

A GIS-based logistic regression model in rock-fall susceptibility mapping along a mountainous road: Salavat Abad case study, Kurdistan, Iran

Ataollah Shirzadi · Lee Saro · Oh Hyun Joo · Kamran Chapi

Received: 29 May 2011 / Accepted: 27 July 2012 / Published online: 1 September 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract This study describes the application of logistic regression to rock-fall susceptibility mapping along 11 km of a mountainous road on the Salavat Abad saddle, in southwest Kurdistan, Iran. To determine the factors influencing rock-falls, data layers of slope degree, slope aspect, slope curvature, elevation, distance to road, distance to fault, lithology, and land use were analyzed by logistic regression analysis. The results are shown as rock-fall susceptibility maps. A spatial database, which included 68 sites (34 rock-fall point cells with value of 1 and 34 no rock-fall point cells with value of 0) was developed and analyzed using a Geographic Information System, GIS. The results are shown as four classes of rock-fall susceptibility. In this study, distance to fault, lithology, slope curvature, slope degree, and distance to road were found to be the most important factors affecting rock-fall. It was concluded that about 76 % of the study area can be classified as having moderate and high susceptibility classes. Rock-fall point cells were used to verify results of the rock-fall susceptibility map using success curve rate and the area under the curve. The verification results showed that the area under the curve for rock-fall susceptibility map is 77.57 %. The results from this study demonstrated that the use of a logistic regression model within a GIS framework is useful and suitable for rock-fall susceptibility mapping. The rock-fall susceptibility map can be used to reduce susceptibility associated with rock-fall.

Keywords Rock-fall · Susceptibility map · Logistic regression · Salavat Abad · Kurdistan · Iran

1 Introduction

Rock-falls occur when masses of rock or other material detach from a steep slope or cliff and descend by free fall, sliding, toppling, or bouncing. Rock-fall is one of the main slope

A. Shirzadi (✉) · K. Chapi
University of Kurdistan, Sanandaj, Kurdistan, Islamic Republic of Iran
e-mail: A.Shirzadi@uok.ac.ir

L. Saro · O. Hyun Joo
Geoscience Information Center, Korea Institute of Geoscience & Mineral Resources (KIGAM),
92 Gwahang-No, Yuseong-Gu, Daejeon 305-350, Korea

movement processes in the natural evolution of rocky mountain slopes (e.g., Selby 1982). On February 10, 1996, a huge rock-fall along highway 229 on Hokkaido Island in Japan killed 20 people and captured much international attention to the rock-fall problem (e.g., Yamagishi 2000). It has been known that 10 % of world population lives in mountainous terrains and 36 % of the lands are mountainous. With the high increase in world population, human activities in mountainous areas have correspondingly increased (e.g., Chau et al. 2003). The intensified human activities in mountainous terrains have led to land cover change, deforestation, and more livestock grazing on steep slopes, which can trigger huge mass movements (e.g., Kelarestaghi and Ahmadi 2009). On March 2008, rock-falls on the Hezar Cham saddle in northern Iran, which was triggered by earthquakes and heavy precipitations, killed 21 people and caused heavy damages to the road (<http://www.ngdir.ir>). On February 14, 2006, a rock-fall on the Gheshm-Maigon road in southern Iran has led to the destruction of 1,000 m of the road and damaged 5 houses. According to the reports by the Transport Office of the Kurdistan province, in 2006, the total costs of the damages from rock-falls to the Salavat Abad saddle (the case study) were about 28,000\$ (Fig. 1).

Studies of rock-falls are often based on field surveys, and susceptibility is estimated either by an empirical assessment of susceptibility to failure, or by the calculation of a safety factor derived from models of rock mechanics (e.g., Hoek and Bray 1981). Once a location is identified with rock-fall risk, the probability of the maximum travel distance and maximum energy of impact of rock-fall events at the location is normally assessed using computer simulations (e.g., Wu 1985; Kobayashi et al. 1990; Azzoni et al. 1995).

There are many different models for study on rock-fall events in three main categories: (1) empirical models based on relationships between topographical factors and the length of the run out of rock-falls (e.g., Keylock and Domaas 1999), (2) process-based models that describe or simulate the modes of motion of falling rocks over slope surfaces (e.g., Kirkby and Statham 1975; Statham 1976; Hungr and Evans 1988; Pfeiffer and Bowen 1989; Kobayashi et al. 1990), and (3) GIS-based models that are running within a GIS environment or they are raster-based models for which input data are provided by GIS analysis (e.g., Evans and Hungr 1993; Hegg and Kienholz 1995; Chau et al. 2004). Little work was done on rock-fall susceptibility mapping based on GIS (e.g., Carrara et al. 1995; Chung et al. 1995; Guzzetti et al. 1999; Suzen and Doyuran 2004a, b; Chau et al. 2004). Susceptibility maps are found to be very useful in estimating, managing, and mitigating mass movement susceptibility for a region (e.g., Corominas and Santacana 2003; Chung and Fabbri 2003; Sassa et al. 2004).

There are various methods to susceptibility mapping including the following: semi-qualitative methods similar to the analytical hierarchy process (AHP) (e.g., Barredol et al. 2000), bivariate statistical analysis (e.g., Kelarestaghi and Ahmadi 2009; Nandi and Shakoor 2009), the probability–frequency ratio model (e.g., Lee and Pradhan 2006), and multivariate regression methods such as logistic regression (e.g., Lee and Sambath 2006; Pradhan 2010; Su and Cui 2010; Choi et al. 2012). The recent and rapid increase in computing capacity has also allowed scientists to treat large sets of data, which is a crucial factor in applying multivariate statistical analysis. Multivariate procedures have long been employed for landslide susceptibility mapping (e.g., Reger 1979; Carrara et al. 1992; Gorseveski et al. 2000; Baeza and Corominas 2001; Lee and Min 2001; Ayenew and Barbieri 2005; Can et al. 2005; Chau and Chan 2005; Greco et al. 2007). Among the various susceptibility mapping methods, logistic regression (LR) presents certain advantages for studying the landslides of soil and/or weathered rocks (e.g., Gorseveski et al. 2000; Dai and Lee 2002; Chau et al. 2004; Ayalew and Yamagishi 2004; Lee and Sambath 2006; Akgun and Bulut 2007; Akgun et al. 2008; Lamelas et al. 2008). Since very few

