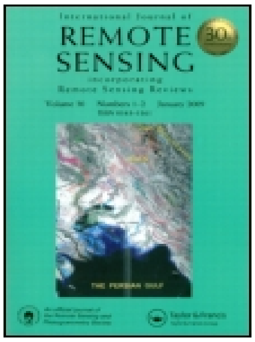

Taylor & Francis
Taylor & Francis Group

International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: http://www.tandfonline.com/loi/tres20

Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data

S. Lee

International Journal of Remote Sensing
Vol. 26, No. 7, 10 April 2005, 1477–1491

Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data

S. LEE
Geoscience Information Center, Korea Institute of Geology, Mining & Materials (KIGAM) 30, Gajeong-Dong, Yuseong-gu, Daejeon, Korea; tel: 82-42-868-3057; fax: 82-42-861-9714; e-mail: leesaro@kigam.re.kr

(*Received 19 April 2004; in final form 20 September 2004*)

The aim of this study is to evaluate the hazard of landslides at Penang, Malaysia, using a Geographical Information System (GIS) and remote sensing. Landslide locations were identified in the study area from interpretation of aerial photographs and from field surveys. Topographical and geological data and satellite images were collected, processed and constructed into a spatial database using GIS and image processing. The factors chosen that influence landslide occurrence were: topographic slope, topographic aspect, topographic curvature and distance from drainage, all from the topographic database; lithology and distance from lineament, taken from the geologic database; land use from Thematic Mapper (TM) satellite images; and the vegetation index value from Système Probatoire de l'Observation de la Terre (SPOT) satellite images. Landslide hazardous areas were analysed and mapped using the landslide-occurrence factors by logistic regression model. The results of the analysis were verified using the landslide location data and compared with probabilistic model. The validation results showed that the logistic regression model is better in prediction than probabilistic model.

1. Introduction

Recently there has been an increasing occurrence of landslides in Malaysia. Most of these landslides occurred on cut slopes or on embankments alongside roads and highways in mountainous areas. Some of these landslides occurred near high-rise apartments and in residential areas, causing great anxiety in many people. A few major and catastrophic landslides have also occurred within the last 10 years. These landslides have resulted in significant damage to both people and property. In the area chosen in this study, Penang in Malaysia, much damage was caused on each of these occasions. Landslides triggered by heavy rainfall are the most common, but few attempts are made to predict them or prevent damage. To remedy this, it is necessary to assess scientifically the area susceptible to landslide. Subsequently, the landslide damage could be greatly decreased. Through scientific analysis of landslides, we can assess and predict landslide-susceptible areas, and thus decrease landslide damage through proper preparation. To achieve this aim, landslide susceptibility analysis techniques have been applied, and verified in the study area. In addition, landslide-related factors were also assessed.

The Penang area has suffered much landslide damage following heavy rains, and was selected as a suitable candidate to evaluate the frequency and distribution of landslides (figure 1). Penang is one of the 13 states of the Federation of Malaysia.

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2005 Taylor & Francis Group Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160412331331012

(*a*)

