

Landslide susceptibility mapping based on GIS and support vector machine models for the Qianyang County, China

Wei Chen¹  · Huichan Chai² · Zhou Zhao¹ · Qiqing Wang² · Haoyuan Hong³ 

Received: 17 March 2015 / Accepted: 10 December 2015 / Published online: 10 March 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract The main purpose of this study is to produce reliable susceptibility maps using GIS-based support vector machine (SVM) models and compare their performances for the Qianyang County of Baoji City, Shaanxi Province, China. In this paper, with kernel classifiers of linear, polynomial, radial basis function and sigmoid, the four various types were applied in landslide susceptibility mapping. The important input parameters for the landslide susceptibility assessment were acquired from different sources. Firstly, 81 landslide sites were obtained by aerial photographs, earlier reports and field surveys. Then, the landslide inventory was randomly classified into two datasets: 70 % (56 landslides) for training the models and 30 % (25 landslides) for validation purpose. Secondly, 15 landslide conditioning factors were selected (i.e., slope angle, slope aspect, altitude, plan curvature, profile curvature, distance to faults, distance to rivers, distance to roads, NDVI, STI, SPI, TWI, geomorphology, rainfall, and lithology). Subsequently, with four types of kernel function classifiers based on landslide conditioning factors, landslide susceptibility parameters were obtained using SVM models. Finally, the rationality of landslide susceptibility maps was verified using the receiver operating characteristics with both success rate curve and

prediction rate curve. The validation results showed that success rates for the four SVM models were 83.15 % (RBF-SVM), 82.72 % (PL-SVM), 81.77 % (LN-SVM), and 79.99 % (SIG-SVM). The prediction rates for the four SVM models were 77.98 % (RBF-SVM), 77.50 % (PL-SVM), 77.07 % (LN-SVM), and 76.08 % (SIG-SVM), respectively. The results showed that the RBF-SVM model had the highest overall performance.

Keywords Landslide susceptibility mapping · support vector machine (SVM) · Geographic information system (GIS) · China

Introduction

Landslides, causing extensive damages to residential regions, economic losses, and human casualties all over the world, are one of the most dangerous geo-hazards in hill and mountain terrains because of the cliffy topography, irrational application of land cover and harmful climatic conditions for landslides (Solaimani et al. 2013; Sujatha et al. 2012; Akgun et al. 2008). Globally, landslides cause almost 1000 deaths per year and property damage of about 4 billion dollars (Lee and Pradhan 2007). There are frequent landslides in China, which often result in many casualties and great quality economic losses. 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>). It is therefore necessary to evaluate the factors that affect instability, study the hazard and forecast of the potential landslides to reduce the damages caused by landslides and evolve rational mitigation method (Sujatha et al. 2012).

✉ Wei Chen
chenwei.0930@163.com

Haoyuan Hong
hong_haoyuan@outlook.com

¹ School of Geology and Environment, Xi'an University of Science and Technology, Xi'an 710054, China

² School of Resources and Earth Science, China University of Mining and Technology, Xuzhou 221116, China

³ Jiangxi Provincial Meteorological Observatory, Jiangxi Meteorological Bureau, Nanchang 330046, China

Currently, there have been many GIS-based methods for assessing landslide susceptibility. Many studies have used probabilistic methods such as the frequency ratio and weight of evidence models (Solaimani et al. 2013; Lee and Pradhan 2007; Regmi et al. 2014; Pradhan et al. 2010; Choi et al. 2012; Vijith and Madhu 2008; Demir et al. 2013; Ozdemir and Altural 2013; Sujatha et al. 2014). Among statistical models, the bivariate and multivariate logistic regression models have also been used for landslide susceptibility mapping (Solaimani et al. 2013; Lee and Pradhan 2007; Akgun 2012; Bai et al. 2010; Ozdemir and Altural 2013; Pradhan and Lee 2010; Choi et al. 2012; Devkota et al. 2013; Yilmaz 2010b; Park et al. 2013; Raman and Punia 2012; Mihaela et al. 2011; Yalcin et al. 2011; Youssef et al. 2015b). In addition, some researchers (Kayastha et al. 2013; Kanungo et al. 2011; Pourghasemi et al. 2012a; Guettouche 2013; Akgun et al. 2012; Sharma et al. 2013; Pradhan 2010, 2011, 2013; Ercanoglu and Gokceoglu 2002, 2004; Oh and Pradhan 2011) have produced the landslide susceptibility maps using the deterministic models such as the analytic hierarchy process (AHP) and fuzzy models. Other new techniques such as fuzzy-logic, artificial neural network (ANN), and neuro-fuzzy models (Park et al. 2013; Pradhan and Buchroithner 2010; Pouydal et al. 2010; Chauhan et al. 2010; Sharma et al. 2013; Guettouche 2013; Choi et al. 2012; Vahidnia et al. 2010; Sezer et al. 2011) also have been used to evaluate landslide susceptibility. In order to determine the better model that is more accurate in landslide susceptibility mapping in a study area, some studies have used two or three models and compared their accuracy, such as probability and statistical analyses, probability and fuzzy-logic analyses, statistical and ANN analyses, analytic hierarchy process, probability and statistical analyses, and probability, statistical, and ANN analyses, etc. (Jaafari et al. 2014; Constantin et al. 2011; Kanungo et al. 2011; Ozdemir and Altural 2013; Park et al. 2013; Pourghasemi et al. 2013a; Pouydal et al. 2010; Solaimani et al. 2013; Yalcin et al. 2011; Youssef et al. 2015a; Demir et al. 2013; Lee and Pradhan 2007; Akgun 2012; Devkota et al. 2013).

The main purpose of this study is to assess the susceptibility of landslides for the Qianyang County of Baoji City, Shanxi Province, China, using a geographical information system (GIS). To achieve this aim, the support vector machine (SVM) with four different kernel functions were used to obtain the landslide susceptibility maps using the ArcGIS 10.0 software.

The SVM means that they have relatively seldom been used for landslide susceptibility mapping. Furthermore, the comparison of different kernel classifiers is rational as SVM could apply various types of kernel functions.

The study area

The study area is located in the City of Baoji, Shanxi Province, China, covering a surface area of about 996.46 km² between latitudes 106°56'15"–107°22'31"E and longitudes 34°34'34"–34°56'56"N. The altitude decreases from North to South and varies in the range from 752 to 1560 m. The climate of the study area is characterized by the warm semi-arid to semi-humid monsoon; the winter is dry and cold, but the summer is hot and rainy. Based on China Meteorological Administration, the temperature of this region varies between –20.6 °C in winter and 40.5 °C in summer with a yearly average of 11.8 °C. The mean relative humidity varies between 59 and 82 %. The mean annual rainfall is around 627.4 mm, and the rainy season is mainly from July to September with the total rainfall accounting to half of the yearly rainfall. Wei and Jing river systems are the main streams in this area and their tributaries shape dendritic drainage systems because of the topographical and geological trait of the area.