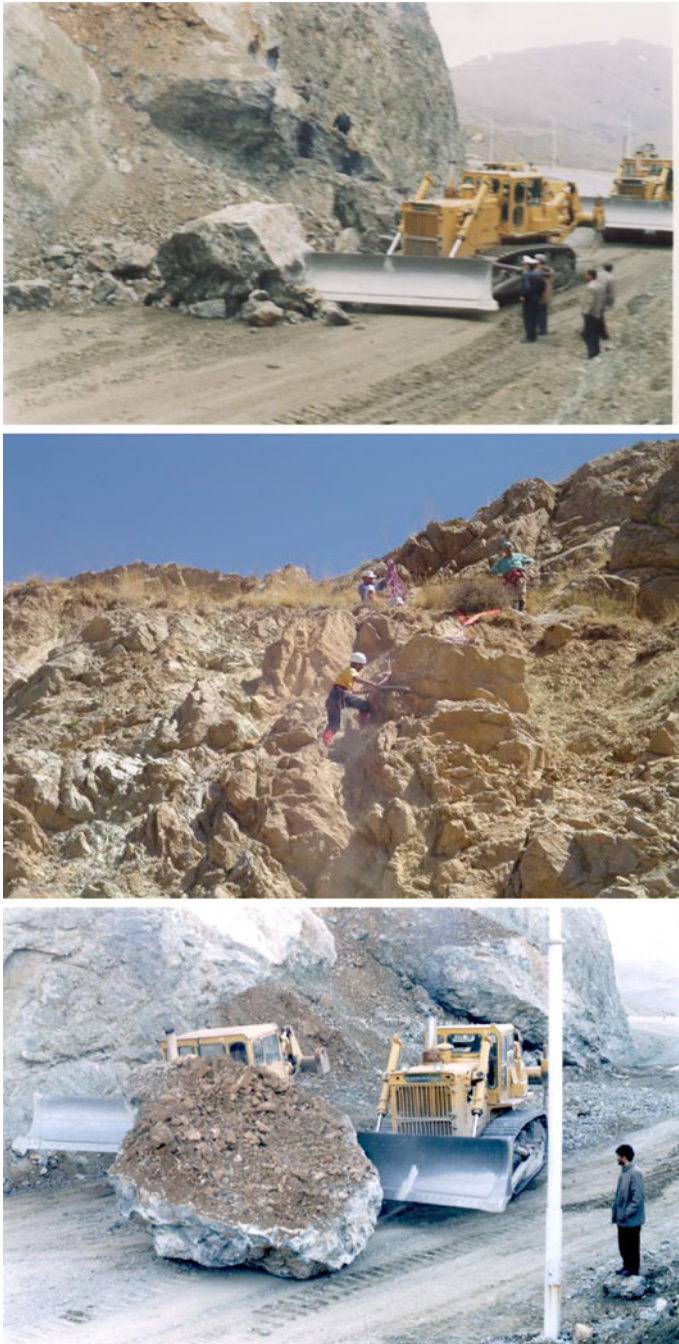


Fig. 1 Rock-falls cleanup on Salavat Abad road at Kurdistan, Iran

studies on rock-fall susceptibility mapping based on GIS and statistical analysis have been noticed (e.g., Chau et al. 2004); we decided to further study this topic. Thus, the aim of this research is to apply and assess a logistic regression model to generate rock-fall susceptibility map along a mountainous road in the western parts of Iran.

2 Materials and methods

2.1 Topographic and geologic setting of the study area

The Salavat Abad saddle is located between longitudes of $47^{\circ}, 07', 48''$ and $47^{\circ}, 08', 57''$ N and latitudes of $35^{\circ}, 15', 16''$ and $35^{\circ}, 19', 40''$ E, in the southwest parts of the Kurdistan province, Iran (Fig. 2). It covers an area of about 1871 hectares. The elevation is varying from 1,699 to 2,500 m above sea level. The study area selected is a part of the Zagros Mountainous region, whose geologic setting resulted from tectonic activities at the end of tertiary period. The Salavat Abad saddle is the effect of internal tectonization by orogen on the Zagros Mountains. Geology units in the study area include the following: North–South trending andesitic and basaltic rocks, limestone and red conglomerates (Fig. 3). The main lithology is the limestone occupying nearly 94 % of the study area. The Salavat Abad saddle has an N–S trend and is thrust over volcanic deposits. The limestone in the study area has been affected by mechanical weathering, snow melt, and frequent freezing so that several rock-falls have taken place at the section of this road.

One step in generating a rock-fall distribution map was taken in the laboratory, and another one was taken in the field. Recognition of rock-falls on aerial photographs is difficult because they are located on steep slopes and artificial slopes and easily confused with man-made objects. Thus, they are not easily identified similarly to the landslides that have occurred at a scale larger than that of the falls. Therefore, the recognition of rock-falls in the study was done directly in the field. The collection of the rock-falls locations was recorded by the Transport Office of the Kurdistan province in 2006. Based on the discontinuity in the rock-falls, the geologists at the Transport Office of the Kurdistan province have recorded the central of each rock-fall slopes as one of the 34 rock-falls in the study area. When the 34 rock-falls were selected, 50 no-rock-falls locations were recorded by the same method as the 34 rock-falls were selected. Finally, 34 out of 50 no-rock-falls were

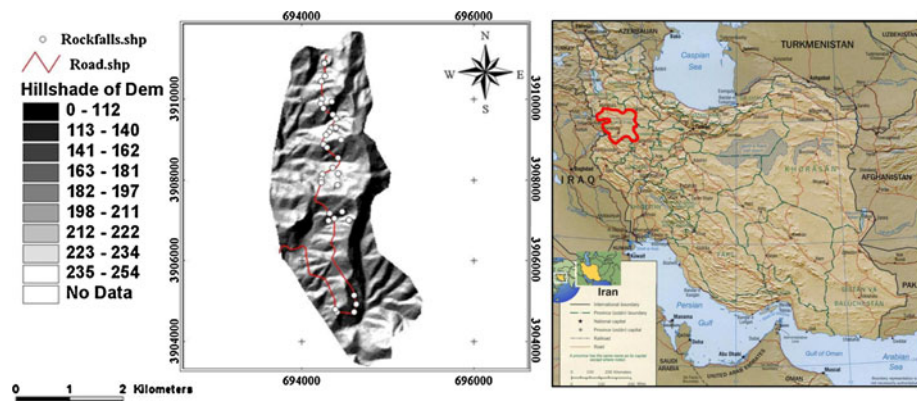


Fig. 2 Location of rock-falls in the study area and Iran

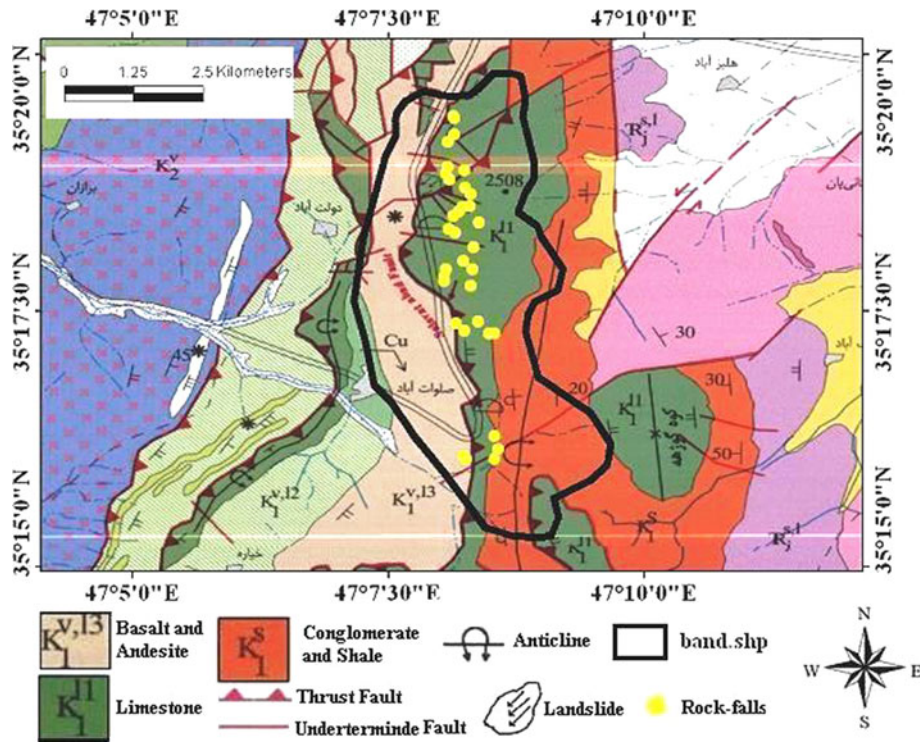


Fig. 3 Geology map of the study area

randomly selected. The data in the report were verified using field surveying and slopes instability observations. The rock-falls that had occurred on the slopes had high-density cracks and joints so that on some of the slopes in the case study area, the dimensions of cracks and joints were comparatively large causing the rocks to fall downward the slope and be collected in its toe (Fig. 1).

