

# Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran

HAMID REZA POURGHASEMI<sup>1</sup>, ABBAS GOLI JIRANDEH<sup>2</sup>, BISWAJEET PRADHAN<sup>3,4,\*</sup>,  
CHONG XU<sup>5</sup> and CANDAN GOKCEOGLU<sup>6</sup>

<sup>1</sup>*Department of Watershed Management Engineering, College of Natural Resources and Marine Sciences, Tarbiat Modares University (TMU), Mazandaran, Iran.*

<sup>2</sup>*Spatial Academy Team, Tehran, Iran.*

<sup>3</sup>*Geospatial Information Science Research Centre (GISRC), University Putra Malaysia, 43400, Serdang, Selangor, Malaysia.*

<sup>4</sup>*Department of Civil Engineering, University Putra Malaysia, 43400, Serdang, Selangor Darul Ehsan, Malaysia.*

<sup>5</sup>*Key Laboratory of Active Tectonics and Volcano, Institute of Geology, China Earthquake Administration, Beijing 100029, People's Republic of China.*

<sup>6</sup>*Applied Geology Division, Department of Geological Engineering, Hacettepe University, Hacettepe, Turkey.*

\*Corresponding author. e-mail: biswajeet24@gmail.com      biswajeet@mailcity.com

The main goal of this study is to produce landslide susceptibility map using GIS-based support vector machine (SVM) at Kalaleh Township area of the Golestan province, Iran. In this paper, six different types of kernel classifiers such as linear, polynomial degree of 2, polynomial degree of 3, polynomial degree of 4, radial basis function (RBF) and sigmoid were used for landslide susceptibility mapping. At the first stage of the study, landslide locations were identified by aerial photographs and field surveys, and a total of 82 landslide locations were extracted from various sources. Of this, 75% of the landslides (61 landslide locations) are used as training dataset and the rest was used as (21 landslide locations) the validation dataset. Fourteen input data layers were employed as landslide conditioning factors in the landslide susceptibility modelling. These factors are slope degree, slope aspect, altitude, plan curvature, profile curvature, tangential curvature, surface area ratio (SAR), lithology, land use, distance from faults, distance from rivers, distance from roads, topographic wetness index (TWI) and stream power index (SPI). Using these conditioning factors, landslide susceptibility indices were calculated using support vector machine by employing six types of kernel function classifiers. Subsequently, the results were plotted in ArcGIS and six landslide susceptibility maps were produced. Then, using the success rate and the prediction rate methods, the validation process was performed by comparing the existing landslide data with the six landslide susceptibility maps. The validation results showed that success rates for six types of kernel models varied from 79% to 87%. Similarly, results of prediction rates showed that RBF (85%) and polynomial degree of 3 (83%) models performed slightly better than other types of kernel (polynomial degree of 2 = 78%, sigmoid = 78%, polynomial degree of 4 = 78%, and linear = 77%) models. Based on our results, the differences in the rates (success and prediction) of the six models are not really significant. So, the produced susceptibility maps will be useful for general land-use planning.

---

**Keywords.** Landslides; support vector machine (SVM); geographical information systems (GIS); remote sensing; Golestan province; Iran.

## 1. Introduction

Landslides are one of the most catastrophic natural hazards occurred in many areas of the world. In Iran, about 187 people have been killed by landslides (Iranian Landslide working party 2007), and losses resulting from mass movements until the end of September 2007 have been estimated at 126,893 billion Iranian Rials (almost \$12,700 million) using the 4900 landslides. The Golestan province, located in the northeast part of Iran, has been heavily affected by landslides in recent years. Landslides are normally triggered by heavy rainfall, but very few attempts have been made to forecast their location or prevent their damage in the Golestan province. Previously only a few investigations of landslide susceptibility analysis have been carried out in Iran (Pourghasemi et al. 2012a, b, c, d, e; Zare et al. 2012). Therefore, in this study, we carried out landslide susceptibility analysis in the Golestan region.

Landslide susceptibility map is one of the crucial information in spatial planning for landslide prone areas. It gives a description about spatial probability of landslide which can be useful in landuse planning. So, the study of landslide susceptibility mapping is rapidly becoming the focus of major scientific research, engineering study and practices, and land-use policy throughout the world. In recent years, different models were used for landslide susceptibility mapping such as probabilistic models (Gokceoglu et al. 2000; Lee and Pradhan 2006; Pradhan et al. 2006, 2010c, d, 2011, 2012; Akgun et al. 2008; Youssef et al. 2009; Pradhan 2010b; Pradhan and Lee 2010a; Pradhan and Youssef 2010; Pourghasemi et al. 2012a, b, e), logistic regression (Guzzetti et al. 1999; Wang and Sassa 2005; Lee and Pradhan 2007; Akgun and Bulut 2007; Pradhan 2010b, c; Pradhan et al. 2008, 2011; Tunusluoglu et al. 2008; Yilmaz 2010b; Akgun 2011; Akgun et al. 2011a; Devkota et al. 2012), geotechnical and the safety factor models (Shou and Wang 2003). All these models provide solutions for integrating information levels and mapping the outputs. Recently, other new methods have been applied for landslide susceptibility evaluation using soft computing and data mining approaches such as fuzzy logic (Ercanoglu and Gokceoglu 2004; Muthu et al. 2008; Pradhan and Buchroithner 2010; Pradhan and Pirasteh 2010; Pradhan 2010a, b, 2011a, b; Pradhan et al. 2009; Akgun et al. 2011b; Pourghasemi et al. 2012c), artificial neural network models (Lee et al. 2004; Gomez and Kavzoglu 2005; Nefeslioglu et al. 2008; Pradhan et al. 2010b; Pradhan and Lee 2009, 2010a, b; Yilmaz 2009a, b, 2010a, b; Zare et al. 2012; Bui et al. 2012b), neuro-fuzzy (Pradhan et al. 2010a; Vahidnia et al.

2010; Sezer et al. 2011; Oh and Pradhan 2011; Bui et al. 2012a), decision tree methods (Akgun et al. 2008; Akgun and Turk 2010; Nefeslioglu et al. 2010; Akgun 2011), spatial multicriteria evaluation (Pourghasemi et al. 2012d), evidential belief function model (Althuwaynee et al. 2012; Bui et al. 2012c), and support vector machine (SVM) (Yao et al. 2008; Yilmaz 2010b; Marjanović et al. 2011; Xu et al. 2012; Bui et al. 2012a) and their performances have been assessed. The spatial results of these approaches are generally appealing, and they give rise to qualitatively and quantitatively map of the landslide hazard areas (Pradhan 2010a). The main difference between the present study and the approaches described in the aforementioned publications is that a GIS-based support vector machine (SVM) by different kernel functions were applied and their results were compared for landslide susceptibility in Kalaleh Township area of Golestan province, Iran. The SVM is a relatively new method that has seldom been applied to landslide susceptibility assessment. Moreover, the comparison of various kernel classifiers is important scientifically because SVM can use several types of kernel functions. This study may contribute on the selection of kernel function for the landslide susceptibility mapping by SVM approach.

## 2. Study area

The study area covers roughly 3430 km<sup>2</sup> between latitudes 37°06'20" to 37°47'58"N, and longitudes 55°18'05" to 56°04'57"E, and is located in the east part of the Golestan province, Iran (figure 1). The altitude of the study area ranges from 0 to 2307 m above m.s.l. The majority of the areas (63.3%) are covered by irrigation agriculture (36.4%) and mixing of range-agriculture (26.9%). The slope angles of the area range from 0° to as much as 89°. The total annual rainfall ranges in different parts of the study region, from 300 to 765 mm. Based on the records from the Iranian Meteorological Department, the maximum rainfall occurs between November and January (Golestan Regional Water Co. 2007). The study area is heterogeneous in terms of terrain complexity (plain and mountainous) and has different climatic regimes (semi-wet in south parts and dry and semi-dry in north parts). The lithology of the study area consists of several geologic formations such as Shemshak, Khoshyeilagh, Abderaz, Chamanbid, Tirgan, Mozduran and Zard (table 1).

## 3. Data inputs

Various thematic data layers representing landslide conditioning factors, such as slope degree, slope

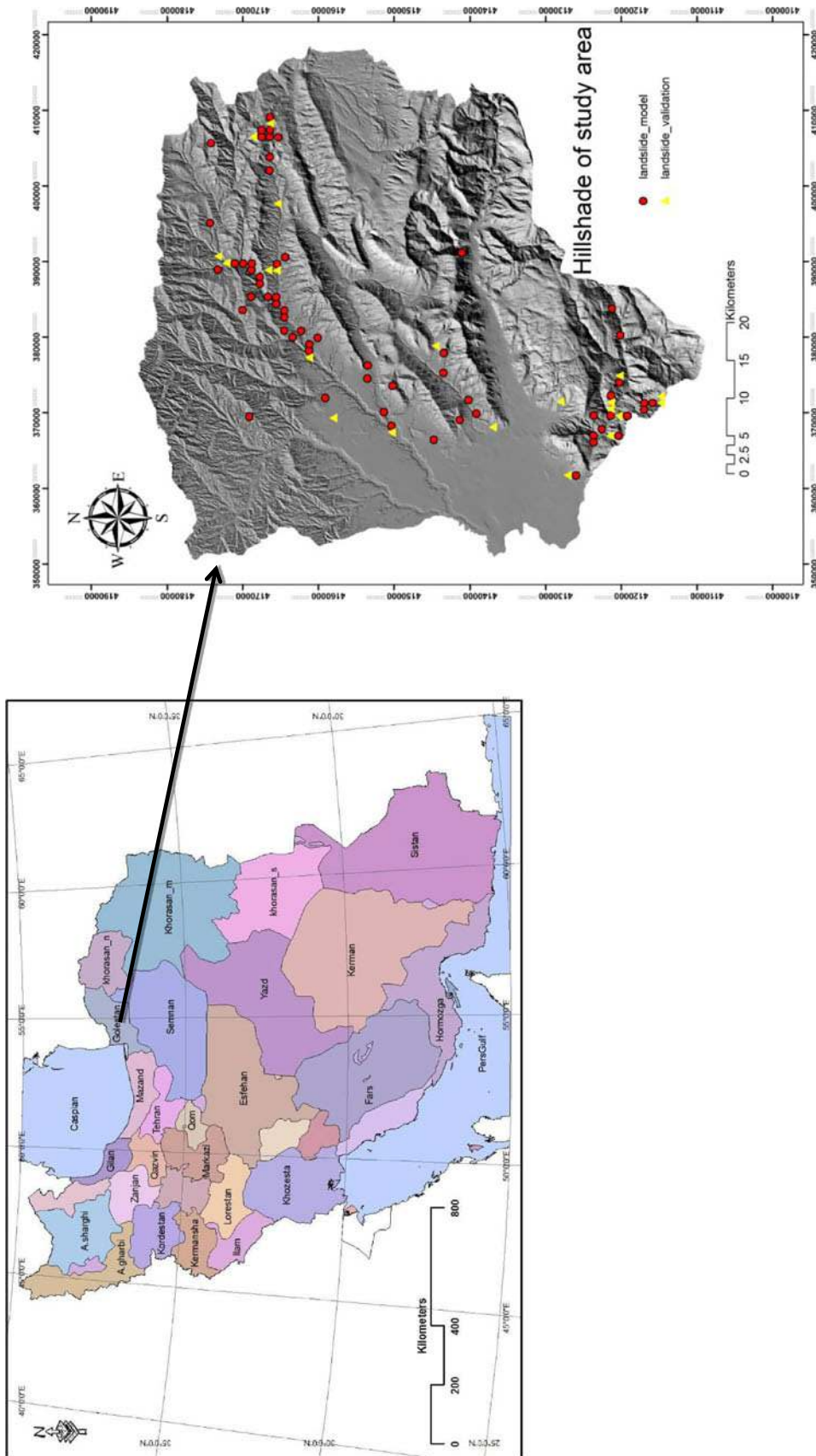


Figure 1. Landslide location map with hill shaded map of study area.