Data preparation

Landslide inventory map

Landslide inventory and mapping is the backbone of landslide susceptibility studies, which can determine the events affecting landslide development in the study region, and the terrain instability factors involved (Guinau et al. 2005; van Westen et al. 2006; Youssef et al. 2015a, b). The first step is to collect all of the available information and data concerning landslides in the area whose liability and accuracy affect the success of the used methodology (Ercanoglu and Gokceoglu 2004; Melchiorre et al. 2011). With collecting the data concerning landslides and study satellite imagery and aerial photographs combining with field surveys using a GPS device, landslide inventory maps can be acquired (Pradhan and Kim 2014). Finally, 81 landslides were acquired by assessing aerial photos of 1:50,000 scale coupled with field surveys in the study area and subsequently digitized for further analysis. Then, it was randomly divided into two parts (70/30), which were used as training and validating purposes, respectively (Fig. 1).

Landslide conditioning factors

With the purpose of applying the SVM model in the study area, 15 landslide conditioning factors, including slope angle, slope aspect, altitude, plan curvature, profile curvature, distance to faults, distance to rivers, distance to roads,

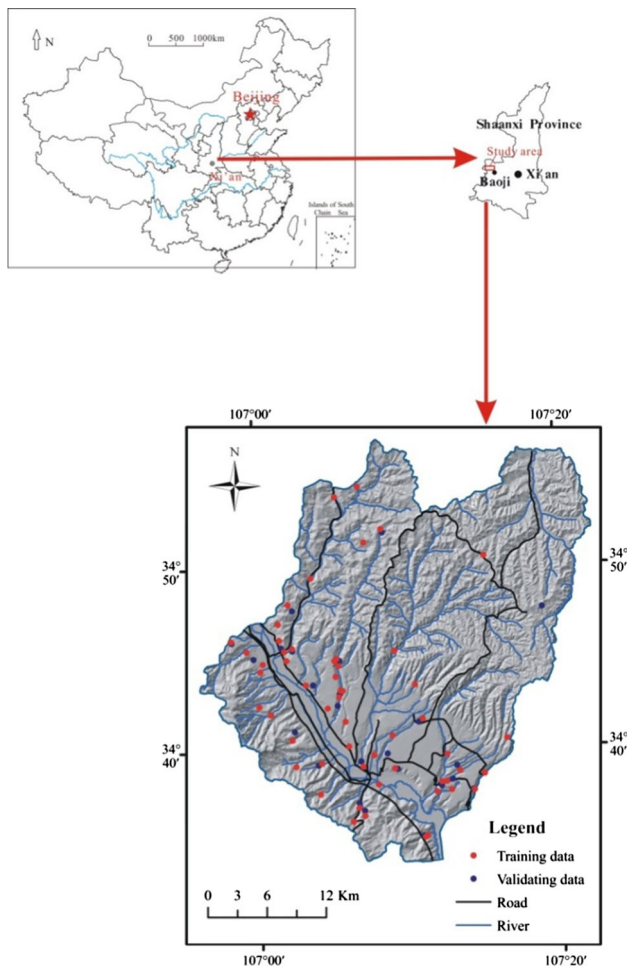


Fig. 1 Location of the study area

NDVI, STI, SPI, TWI, geomorphology, rainfall, and lithology were used. All of the data were converted into raster format with a pixel size of 50×50 m. These conditioning factors were classified into four groups: topographic factors (i.e., slope angle, slope aspect, altitude, plan curvature, profile curvature, STI, SPI, TWI), distance related factors (i.e., distance to roads, distance to rivers, and distance to faults), ground conditions (i.e., geomorphology, NDVI, and lithology) and triggering factors (i.e., rainfall) (Akgun et al. 2012; Melchiorre et al. 2011).

Topographic factors including slope angle, slope aspect, altitude, plan curvature, profile curvature, STI, SPI, TWI were mainly produced from the DEM of the study area. The slope angle, directly related to landslide incidence, is frequently applied in landslide susceptibility studies (He et al. 2012; Dai and Lee 2001). Slope angles in the study area ranged from 0° to 38° , and were reclassified into five classes, i.e., 0° – 7° , 7° – 14° , 14° – 21° , 21° – 28° , and 28° – 38° (Fig. 2a). Aspect, describing the direction of slope, is also an important factor for landslide susceptibility analysis, as aspect controls the formation of the landslide such as

lineaments, rainfalls, wind effects, and exposure to sunshine (Yalcin and Bulut 2007; Pourghasemi et al. 2012a; He et al. 2012). Aspect in the study area was classified into nine categories: flat (-1), north (337.5° – 360° , 0° – 22.5°), northeast (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°) (Fig. 2b). Altitude or elevation, controlled by several geologic and geomorphologic processes, is also frequently used in landslide susceptibility mapping (Pourghasemi et al. 2012b; Pradhan and Kim 2014). Elevation values in the study area ranged from 720 to 1560 m, and five categories of elevations were identified, i.e., 720–850, 850–1000, 1000–1150, 1150–1300, and 1300–1560 m (Fig. 2c). The plan curvature influences the convergence and divergence of flow across a surface. The profile curvature, the vertical plane parallel to the slope direction, affects the acceleration and deceleration of down slope flows, and as a result, influences erosion and deposition (He et al. 2012; Kritikos and Davies 2014; Kannan et al. 2013). In this study, plan curvature and profile curvature were calculated in GIS software of Arc GIS 10.0; and they were divided into three classes: <-0.05 , -0.05 to 0.05 , and >0.05 , respectively (Fig. 2d, e). The sediment transport index (STI) reflects the process of erosion and deposition (Devkota et al. 2013). In the study, STI was classified into four classes: <3 , 3–9, 9–15, and >15 (Fig. 2f). The stream power index (SPI), describing erosion capability of water flow, is also considered as a factor influencing the stability in the study region (Regmi et al. 2014; Conforti et al. 2011). The SPI map was grouped into four different classes: <5 , 5–10, 10–40, and >40 (Fig. 2g). The topographic wetness index (TWI) describes the effect of topography on the location and size of saturated source areas of runoff generation, and was considered as another contributing factor (Pourghasemi et al. 2013b; Pradhan and Kim 2014). The TWI values of this area were arranged in four classes: <7 , 7–10, 10–13, and >13 , respectively (Fig. 2h).

Faults are responsible for triggering a large number of landslides due to the tectonic breaks that usually decrease the rock strength (Devkota et al. 2013). Therefore, the distance to faults was also a necessary parameter in the susceptibility analysis. In the study area, the distance to faults map was reclassified into five divisions, such as 0–2000, 2000–4000, 4000–6000, 6000–8000, and >8000 m, respectively (Fig. 2i).

The distance to rivers, controlling the stability of a slope, is another important factor for landslide susceptibility analysis. On the basis of rivers and streams, a map of proximity to drainage was generated using ArcGIS 10.0 and was divided into five categories, such as 0–200, 200–400, 400–600, 600–800, and >800 m (Fig. 4j).