The most significant fault in the study area is the Salavat Abad deep active fault. It passes through the entire area from the south to the north. The distances of all rock-falls to the fault section were calculated and classified for the convenience of interpretation. Generally, rock masses near the Salavat Abad fault zone are less intact and severely fractured. According to our investigation, the Salavat Abad deep fault is a dominant factor controlling the development of rock-falls (Fig. 4).

2.2 Data and methods

First of all, eight effective factors on rock-falls were recognized based on the field work including social research techniques, that is, local interviews and comprehensive literature reviews. They are as follows: slope degree, slope aspect, slope curvature, elevation, distance to road, distance to fault, lithology, and land use. The dependent variable was the presence of rock-falls at the thirty-four locations used as sites considered as susceptible to rock-fall along the 11 km of a road in the Salavat Abad saddle. The Sanandaj and Halizabad 1:25,000 topography maps were geo-referenced with the Erdas Imagine 8.5 Software. Contour lines were obtained from the 1:25,000 topographic maps with interval of

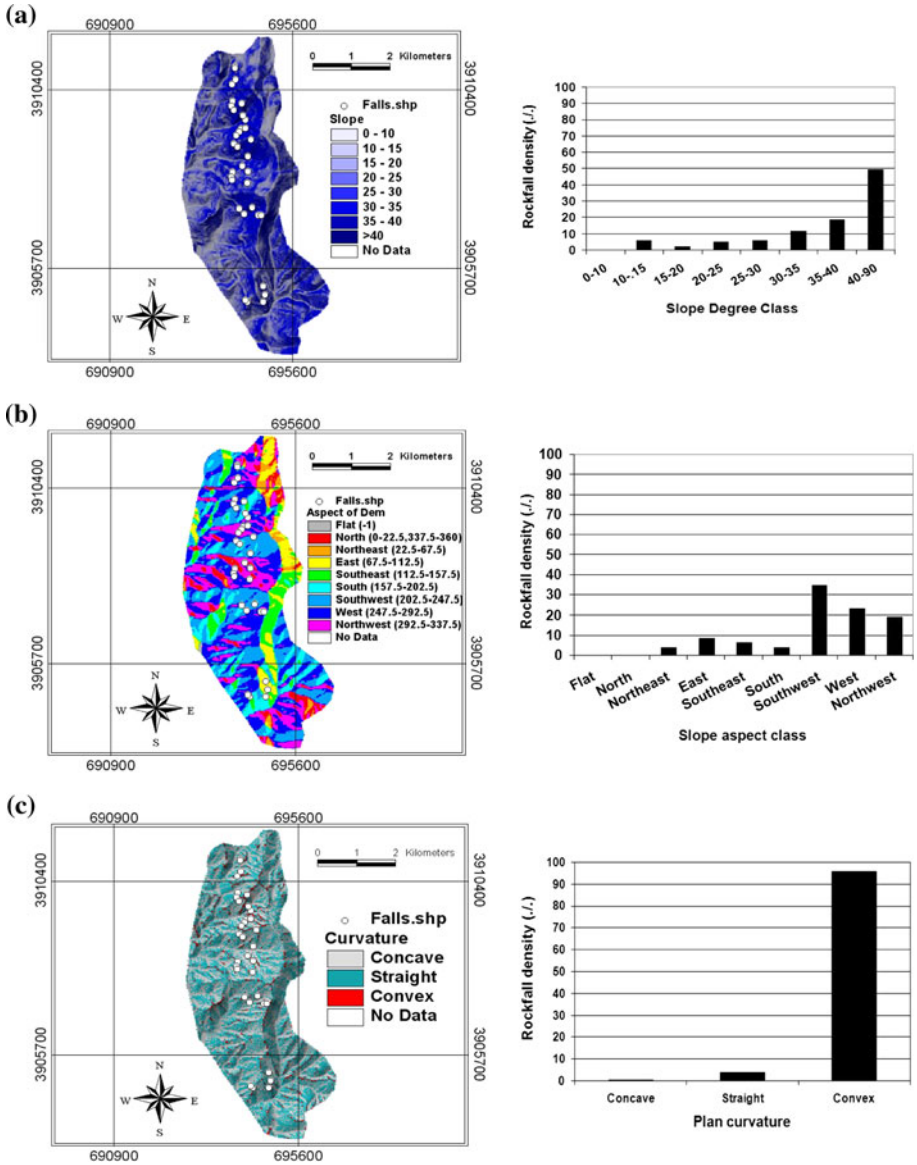


Fig. 4 Independent variables maps and histogram of rock-fall density for each variable including the following: **a** slope angle, **b** slope aspect, **c** slope curvature, **d** elevation above sea level, **e** distance to road, **f** distance to fault, **g** lithology

10 m using the Arc View 3.2 software. The digital elevation model has been produced from the digitized contour lines with 20 × 20 m pixel resolution, using a triangular irregular network (TIN) model. Slope degree, slope aspect, slope curvature, and elevation maps were derived from the digital elevation model (DEM). Distance to road and fault were calculated using conventional GIS functions. Lithology map with the same resolution was obtained from a 1:100,000 geology map (Table 1). The 34 rock-fall areas were