Table 1. Description of geological units of the study area.

Symbol	Formation	Lithology
Group 1		
Q <sup>al</sup>	–	Recent river deposits
Q <sup>t2</sup>	–	Young alluvial terraces
Q <sup>t1</sup>	–	Old alluvial terraces
Group 2		
P1 <sup>c</sup>	–	Conglomerate with intercalations of sandstone and clay
Pe <sup>m</sup>	–	Gray to white marl
Js	Shemshak	Sandstone, shale and conglomerate
D <sub>kh</sub>	Khosh Yeilagh	Dark grey shale, limestone, dolomite, sandstone, marl and diabase
E1m	–	Red marl with limestone and gypsum
Group 3		
K <sub>ad</sub>	Abderaz	Argillaceous limestone and chert
S <sub>n</sub>	–	Coral limestone, shale and dolomite
E <sup>1</sup>	–	Buff and cream medium to thick-bedded argillaceous limestone
J <sub>cb</sub>	Chamanbid	Pale-grey to cream, thin to medium-bedded argillaceous limestone with marl intercalations
Group 4		
K <sub>at</sub>	–	Andesite tuff
Group 5		
K <sub>tr</sub>	Tirgan	Oolite limestone and bioclastic
J <sub>mz</sub>	Mozduran	Pale-grey, thick-bedded to massive cherty limestone
J1	–	Massive light grey limestone
Tre2	–	Massive dolomite
Group 6		
E <sub>kh</sub>	–	Shale and sandstone
K <sub>Z</sub>	Zard	Alternation of, yellow to pale green, calcareous sandstone, sandy limestone and marlstone

aspect, altitude, plan curvature, profile curvature, tangential curvature, surface area ratio (SAR), lithology, land use, distance from faults, distance from rivers, distance from roads, topographic wetness index (TWI) and stream power index (SPI) were prepared. These factors fall under the category of conditioning factors, which make the area susceptible to movement without actually initiating a landslide; thus, these factors are considered to be responsible for the occurrence of landslides in the regions for which pertinent data can be collected from available resources and from the field. The triggering factors, such as rainfall and earthquake, set the movement off by shifting the slope from a marginally stable to an actively unstable area. Furthermore, the attributes of the ground in terms of landslide susceptibility were considered in the present study.

### 3.1 Landslide inventory map

The mapping of actual landslides in the study area is essential for describing the relationship between the landslide distribution and the conditioning factors. To produce a detailed and reliable landslide

inventory map, extensive field surveys and observations were performed in the study area. A total of 82 landslides were identified and mapped in the study area by evaluating aerial photos in 1:25,000 scale supported by field survey (figure 1). Most of the landslides are shallow rotational with a few translational. However, during the analyses performed in the present study, only rotational failure according to the landslide classification system proposed by Varnes (1978) was considered and translational slides were eliminated because its occurrence is rare. The minimum and maximum size of landslides is 450 and 5,700 m<sup>2</sup>, respectively. All the 82 landslides that were included in the study occurred over the past 20 years (1990–2010). Of the 82 landslides identified, 75% (61 landslide locations) locations were chosen randomly for the landslide susceptibility modeling, while the remaining 25% (21 landslide locations) cases were used for the model validation.

### 3.2 Conditioning factors

A digital elevation model (DEM), with a resolution of 15 m, was generated from topographic maps in

1:25,000 scale having a contour interval of 10 m. Using this DEM, slope degree, slope aspect, altitude, plan curvature, profile curvature, tangential curvature, and surface area ratio (SAR) were produced (figure 2a–g). The slope map was reclassified into four categories namely: (1)  $0^{\circ}$ – $5^{\circ}$ , (2)  $6^{\circ}$ – $15^{\circ}$ , (3)  $16^{\circ}$ – $30^{\circ}$  and (4)  $> 30^{\circ}$  (figure 2a). The aspect map is grouped into 9 classes including eight directions and flat (figure 2b). Altitude is also a significant landslide conditioning factor because it is controlled by several geologic and geomorphological processes (Gritzner *et al.* 2001; Dai and Lee 2002; Ayalew *et al.* 2005). An altitude map is

prepared from the  $15 \times 15$  m digital elevation model (1: 25,000 scale with 10 m contour intervals) and grouped into 5 classes (figure 2c). Plan curvature is described as the curvature of a contour line formed by intersecting a horizontal plane with the surface (figure 2d). The influence of plan curvature on the slope erosion processes is the convergence or divergence of water during downhill flow (Ercanoglu and Gokceoglu 2002; Oh and Pradhan 2011). The plan curvature map was produced using a system for automated geoscientific analyses (SAGA) GIS. The profile curvature or vertical is curvature of corresponding normal section, which

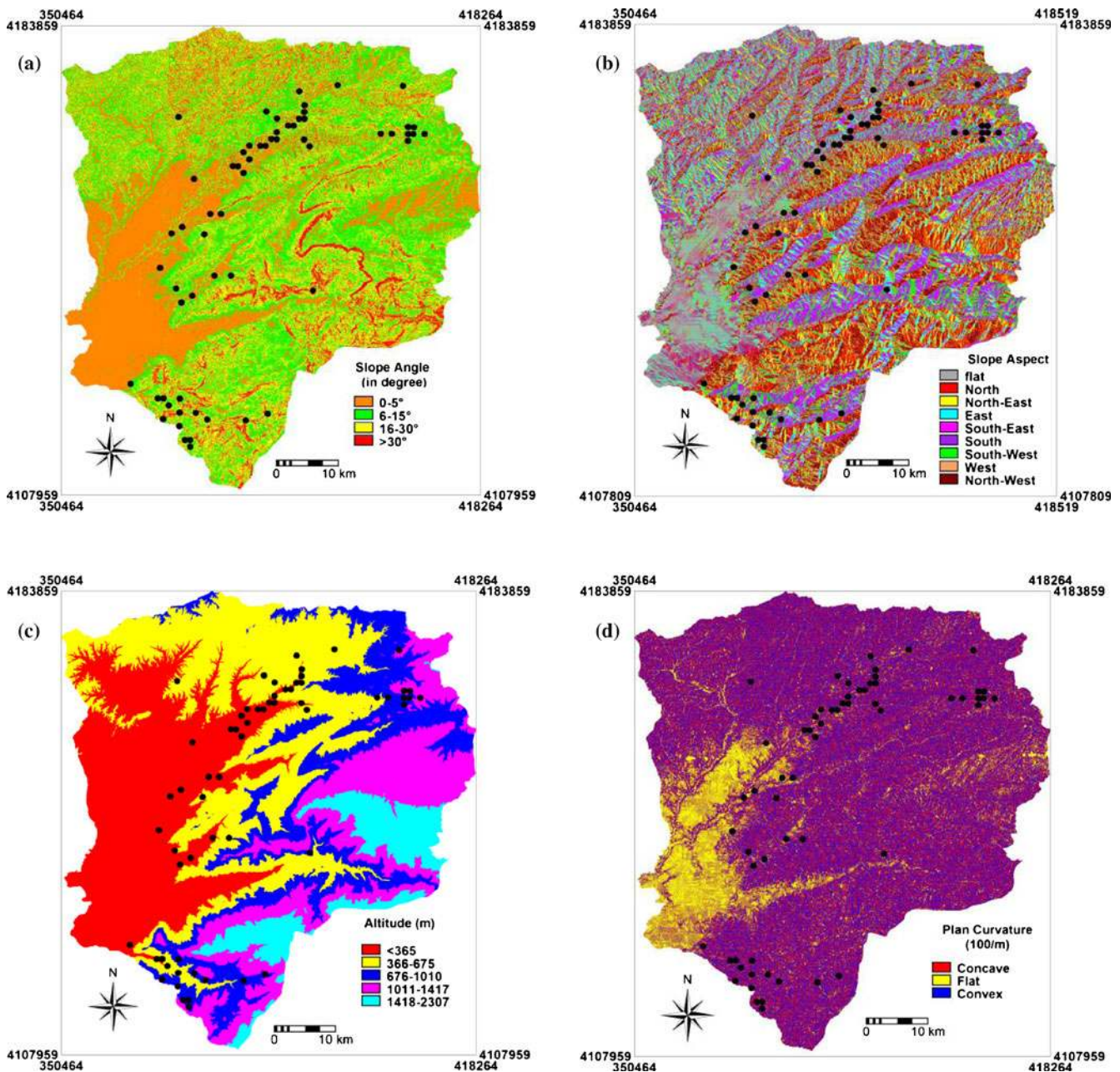


Figure 2. Topographical parameter maps of the study area: (a) slope degree, (b) aspect, (c) altitude, (d) plan curvature, (e) profile curvature, (f) tangential curvature and (g) surface area ratio (SAR).

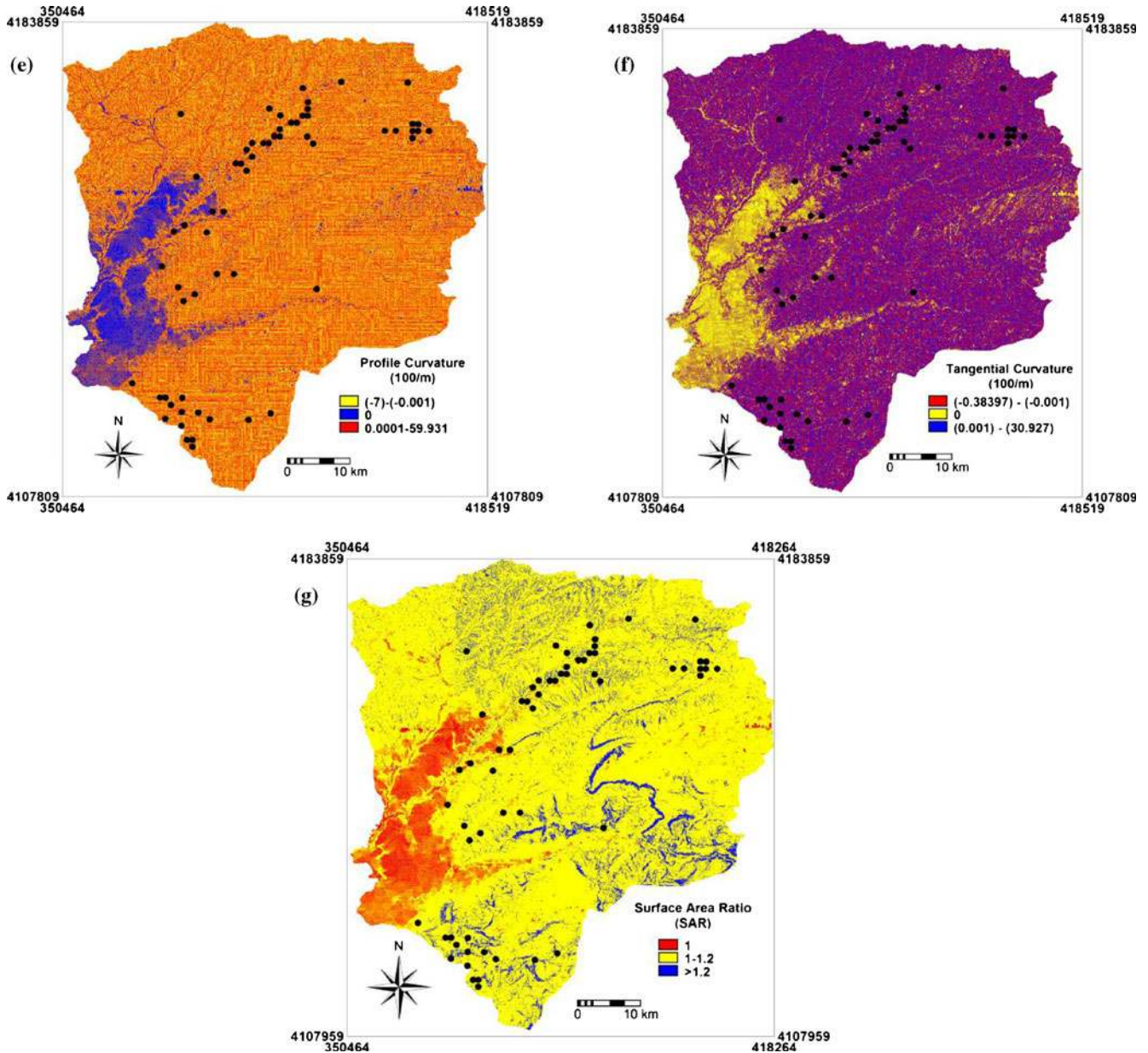


Figure 2. (Continued)