Fig. 2 Landslide conditioning factors of the study area: **a** slope angle, **b** slope aspect, **c** elevation, **d** plan curvature, **e** profile curvature, **f** STI, **g** SPI, **h** TWI, **i** distance to faults, **j** distance to rivers, **k** distance to roads, **l** geomorphology, **m** NDVI, **n** rainfall, and **o** lithology

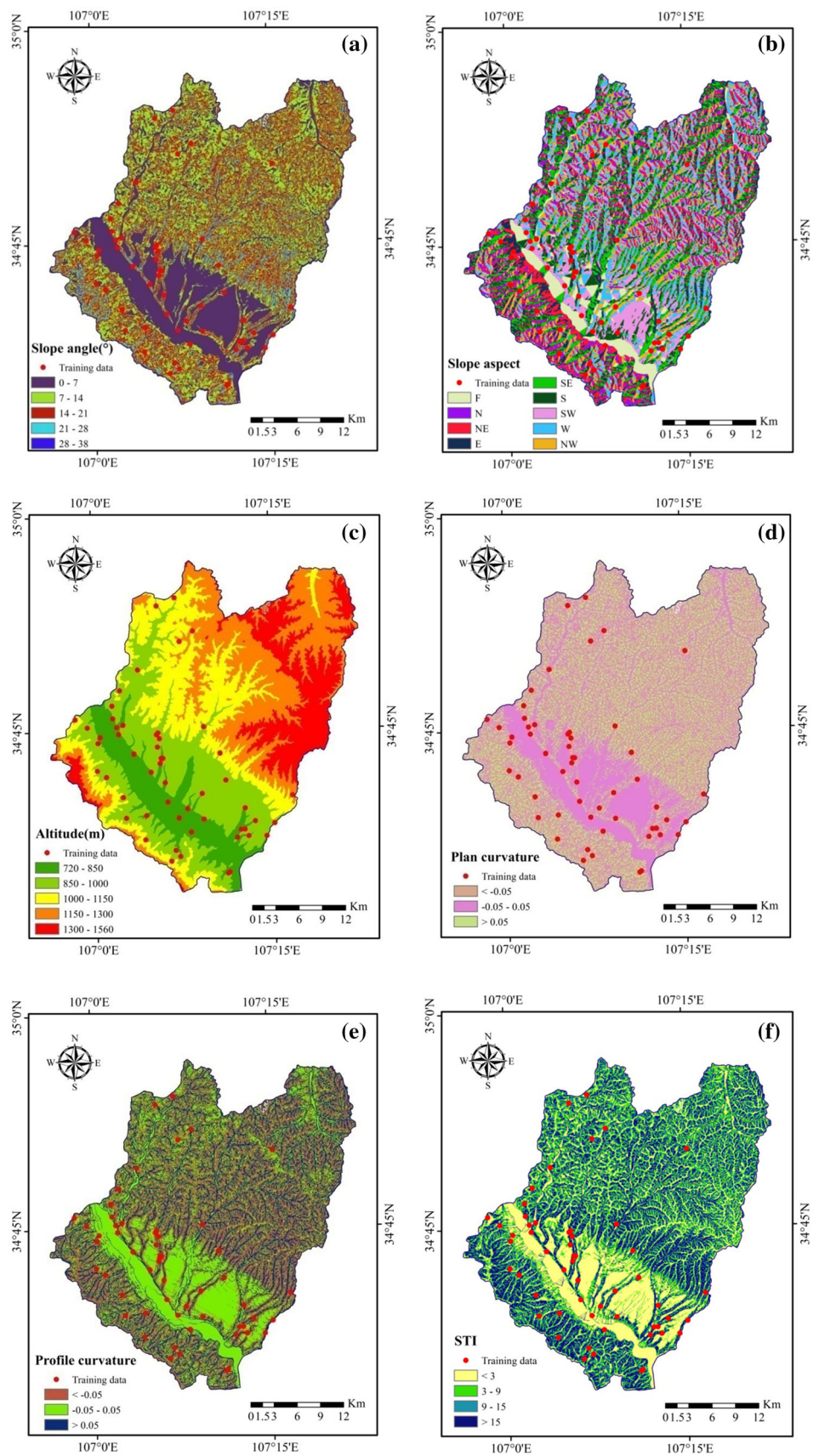


Fig. 2 continued

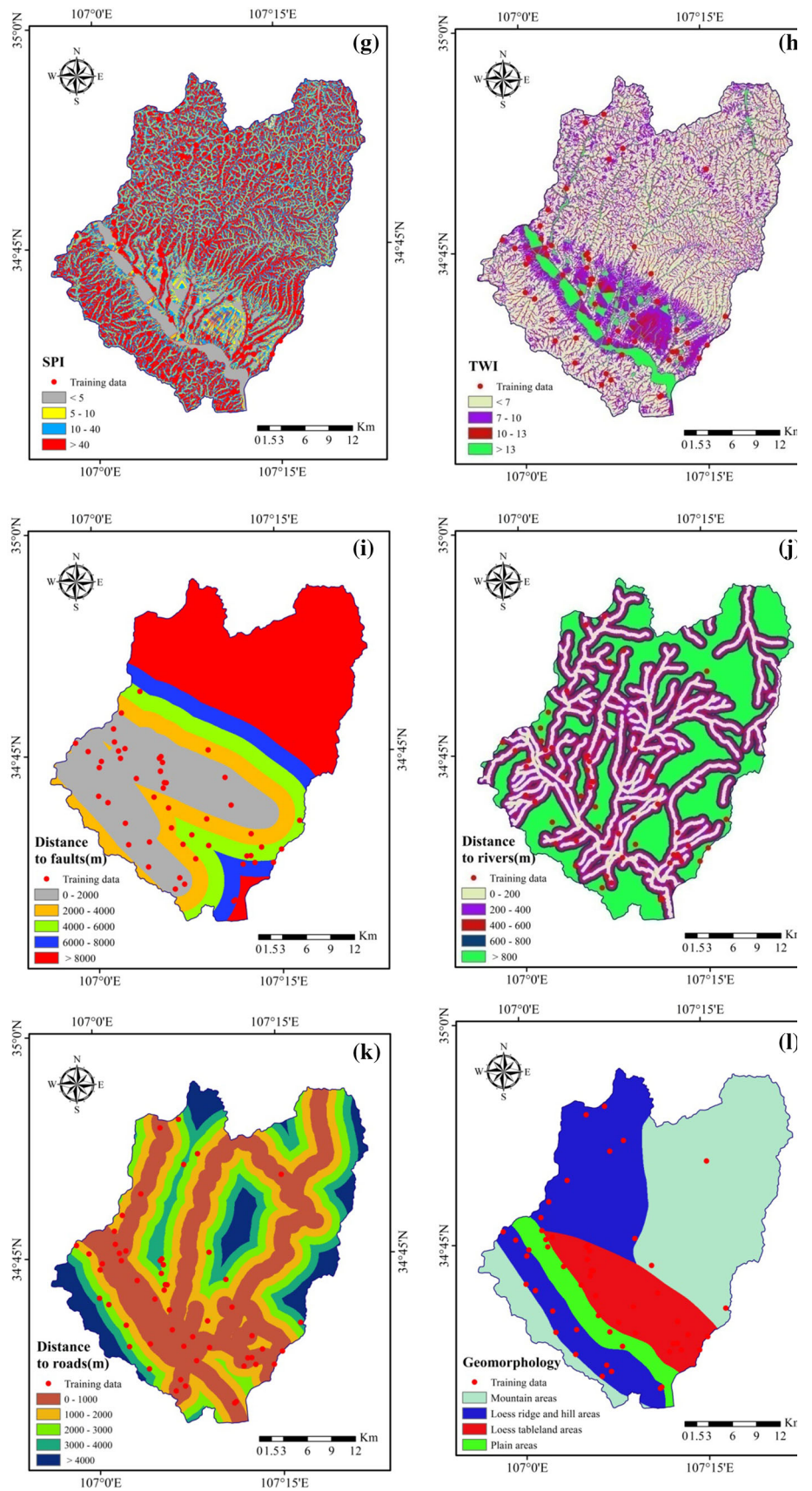
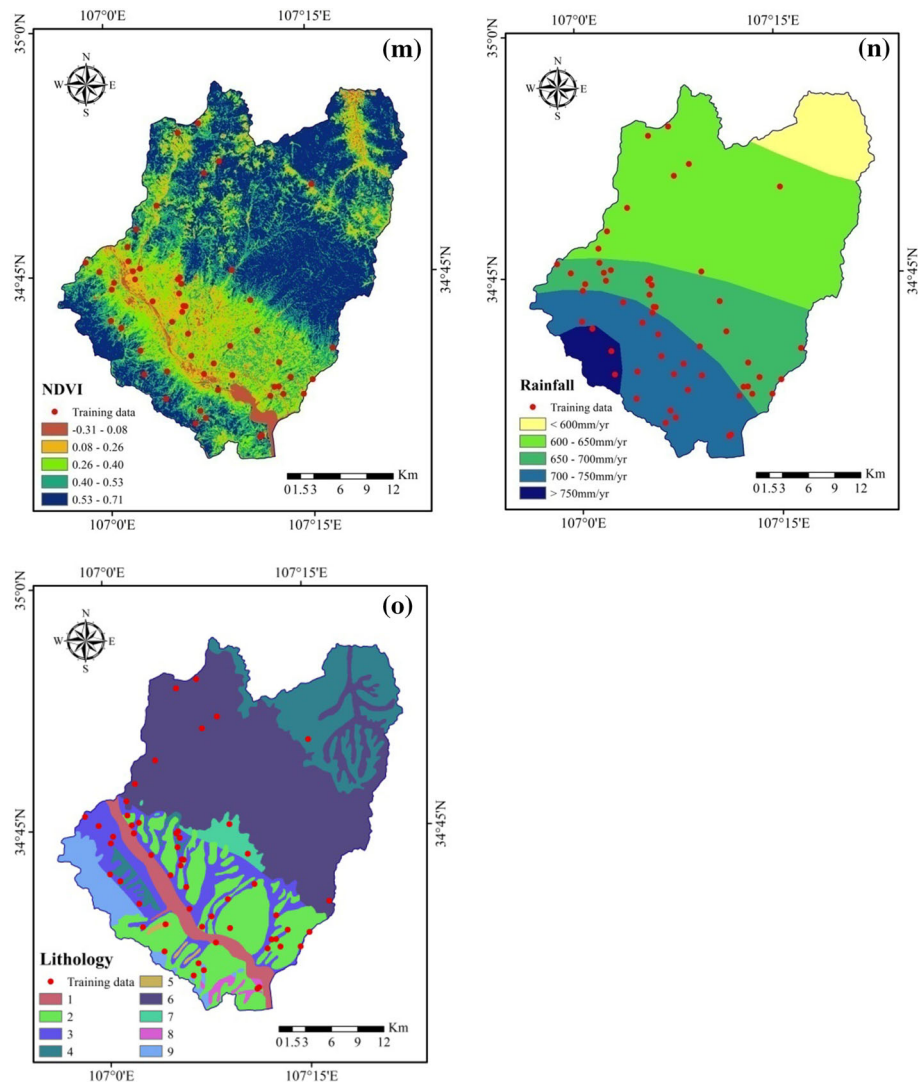


Fig. 2 continued



The distance to roads is an important anthropogenic factor influencing landslides occurrence. In the present study, the distance to roads was calculated and reclassified the resultant map into five classes: 0–1000, 1000–2000, 2000–3000, 3000–4000, and >4000 m, respectively (Fig. 2k).

Geomorphology is an important factor which is closely related to landslide occurrence (Kannan et al. 2013). Four geomorphologic units of the study area can be identified, i.e., mountain areas, loess ridge and hill areas, loess tableland areas and plain areas (Fig. 2l).

The NDVI is also considered as a conditioning factor related to landslide occurrence. In general, the higher the value of NDVI is, the larger the area that is covered by vegetation (He et al. 2012). In this study, the NDVI map was obtained from Landsat satellite image and reclassified into five classes, i.e., –0.31 to 0.08, 0.08 to 0.26, 0.26 to 0.40, 0.40 to 0.53, and 0.53 to 0.71, respectively (Fig. 2m).