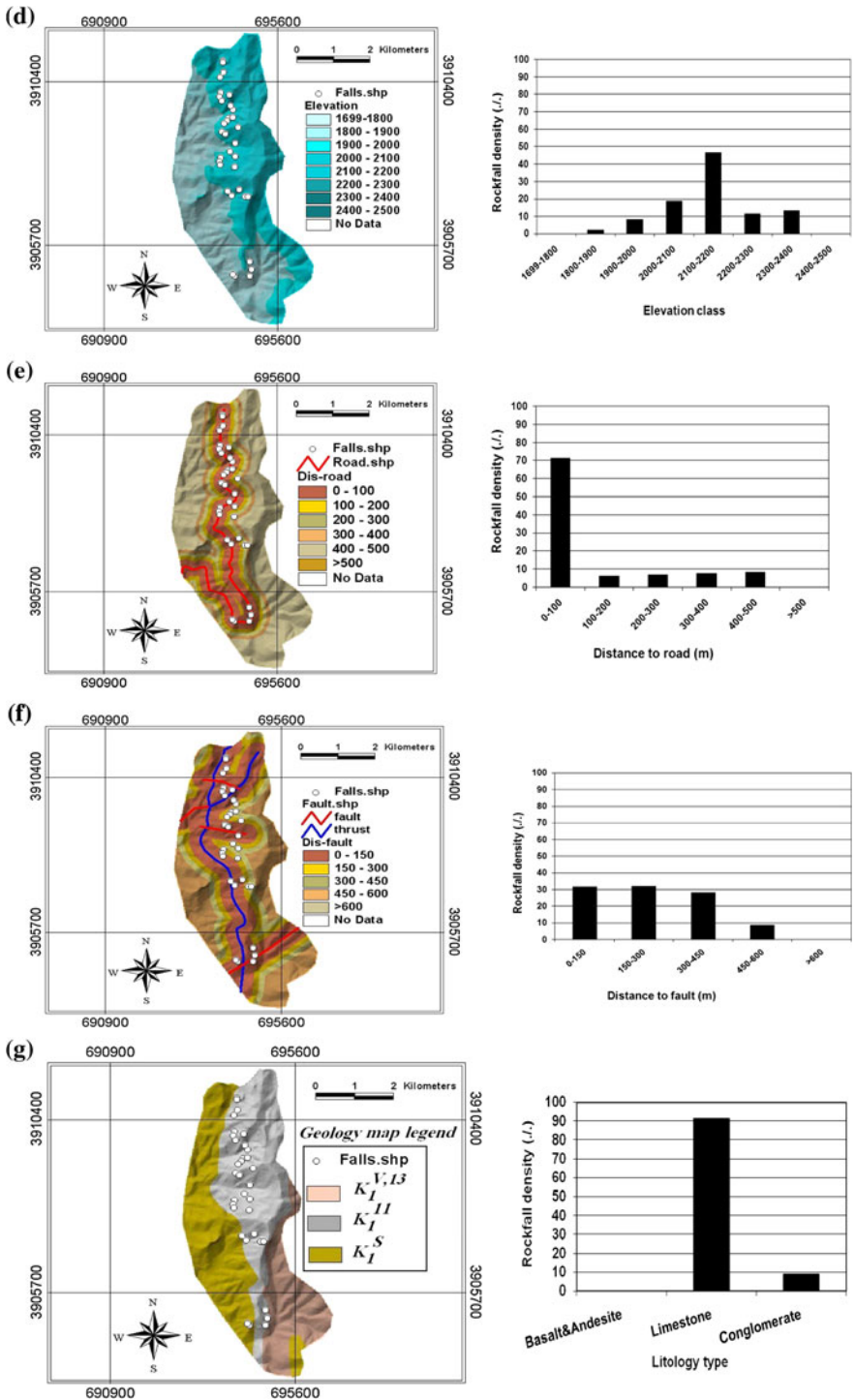


Fig. 4 continued

Table 1 Data layer of study area

Classification	Sub-classification	Extracted factor	GIS data type	Scale	References
Basic map	Rock-fall	Rock-falls inventory	Point coverage	1:25,000	Field works
	Topography	Slope gradient	Grid	20 m × 20 m	Iran Cartographic Organization
		Slope aspect	Grid	20 m × 20 m	
		Slope curvature	Grid	20 m × 20 m	
	Geology	Lithology	Polygon coverage	1:100,000	Iran Geological Organization
		Fault	Line coverage		
	Land Sat ETM image	Land use	Grid	20 m × 20 m	Iran Space Agency

digitized to produce a rock-fall distribution map. In this study, each 20-m pixel represents a rock-fall. As mentioned before, 34 pixels (rock-fall or slope) were selected to establish the relationship between rock-fall presence and the presence of factor. All data layers were rasterized to the DEM resolution using Arc View 3.2. The rock-fall locations that are used in present study were mainly extracted from the Transport Office of the Kurdistan province as well as from past reports and field verification. In this study, we have used a basic data set prepared by the Iran Cartographic Center, the Iran Geological Organization and the Iran Space Agency. The data are of relatively high quality. Table 1 shows some detailed information about the data used.

2.3 Logistic multiple regression

Logistic multiple regression is a multivariate technique, which consider several physical parameters that may affect the probability of rock-fall occurrence. The regression is considered here ideal to integrate both the binary and the scalar values as the independent variables, which allows for the use of variables that are not continuous or qualitatively derived. In the present study, the dependent variable is a binary variable representing the presence or absence of rock-falls. The technique of logistic multiple regression yields coefficients for each variable based on data derived from samples taken by the study area. These coefficients serve as weights in an algorithm, which can be used in the GIS database to produce a map depicting the probability of rock-fall occurrence. Quantitatively, the relationship between the occurrence and its dependency on several variables can be expressed as:

$$P(\text{event}) = \frac{1}{(1 + \exp^{-z})} \quad (1)$$

where p (event) is probability of an event occurring. In the present, p (event) is the estimated probability of rock-fall occurrence. As Z varies from $-\infty$ to $+\infty$, the probability varies from 0 to 1. Z is the linear combination:

$$z = B_0 + B_1X_1 + B_2X_2 + \cdots + B_nX_n \quad (2)$$

General equation:

$$Y = \log it(p) = \ln\left(\frac{P}{1-p}\right) = B_0 + B_1X_1 + B_2X_2 + \dots + B_nX_n \tag{3}$$

where Y is probability of rock-fall occurrence, B_n ($i = 0, 1, \dots, n$) is the coefficient estimated from the sample data, n is the number of independent variables, and X_n ($i = 0, 1, \dots, n$) is the independent variables. In logistic multiple regression, a coding scheme should be selected for the categorical variables that by creating a new set of variables that correspond in some way to the original categories. The number of new variables required to present a categorical variables is one less than that of the number of categories. The coefficients of the logistic multiple regression models are estimated using the maximum-likelihood method. In other words, the coefficients that make the observed results most “likely” are selected. Since the relationship between the independent variables and the probability is nonlinear in the logistic multiple regression model, an iterative algorithm is necessary for parameter estimation. Logistic multiple regression modeling is intended to describe the likelihood of rock-fall occurrence on a regional scale and is very suitable for the assessment of slope instability, since the observed data consist of locations (points) or cells with a value of 0 (absence of rock-fall) or 1 (presence of rock-fall). This method allows a spatial distribution of probabilities or susceptibility values to calculate within the GIS environment.

3 Results

3.1 Rock-fall susceptibility modeling

A multiple logistic regression model was constructed initially based on the physical parameters (Table 2). In the present analysis, the likelihood-ratio test is used for determining whether variables should be added to the model. If the observed significance level is greater than the probability for remaining in the model (0.1 in this study), the variable is excluded from the process and the processing statistics are recalculated to see whether any other variables are eligible for excluding. Both slope aspect and land use were found to be not significant and were thus eliminated from the stepwise procedure. Not that all the variables in the analysis are binary variables representing presence or absence of the corresponding mapping units or intervals. For each variable, the last category is used as the default reference category, and coefficient of that map unit or category is thus overridden. Coefficients, significance, and Exp (β) are shown in Table 3. Logistic regression equations, Eq. (4) for the Salavat Abad saddle were obtained:

$$p(\text{Rockfall-Salvatabad}) = \frac{1}{1 + \exp^{-(2.189-2.705(A3)+3.135(E2)-2.832(F2)+4.976(G4)-1.672(H3)}} \tag{4}$$

For the rock-fall susceptibility map (Fig. 6), in the study area, the values obtained from applying the logistic multiple regressions (Eq. 4) were transferred to the Arc View GIS 3.2 Software system and applied to the independent variables representing the present condition for each pixel or cell within the study area. At last, the range of values of susceptibility to rock-falls in the susceptibility map was classified into 4 categories: very low (0–0.25), low (0.25–0.35), moderate (0.35–0.5), and high (>0.5). The ranges of the

Table 2 List of independent variables used in logistic regression

Slope degree	Symbol	Slope aspect	Symbol	Elevation	Symbol	Land use	Symbol
0–10	A1	Flat	B1	1,699–1,800	C1	Garden	D1
10–15	A2	North	B2	1,800–1,900	C2	Garden and rang	D2
15–20	A3	East north	B3	1,900–2,000	C3		D3
20–25	A4	East	B4	2,000–2,100	C4	Rocky area	D4
25–30	A5	East south	B5	2,100–2,200	C5	Salavat Abad village	D5
30–35	A6	South	B6	2,200–2,300	C6		
35–40	A7	West south	B7	2,300–2,400	C7	Semi-density range	
>40	A8	West	B8	2,400–2,500	C8		
		West north	B9				