is tangential to a flow-line (figure 2e). It is negative when the normal section concavity is directed up, and positive in the opposite case (Hengl *et al.* 2003). It shows the flow acceleration, erosion (negative values)/deposition (positive values) rate and gives a basic idea of geomorphology (Yesilnacar 2005). In addition, the profile curvature is important because it controls the change of velocity of mass flowing down the slope (Talebi *et al.* 2007). Also, tangential curvature (primary topographic attribute) is a result of plan curvature and slope (figure 2f). It is measured in the direction of a tangent to a contour line. Tangential curvature represents local changes in the direction of the contour, which are related to the gravitational flux (Yesilnacar 2005). Values less than zero and

greater than zero indicate concavity (deposition) and convexity (erosion) for study area. Surface area ratio is a basis for useful measure of landscape topographic roughness and convolutedness (figure 2g). The surface area ratio of any particular region on the landscape can be calculated based on the following equation (Jenness 2002):

$$\text{SAR} = (S/A_S) \quad (1)$$

where  $S$  is the surface area of that region and  $A_S$  is the planimetric area. The value one represents smooth areas and higher values represent rough parts.

The geological map of the study area was prepared by the Geological Survey of Iran (GSI) at 1:100,000 scale, and was digitized in GIS.

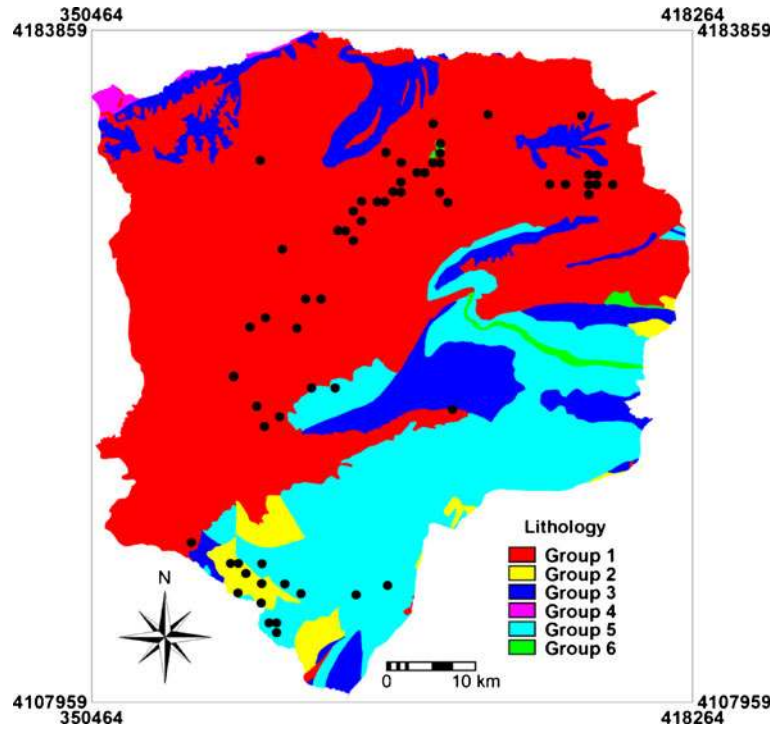


Figure 3. The lithology map of the study area.

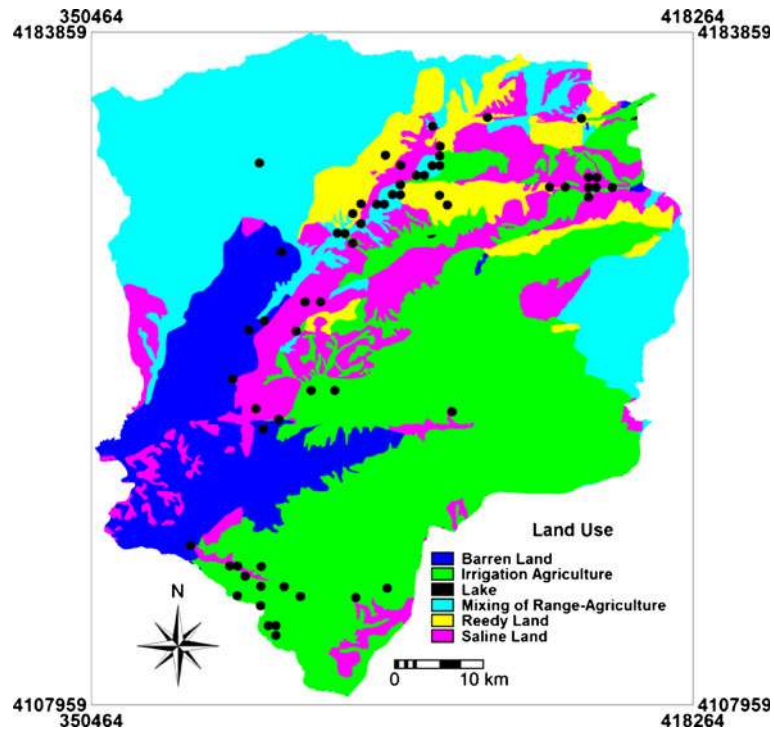


Figure 4. The land use map of the study area.

The study area is covered with various types of lithological formations. The general geological setting of the area is shown in figure 3, and the lithological properties are summarized in table 1. In this study, a Landsat/ETM<sup>+</sup> satellite image was

used for the year 2010 and a supervised classification approach was employed. Six different types of land use were classified. These classes are barren land, irrigation agriculture, lake, mixing of range-agriculture, reedy land and saline land areas

(figure 4). Most of the study area is covered by irrigation agriculture (36.4%) and mixing of range-agriculture (26.9%). The distance from the faults is calculated at 100 m intervals using the geological map (figure 5a). An important parameter that controls the stability of a slope is the saturation degree of the material on the slope (Yalcin and Bulut 2007). The closeness of the slope to drainage structures is another important factor in terms of stability. Five different buffer zones were created within the study area to determine the degree to which the streams affected the slopes (figure 5b). Similar to the effect of the distance to streams, landslides may occur on the road and on the side of slopes affected by roads (Pachauri et al. 1998; Ayalew and Yamagishi 2005). Five different buffer

zones are created on the path of the road to determine the effect of the road on the stability of slope (figure 5c).

Another topographic factor within the runoff model is the topographic wetness index (TWI). A topographic wetness index measures the degree of accumulation of water at a site (figure 6a). It is defined as (Beven and Kirkby 1979; Moore et al. 1991):

$$TWI = \ln(A_S / \tan \beta) \quad (2)$$

where  $A_S$  is the catchment area and  $\beta$  is the slope angle in degrees. The stream power index (SPI) is a compound topographic attribute. It is a measure of the erosive power of flowing water based on the assumption that discharge is proportional to

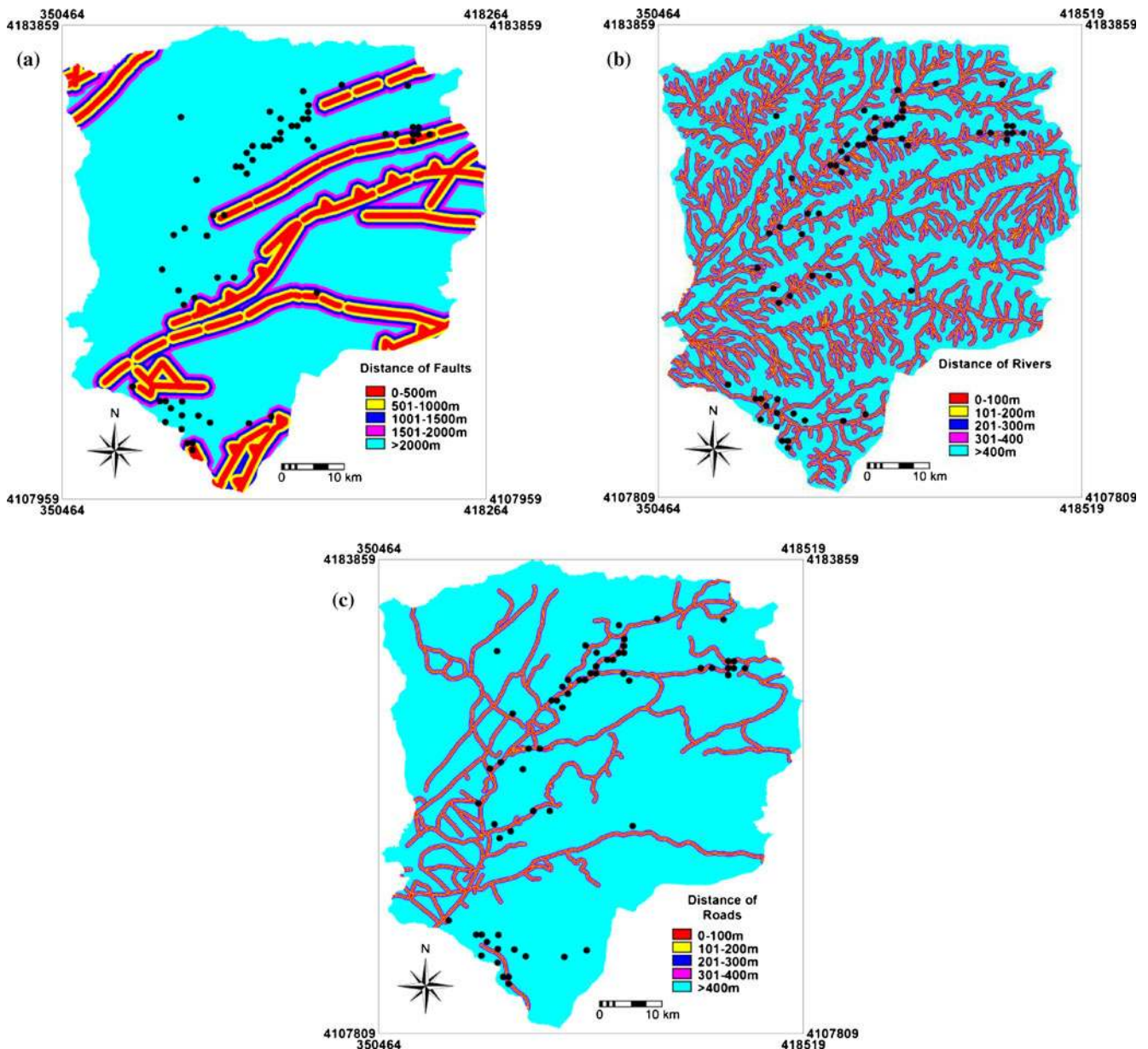


Figure 5. Distances of (a) faults, (b) rivers and (c) roads.