The rainfall, closely associated with landslide initiation, is one of the main parameters in landslide susceptibility mapping. The annual rainfall of the study area was classified into five classes: <600, 600–650, 650–700, 700–750, and >750 mm/year, respectively (Fig. 2n).

Lithology is one of the most common determinant factors in most landslide stability studies. The geological map of the study area is compiled from existing geological maps and publications in Arc GIS 10.0. The lithological units of the study area are shown in Table 1, and the general geological setting of the area is shown on the source map (Fig. 2o).

Support vector machines

Support vector machine (SVM), a supervised learning method, is established on the basis of statistical learning theory. With the purpose to search an optimal separating

Table 1 Description of geological units of the study area

No.	Code	Formation	Lithology	Geological age
1	Q ₄	–	Sand, gravel, silty clay	Quaternary
2	Q ₃	–	Loess	Quaternary
3	Q ₂	–	Loess	Quaternary
4	Q ₁	–	Loess	Quaternary
5	N ₂	–	Glutenite, sandstone, siltstone	Neogene
6	K _{1Lh}	Luohandong	Sandstone, siltstone, mudstone	Early Cretaceous
	K _{1h}	Huanhehuachi	Sandstone, siltstone, mudstone	Early Cretaceous
	K _{1L}	Luohe	Glutenite, sandstone	Early Cretaceous
	K _{1y}	Yijun	Glutenite	Early Cretaceous
7	J _{3f}	Fenfanghe	Glutenite, sandstone, siltstone	Late Jurassic
	J _{2z}	Zhiluo	Glutenite, siltstone, shale	Middle Jurassic
8	P ₁	–	Sandstone, shale	Early Permian
	O _{1–2}	–	Limestone, dolomite	Early–Middle Ordovician

hyperplane, the theory changed original import space into a dimensional feature space (Vapnick 1998; Xu et al. 2012; Bui et al. 2015).