Lithology	Symbol	Distance to road	Symbol	Distance to fault	Symbol	Slope curvature	Symbol
Basalt and andesite	E1	0–100	F1	0–150	G1	Concave	H1
	E2	100–200	F2	150–300	G2	Straight (flat)	H2
Limestone	E3	200–300	F3	300–450	G3		H3
Conglomerate and shale		300–400	F4	450–600	G4	Convex	
		400–500	F5	>600	G5		
		>500	F6				

individual categories were derived based on histogram of the estimated susceptibility to rock-fall shown in Fig. 5 (Dai and Lee 2002). At first, the rock-fall susceptibility map was divided into 10 categories. This map was then overlaid with presence and absence of rock-falls, and this histogram was obtained. Two categories from 0 to 0.5 (left side as low susceptibility) and from 0.5 to 1 (right side as high susceptibility) are shown. The left side was divided into three categories based on absence of rock-falls, and the right left was divided into one category. Finally, the susceptibility map was classified into 4 categories: very low (0–0.25), low (0.25–0.35), moderate (0.35–0.5), and high (>0.5). The rock-fall susceptibility map was divided into 10 categories. Theoretically, if we have a model that successfully distinguishes the two groups based on a classification cutoff value of 0.5, the cases for which rock-falls have occurred should be to the right of 0.5, while the cases for which rock-falls have not occurred should be to the left of 0.5.

3.2 Susceptibility map and its reliability

The actual calculation was done using the “Map Calculator” and “Map Query” functions in Arc/view 3.2. The β coefficients had been obtained from the statistical analysis. Then, β values were assigned to each category of variable (Table 3). After assignment of β value to each category, we can combine all layers (Fig. 4) by using Eqs. (1) and (2) to yield the susceptibility map (Fig. 6). The fitted logistic regression model can be used to calculate the probability of rock-fall occurrence of both the rock-fall and the non-rock-fall. Whenever the probability of a rock-fall is larger than 0.5, the rock-fall is considered likely; and whenever the probability of a rock-fall is less than 0.5, that rock-fall is considered unlikely. Thus, when the probability in (1) for the any of original rock-fall is larger than 0.5, the prediction is considered acceptable. Similarly, if $p < 0.5$ for the case of non-rock-fall, the

Table 3 The coefficients, significance, and Exp (β) for logistic regression in this study

Independent parameter	Class	Coefficient	Significant	Exp (β)
Slope angle	15–20°	−2.705	0.033	0.067
Geology	Limestone	3.135	0.000	22.978
Distance to road	100–200 m	−2.832	0.027	0.059
Distance to fault	450–600 m	4.976	0.011	144.825
Slope curvature	Convex	−1.672	0.025	0.188
Constant		−2.189	0.005	0.112

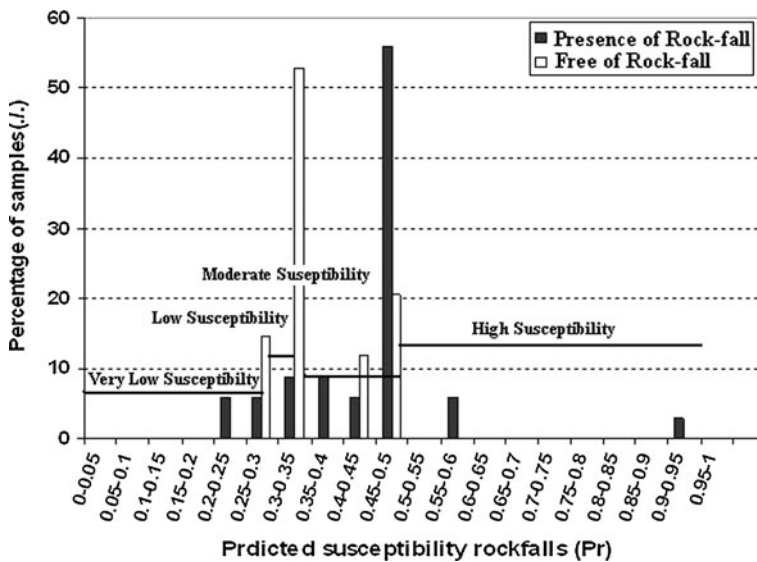


Fig. 5 Histogram of the predicted rock-fall susceptibility for the samples used in this study

prediction is viewed as successful. In validation, the accurate rate is 85.3 % for rock-fall group and 75.1 % for non-rock-fall group using error matrix method. The total accurate rate is 79.1 %, which is considered acceptable (Table 4).

In the other method, we have conducted another series of logistic regression analysis by excluding each of the categories of slope degree, slope aspect, slope curvature, elevation, distance to road, distance to fault, lithology, and land use. This is a standard technique in logistic regression analysis to examine the statistical significance of each of the selected independent variables (e.g., Hosmer and Lemeshow 2000; Menard 2002). It was found that $-2LL$ is the smallest if all variables are included. The $-2LL$ can be considered as a goodness of fit for the model. Since the likelihood is between 0 and 1, log likelihood ranges from negative infinity to zero, and in turn, the $-2LL$ would be from 0 to infinity. Therefore, the original analysis of incorporating all variables should be used. Table 5 has shown that $-2LL$ for all eight variables is 57.445, and for slope aspect, elevation and land use, this value is the same. In other words, these three variables are not significant at 99 % level of probability with rock-falls and excluded from the final model. However, there is no universal standard for this parameter to be used for validity check. Another approach is to

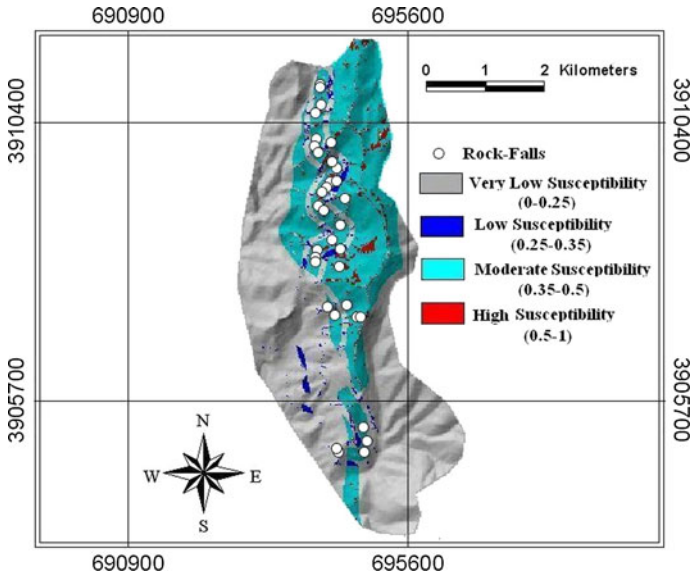


Fig. 6 Rock-fall susceptibility map of the Salavat Abad saddle

Table 4 Classification table and predicted percentage correct test

Observed	Predicted		Correction percentage (%)
	Absence of rock-fall (0)	Presence of rock-fall (1)	
Absence of rock-fall (0)	24	10	70.6
Presence of rock-fall (1)	5	29	85.3
			77.9

use the coefficient of correlation. Similar to the R^2 in linear regression, there are also correlation coefficients for logistic regression analysis and they are called Cox and Snell R^2 (e.g., Cox and Snell 1989) and Nagelkerke R^2 (e.g., Nagelkerk 1991). For the present model, they are 0.418 and 0.558, respectively (Table 5). The theoretical values of these coefficients are again from 0 to 1. Unlike the linear regression, these coefficients can be relatively small (this does not necessarily invalidate the model in the case of logistic regression) and there is again no universal standard of what value of Cox and Snell R^2 and Nagelkerke R^2 should be obtained in the regression to be acceptable, that $R^2 > 0.9$ is normally considered as a good indicator of a reasonable fit.