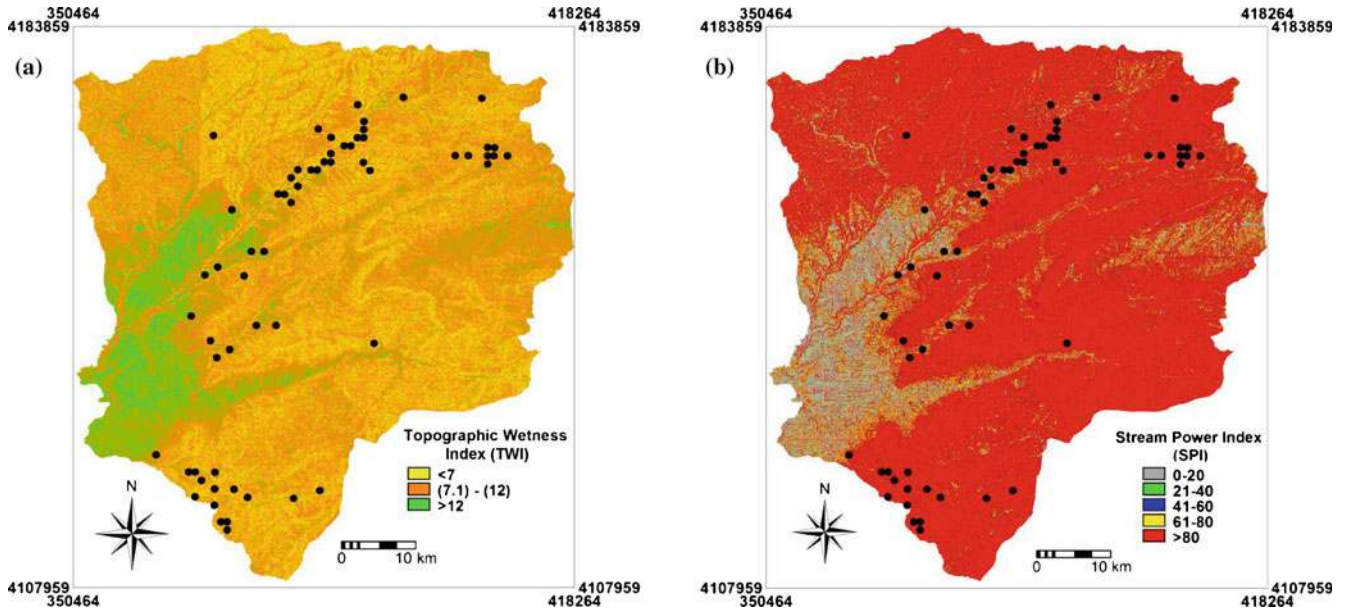


Figure 6. Hydrological based terrain maps of the study area: (a) TWI and (b) SPI.

specific catchment area (figure 6b). Stream power index was calculated based on the formula given by Moore *et al.* (1991).

$$\text{SPI} = A_S \times \tan \beta \quad (3)$$

where  $A_S$  is the specific catchment's area and  $\beta$  is the local slope gradient measured in degrees.

#### 4. Support vector machine

Support vector machine (SVM) is a classification system derived from statistical learning theory that can be used for both classification and regression (Vapnik 1995; Christianini and Shawe-Taylor 2000). It separates the classes with a decision surface that maximizes the margin between the classes. The surface is often called the optimal hyper-plane, and the data points closest to the hyper-plane are called support vectors. The support vectors are the critical elements of the training set. SVM algorithms, the available decision rules, provide a general method of function estimation that is performed by solving only a convex (quadratic) optimization problem. The aim of mentioned classification approach is not simply to separate the two classes, but to maximize the margin between them by minimizing  $\|W\|$ . Kanevski *et al.* (2009) showed that hyper-plane with a large margin should be more resistant to noise and possess better generalization than a hyper-plane with a small margin. Hence, SVMs have an important advantage compared to other machine learning algorithms (i.e., multi-layer perceptrons): the uniqueness of the solution (Micheletti 2011). So,

consider a set of linear separable training vectors  $X_i$  ( $i = 1, 2, \dots, n$ ). The training vectors consist of two classes, which are denoted as  $Y_i = \pm 1$ . The goal of SVM is to search an  $n$ -dimensional hyper-plane differentiating the two classes by their maximum gap. Mathematically, it can be expressed as (Yao *et al.* 2008; Xu *et al.* 2012):

$$1/2 \|W\|^2 \quad (4)$$

$$Y_i ((W * X_i) + b) \geq 1 \quad (5)$$

where  $\|W\|$  is the norm of the normal of the hyper-plane, and  $b$  is a scalar base.

A Lagrangian formulation is introduced to solve the problem (equation 6). It consists in the adjunction of Lagrange multipliers  $\lambda_i$  in constraint. Thus, the goal is now to minimize the Lagrangian  $L$  with to  $W$  and  $b$  and maximize with respect to  $\lambda_i$ . For this reason, we used of following equation:

$$L = 1/2 \|W\|^2 - \sum_{i=1}^n \lambda_i (Y_i ((W * X_i) + b) - 1). \quad (6)$$

To perform the landslide susceptibility mapping using SVM, we used remote sensing software namely the environment for visualizing Images (ENVI 4.3). The ENVI 4.3 SVM classifier provides four types of kernels: linear, polynomial, radial basis function (RBF) and sigmoid. The default is the radial basis function kernel, which works well in most cases (ENVI 2006). The mathematical representation of each kernel (linear, polynomial,

radial basis function, and sigmoid) is listed below, respectively (ENVI 2006):

$$K(X_i, X_j) = X_i^T X_j \quad (7)$$

$$K(X_i, X_j) = (\gamma X_i^T X_j + r)^d, \quad \gamma > 0 \quad (8)$$

$$K(X_i, X_j) = (-\gamma \|X_i - X_j\|), \quad \gamma > 0 \quad (9)$$

$$K(X_i, X_j) = \tanh(\gamma X_i^T X_j + r) \quad (10)$$

where  $\gamma$  is the gamma term in the kernel function for all kernel types except linear.

$d$  is the polynomial degree term in the kernel function for the polynomial kernel.

$r$  is the bias term in the kernel function for the polynomial and sigmoid kernels.

$\gamma$ ,  $d$ , and  $r$  are user-controlled parameters, as their correct definition significantly increases the accuracy of the SVM solution.

## 5. Results and discussion

In this research, a support vector machine with six types of kernel classifiers such as linear, polynomial degree of 2, polynomial degree of 3, polynomial degree of 4, radial basis function (RBF) and sigmoid were used in GIS for landslide susceptibility mapping in the study area. A total of 82 landslides were mapped using aerial photographs and subsequently verified through the field checking. Fourteen landslide conditioning factors were considered that include: slope degree, slope aspect, altitude, plan curvature, profile curvature, tangential curvature, surface area ratio (SAR), lithology, land use, distance from faults, distance from rivers, distance from roads, topographic wetness index (TWI) and stream power index (SPI).

The results of spatial relationship between landslides and conditioning factors using frequency ratio model is shown in table 2. In table 2, for the slope degree between  $16^\circ$  and  $30^\circ$ , the frequency ratio was 0.7 which indicates a very low probability of landslide occurrence. Similarly, for the slope angle between  $6^\circ$  and  $15^\circ$ , the ratio was  $>1$  (1.2), which indicates a high probability of landslide occurrence. Gentle slopes are expected to have a low frequency of landslides because of the generally lower shear stresses associated with low gradients. Steep natural slopes resulting from outcropping bedrock, however, may not be susceptible to shallow landslides. This may be due to the

decrease of soil accumulation in steeper slope and anthropogenic (human-caused) interferences such as land use change in the middle slopes. The frequency ratio for the slope aspect was high for southwest-facing and south facing slopes (FR value of 1.6 and 1.4, respectively) but was low for flat slopes (0.3). The frequency ratio between landslide occurrence and altitude showed that the altitude between 676 and 1010 m ranges had the highest value 1.4 and 1418 and 2307 m class had the lowest value (0.0). The curvature values represent the morphology of the topography. A positive curvature is an upwardly convex cell, and a negative curvature is upwardly concave cell. As shown in table 2, from the results of the frequency ratio model, the more positive or negative the curvature value, the higher the probability of landslide occurrence. Flat areas had a low curvature value of 0.8, whereas convex shape had highest value of 1.3. The reason for this is that a convex rounded hilltop slope could expose to repeated dilation and contraction of loose debris on an inclined surface that might induce a creeping or mudslide due to heavy rainfall. In the case of profile curvature, most of the landslides occurred in  $(-7)$  and  $(-0.0001)$  class with FR value of 1.1. This means that the landslide probability is higher in this class. In profile curvatures, slope stability slightly increases when plan shape changes from concave to convex. However, this effect is more pronounced when it changes from straight to convex. In the case of tangential curvature, the frequency ratio was higher (1.3) in 0.001 and 30.93 class and lower (0.8) in flat areas. In the case of surface area ratio, the value one represents smooth areas and higher values represent rough parts. So, the frequency ratio for the SAR was high in  $>1.2$  class, which indicates a high probability of landslide occurrence. There were six groups of lithological units within the study area. The frequency ratio between landslide occurrence and lithology suggests that the group 2 (i.e., P1c (conglomerate with intercalations of sandstone and clay), Pem (gray to white marl), Js (sandstone, shale and conglomerate), Dkh (dark grey shale, limestone, dolomite, sandstone, marl and diabase) and E1m (red marl with limestone and gypsum)) had the highest value (3.5), whereas the group 4 ( $K_{at}$ ) with Andesite tuff formation had the lowest value (0.0). When comparing the relationship between landslides and land use, the ratio value was high in reedy land area (2.6); however, lake and saline land had a lower frequency ratio (0.0), indicating a lesser chance of landslide occurrence. In addition, we found that landslides at a distance of 1501 and 2000 m from a fault had a higher frequency ratio (1.4), indicating a higher probability of a landslide occurrence. The influence of river systems and roads on landslide susceptibility

Table 2. *Spatial relationship between each landslide conditioning factor and landslides by frequency ratio.*