For example, consider a training dataset of instance-label pairs (x_i, y_i) , where $i = 1, 2, \dots, n$, x_i is an input vector that includes 15 landslide conditioning factors, $y_i \in \{1, -1\}$ is its corresponding two output classes, i.e., landslide and non-landslide, n is the number of training samples. The aim of SVM is to search an n -dimensional hyperplane differentiating the two types by their maximum gap. Its mathematical expression is as follows (Yao et al. 2008; Xu et al. 2012; Tehrani et al. 2015):

$$1/2\|w\|^2 \tag{1}$$

$$y_i((w \times x_i) + b) \geq 1 \tag{2}$$

where $\|w\|$ is the norm of the normal of the hyperplane, b is a constant. Introducing the Lagrangian multiplier (λ_i), the cost function can be defined as:

$$L = 1/2\|w\|^2 - \sum_{i=1}^n \lambda_i(y_i((w \times x_i) + b) - 1) \tag{3}$$

For non-separable case, introducing slack variables ξ_i (Vapnik 1995), Eq. (2) can be modified as:

$$y_i((w \times x_i) + b) \geq 1 - \xi_i \tag{4}$$

then, introducing $\nu(0, 1)$ to express misclassification (Schölkopf et al. 2000; Wu et al. 2014), Eq. (1) can be defined as:

$$L = \frac{1}{2}\|w\|^2 - \frac{1}{\nu n} \sum_{i=1}^n \xi_i \tag{5}$$

Besides that, a kernel function $K(x_i, x_j)$ is applied to account for nonlinear decision boundary (Vapnik 1995). In the study, the following four types of kernel function were

applied to examine the efficiency of each kernel function in landslide susceptibility mapping (Xu et al. 2012; Pourghasemi et al. 2013b):

Linear: $K(x_i, x_j) = x_i^T \times x_j$ (6)

Polynomial: $K(x_i, x_j) = (\gamma \times x_i^T \times x_j + r)^d, \gamma > 0$ (7)

Sigmoid: $K(x_i, x_j) = \tanh(\gamma \times x_i^T \times x_j + r)$ (8)

Radial basis function: $K(x_i, x_j) = (-\gamma(x_i - x_j))$, $\gamma > 0$ (9)

where d , r , and γ are parameters of the kernel functions (Pourghasemi et al. 2013b).

Results and discussion

The results of spatial relationship between landslides and conditioning factors using frequency ratio are shown in Table 2. In Table 2, landslides were most abundant in the class 14°–21°, indicating the highest probability of landslide occurrence in this group, followed by slope category 21°–28°. For the slope aspect, the frequency ratio was highest for north-facing slopes (FR value of 1.67) and lowest for flat slopes (0.0). For the elevation, the frequency ratio was highest for the class 720–850 m. In the case of plan curvature, the frequency ratio was 1.16 for class –0.05 to 0.05, indicating a very high probability of landslide occurrence. Similarly, for the profile curvature in the class >0.05, the frequency ratio was 1.44, which indicates a high probability of landslide occurrence. In the case of STI, SPI and TWI most of the landslides occurred in the class >15, >40 and 10–13, respectively. The relationship between landslides and their distance to faults, rivers and roads shows that when distance to a fault, river or road line increases, the probability of landslide occurrence

Table 2 Spatial relationship between each landslide conditioning factor and landslide using frequency ratio model

Factors	Class	No. of landslide	Percentage of landslide	No. of pixels in domain	Percentage of domain	Frequency ratio
Slope angle (°)	0 to 7	17	30.36	130,540	32.51	0.93
	7 to 14	19	33.93	142,908	35.59	0.95
	14 to 21	17	30.36	105,127	26.18	1.16
	21 to 28	3	5.36	21,767	5.42	0.99
	28 to 38	0	0.00	1191	0.30	0.00
Slope aspect	Flat	0	0.00	24,750	6.16	0.00
	N	9	16.07	38,630	9.62	1.67
	NE	7	12.50	41,060	10.23	1.22
	E	8	14.29	45,023	11.21	1.27
	SE	6	10.71	53,668	13.37	0.80
	S	4	7.14	39,222	9.77	0.73
	SW	9	16.07	52,826	13.16	1.22
	W	3	5.36	53,388	13.30	0.40
	NW	10	17.86	52,966	13.19	1.35
Elevation (m)	720 to 850	23	41.07	47,992	11.95	3.44
	850 to 1000	26	46.43	100,242	24.96	1.86
	1000 to 1150	4	7.14	91,991	22.91	0.31
	1150 to 1300	2	3.57	102,322	25.48	0.14
	1300 to 1560	1	1.79	58,986	14.69	0.12
Plan curvature	<-0.05	16	28.57	109,181	27.19	1.05
	-0.05 to 0.05	28	50.00	172,650	43	1.16
	>0.05	12	21.43	119,702	29.81	0.72
Profile curvature	<-0.05	15	26.79	122,427	30.49	0.88
	-0.05 to 0.05	15	26.79	149,516	37.24	0.72
	>0.05	26	46.43	129,590	32.27	1.44
STI	<3	7	12.50	98,781	24.60	0.51
	3 to 9	16	28.57	107,663	26.81	1.07
	9 to 15	9	16.07	80,233	19.98	0.80
	>15	24	42.86	114,856	28.60	1.50
TWI	<7	31	55.36	220,876	55.01	1.01
	7 to 10	15	26.79	120,942	30.12	0.89
	10 to 13	6	10.71	25,225	6.28	1.71
	>13	4	7.14	34,490	8.59	0.83
SPI	<5	3	5.36	59,426	14.80	0.36
	5 to 10	3	5.36	43,551	10.85	0.49
	10 to 40	18	32.14	138,723	34.55	0.93
	>40	32	57.14	159,833	39.81	1.44
Distance to faults (m)	0 to 2000	29	51.79	107,371	26.72	1.94
	2000 to 4000	9	16.07	56,616	14.09	1.14
	4000 to 6000	8	14.29	43,323	10.78	1.33
	6000 to 8000	5	8.93	31,701	7.89	1.13
	>8000	5	8.93	162,839	40.52	0.22
Distance to rivers (m)	0 to 200	26	46.43	89,465	22.26	2.09
	200 to 400	8	14.29	72,371	18.01	0.79
	400 to 600	5	8.93	56,804	14.14	0.63
	600 to 800	4	7.14	45,418	11.30	0.63
	>800	13	23.21	137,792	34.29	0.68