For validation of rock-fall susceptibility calculation models, two basic assumptions are needed. Firstly, rock-falls are related to spatial information such as topography, geology, road, and fault, and secondly, future rock-falls will be precipitated by a specific impact factor such as earthquake. In this study, the two assumptions are satisfied because the rock-falls were related to the spatial information, and the rock-falls were precipitated by one road and fault in the study area. Also, the susceptibility map was verified using existing rock-fall location. For this, the logistic regression analysis result in modeling rock-fall susceptibility in the study area was evaluated by calculating the success curve rate (SCR) and the percentage of known rock-falls in the various susceptibility categories. For the

Table 5 Some statistics and map accuracy evaluation

Independent variable	−2log likelihood (−2LL)	Cox and Snell R^2	Nagelkerke R^2	(AUC %)
All variables	57.445	0.418	0.558	77.57
Without slope degree	59.904	0.397	0.529	71.80
Without slope curvature	62.985	0.396	0.492	73.97
Without elevation	57.445	0.418	0.558	73.11
Without distance to road	72.230	0.277	0.369	66.17
Without distance to fault	72.390	0.275	0.367	70.24
Without lithology	62.167	0.376	0.502	72.93

verification by the SCR method, the area under the SCR curve (AUC) model was used (e.g., Lee and Sambath 2006; Lee and Pradhan 2006; Oh et al. 2009; Pradhan and Lee 2009, 2010; Nandi and Shakoor 2009; Pradhan 2010; Oh and Lee 2011). To obtain the SCR, the calculated rock-fall susceptibility index values of all pixels in the study area were ranked in a descending order. Then, the ranked pixel values were divided into 100 classes with accumulated 1 % equal area intervals. The rate curves explain how well the method and factors classify rock-falls. The area under the curve (AUC) can be used to assess the prediction of the model. Total area = 1 denotes perfect prediction accuracy. The rate verification results appear as graph in Fig. 7. For the case of this study, the area under the curve was 0.7757 and the prediction accuracy was 77.57 % (Table 5; Fig. 7). For example, in the case of our application to the study area, the 90–100 % class with the highest 10 % of probability of a rock-fall contains 15 % of the rock-falls in that area. The 80–100 % class with the highest 20 % of probability of a rock-fall contains 53 % of the rock-falls in the area. The 70–100 % class with the highest 30 % of probability of a rock-fall contains 71 % of the rock-falls in that area (Fig. 7).

In addition to success rate and AUC to investigate the reliability of the hazard analysis, another way is to consider the ratio of the rock-fall location points versus the rock-fall susceptibility. Lee and Min (2001) suggested and used this method for landslide

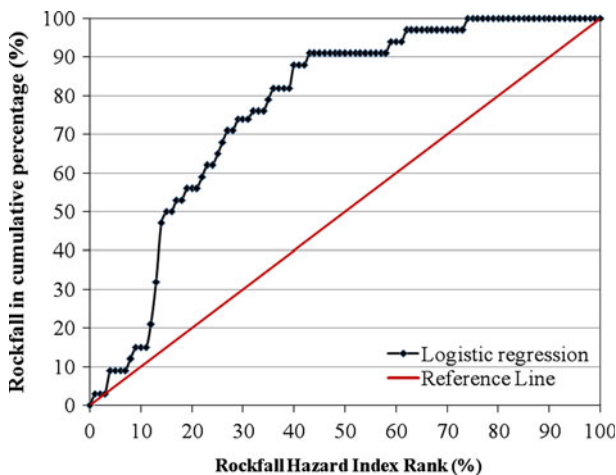


Fig. 7 Cumulative frequency diagram showing rock-fall susceptibility index rank occurring in cumulative percent of rock-fall occurrence (AUC = 77.57 %)

Table 6 Comparison of rock-fall occurrence and rock-fall susceptibility map using logistic regression method for the case study

Range of susceptibility map	Average of susceptibility map	Number of pixels	Percentage of pixels (a, %)	Number of rock-falls	Percentage of rock-falls (b, %)	b/a
0–0.228	0.266	15,782	33.74	4	11.764	0.348
0.228–0.380	0.341	15,021	32.11	3	8.823	0.274
0.380–0.458	0.419	4,247	9.08	3	8.823	0.971
0.458–0.535	0.496	10,563	22.58	21	61.764	2.73
0.535–0.613	0.574	671	1.43	2	5.882	4.11
0.613–0.690	–	–	–	–	–	–
0.690–0.768	–	–	–	–	–	–
0.768–0.845	–	–	–	–	–	–
0.845–0.923	–	–	–	–	–	–
0.923–1	0.961	497	1.06	1	2.944	2.77
Sum		46,781	100	34	100	

susceptibility (e.g., Chau and Chan 2005). At this way, rock-fall probability should first be divided into intervals such that within each of them the amount of coverage area is about the same (Table 6). As shown in Table 6, the susceptibility and probability of larger than 0.5 roughly corresponds to $b/a > 1$. For example; in the first, second, and third class of the susceptibility and probability less than 0.5, the amount of $b/a < 1$, and for the other classes, the susceptibility and probability more than 0.5, the amount of $b/a > 1$, which indicates that the present logistic regression provides a reliable hazard or susceptibility map for the case study.

The accuracy of the rock-fall susceptibility map provided by the model is further checked by calculating Seed Cell Area Index (SCAI), which was suggested by Suzen and Doyuran (e.g., Suzen and Doyuran 2004a, b) as follows:

$$\text{SCAI} = \frac{\text{Area extent of susceptibility class (\%)}}{\text{rockfall in each susceptibility class (\%)}}$$

The SCAI lies in the correct classification of seed cells within a very conservative areal extent, and it is expected that the high and very high susceptibility classes should have very small SCAI values, and that low and very low susceptibility classes should have much higher SCAI values (e.g., Aykut and Necdet 2011). The resulting values are the SCAI densities of rock-fall among the class that are shown in Table 7.

