazard assessment using a three-dimensional deterministic model in a mountainous area. Comput Geotech. 45:1–10.
Kavzoglu T, Sahin EK, Colkesen I. 2014. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. Landslides. 11:425–439.
Kawabata D, Bandibas J. 2009. Landslide susceptibility mapping using geological data, a DEM from ASTER images and an artificial neural network (ANN). Geomorphology. 113:97–109
Kayastha P, Dhital MR, De Smedt F. 2012. Landslide susceptibility mapping using the weight of evidence method in the Tinau watershed, Nepal. Nat Hazards. 63:479–498.
Kayastha P, Dhital MR, De Smedt F. 2013. Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: a case study from the Tinau watershed, west Nepal. Comput Geosci. 52:398–408.
Kornejady A, Ownegh M, Rahmati O, Bahremand A. 2017. Landslide susceptibility assessment using three bivariate models considering the new topo-hydrological factor: hand. Geocarto Int. 1–31.
Lee S, Choi J, Woo I. 2004. The effect of spatial resolution on the accuracy of landslide susceptibility mapping: a case study in Boun, Korea. Geosci J. 8:51–60
Lee S, Lee MJ. 2006. Detecting landslide location using KOMPSAT 1 and its application to landslide-susceptibility mapping at the Gangneung area, Korea. Adv Space Res. 38:2261–2271.
Meten M, PrakashBhandary N, Yatabe R. 2015. Effect of landslide factor combinations on the prediction accuracy of landslide susceptibility maps in the Blue Nile Gorge of Central Ethiopia. Geoenviron Disasters. 2:1–17.
Mohammady M, Pourghasemi HR, Pradhan B. 2012. Landslide susceptibility mapping at Golestan Province, Iran: a comparison between frequency ratio, Dempster-Shafer, and weights-of-evidence models. J Asian Earth Sci. 61:221–236.
Nefeslioglu HA, San BT, Gokceoglu C, Duman TY. 2012. An assessment on the use of Terra ASTER L3A data in landslide susceptibility mapping. Int J Appl Earth Obs Geoinf. 14:40–60.
Oh HJ, Pradhan B. 2011. Application of a neuro-fuzzy model to landslide susceptibility mapping for shallow landslides in tropical hilly area. Comput Geosci. 37:1264–1276.

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

Palamakumbure D, Flentje P, Stirling D. 2015. Consideration of optimal pixel resolution in deriving landslide susceptibility zoning within the Sydney Basin, New South Wales, Australia. Comput Geosci. 82:13–22

Piacentini D, Troiani F, Soldati M, Notarnicola C, Savelli D, Stefan S, Schneiderbauer S, Strada C. 2012. Statistical analysis for assessing shallow-landslide susceptibility in South Tyrol (south-eastern Alps, Italy). Geomorphology. 151–152:196–206.

Pourghasemi HR, Mohammady M, Pradhan B. 2012. Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. Catena. 97:71–84.

Pradhan AMS, Kim YT. 2016. Evaluation of a combined spatial multi-criteria evaluation model and deterministic model for landslide susceptibility mapping. Catena. 140:125–139.

Regmi NR, Giardino JR, Vitek JD. 2010. Assessing susceptibility to landslides: using models to understand observed changes in slopes. Geomorphology. 122:25–38.

Remondo J, González A, Díaz De Terán JR, Cendrero A, Fabbri A, Chung CF. 2003. Validation of landslide susceptibility maps; examples and applications from a casestudy in Northern Spain. Nat Hazards. 30:437–449.

Schicker R, Moon V. 2012. Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. Geomorphology. 161–162:40–57.

Sezer EA, Pradhan B, Gokceoglu C. 2011. Manifestation of an adaptive neuro-fuzzy model on landslide susceptibility mapping: Klang valley, Malaysia. Expert Syst Appl. 38:8208–8219.

Shahabi H, Ahmad BB, Khezri S. 2013. Evaluation and comparison of bivariate and multivariate statistical methods for landslide susceptibility mapping (case study: Zab basin). Arabian J Geosci. 6:3885–3907.

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

Song KY, Oh HJ, Choi J, Park I, Lee C, Lee S. 2012. Prediction of landslides using ASTER imagery and data mining models. Adv Space Res. 49:978–993.

Stumpf A, Kerle N. 2011a. Combining Random Forests and object-oriented analysis for landslide mapping from very high resolution imagery. Procedia Environ Sci. 3:123–129.

Stumpf A, Kerle N. 2011b. Object-oriented mapping of landslides using Random Forests. Remote Sens Environ. 115:2564–2577.

Tangestani MH. 2009. A comparative study of Dempster-Shafer and fuzzy models for landslide susceptibility mapping using a GIS: An experience from Zagros Mountains, SW Iran. J Asian Earth Sci. 35:66–73.

Timilsina M, Bhandary NP, Dahal RK, Ryuichi Y. 2014. Distribution probability of large-scale landslides in central Nepal. Geomorphology. 226:236–248.

Umar Z, Pradhan B, Ahmad A, Jebur M N, Tehrany M S. 2014. Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. Catena. 118:124–135.

Van Western C J. 2002. Use of weights of evidence modeling for landslide susceptibility mapping. International Institute for Geoinformation Science and Earth Observation: Enschede, The Netherlands. p. 21.

Wang L J, Guo M, Sawada K, Lin J, Zhang J. 2015b. Landslide susceptibility mapping in Mizunami City, Japan: a comparison between logistic regression, bivariate statistical analysis and multivariate adaptive regression spline models. Catena. 135:271–282.

Wang Q, Wang D, Huang Y., Wang Z, Zhang L, Guo Q, Sang M. 2015a. Landslide susceptibility mapping based on selected optimal combination of landslide predisposing factors in a large catchment. Sustainability. 7:16653–16669.

Xie M, Tetsuro E, Qiu C, Jia L. 2007. Spatial three-dimensional landslide susceptibility mapping tool and its applications. Earth Sci Front. 14:73–84.

Xu C, Dai F, Xu X, Yuan HL. 2012. GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. Geomorphology. 145–146:70–80.

Yalcin A. 2008. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. Catena. 72:1–12.

Yao X, Tham LG, Dai FC. 2008. Landslide susceptibility mapping based on support vector machine: a case study on natural slopes of Hong Kong, China. Geomorphology. 101:572–582.

Zhang G, Cai Y, Zheng Z, Zhen J, Liu Y, Huang K. 2016. Integration of the statistical index method and the analytic hierarchy process technique for the assessment of landslide susceptibility in Huizhou, China. Catena. 142:233–244.