Table 2 continued

Factors	Class	No. of landslide	Percentage of landslide	No. of pixels in domain	Percentage of domain	Frequency ratio
Distance to roads (m)	0 to 1000	30	53.57	147,594	36.73	1.46
	1000 to 2000	8	14.29	96,043	23.90	0.60
	2000 to 3000	13	23.21	71,336	17.75	1.31
	3000 to 4000	5	8.93	47,452	11.81	0.76
	>4000	0	0.00	39,425	9.81	0.00
Geomorphology	Mountain areas	3	5.36	180,810	45.00	0.12
	Loess ridge and hill areas	23	41.07	129,203	32.15	1.28
	Loess tableland areas	22	39.29	62,630	15.59	2.52
	Plain areas	8	14.29	29,173	7.26	1.97
NDVI	-0.31 to 0.08	1	1.79	10,985	2.73	0.65
	0.08 to 0.26	13	23.21	44,999	11.20	2.07
	0.26 to 0.40	21	37.50	81,702	20.34	1.84
	0.40 to 0.53	15	26.79	107,741	26.82	1.00
	0.53 to 0.71	6	10.71	156,351	38.91	0.28
Rainfall	<600 mm/year	0	0.00	33,234	8.27	0.00
	600 to 650 mm/year	9	16.07	184,811	45.99	0.35
	650 to 700 mm/year	24	42.86	84,360	20.99	2.04
	700 to 750 mm/year	19	33.93	85,366	21.25	1.60
	>750 mm/year	4	7.14	14,045	3.50	2.04
Lithology	1	2	3.57	17,667	4.41	0.81
	2	17	30.36	68,524	17.11	1.77
	3	22	39.29	44,002	10.99	3.58
	4	2	3.57	57,235	14.29	0.25
	5	1	1.79	1260	0.31	5.68
	6	7	12.50	182,987	45.68	0.27
	7	2	3.57	8518	2.13	1.68
	8	1	1.79	2853	0.71	2.51
	9	2	3.57	17,509	4.37	0.82

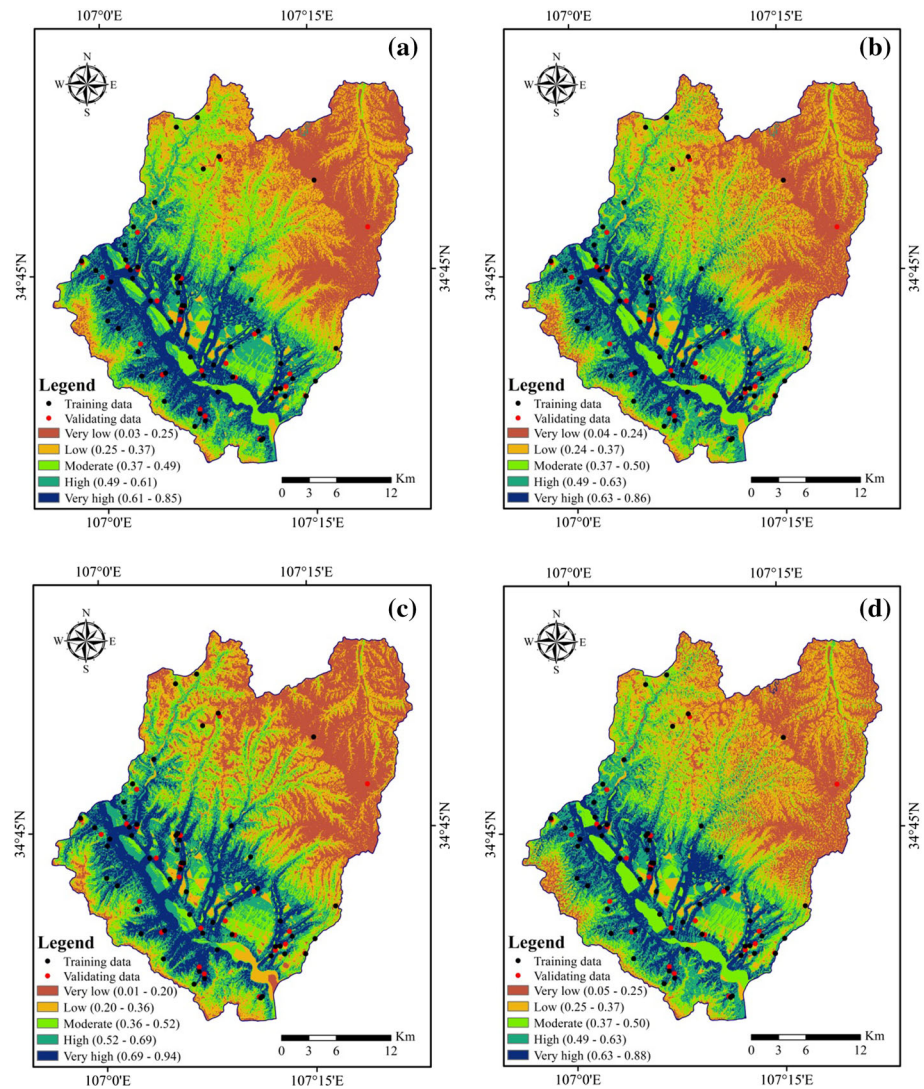
decreases. The frequency ratio between landslide occurrence and geomorphology showed that the loess table land areas had the highest value 2.52 and mountain areas had the lowest value (0.12). The frequency ratio for the NDVI was high between 0.08 and 0.26, which indicates a very high probability of landslide occurrence. As shown in Table 1, it can be observed that as rainfall increases, the landslide frequency generally increases.

In this study, the SVM model with four types of kernel classifiers such as linear (LN), polynomial degree of 2 (PL), sigmoid (SIG), and radial basis function (RBF) were trained using the ENVI5.1 software. The probability of landslide occurrence falls in the range between 0 and 1. The results were then exported into the ArcGIS 10.0 software for visualization. Finally, the LSI of the produced maps was grouped into five classes (very low, low, moderate, high, and very high) using the natural break method. The four landslide susceptibility maps are shown in Fig. 3.

Validation and comparison of susceptibility maps

Validation is an absolutely essential component in the development of landslide susceptibility and determination of its quality (Pourghasemi et al. 2012c). Landslide susceptibility maps are meaningless without validation (Chung and Fabbri 2003). In the study, the receiver operating characteristics (ROC) curve was used to assess the overall performance of the four used models, because the ROC curve is helpful for representing the quality of deterministic and probabilistic forecast systems (Akgun et al. 2012; Youssef et al. 2015a, b). The ROC curve plots the false positive rate on the X-axis and true positive rate on the Y-axis, which shows the trade-off between the two rates (Pradhan 2013). The area under the curve (AUC) represents the quality of the probabilistic model to reliably predict the occurrence or non-occurrence of landslides (Youssef et al. 2015a, b). The success rate was obtained

Fig. 3 Landslide susceptibility maps: **a** the LN-SVM model; **b** the PL-SVM model; **c** the SIG-SVM model; **d** the RBF-SVM model



using the training dataset. As shown in Fig. 4a, the RBF-SVM model represented the highest value of success rate (83.15 %), followed by the PL-SVM model (82.72 %), LN-SVM model (81.77 %) and the SIG-SVM (79.99 %).

The prediction capability of the four landslide susceptibility maps was obtained using the validation dataset. The result is shown in Fig. 4b. It can be observed that the RBF-SVM model had the highest prediction rate (77.98 %). Moreover, the prediction rates were 77.07, 77.50 and 76.08 % for LN-SVM model, PL-SVM model, and SIG-SVM model, respectively.

Discussion and conclusions

The preparation of landslide susceptibility maps is a crucial step that can help planners, local administrations, and decision makers in disaster planning. Accuracy of the landslide susceptibility maps is important for reducing the

losses of life and property (Kavzoglu et al. 2014). Landslide susceptibility can be assessed using different methods and many research papers were published in order to solve the deficiencies and difficulties in the landslide susceptibility mapping (Yilmaz 2010a). The main objective of this research is to produce landslide susceptibility maps for the Qianyang County, China, using SVM based on four types of kernel classifiers such as linear, polynomial, sigmoid and radial basis function.