4 Discussion and conclusion

The aim of this study was to assess efficiency of logistic regression model for rock-fall susceptibility mapping along a mountainous road in the Kurdistan province, Iran. Rock-falls are natural phenomena that often have detrimental consequences. In susceptibility management, rock-fall susceptibility map can help to effectively prevent and manage susceptibility. Many qualitative and quantitative techniques are useful for analyzing the relationship between rock-falls and their affective parameters. In this research, we attempted to provide rock-fall susceptibility maps using the relationship between rock-fall

Table 7 Rock-fall distribution in predicted rock-fall susceptible zone

Rock-fall susceptible zone	%Area of predicted zone	%Area of observed rock-fall per class	SCAI (rock-fall density)
Very low	42.12	14.72	2.8614
Low	13.21	8.82	1.4977
Moderate	38.15	64.67	0.5899
High	6.52	11.79	0.5530

locations and determining parameters. The logistic regression model was applied to study the impact of different parameters on rock-fall and susceptibility map of the area. The first results of the logistic regression were the model statistics and coefficients, which were useful to assess the accuracy of the regression function and the role of parameters on the presence or absence of landslides. The forward condition stepwise method was applied for statistical analyses. Slope aspect, land use, and elevation were considered to be not significant for predicting rock-fall and were excluded from the final model. For the study area, in particular, the most important effective factor is 450–600 m distance to fault with coefficient of 4.976 and 0.011. p Value for significance and $\text{Exp}(\beta)$ equal to 144.825, which means that if distance to fault increases by one unit, the value for rock-fall occurrence will increase 144.825 in time. Other factors are geology (limestone), slope curvature (convex), slope angle (15–20°), and distance to road (100–200 m), respectively. Distance to fault and geology by having positive coefficient and high value for $\text{Exp}(\beta)$ has more influence than other factors. The results of logistic regression model were validated using some kinds of validation strategies and were accepted. The results are in line with those of Chau et al. (2004) and Chau and Chan (2005) in their study. Considering coefficients estimated for the logistic regression (Table 3), the “closeness to roads” parameter was found to have the strongest relationship with rock-fall occurrence. Ayalew et al. (2005) have introduced “proximity to roads” parameter as the most important factor on landslide occurrence in Kakuda-Yahico, central Japan. They declared that most of the landslides located in the range 0–100 m from roads. Also, Lee and Sambath (2006), Greco et al. (2007) and Kelarestaghi and Ahmadi (2009) have emphasized adverse effect of road construction on landslide occurrence in their studies. Table 2 indicates that geological units belonging to the Cretaceous era and consisting mostly of limestone are more susceptible to rock-fall. Can et al. (2005) and Nefeslioglu et al. (2008) emphasized causative role of geological units on mass movements in their researches. Land use was excluded from the logistic regression model runs because land use units do not show much change in the study area and most of the rock-falls have occurred in rocky areas. In the Salavat Abad saddle, the Salavat Abad fault by a thrust mechanism led to the generation of joints and cracks of different sizes so that by continued physical weathering and thawing, they have caused rock-falls due to gravity toward the toe of the slope. The road in the region has been encompassed by this fault spiral form, and most of the rock-falls have occurred around it. Therefore, the presence of the fault is one of the most important effective factors for rock-falls occurrence in the study area. The removal of the slope aspect from the logistic regression model run is caused by the presence of the Salavat Abad fault. The role of the Salavat Abad fault was very predominating so that slope aspect did not show any statistical relationship with rock-falls: some of the rock-falls occurred on the west slope instead of on the east slope. The rock-fall susceptibility map was classified into four categories as follows: very low, low, moderate, and high. The rock-fall susceptibility map was assessed

using the area under the curve (AUC) in ROC value (Fig. 7). This diagram shows the 77.57 % as the value of the area under the curve (AUC), which indicates that in the study area; the rock-fall susceptibility map is highly accurate.

To quantitatively compare the result, the areas under the curves (AUC) were individually recalculated for all factors. The results showed that with all effective factors in the logistic regression model, the AUC value rises to 77.57 %, and the values obtained for the other variables are as follows: 71.80 %, slope aspect, 73.97 %, slope curvature, 73.11 %, elevation above the sea level, 66.17 %, geology, 70.24 %, distance to road, and 72.93 %, distance to fault. So, all variables in the rock-fall susceptibility map appear effective, and the prediction of the susceptibility map was measured in this way. These results are in agreement with the results of Lee and Sambath (2006), Lee and Pradhan (2006, 2007), Oh et al. (2009), Jadda et al. (2009), Nandi and Shakoor (2009), Oh et al. (2009), Pradhan (2010), Pradhan and Lee (2010), Chauchan et al. (2010) and Oh and Pradhan (2011) to apply the success rate curve and AUC to investigate the reliability of the landslides susceptibility map. We have also found that 14.72, 8.82, 64.67, and 11.79 % of the area is located at very low, low, moderate, and high susceptible zones, respectively.

Table 6 indicates that rock-fall susceptibility and probability has good accordance with rock-falls because in the classes with the susceptibility and probability less than 0.5, the amount of $b/a < 1$, and in the classes with the susceptibility and probability more than 0.5, the amount of $b/a > 1$. This result is in line with the result of Chau and Chan in evaluating landslide susceptibility map using logistic regression.

The SCAI values given in Table 7 show that the map generated is adequate because the high and moderate susceptibility classes have very low SCAI values, whereas the SCAI values of the very low and low susceptibility classes are very high (Table 7).

In general, in the study area, in addition to natural parameters including “slope gradient,” “slope curvature,” “geology,” and “closeness to fault,” human activities have played a major role on rock-falls. The susceptibility map produced is here considered as acceptable as a basis for studies on mass movement risk management in the study area. If the susceptibility map was overlaid with a vulnerability map (an inventory of building, infrastructure, and other elements at risk and of their expected losses), an objective risk assessment could be achieved. The information derived from this map can help citizens, planners, and engineers to reduce losses caused by existing and future rock-falls by means of prevention and mitigation.

Acknowledgments We wish to thank Prof. Yamagishi Hiromitsu, Ehime University, Japan and Prof. Kam Tim Chau, the Hong Kong Polytechnic University of China, for their advices and suggestions on an earlier version of this manuscript. The authors also wish to thank the Transport Office of the Kurdistan province for the report of rock-falls location in the study area, and the University of Kurdistan, and the University of Mazandaran for their financial supports.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Akgun A, Bulut F (2007) GIS-based landslide susceptibility for Arsin-Yomra (Trabzon, North Turkey) region. *Environ Geol* 51:1377–1387
- 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* 54:1127–1143