Class	Percentage of domain (a)	No. of landslide	Percentage of landslide (b)	Frequency ratio (b/a)
Slope degree				
0–5°	27.63	19	31.15	1.13
6–15°	32.39	23	37.70	1.16
16–30°	29.45	12	19.67	0.67
>30°	10.53	7	11.48	1.09
Slope aspect				
Flat	5.26	1	1.64	0.31
North	13.68	8	13.11	0.96
Northeast	8.89	6	9.84	1.11
East	11.05	6	9.84	0.89
Southeast	10.27	3	4.92	0.48
South	11.78	10	16.39	1.39
Southwest	10.29	10	16.39	1.59
West	14.34	7	11.48	0.80
Northwest	14.45	10	16.39	1.13
Altitude (m)				
<365	29.55	25	40.98	1.39
366–675	28.57	17	27.87	0.98
676–1010	17.09	15	24.59	1.44
1011–1417	17.02	4	6.56	0.39
1418–2307	7.78	0	0	0
Plan curvature (100/m)				
Concave	41.66	21	34.43	0.83
Flat	15.00	7	11.48	0.77
Convex	43.33	33	54.10	1.25
Profile curvature (100/m)				
(–7)–(–0.0001)	45.38	30	49.18	1.08
0	9.09	5	8.20	0.90
(0.0001)–(59.931)	45.52	26	42.62	0.94
Tangential curvature (100/m)				
(–0.384)–(–0.001)	41.66	21	34.43	0.83
0	15.01	7	11.48	0.76
(0.001)–(30.93)	43.33	33	54.10	1.25
Surface area ratio (SAR)				
1	7.34	3	4.92	0.67
1–1.2	84.93	52	85.25	1.00
>1.2	7.73	6	9.83	1.27
Lithology				
Group 1	61.58	44	72.13	1.17
Group 2	2.79	6	9.84	3.53
Group 3	12.14	1	1.64	0.14
Group 4	0.47	0	0	0
Group 5	22.28	9	14.75	0.66
Group 6	0.74	1	1.64	2.22
Land use				
Barren land	13.80	6	9.84	0.71
Irrigation agriculture	36.26	17	27.87	0.77
Lake	0.10	0	0	0
Mixing of range–agriculture	26.94	10	16.39	0.61
Reedy land	7.16	9	14.75	2.06
Saline land	15.74	19		0
Distance from Faults (m)				
0–500 m	9.18	4	6.56	0.71
501–1000 m	8.78	3	4.92	0.56
1001–1500 m	7.90	5	8.20	1.04
1501–2000 m	6.92	6	9.83	1.42
>2000 m	67.24	43	70.49	1.05

Table 2. (Continued)

Class	Percentage of domain (a)	No. of landslide	Percentage of landslide (b)	Frequency ratio (b/a)
Distance from rivers (m)				
0–100 m	13.40	8	13.11	0.98
101–200 m	12.87	14	22.95	1.78
201–300 m	12.13	7	11.48	0.95
301–400 m	10.63	10	16.39	1.54
>400 m	50.98	22	36.07	0.71
Distance from roads (m)				
0–100 m	4.22	11	18.03	4.27
101–200 m	4.12	6	9.84	2.39
201–300 m	4.04	6	9.84	2.44
301–400 m	3.74	2	3.28	0.88
>400 m	83.89	36	59.02	0.70
Topographic wetness index (TWI)				
<7	19.04	12	19.67	1.03
7.1–12	70.38	39	63.94	0.91
>12	10.58	10	16.39	1.55
Stream power index (SPI)				
0–20	8.45	4	6.56	0.78
21–40	4.03	4	6.56	1.63
41–60	3.89	3	4.92	1.26
61–80	7.01	6	9.83	1.40
>80	76.64	44	72.13	0.94

was also analyzed by identifying the river and road line buffering (table 2). The relationship between a landslide and its distance from river and road lines showed that when distance from a river system and road line increases, the landslide occurrence probability decreases. There were high frequency

ratios close to the river channels and roads, indicating a higher susceptibility. At a distance of 101 and 200 m, the ratio was  $>1$ , indicating a high probability of landslide occurrence, and at distances  $>400$  m, the ratio was  $<1$ , indicating very less probability. This can be attributed

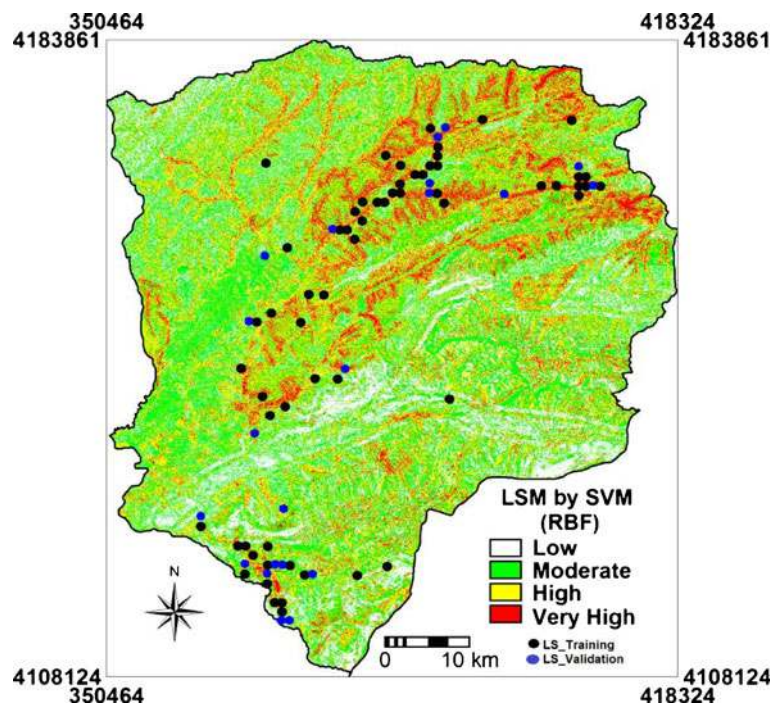


Figure 7. Landslide susceptibility map based on SVM (RBF).

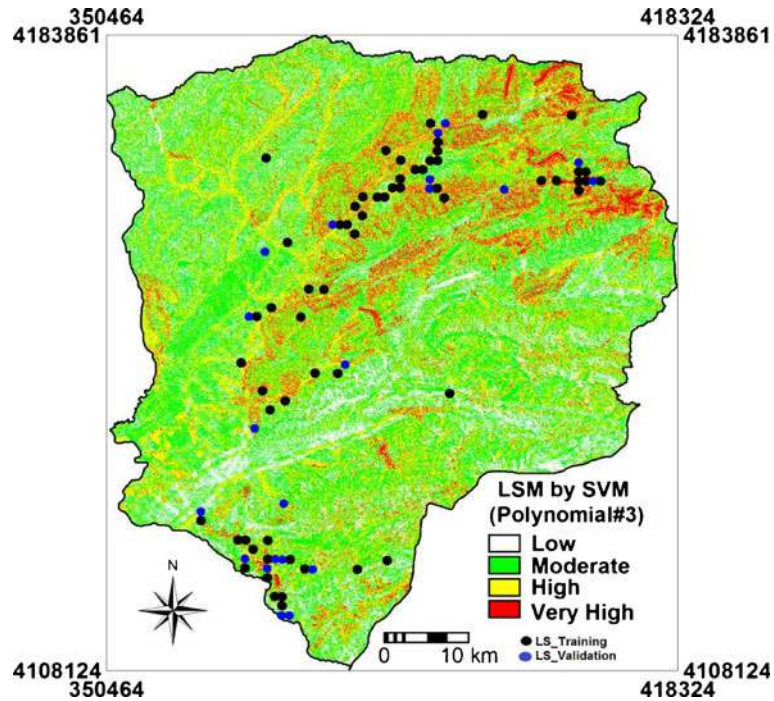


Figure 8. Landslide susceptibility map based on SVM (polynomial#3).

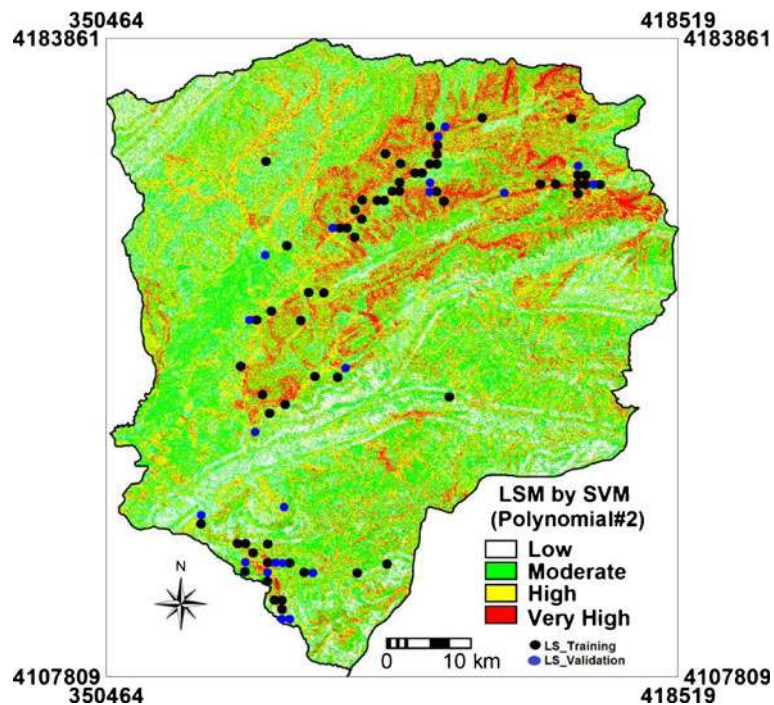


Figure 9. Landslide susceptibility map based on SVM (polynomial#2).

to the fact that terrain modification caused by gully erosion may influence the initiation of landslides. Similarly, the proximity of landslide to road was calculated by buffering (table 2). It can be observed that as the distance from road decreases, the landslide frequency generally decreases. At distances between 0 and 100 m, 100 and 200 m,

and 3000 m, the ratios were  $>1$  (4.3, 2.4 and 2.4, respectively), indicating a very high probability of landslide occurrence. The frequency ratio for the TWI and SPI were high between the range  $>12$  and 61 and 80, respectively. The frequency ratio value remained high, where the SPI values were low.

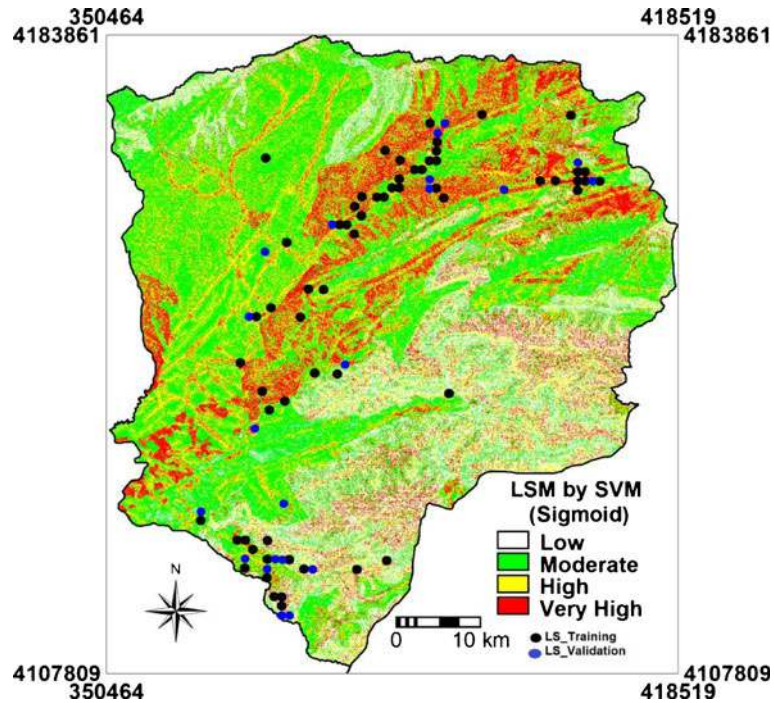


Figure 10. Landslide susceptibility map based on SVM (sigmoid).