As the first step, a reliable landslide inventory map is necessary for landslide susceptibility mapping. In the study, 70 % of landslides were used for training the models and the others were used for validation purpose.

Secondly, 15 landslide conditioning factors such as slope angle, slope aspect, altitude, plan curvature, profile curvature, distance to faults, distance to rivers, distance to roads, NDVI, STI, SPI, TWI, geomorphology, rainfall, and lithology were constructed and used for producing landslide susceptibility maps.

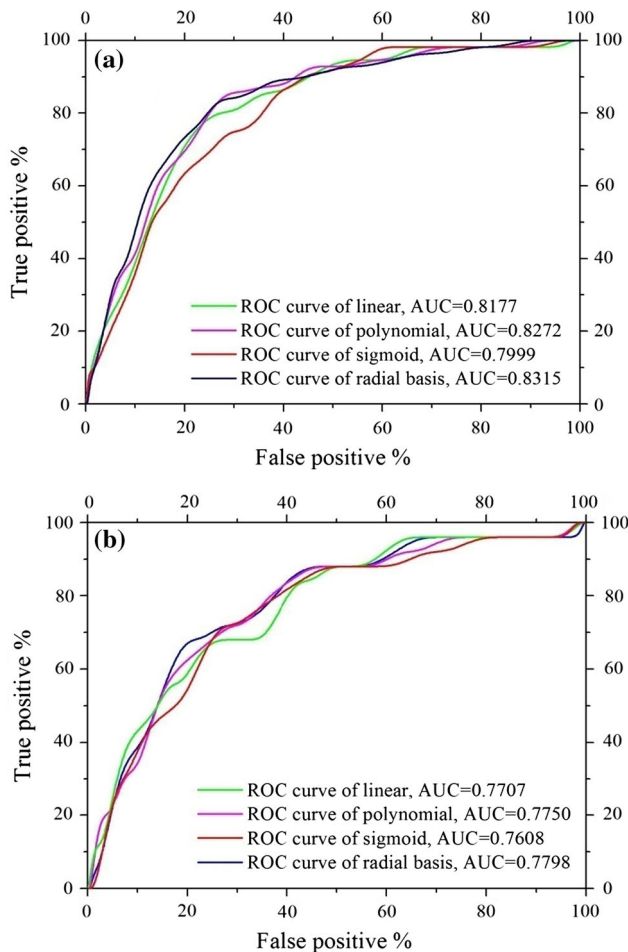


Fig. 4 **a** Success rate and **b** prediction rate for the LN-SVM, the PL-SVM, the SIG-SVM, and the RBF-SVM models

Finally, five landslide susceptibility classes, i.e., very low, low, moderate, high, and very high susceptible for landsliding, were derived with natural break method. The spatial performances of the obtained landslide susceptibility maps were compared using ROC curves.

The validation results showed that the landslide susceptibility map generated by RBF-SVM model had the highest prediction rate (77.98 %), followed by the PL-SVM model (77.50 %), the LN-SVM (77.07 %), and the SIG-SVM (76.08 %). Success rate curves gave similar results, with RBF-SVM model the highest AUC value (83.15 %), followed by the PL-SVM model (82.72 %), the LN-SVM model (81.77 %), and the SIG-SVM model (79.99 %).

SVM model has been used in many literatures. Brenning (2005) obtained sufficiently smooth prediction surfaces for creating susceptibility map by using SVM. Yao et al. (2008) used the SVM in landslide susceptibility mapping, they found that SVM was a useful tool in landslide susceptibility assessment, and they found that the SVM had

better prediction efficiency than LR. Marjanović et al. (2011) commented on the strengths and weaknesses of the SVM model, and indicated that SVM models do not need any feature selection technique as opposed to some other methods such as decision trees. Xu et al. (2012) found that the radial basis and polynomial kernel functions were suitable for modeling any input training data. San (2014) used SVM to generate medium scale landslide susceptibility maps; they also found that SVM presented high classification accuracy.

The results of the present study show that the SVM, based on four types of kernel classifiers, have been applied successfully to the production of landslide susceptibility maps. The landslide susceptibility maps provide valuable information on the slope stability in the study area, which could be of benefit to infrastructure planning, land use, engineering and hazard mitigation design.

Acknowledgments The authors want to express their gratitude to anonymous reviewers and editors for their valuable comments which were very useful in bringing the manuscript into the present form. The study is supported by the National Science Foundation of China (Grant No. 41302276) and Natural Science Foundation of Shaanxi Province (Grant No. 2012JM5008).

References

Akgun A (2012) A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at İzmir, Turkey. *Landslides* 9:93–106

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

Akgun A, Sezer EA, Nefeslioglu HA, Gokceoglu C, Pradhan B (2012) An easy-to-use MATLAB program (MamLand) for the assessment of landslide susceptibility using a Mamdani fuzzy algorithm. *Comput Geosci* 38(1):23–34

Bai SB, Wang J, Lü GN, 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(1):23–31

Brenning A (2005) Spatial prediction models for landslide hazards: review, comparison and evaluation. *Nat Hazards Earth Syst Sci* 5(6):853–862

Bui DT, Tuan TA, Klempe H, Pradhan B, Revhaug I (2015) Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides*. doi:10.1007/s10346-015-0557-6

Chauhan S, Sharma M, Arora M, Gupta N (2010) Landslide susceptibility zonation through ratings derived from artificial neural network. *Int J Appl Earth Obs Geoinf* 12:340–350

Choi J, Oh HJ, Lee HJ, Lee C, Lee S (2012) Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS. *Eng Geol* 124:12–23

Chung CJF, Fabbri AG (2003) Validation of spatial prediction models for landslide hazard mapping. *Nat Hazards* 30(3):451–472