- Ayalew L, Yamagishi H (2004) The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakud-Yahiko Mountains, Central Japan. *Geomorphology* 65:15–31
- Ayalew L, Yamagishi H, Marui H, Kanno T (2005) Landslides in Sado Island of Japan: Part II. GIS-based susceptibility mapping with comparisons of results from two methods and verifications. *Eng Geol* 81:432–445
- Ayeneu T, Barbieri G (2005) Inventory of rock-fall and susceptibility mapping in the Dessie area, northern Ethiopia. *Eng Geol* 77:1–15
- Aykut A, Necdet T (2011) Mapping erosion susceptibility by a multivariate statistical method: a case study from the Ayvalik region, NW Turkey. *Comput Geosci* 37:1515–1524
- Azzoni A, Barbera GL, Zaninetti A (1995) Analysis and prediction of rock-falls using a mathematical model. *J Rock Mech Min Sci Geomech Abstr* 32:709–724
- Baeza C, Corominas J (2001) Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surf Proc Land* 26:1251–1263
- Barredol JI, Benavides A, Herhl J, Van Westen CJ (2000) Comparing heuristic landslide susceptibility assessment techniques using GIS in the Tirajana basin, Gran Canaria Island, Spain. *J Appl Earth Observ Geoinf* 2:9–23
- Can T, Nefeslioglu HA, Gokceoglu C, Sonmez H, Duman Y (2005) Susceptibility assessments of shallow earth flows triggered by heavy rainfall at three catchments by logistic regression analysis. *Geomorphology* 82:250–271
- Carrara A, Cardinali M, Guzzetti F (1992) Uncertainty in assessing landslide susceptibility and risk. *J Interdenominat Theol Center (ITC)* 2:172–183
- Carrara A, Cardinali M, Guzzetti F, Reichenbach P (1995) GIS technology in mapping landslide susceptibility. In: Carrara A, Guzzetti F (eds) *Geographical information systems in assessing natural hazards*. Kluwer, Dordrecht, pp 135–175
- Chau KT, Chan JE (2005) Regional bias of landslide data in generating susceptibility maps using logistic regression for Hong Kong Island. *J Rock Mech Rock Eng* 2(4):280–290
- Chau KT, Wong RHC, Lin J, Lee CF (2003) Rock-fall susceptibility analysis for Hong Kong based on Rock-fall Inventory. *J Rock Mech Rock Eng* 36(5):383–408
- Chau KT, Tang YF, Wong RHC (2004) GIS-Based Rock-fall susceptibility map for Hong Kong. *J Rock Mech Min Sci* 41(3):530–542
- Chauchan Sh, Sharma M, Arora MK, Gupta NK (2010) Landslide susceptibility zonation through rating derived from artificial neural network. *J Appl Earth Observ Geoinf* 12:340–350
- Choi J, Oh HJ, Lee HJ, Lee Ch, 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 CF, Fabbri AG (2003) Validation of spatial prediction models for landslide hazard mapping. *Nat Hazards* 30:451–472
- Chung CF, Fabbri AG, Van Westen CJ (1995) Multivariate regression analysis for landslide susceptibility zonation. In: Carrara A, Guzzetti F (eds) *Geographical information systems in assessing natural hazards*. Kluwer, Dordrecht, pp 107–133
- Corominas J, Santacana N (2003) Stability analysis of the Vallcebre translational slide Eastern Pyrenees (Spain) by means of a GIS. *Nat Hazard* 30(3):473–485
- Cox DR, Snell EJ (1989) *Analysis of binary data*, 2nd edn. Chapman and Hall, London
- Dai FC, Lee CF (2002) Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology* 42:213–228
- Evans SG, Hungr O (1993) The assessment of rock-fall hazard at the base of talus slopes. *Can Geotech J* 30:620–636
- Gorseveski PV, Gessler P, Foltz RB (2000) Spatial prediction of Landslide susceptibility using logistic regression and GIS. In: 4th International conference on integrating GIS and environmental modeling, Alberta, 9 pp
- Greco R, Sorriso-Valvo M, Catalano E (2007) Logistic regression analysis in the evaluation of mass movement's susceptibility case study: Calabria, Italy. *Eng Geol* 89:47–66
- Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Rock-fall susceptibility evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology* 31:181–216
- Hegg C, Kienholz H (1995) Determining paths of gravity-driven slope processes—the ‘Vector Tree Model’. In: Carrara A, Guzzetti F (eds) *Geographic information systems in assessing natural hazards*. Kluwer Academic Publishers, Dordrecht, pp 79–92
- Hoek E, Bray W (1981) *Rock slope engineering* (3rd edn) Institution of Mining and Metallurgy, London. 527 pp. International Society for Rocks Mechanics. 1978. Suggested methods for the quantitative description of discontinuities in rock masses. *J Rock Mech Min* 15:319–368

- Hosmer DW, Lemeshow S (2000) Applied logistic regression, 2nd edn. Wiley, New York
- Hungro O, Evans SG (1988) Engineering evaluation of fragmental rock-fall hazards. Proceedings of the 5th international symposium on landslides in Lausanne, Balkema, Rotterdam, pp 685–690
- Jadda M, Shafri HZM, Mansor ShB, Sharifikia M, Pirasteh S (2009) Landslide susceptibility evaluation and factor effect analysis using probabilistic-frequency ratio model. *Eur J Sci Res* 33(4):654–668
- Kelarestaghi A, Ahmadi H (2009) Landslide susceptibility analysis with a bivariate approach and GIS in northern Iran. *Arab J Geosci* 2:95–101
- Keylock C, Domaas U (1999) Evaluation of topographic models of rock-fall travel distance for use in hazard applications. *Artic Antarctic Alp Res* 31(3):312–320
- Kirkby MJ, Statham I (1975) Surface stone movement and scree formation. *J Geol* 83:349–362
- Kobayashi Y, Harp EL, Kagawa T (1990) Simulation of rock-falls triggered by earthquakes. *J Rock Mech Rock Eng* 23:1–20
- Lamelas MT, Marinoni O, Hoppe A, Riva J (2008) Doline probability map using logistic regression and GIS technology in the central Ebro Basin (Spain). *Environ Geol* 54:963–977
- Lee S, Min K (2001) Statistical analysis of Landslide susceptibility at Yongin, Korea. *Environ Geol* 40:1095–1113
- Lee S, Pradhan B (2006) Probabilistic landslide susceptibility s and risk mapping on Penang Island, Malaysia. *J Earth Syst Sci* 115(6):661–672
- Lee S, Sambath T (2006) Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. *Environ Geol* 50:847–855
- Menard SW (2002) Applied logistic regression analysis, 2nd edn. Sage, Thousand Oaks, 111 pp. Network model. *Environmental Earth Sciences*. doi:10.1007/s12665-009-0245-8 (on-line first)
- Nagelkerk E (1991) A note on a general definition of the coefficient of determination. *Biometrika* 78(3):691–692
- Nandi A, Shakoor A (2009) A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses. *Eng Geol* 110:11–20
- Nefeslioglu HA, Gokceoglu C, 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
- Oh HJ, Lee S (2011) Landslide susceptibility mapping on Panaon Island, Philippines using a geographic information system. *Environ Earth Sci* 62:935–951
- 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
- Oh HJ, Lee S, Chotikasathien W, Kim CH, Kwon JH (2009) Predictive landslide susceptibility mapping using spatial information in the Pechabun area of Thailand. *Environ Geol* 57:641–651
- Pfeiffer TJ, BowenTD (1989) Computer simulation of rock-falls. *Bull Assoc Eng Geol* 26(1):135–146
- Pradhan B (2010) Remote sensing and GIS-based landslide susceptibility analysis and cross-validation using multivariate logistic regression model on three test areas in Malaysia. *Adv Space Res* 45:1244–1256
- Pradhan B, Lee S (2009) Landslide risk analysis using artificial neural network model focusing on different training sites. *Int J Phys Sci* 3(11):1–15
- Pradhan B, Lee S (2010) Landslide susceptibility assessment and factor effect: back-propagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modeling. *Environ Model Softw* 25:747–759
- Reger JP (1979) Discriminate analysis as a possible tool in landslide investigations. *Earth Surf Proc Land* 4:267–273
- Sassa K, Wang G, Fukuoka H, Wang F, Ochiai T, Sugiyama M, Sekiguchi T (2004) Landslide risk evaluation and susceptibility zoning for rapid and long-travel rock-fall in urban development areas. *Landslides* 1(3):221–235
- Selby MJ (1982) Hill slope Material and process. Oxford university press, Oxford
- Statham I (1976) A scree slope rock-fall model. *Earth Surf Process* 1:43–62
- Su F, Cui P (2010) Susceptibility assessment of landslides caused by the wenchuan earthquake using a logistic regression model. *J Mt Sci* 7:234–245
- Suzen ML, Doyuran V (2004a) HA comparison of the GIS based landslide susceptibility assessment methods: multivariate versus bivariate. *Environ Geol* 45:665–679
- Suzen ML, Doyuran V (2004b) Data driven bivariate landslide susceptibility assessment using geographical information systems: a method and application to Asarsuyu catchment, Turkey. *Eng Geol* 71:303–352
- Wu SS (1985) Rock fall evaluation by computer simulation. *Transp Res Rec* 1031:1–5
- Yamagishi H (2000) Recent rock-fall in western Hokkaido, Japan. *Pure Appl Geophys* 157:1115–1134