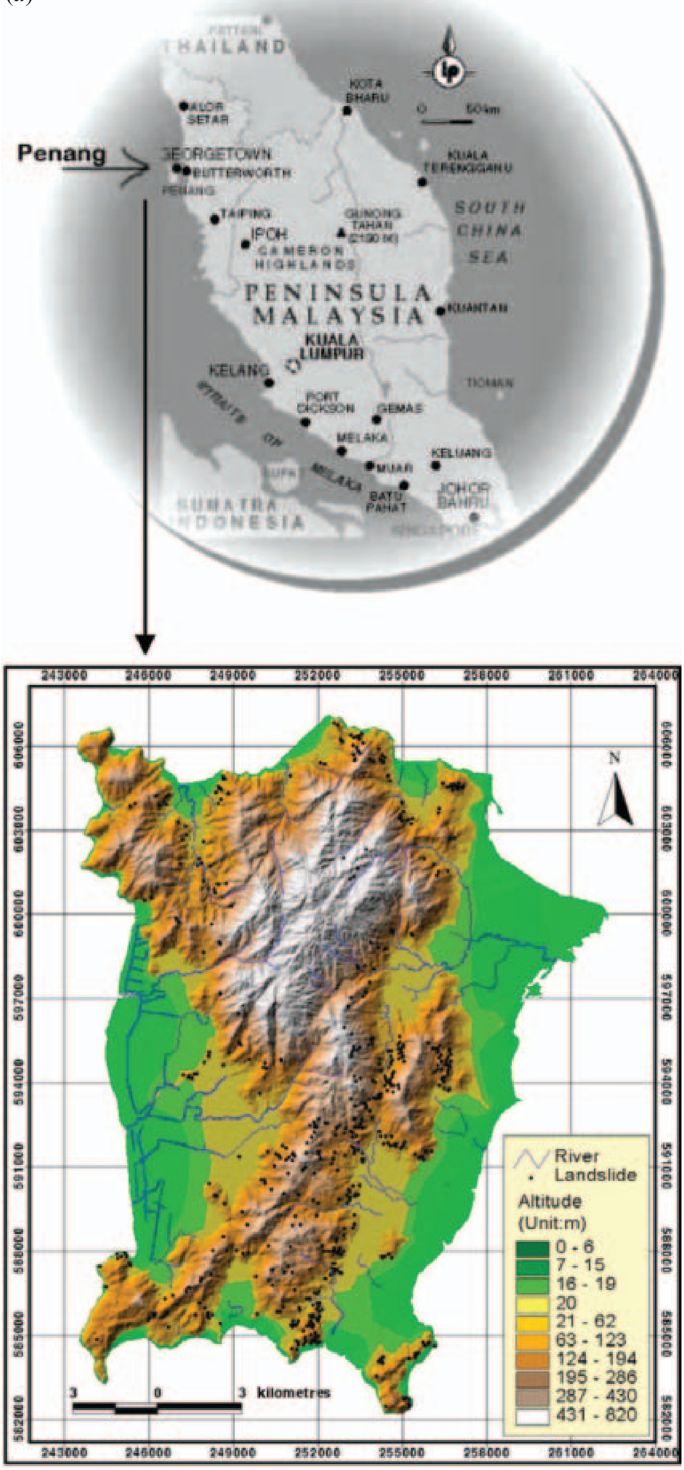

Figure 1. Landslide location map with hillshaded map (*a*) and SPOT satellite image (bands 4, 3, 2) (*b*) of study area.

(*b*)

Figure 1. *Continued.*

The Penang area is located on the north-west coast of the Malaysian peninsular. It is bounded to the north and east by the state of Kedah, to the south by the state of Perak, and to the west by the Straits of Malacca and by Sumatra (Indonesia). Penang consists of the island of Penang, and a coastal strip on the mainland, known as Province Wellesley. The island covers an area of 285 km^2, and is separated from the mainland by a channel. The rainfall is quite evenly distributed throughout the year, with more rain occurring from September to November. Penang has a population of approximately one million people. The bedrock geology of the study area consists mainly of granite.

There have been many studies carried out on landslide hazard evaluation using a Geographical Information System (GIS); for example, Guzzetti *et al.* (1999) summarized many landslide hazard evaluation studies. Recently, there have been studies on landslide hazard evaluation using GIS, and many of these studies have applied probabilistic methods (Rowbotham and Dudycha 1998, Jibson *et al.* 2000, Luzi *et al*. 2000, Parise and Jibson 2000, Rautelal and Lakhera 2000, Baeza and Corominas 2001, Lee and Min 2001, Temesgen *et al*. 2001, Clerici *et al*. 2002, Donati and Turrini 2002, Lee *et al*. 2002a, b, 2004a, Zhou *et al*. 2002, Lee and Choi 2003). One of the statistical methods available, the logistic regression method, has also been applied to landslide hazard mapping (Atkinson and Massari 1998, Dai *et al*. 2001, Dai and Lee 2002, Ohlmacher and Davis 2003), as has the geotechnical method and the safety factor method (Gokceoglu *et al*. 2000, Romeo 2000, Refice and Capolongo 2002, Carro *et al*. 2003, Shou and Wang 2003, Zhou *et al*. 2003). As a new approach to landslide hazard evaluation using GIS, data mining using fuzzy logic, and artificial neural network methods have been applied (Ercanoglu and Gokceoglu 2002; Pistocchi *et al*. 2002, Lee *et al*. 2003a, b, 2004b).

A key assumption using this approach is that the potential (occurrence possibility) of landslides will be comparable to the actual frequency of landslides. Landslide occurrence areas were detected in the Penang area, Malaysia by interpretation of aerial photographs and field surveys. A map of landslides was prepared from aerial photographs, in combination with the GIS, and this was used to evaluate the frequency and distribution of shallow landslides in the area. Topography and lithology databases were constructed and lineament, land use and vegetation index value extracted from Landsat Thematic Mapper (TM) and Système Probatoire de l'Observation de la Terre (SPOT) XS satellite image for the analysis. Then, the calculated and extracted factors were converted to a $10\,\mathrm{m} \times 10\,\mathrm{m}$ grid (ARC/INFO GRID type). Then, using a logistic multiple-regression model, the spatial relationships between the landslide location and each landslide-related factor were analysed, and a formula of landslide occurrence possibility was extracted using the relationships in the SPSS statistical program. This formula was used to calculate the landslide susceptibility index and the index was mapped to represent landslide susceptibility. Finally, the susceptibility map was verified using known landslide locations and success rates were calculated for quantitative validation. To compare the validation result, probabilistic model, frequency ratio, was applied using the same database. Using the frequency ratio models, the spatial relationships between the landslide location and each landslide related factor was extracted. Then, the relationship was used as each factor's rating in the overlay analysis. In the study, GIS software, ArcView 3.2, and ARC/INFO 8.1 NT version software packages were used as the basic analysis tools for spatial management and data manipulation.