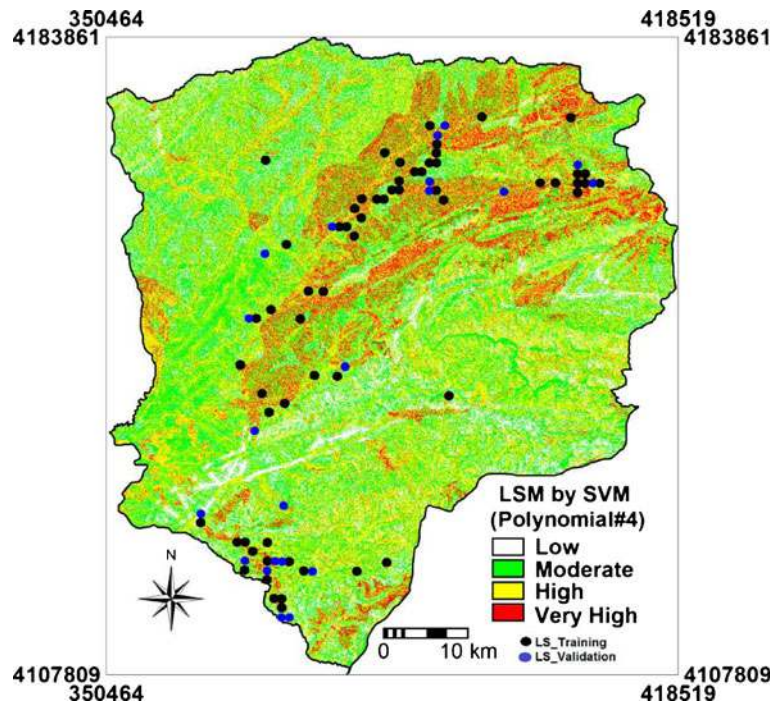


Figure 11. Landslide susceptibility map based on SVM (polynomial#4).

The SVM classification output or result is the decision values of each pixel for each class that were used. The probability values in ENVI software as rule image represent ‘true’ probability in the sense that each probability falls in the range of 0 to 1, and the sum of these values for each pixel equals 1.

The final ‘output rule images’ is the landslide susceptibility index file with two values, i.e., 0 represents stable and 1 represents landslide. Finally, the landslide susceptibility maps were reclassified into four susceptibility classes: high, moderate, low and very low (figures 7–12). For the classification of the

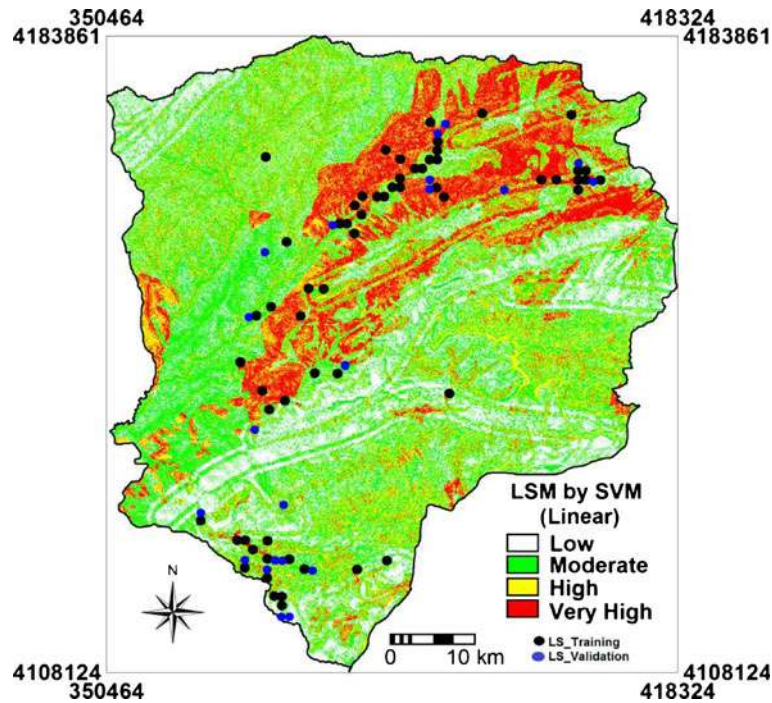


Figure 12. Landslide susceptibility map based on SVM (linear).

landslide susceptibility zones, a natural break classification scheme was applied (Pourghasemi *et al.* 2012c, e).

## 6. Validation and comparison of landslide susceptibility maps

### 6.1 Success and prediction rate for landslide susceptibility maps

The critical strategy in prediction models is the task of validating the predicted results that can provide meaningful interpretation of the results. To apply this validation, we used success rate and relative operating characteristics curve (ROC) by comparing the existing landslide locations with the six landslide susceptibility maps. Using the landslide grid cells in the training dataset, the success

rate results were obtained. The success rate curves were obtained by using the training dataset 75% (61 landslide locations). Table 3 shows the success rate and prediction rate curves for types of kernel models of support vector machine. The validation results showed that success rates for six types of kernel models vary from 79.3% to 87.8%. Consequently, the RBF and polynomial degree of 3 models have the highest area under the curve (AUC) value (0.88 and 0.87) but linear model has the less AUC with value of 0.79. Since the success rate method used the landslide pixels that have already been used for building the landslide models, the success rate is not a suitable method for assessing the prediction capability of the models. However, the success rate method may help to determine how well the resulting landslide susceptibility maps have classified the areas of existing landslides (Bui *et al.* 2011). The prediction rate (ROC curve) explains how well the model and predictor variables predict the landslide (Bui *et al.* 2011; Pourghasemi *et al.* 2012e). This method is already widely used as a measure of performance of a predictive rule (Yesilnacar and Topal 2005; Van Den Eeckhaut *et al.* 2006; Pradhan *et al.* 2010a, b). ROC plots the different accuracy values obtained against the whole range of possible threshold values of the functions, and the AUC serves as a global accuracy statistic for the model, regardless of a specific discriminate threshold. This curve is obtained by plotting all combinations of sensitivities and proportions of false negatives (1-specificity), which may be obtained by varying the decision threshold.

Table 3. Results of success rate and prediction rate curves for SVM types of kernel models.

Model (kernel type)	Success rate	Prediction rate
Linear	0.7935	0.771
Radial basis function (RBF)	0.878	0.8537
Sigmoid	0.8030	0.7834
Polynomial #2	0.8034	0.7857
Polynomial #3	0.8679	0.8333
Polynomial #4	0.8025	0.7823

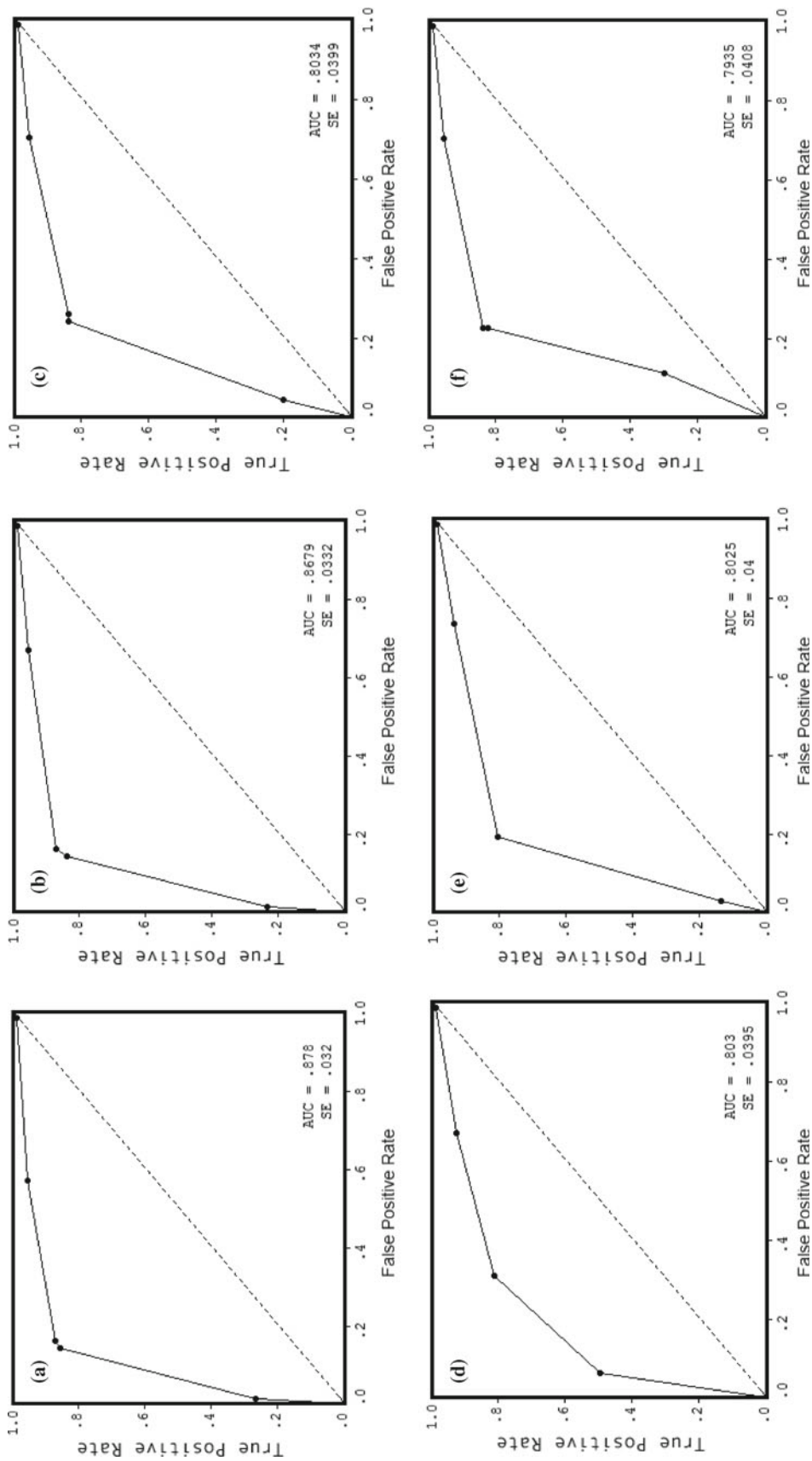


Figure 13. Success rate curves: (a) RBF, (b) polynomial degree of 3, (c) polynomial degree of 4, and (d) sigmoid, (e) polynomial degree of 2, and (f) linear.

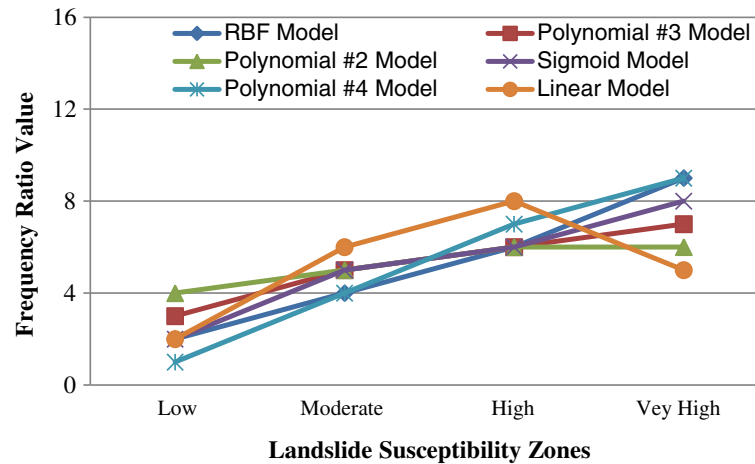


Figure 14. Frequency ratio plots of four landslide susceptibility zones of the SVM models.