- Conforti M, Aucelli PP, Robustelli G, Scarciglia F (2011) Geomorphology and GIS analysis for mapping gully erosion susceptibility in the Turbolo stream catchment (Northern Calabria, Italy). *Nat Hazards* 56(3):881–898
- Constantin M, Bednarik M, Jurchescu MC, Vlaicu M (2011) Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). *Environ Earth Sci* 63(2):397–406
- Dai FC, Lee CF (2001) Terrain-based mapping of landslide susceptibility using a geographical information system: a case study. *Can Geotech J* 38(5):911–923
- 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, Dhital MR, 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
- Ercanoglu M, Gokceoglu C (2002) Assessment of landslide susceptibility for a landslide-prone area (north of Yenice, NW Turkey) by fuzzy approach. *Environ Geol* 41:720–730
- Ercanoglu M, Gokceoglu C (2004) Use of fuzzy relations to produce landslide susceptibility map of a landslide prone area (West Black Sea Region, Turkey). *Eng Geol* 75(3):229–250
- Guettouche MS (2013) Modeling and risk assessment of landslides using fuzzy logic. Application on the slopes of the Algerian Tell (Algeria). *Arab J Geosci* 6:3163–3173
- Guinau M, Pallás R, Vilaplana JM (2005) A feasible methodology for landslide susceptibility assessment in developing countries: a case-study of NW Nicaragua after Hurricane Mitch. *Eng Geol* 80(3):316–327
- He S, Pan P, Dai L, Wang H, Liu J (2012) Application of kernel-based Fisher discriminant analysis to map landslide susceptibility in the Qingan River delta, Three Gorges, China. *Geomorphology* 171:30–41
- 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, Anabalagan 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
- Kanungo DP, Sarkar S, Sharma S (2011) Combining neural network with fuzzy, certainty factor and likelihood ratio concepts for spatial prediction of landslides. *Nat Hazards* 59(3):1491–1512
- 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
- Kayastha P, Dhital MR, De Smedt F (2013) Application of analytical hierarchy process (AHP) for landslide susceptibility mapping: a case study from the Tinau watershed, west Nepal. *Comput Geosci* 52:398–408
- Kritikos T, Davies T (2014) Assessment of rainfall-generated shallow landslide/debris-flow susceptibility and runoff using a GIS-based approach: application to western Southern Alps of New Zealand. *Landslides*. doi:10.1007/s10346-014-0533-6
- Lee S, Pradhan B (2007) Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* 4(1):33–41
- Marjanović M, Kovačević M, Bajat B, Voženilek V (2011) Landslide susceptibility assessment using SVM machine learning algorithm. *Eng Geol* 123:225–234
- Melchiorre C, Abella EC, van Westen CJ, Matteucci M (2011) Evaluation of prediction capability, robustness, and sensitivity in non-linear landslide susceptibility models, Guantánamo, Cuba. *Comput Geosci* 37(4):410–425
- Mihaela C, Martin B, Marta CJ, Marius V (2011) Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). *Environ Earth Sci* 63:397–406
- Oh HJ, Pradhan B (2011) Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. *Comput Geosci* 37(9):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
- Park S, Choi C, Kim B, Kim J (2013) Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea. *Environ Earth Sci* 68:1443–1464
- Pourghasemi HR, Pradhan B, Gokceoglu C (2012a) Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Nat Hazards* 63:965–996
- Pourghasemi HR, Pradhan B, Gokceoglu C, Moezzi KD (2012b) Landslide susceptibility mapping using a spatial multicriteria evaluation model at Haraz Watershed, Iran. *Terrigenous mass movements*. Springer, Berlin Heidelberg, pp 23–49
- Pourghasemi HR, Mohammady M, Pradhan B (2012c) 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, Mohammadi M, Moradi HR (2013a) Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran. *Arab J Geosci* 6(7):2351–2365
- Pourghasemi HR, Jirandeh AG, Pradhan B, Xu C, Gokceoglu C (2013b) Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran. *J Earth Syst Sci* 122(2):349–369
- Pouydl CP, Chang C, Oh HJ, Lee S (2010) Landslide susceptibility maps comparing frequency ratio and artificial neural networks: a case study from the Nepal Himalaya. *Environ Earth Sci* 61:1049–1064
- 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 B (2011) Manifestation of an advanced fuzzy logic model coupled with geoinformation techniques coupled with geoinformation techniques for landslide susceptibility analysis. *Environ Ecol Stat* 18(3):471–493
- Pradhan B (2013) A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput Geosci* 51:350–365
- Pradhan B, Buchroithner MF (2010) Comparison and validation of landslide susceptibility maps using an artificial neural network model for three test areas in Malaysia. *Environ Eng Geosci* 16(2):107–126
- 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: back-propagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modeling. *Environ Modell Softw* 25(6):747–759
- Pradhan B, Oh HJ, Buchroithner M (2010) Weights-of-evidence model applied to landslide susceptibility mapping in a tropical hilly area. *Geomat Nat Hazards Risk* 1(3):199–223
- Raman R, Punia M (2012) The application of GIS-based bivariate statistical methods for landslide hazards assessment in the upper Tons river valley, Western Himalaya, India. *Georisk: Assess Manag Risk Eng Syst Geohazards* 6(3):145–161
- 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
- San BT (2014) An evaluation of SVM using polygon-based random sampling in landslide susceptibility mapping: The Candir catchment area (western Antalya, Turkey). *Int J Appl Earth Obs Geoinf* 26:399–412
- Schölkopf B, Smola AJ, Williamson RC, Bartlett PL (2000) New support vector algorithms. *Neural Comput* 12(5):1207–1245
- 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
- Sharma LP, Patel Nilanchal, Ghose MK, Debnath P (2013) Synergistic application of fuzzy logic and geo-informatics for landslide vulnerability zonation—a case study in Sikkim Himalayas, India. *Appl Geomat* 5:271–284
- 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
- Sujatha ER, Rajamanickam GV, Kumaravel P (2012) Landslide susceptibility analysis using probabilistic certainty factor approach: a case study on Tevankarai stream watershed, India. *J Earth Syst Sci* 121(5):1337–1350
- Sujatha ER, Kumaravel P, Rajamanickam GV (2014) Assessing landslide susceptibility using Bayesian probability-based weight of evidence model. *Bull Eng Geol Environ* 73(1):147–161
- Tehrany MS, Pradhan B, Mansor S, Ahmad N (2015) Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. *Catena* 125:91–101
- Vahidnia MH, Alesheikh AA, Alimohammadi A, Hosseinali F (2010) A GIS-based neuro-fuzzy procedure for integrating knowledge and data in landslide susceptibility mapping. *Comput Geosci* 36(9):1101–1114
- 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
- Vapnick VN (1998) *Statistical learning theory*. Wiley, New York
- Vapnik VN (1995) *The nature of statistical learning theory*. Springer, New York
- Vijith H, Madhu G (2008) Estimating potential landslide sites of an upland sub-watershed in Western Ghat's of Kerala (India) through frequency ratio and GIS. *Environ Geol* 55(7):1397–1405
- Wu X, Ren F, Niu R (2014) Landslide susceptibility assessment using object mapping units, decision tree, and support vector machine models in the Three Gorges of China. *Environ Earth Sci* 71(11):4725–4738
- Xu C, Dai F, Xu X, Lee YH (2012) GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. *Geomorphology* 145:70–80
- Yalcin A, Bulut F (2007) Landslide susceptibility mapping using GIS and digital photogrammetric techniques: a case study from Ardesen (NE-Turkey). *Nat Hazards* 41(1):201–226
- Yalcin A, Reis S, Cagdasoglu A, 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* 85:274–287
- 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(4):572–582
- Yilmaz I (2010a) Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. *Environ Earth Sci* 61(4):821–836
- Yilmaz I (2010b) The effect of the sampling strategies on the landslide susceptibility mapping by conditional probability and artificial neural networks. *Environ Earth Sci* 60(3):505–519
- Youssef AM, Al-Kathery M, Pradhan B (2015a) 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
- Youssef AM, Pradhan B, Jebur MN, El-Harbi HM (2015b) Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia. *Environ Earth Sci* 73(7):3745–3761