2. Data using GIS and remote sensing

Accurate detection of the location of landslides is very important for probabilistic landslide susceptibility analysis. The application of remote sensing methods, such as aerial photographs and satellite images, are used to obtain significant and cost-effective information on landslides. In this study, 1 : 10 000–1 : 50 000-scale aerial photographs were used to detect the landslide locations. These photographs were taken during the period 1981–2000, and the landslide locations were detected by photo interpretation and the locations verified by fieldwork. Recent landslides were observed in aerial photographs from breaks in the forest canopy, bare soil, or other geomorphic characteristics typical of landslide scars, for example, head and side scarps, flow tracks, and soil and debris deposits below a scar. To assemble a database to assess the surface area and number of landslides in each of three study areas, a total of 541 landslides were mapped in a mapped area of $293\,\mathrm{km}^2$.

Table 1. Data layer of study area.

Classification	Sub-classification	GIS data type	Scale
Geological hazard	Landslide	Point coverage	1 : 50 000
Basic map	Topographic map	Line and point coverage	1 : 50 000
	Geological map	Polygon coverage	1 : 50 000
	Land use	GRID	$30\,\mathrm{m} \times 30\,\mathrm{m}$
	Vegetation index (NDVI)	GRID	$10\,\mathrm{m} \times 10\,\mathrm{m}$

Identification and mapping of a suitable set of instability factors having a relationship with the slope failures requires an *a priori* knowledge of the main causes of landslides (Guzzetti *et al.* 1999). These instability factors include surface and bedrock lithology and structure, bedding altitude, seismicity, slope steepness and morphology, stream evolution, groundwater conditions, climate, vegetation cover, land use and human activity. The availability of thematic data varies widely, depending on the type, scale and method of data acquisition. To apply the probabilistic method, a spatial database that considers landslide-related factors was designed and constructed. These data are available in Malaysia either as paper or as digital maps. The spatial database constructed is shown in table 1.

There were eight factors considered in calculating the probability, and the factors were extracted from the constructed spatial database. The factors were transformed into a vector-type spatial database using the GIS, and landslide-related factors were extracted using our database. Using the topographic database, a digital elevation model (DEM) was created first. Contour and survey base points that had elevation values from the 1 : 50 000-scale topographic maps were extracted, and a DEM was constructed with a resolution of 10 m. Using this DEM, the slope angle, slope aspect and slope curvature were calculated. In addition, the distance from drainage was calculated using the topographic database. The drainage buffer was calculated in 100 m intervals. Using the geology database, the lithology map was obtained from a 1 : 50 000-scale geological map, and the distance from lineament calculated. The lineament buffer was calculated in 100 m intervals. Land use data were classified using a Landsat TM image employing an unsupervised classification method and field survey. The 11 classes identified, such as urban, water, forest, agricultural area, and barren area, were extracted for land use mapping. Finally, the Normalized Difference Vegetation Index (NDVI) value was calculated using the formula $\mathrm{NDVI}=(\mathrm{IR}-\mathrm{R})/(\mathrm{IR}+\mathrm{R})$, where IR value is the infrared portion of the electromagnetic spectrum, and R value is the red portion of the electromagnetic spectrum. The NDVI value denotes areas of vegetation in an image.

The factors were converted to a raster grid with $10\,\mathrm{m} \times 10\,\mathrm{m}$ cells for application of the logistic regression and frequency ratio model. The area grid was 2490 rows by 1884 columns (i.e. total number is 4 691 160) and 541 cells had landslide occurrences.

3. Logistic regression model and its application

Logistic regression allows one to form a multivariate regression relation between a dependent variable and several independent variables. Logistic regression, which is one of the multivariate analysis models, is useful for predicting the presence or absence of a characteristic or outcome based on values of a set of predictor variables. The advantage of logistic regression is that, through the addition of an appropriate link function to the usual linear regression model, the variables may be either continuous or discrete, or any combination of both types and they do not necessarily have normal distributions. In the case of multi-regression analysis, the factors must be numerical, and in the case of a similar statistical model, discriminant analysis, the variables must have a normal distribution. In the present situation, the dependent variable is a binary variable representing presence or absence of landslide. Where the dependent variable is binary, the logistic link function is applicable (Atkinson and Massari 1998). For this study, the dependent variable must be input as either 0 or 1, so the model applies well to landslide possibility analysis. Logistic regression coefficients can be used to estimate ratios for each of the independent variables in the model.