The range of values of the ROC curve area is 0.5–1 for a good fit, while values below 0.5 represent a random fit. The results of the ROC curve test are illustrated in figure 13 and table 3. These curves indicate that RBF and polynomial degree of 3 has relatively higher prediction performance than other types of kernel models.

As can be seen from the AUC values, the RBF yielded the most accurate results when compared to the kernel models considered in the study. The main reason for this result is sourced from the nature of RBF. In landslide susceptibility assessment, the approximation capacity of the RBF is better than the kernel functions. When only the kernel functions are considered, the function having the polynomial degree of 3 exhibited the best performance. This result can be considered as normal because polynomial degree of 3 yielded as optimum result. If the polynomial degree increases, the overfitting problem arises. For this reason, the performance of the kernel function having the polynomial degree of 4 is worse than that of the kernel function having the polynomial degree of 3. As a main conclusion obtained from this assessment, it is possible to say that the landslide susceptibility problem is a nonlinear problem. For this reason, using the nonlinear approaches gives the better results than that of linear approaches. However, it

should be abstained from the overfitting problem when using the nonlinear approaches.

## 6.2 Frequency ratio analysis for landslide susceptibility maps

Relative frequency ratio analysis was performed on the classification results by overlaying the landslide grid cells of the six landslide susceptibility classes for each landslide susceptibility map. In the next stage, the distribution of the actual landslide areas is determined based on the landslide susceptibility zones. Finally, all the landslide grid cells were overlaid on landslide susceptibility zone (low, moderate, high and very high) in GIS, and frequency ratio was calculated for each of the susceptibility zones (Pradhan 2010a, b; Pourghasemi *et al.* 2012c, d). In an ideal landslide susceptibility map, the frequency ratio value increases from a low to a very high susceptibility zone (Pradhan and Lee 2010b; Pourghasemi *et al.* 2012c, d). A plot of the frequency ratio for the four landslide susceptibility classes of the six landslide susceptibility models is shown in figure 14. The results showed that the frequency ratio is gradually increased from the low to the very high susceptibility zone in study area. Also, descriptions of the four susceptibility zones of

Table 4. Descriptions of the four susceptibility zones of the six SVM models.

Landslide susceptibility zone	Percentage (%) of landslide					
	RBF	Polynomial #3	Polynomial #2	Sigmoid	Polynomial #4	Linear
Low	20.17	18.62	21.38	23.63	26.1	27.9
Moderate	36.77	41.09	38.86	36.93	36.15	36.5
High	31.32	31.36	27.58	24.34	26.94	21.67
Very high	11.74	8.93	12.17	15.1	10.81	13.93
Sum of high and very high	43.06	40.29	39.75	39.44	37.75	35.60

the six SVM models are shown in table 4. It could be observed that the sum percentage of existing landslide pixels for high and very high susceptibility zones (3 and 4) are 43% (RBF), 40% (polynomial #3), 39% (polynomial #2), 39% (sigmoid), 37% (polynomial #4) and 35% (linear), respectively. So, this result indicates that the extent of the landslide zones located in the high and very high susceptibility class for RBF model is higher than the other models. It is interesting to note that the frequency ratio value for linear model dramatically decreases for 'very high' landslide susceptibility zone. The main reason of this anomaly is sourced from the accuracy of the model. In fact, the accuracy of the linear model is highly low when compared to the other models employed in the present study. Consequently, 36%, 41%, 38%, 36%, 36% and 36% of the landslide zones coincided with the moderate susceptibility (zone 2) class of landslide susceptibility maps produced by RBF, polynomial #3, polynomial #2, sigmoid, polynomial #4 and linear models, respectively.

## 7. Concluding remarks

The preparation of landslide susceptibility maps is one of great interest to planning agencies for preliminary hazard studies, especially when a regulatory planning policy is to be implemented. Application of the SVM has been successfully demonstrated in this paper for the spatial prediction of landslide susceptibility. The main goal of the current study was to produce landslide susceptibility map for the study area using support vector machine (SVM) using six types of kernel classifiers such as linear, polynomial degree of 2, polynomial degree of 3, polynomial degree of 4, radial basis function (RBF) and sigmoid. At first stage, through multiple field investigations and aerial photo interpretation, a landslide inventory map was created. Subsequently, 14 data layers, as the landslide conditioning factors, are exploited to establish the spatial relationship between the landslides and the factor to detect the most susceptible areas. These factors are slope degree, slope aspect, altitude, plan curvature, profile curvature, tangential curvature, surface area ratio (SAR), lithology, land use, distance from faults, distance from rivers, distance from roads, topographic wetness index (TWI) and stream power index (SPI). The image processing software ENVI was used for the application of SVM model with different kernels in landslide susceptibility mapping. For result validation, both success rate and prediction rate curves were used. The validation results showed that the SVM model using RBF and polynomial degree of 3 have slightly better prediction accuracy (85%)

and (83%) than the others types of kernel models employed in this study. Based on our results, the differences in the rates (success and prediction) of the six models are not really significant. As a conclusion, the selection of the kernel function and its parameters play an important role in landslide susceptibility assessment. Also, it is clear that the degree of polynomial function had significant effect in the model.

Landslide susceptibility maps are of great help to planners and engineers for choosing suitable locations to implement developments. These results can be used as basic data to assist slope management and land-use planning. Moreover, this approach can be replicated in other study areas in order to compare their results and efficiency. It is important to note how, in SVMs algorithms, the available decision rules provide a general method of function estimation that is performed by solving only a convex (quadratic) optimization problem. Hence, SVMs have an important advantage compared to other machine learning algorithms: the uniqueness of the solution. As a final conclusion, the analysed results obtained from the study can provide very useful information for decision making and policy planning in landslide areas.

## Acknowledgements

Authors would like to thank two anonymous reviewers for their helpful reviews on the earlier version of the manuscript. Thanks to Dr Nibir Mandal for his valuable suggestions that improved the paper.

## References

- Akgun A 2011 A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: A case study at İzmir, Turkey; *Landslides*, doi: [10.1007/s10346-011-0283-7](https://doi.org/10.1007/s10346-011-0283-7).
- Akgun A and Bulut F 2007 GIS-based landslide susceptibility for Arsin-Yomra (Trabzon, North Turkey) region; *Environ. Geol.* **51**(8) 1377–1387.
- Akgun A and Turk N 2010 Landslide susceptibility mapping for Ayvalik (Western Turkey) and its vicinity by multi criteria decision analysis; *Environ. Earth Sci.* **61**(3) 595–611.
- Akgun A, Dag S and 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, Kincal C and Pradhan B 2011a Application of remote sensing data and GIS for landslide risk assessment as an environmental threat to Izmir city (west Turkey); *Environ. Monit. Assess.*, doi: [10.1007/s10661-011-2352-8](https://doi.org/10.1007/s10661-011-2352-8).
- Akgun A, Sezer E A, Nefeslioglu H A, Gokceoglu C and Pradhan B 2011b An easy-to-use MATLAB program

- (MamLand) for the assessment of landslide susceptibility using a Mamdani fuzzy algorithm; *Comput. Geosci.* **38**(1) 23–34.
- Althuwaynee O F, Pradhan B and Lee S 2012 Application of an evidential belief function model in landslide susceptibility mapping; *Comput. Geosci.* **44** 120–135, doi: [10.1016/j.cageo.2012.3](https://doi.org/10.1016/j.cageo.2012.3).
- Ayalew L and Yamagishi H 2005 The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan; *Geomorphology* **65**(1/2) 15–31.
- Ayalew L, Yamagishi H, Marui H and Kanno T 2005 Landslide in Sado Island of Japan: Part II. GIS-based susceptibility mapping with comparison of results from two methods and verifications; *Eng. Geol.* **81** 432–445.
- Beven K and Kirkby M J 1979 A physically based, variable contributing area model of basin hydrology; *Hydrol. Sci. Bull.* **24** 43–69.
- Bui D T, Pradhan B, Lofman O, Revhaug I and Dick O B 2011 Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS; *Comput. Geosci.*, doi: [10.1016/j.cageo.2011.10.031](https://doi.org/10.1016/j.cageo.2011.10.031).
- Bui D T, Pradhan B, Lofman O and Revhaug I 2012a Landslide susceptibility assessment in Vietnam using support vector machines, decision tree and Naïve Bayes models; *Math. Problems in Eng.*, <http://www.hindawi.com/journals/mpe/aip/974638/>.
- Bui D T, Pradhan B, Lofman O, Revhaug I and Dick O B 2012b Landslide susceptibility assessment in the Hoa Binh province of Vietnam using Artificial Neural Network; *Geomorphology*, doi: [10.1016/j.geomorph.2012.04.023](https://doi.org/10.1016/j.geomorph.2012.04.023).
- Bui D T, Pradhan B, Lofman O, Revhaug I and Dick O B 2012c Spatial prediction of landslide hazards in Vietnam: A comparative assessment of the efficacy of evidential belief functions and fuzzy logic models; *Catena* **96** 28–40.
- Christianini N and Shawe-Taylor J 2000 *An introduction to Support Vector Machines*; Cambridge University Press; ISBN 0521780195.
- Dai F C and Lee C F 2002 Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong; *Geomorphology* **42** 213–228.
- Devkota K C, Regmi A D, Pourghasemi H R, Yoshida K, Pradhan B, Ryu I C, Dhital M R and Althuwaynee O F 2012 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*, doi: [10.1007/s11069-012-0347-6](https://doi.org/10.1007/s11069-012-0347-6).
- ENVI Version 4.3, 2006. ITT industries, Inc. <http://www.RSInc.com/envi>.
- Ercanoglu M and 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 and 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** 229–250.
- Gokceoglu C, Sonmez H and Ercanoglu M 2000 Discontinuity controlled probabilistic slope failure risk maps of the Altindag (settlement) region in Turkey; *Eng. Geol.* **55** 277–296.
- Golestan Regional Water Co. 2007 Golestan province meteorological information report.
- Gomez H and Kavzoglu T 2005 Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela; *Eng. Geol.* **78** 11–27.
- Gritzner M L, Marcus W A, Aspinall R and Custer S G 2001 Assessing landslide potential using GIS, soil wetness modeling and topographic attributes, Payette River, Idaho; *Geomorphology* **37** 149–165.
- Guzzetti F, Carrara A, Cardinali M and Reichenbach P 1999 Landslide hazard evaluation: A review of current techniques and their application in a multi-case study, central Italy; *Geomorphology* **31** 181–216.
- Hengl T, Gruber S and Shrestha D P 2003 Digital Terrain Analysis in ILWIS. International Institute for Geo-Information Science and Earth Observation Enschede (The Netherlands), 62p.
- Iranian Landslide working party 2007 Iranian landslides list, Forest, Rangeland and Watershed Association, Iran, p. 60.
- Jenness J 2002 Surface Areas and Ratios from Elevation Grid, Jenness Enterprises, [http://www.jennessent.com/arcview/surface\\_areas.htm](http://www.jennessent.com/arcview/surface_areas.htm) (connected: 10.08.2003).
- Kanevski M, Pozdnoukhov A and Timonin V 2009 *Machine Learning for Spatial Environmental Data: Theory, Applications and Software*; CRC Press.
- Lee S and Pradhan B 2006 Probabilistic landslide risk mapping at Penang Island, Malaysia; *J. Earth Syst. Sci.* **115**(5) 661–672.
- Lee S and Pradhan B 2007 Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models; *Landslides* **4** 33–41.
- Lee S, Ryu J H, Won J S and Park H J 2004 Determination and application of the weights for landslide susceptibility mapping using an artificial neural network; *Eng. Geol.* **71** 289–302.
- Marjanović M, Kovačević M, Bajat B and Voženilek V 2011 Landslide susceptibility assessment using SVM machine learning algorithm; *Eng. Geol.* **123** 225–234.
- Micheletti N 2011 Landslide susceptibility mapping using adaptive support vector machines and feature selection, A Master Thesis submitted to University of Lausanne Faculty of Geosciences and Environment for the Degree of Master of Science in Environmental Geosciences, 99p.
- Moore I D, Grayson R B and Ladson A R 1991 Digital terrain modeling: A review of hydrological, geomorphological, and biological applications; *Hydrol. Proc.* **5** 3–30.
- Muthu K, Petrou M, Tarantino C and Blonda P 2008 Landslide possibility mapping using fuzzy approaches; *IEEE Trans. Geosci. Remote Sens.* **46** 1253–1265.
- Nefeslioglu H A, Gokceoglu C and Sonmez H 2008 An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps; *Eng. Geol.* **97** 171–191.
- Nefeslioglu H A, Sezer E, Gokceoglu C, Bozkir A S and Duman T Y 2010 Assessment of landslide susceptibility by decision trees in the Metropolitan Area of Istanbul, Turkey; *Math. Problems Eng. Volume*, doi: [10.1155/2010/901095](https://doi.org/10.1155/2010/901095).
- Oh H J and Pradhan B 2011 Application of a neuro-fuzzy model to landslide susceptibility mapping for shallow landslides in tropical hilly area; *Comput. Geosci.* **37**(9) 1264–1276.
- Pachauri A K, Gupta P V and Chander R 1998 Landslide zoning in a part of the Garhwal Himalayas; *Environ. Geol.* **36**(3–4) 325–334.
- Pourghasemi H R, Pradhan B, Gokceoglu C, Mohammadi M and Moradi H R 2012a Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran; *Arab. J. Geosci.*, doi: [10.1007/s12517-012-0532-7](https://doi.org/10.1007/s12517-012-0532-7).

- Pourghasemi H R, Pradhan B, Gokceoglu C and Deylami Moezzi K 2012b A comparative assessment of prediction capabilities of Dempster-Shafer and Weights-of-evidence models in landslide susceptibility mapping using GIS; *Geomatics, Nat. Hazards Risk*, doi: [10.1080/19475705.2012.662915](https://doi.org/10.1080/19475705.2012.662915).
- Pourghasemi H R, Pradhan B and Gokceoglu C 2012c Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran; *Nat. Hazards* **63**(2) 965–996.
- Pourghasemi H R, Gokceoglu C, Pradhan B and Deylami Moezzi K 2012d Landslide susceptibility mapping using a spatial multicriteria evaluation model at Haraz Watershed, Iran; (eds) Pradhan B and Buchroithner M, *Terrigenous Mass Movements* (Berlin Heidelberg: Springer-Verlag), pp. 23–49, doi: [10.1007/978-3-642-25495-6-2](https://doi.org/10.1007/978-3-642-25495-6-2).
- Pourghasemi H R, Mohammady M and Pradhan B 2012e Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran; *Catena* **97** 71–84.
- Pradhan B 2010a Application of an advanced fuzzy logic model for landslide susceptibility analysis; *Int. J. Comput. Intel. Sys.* **3**(3) 370–381.
- Pradhan B 2010b Landslide susceptibility mapping of a catchment area using frequency ratio, fuzzy logic and multivariate logistic regression approaches; *J. Indian Soc. Rem. Sen.* **38**(2) 301–320.
- Pradhan B 2010c Remote sensing and GIS-based landslide hazard analysis and cross-validation using multivariate logistic regression model on three test areas in Malaysia; *Adv. Space Res.* **45**(10) 1244–1256.
- Pradhan B 2011a 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, doi: [10.1007/s10651-010-0147-7](https://doi.org/10.1007/s10651-010-0147-7).
- Pradhan B 2011b Use of GIS-based fuzzy logic relations and its cross application to produce landslide susceptibility maps in three test areas in Malaysia; *Environ. Earth Sci.* **63**(2) 329–349.
- Pradhan B and Buchroithner M F 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, doi: [10.2113/gsegeosci.16.2.107](https://doi.org/10.2113/gsegeosci.16.2.107).
- Pradhan B and Lee S 2009 Landslide risk analysis using artificial neural network model focusing on different training sites; *Int. J. Physic. Sci.* **3**(11) 1–15.
- Pradhan B and Lee S 2010a 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** 747–759.
- Pradhan B and Lee S 2010b Delineation of landslide hazard areas using frequency ratio, logistic regression and artificial neural network model at Penang Island, Malaysia; *Environ. Earth Sci.* **60** 1037–1054.
- Pradhan B and Pirasteh S 2010 Comparison between prediction capabilities of neural network and fuzzy logic techniques for landslide susceptibility mapping; *Disaster Adv.* **3**(2) 26–34.
- Pradhan B and Youssef A M 2010 Manifestation of remote sensing data and GIS on landslide hazard analysis using spatial-based statistical models; *Arab. J. Geosci.* **3**(3) 319–326.
- Pradhan B, Singh R P and Buchroithner M F 2006 Estimation of stress and its use in evaluation of landslide prone regions using remote sensing data; *Adv. Space Res.* **37** 698–709.
- Pradhan B, Lee S, Mansor S, Buchroithner M F, Jallaluddin N and Khujaimah Z 2008 Utilization of optical remote sensing data and geographic information system tools for regional landslide hazard analysis by using binomial logistic regression model; *Appl. Remote Sens.* **2** 1–11.
- Pradhan B, Lee S and Buchroithner M B 2009 Use of geospatial data for the development of fuzzy algebraic operators to landslide hazard mapping: A case study in Malaysia; *Appl. Geomatics* **1** 3–15.
- Pradhan B, Sezer E, Gokceoglu C and Buchroithner M F 2010a Landslide susceptibility mapping by neuro-fuzzy approach in a landslide prone area (Cameron Highland, Malaysia); *IEEE Trans. Geosci. Remote Sens.* **48**(12) 4164–4177.
- Pradhan B, Youssef A M and Varathrajoo R 2010b Approaches for delineating landslide hazard areas using different training sites in an advanced artificial neural network model; *Geospatial Information Sci.* **13**(2) 93–102.
- Pradhan B, Lee S and Buchroithner M F 2010c Remote sensing and GIS based landslide susceptibility analysis and its cross-validation in three test areas using a frequency ratio model; *Photogramm. Fernerkundung Geoinformation* **1** 17–32, doi: [10.1127/14328364/2010/0037](https://doi.org/10.1127/14328364/2010/0037).
- Pradhan B, Oh H J and Buchroithner M 2010d Weights-of-evidence model applied to landslide susceptibility mapping in a tropical hilly area; *Geomatics, Nat. Hazards Risk* **1**(3) 199–223, doi: [10.1080/19475705.2010.498151](https://doi.org/10.1080/19475705.2010.498151).
- Pradhan B, Mansor S, Pirasteh S and Buchroithner M 2011 Landslide hazard and risk analyses at a landslide prone catchment area using statistical based geospatial model; *Int. J. Remote Sens.* **32**(14) 4075–4087, doi: [10.1080/01431161.2010.484433](https://doi.org/10.1080/01431161.2010.484433).
- Pradhan B, Chaudhari A, Adinarayana J and Buchroithner M F 2012 Soil erosion assessment and its correlation with landslide events using remote sensing data and GIS: A case study at Penang Island, Malaysia; *Environ. Monit. Assess.* **184**(2) 715–727.
- Sezer E A, Pradhan B and Gokceoglu C 2011 Manifestation of an adaptive neuro-fuzzy model on landslide susceptibility mapping: Klang valley, Malaysia; *Expert Syst. Appl.* **38**(7) 8208–8219.
- Shou K J and Wang C F 2003 Analysis of the Chiufengshan landslide triggered by the 1999 Chi-Chi earthquake in Taiwan; *Eng. Geol.* **68** 237–250.
- Talebi A, Uijlenhoet R and Troch P A 2007 Soil moisture storage and hillslope stability; *Nat. Hazards Earth Syst. Sci.* **7** 523–534.
- Tunusluoglu M C, Gokceoglu C, Nefeslioglu H A and Sonmez H 2008 Extraction of potential debris source areas by logistic regression technique: A case study from Barla, Besparmak and Kapi mountains (NW Taurids, Turkey); *Environ. Geol.* **54** 9–22.
- Vahidnia M H, Alesheikh A A, Alimohammadi A and Hosseinali F 2010 A GIS-based neurofuzzy procedure for integrating knowledge and data in landslide susceptibility mapping; *Comput. Geosci.*, doi: [10.1016/j.cageo.2010.04.004](https://doi.org/10.1016/j.cageo.2010.04.004).
- Van Den Eeckhaut M, Vanwalleghem T, Poesen J, Govers G, Verstraeten G and Vandekerckhove L 2006 Prediction of landslide susceptibility using rare events logistic regression: A case study in the Flemish Ardennes (Belgium); *Geomorphology* **76** 392–410.
- Varnes D J 1978 Slope movement types and processes; In: *Landslide analysis and control* (eds) Schuster R L and Krizek R J, Special Report. Transportation Research Board, National Academy of Sciences, New York, **176** 12–33.
- Vapnik V N 1995 *The Nature of Statistical Learning Theory* (New York: Springer Verlag).

- Wang H B and Sassa K 2005 Comparative evaluation of landslide susceptibility in Minamata area, Japan; *Environ. Geol.* **47** 956–966.
- Xu C, Dai F, Xu X and Lee Y H 2012 GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China; *Geomorphology*, doi: [10.1016/j.geomorph.2011.12.040](https://doi.org/10.1016/j.geomorph.2011.12.040).
- Yalcin A and Bulut F 2007 Landslide susceptibility mapping using GIS and digital photogrammetric techniques: A case study from Ardesen (NE-Turkey); *Nat. Hazards* **41** 201–226.
- Yao X, Tham L G and Dai F C 2008 Landslide susceptibility mapping based on Support Vector Machine: A case study on natural slopes of Hong Kong, China; *Geomorphology* **101** 572–582.
- Yesilnacar E K 2005 The Application of Computational Intelligence to Landslide Susceptibility Mapping in Turkey; Ph.D Thesis, Department of Geomatics, University of Melbourne, 423p.
- Yesilnacar E and Topal T 2005 Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey); *Eng. Geol.* **79**(3–4) 251–266.
- Yilmaz I 2009a A case study from Koyulhisar (Sivas–Turkey) for landslide susceptibility mapping by artificial neural networks; *Bull. Eng. Geol. Environ* **68**(3) 297–306.
- Yilmaz I 2009b Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat-Turkey); *Comput. Geosci.* **35**(6) 1125–1138.
- Yilmaz I 2010a The effect of the sampling strategies on the landslide susceptibility mapping by conditional probability (CP) and artificial neural network (ANN); *Environ. Earth Sci.* **60** 505–519.
- Yilmaz I 2010b Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: Conditional probability, logistic regression, artificial neural networks, and support vector machine; *Environ. Earth Sci.* **61** 821–836.
- Youssef A M, Pradhan B, Gaber A F D and Buchroithner M F 2009 Geomorphological hazard analysis along the Egyptian Red Sea coast between Safaga and Quseir; *Nat. Hazards Earth Syst. Sci.* **9** 751–766.
- Zare M, Pourghasemi H R, Vafakhah M and Pradhan B 2012 Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: A comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms; *Arab. J. Geosci.*, doi: [10.1007/s12517-012-0610-x](https://doi.org/10.1007/s12517-012-0610-x).