Quantitatively, the relationship between the occurrence and its dependency on several variables can be expressed as:

$$p=1/(1+e^{-z}) \tag{1}$$

where p is the probability of an event occurring. In the present situation, the value p is the estimated probability of landslide occurrence. The probability varies from 0 to 1 on an S-shaped curve and z is the linear combination. It follows that logistic regression involves fitting an equation of the following form to the data:

$$z=b_0+b_1x_1+b_2x_2+\cdots+b_nx_n \tag{2}$$

where b_0 is the intercept of the model, the b_i (i=0, 1, 2, … , n) are the slope coefficients of the logistic regression model, and the x_i (i=0, 1, 2, … , n) are the independent variables. The linear model formed is then a logistic regression of presence or absence of landslides (present conditions) on the independent variables (pre-failure conditions).

Using the logistic regression model, the spatial relationship between landslide-occurrence and factors influencing landslides was assessed. The spatial databases of each factor were converted to ASCII format files for use in the statistical package, and the correlations between landslide and each factor were calculated. There are two cases. In the first case, only one factor was used. In this case, logistic regression formulae were created for each case. The coefficient is shown in table 2. Finally, the probability that predicts the possibility of landslide-occurrence was calculated using the spatial database, data from table 2, equations (1) and (2). In the second case, all factors were used. In this case, logistic regression formulae were created as shown in equations (1) and (3) for each case. The coefficient is shown in table 2.

$$\begin{aligned} z_9 = {} & (0.0308\times \mathrm{SLOPE})+\mathrm{ASPECT}+(-0.0000512\times \mathrm{CURVATURE}) \\ & +(0.0006\times \mathrm{DRAINAGE})+\mathrm{LITHOLOGY}_b+(-0.0012\times \mathrm{FAULT}) \\ & +(-0.0053\times \mathrm{NDVI})+\mathrm{LAND\ USE}_b-11.7303 \end{aligned} \tag{3}$$

where SLOPE is slope value; CURVATURE is curvature value; DRAINAGE is distance from drainage value; FAULT is distance from fault value; NDVI is NDVI value; ASPECT_c, $\mathrm{LITHOLOGY}_c$, $\mathrm{LAND\ USE}_c$ are logistic regression coefficient value listed in table 2; and z_n is a parameter).

Table 2. Coefficients value for logistic regression in the case of each factor.

Case of each factor used			Case of all factors used	
Factor	Layer	B	Layer	B
SLOPE	Constant	0.0399	SLOPE	0.0308
		−9.2011	Flat area	
ASPECT	Flat area	−0.7781	N	−0.5076
	N	0.0880	NE	−0.0902
	NE	0.6177	E	0.4824
	E	−0.0811	SE	−0.0036
	SE	0.2751	S	0.1476
	S	0.7497	SW	0.5037
	SW	0.1770	W	0.2502
	W	−0.0765	NW	0.0657
	NW	0.0000	CURVA	5.12E-05
	Constant	−8.6546	Alluvium	−0.0601
CURVA		9.12E-05	DRAINAGE	0.0006
	Constant	−8.5962	FAULT	−0.0012
LITHOLOGY	Alluvium	−.8658	NDVI	−0.0053
	Granite	0.0000	LAND USE	
	Constant	−8.3756	No data	3.6512
DRAINAGE		−0.0011	Urban	3.9735
	Constant	−8.4254	Mixed	3.5174
FAULT		−0.0009	Forest	1.9491
	Constant	−8.0873	Scrub	5.0356
NDVI		0.0086	Aquaculture	3.4590
	Constant	−8.8239	Swamp	0.6215
LAND USE	No data	3.1432	Rubber	3.4325
	Urban	3.3120	Rice	3.8422
	Mixed	4.1061	Coconut	2.2737
	Forest	2.9037	Barren	−0.7109
	Scrub	5.3331	Constant	−11.7303
	Aquaculture	3.3222		
	Swamp	0.0000		
	Rubber	4.2736		
	Rice	1.9970		
	Coconut	0.0000		
	Barren	0.0000		
	Oil palm	0.0000		
	Constant	−12.2029		

Using equations (1) and (2), the possibility of landslide occurrence was calculated and using the possibility, eight landslide susceptibility maps were made. Among the eight cases maps, case of the usage of all factors, which use equations (1) and (3), is shown in figure 2.

4. Frequency ratio model and its application

Frequency ratio approaches are based on the observed relationships between distribution of landslides and each landslide-related factor, to reveal the correlation between landslide locations and the factors in the study area. Using the frequency ratio model, the spatial relationships between landslide-occurrence location and each factor's contributing landslide occurrence were derived. The frequency is calculated from analysis of the relation between landslides and the considered

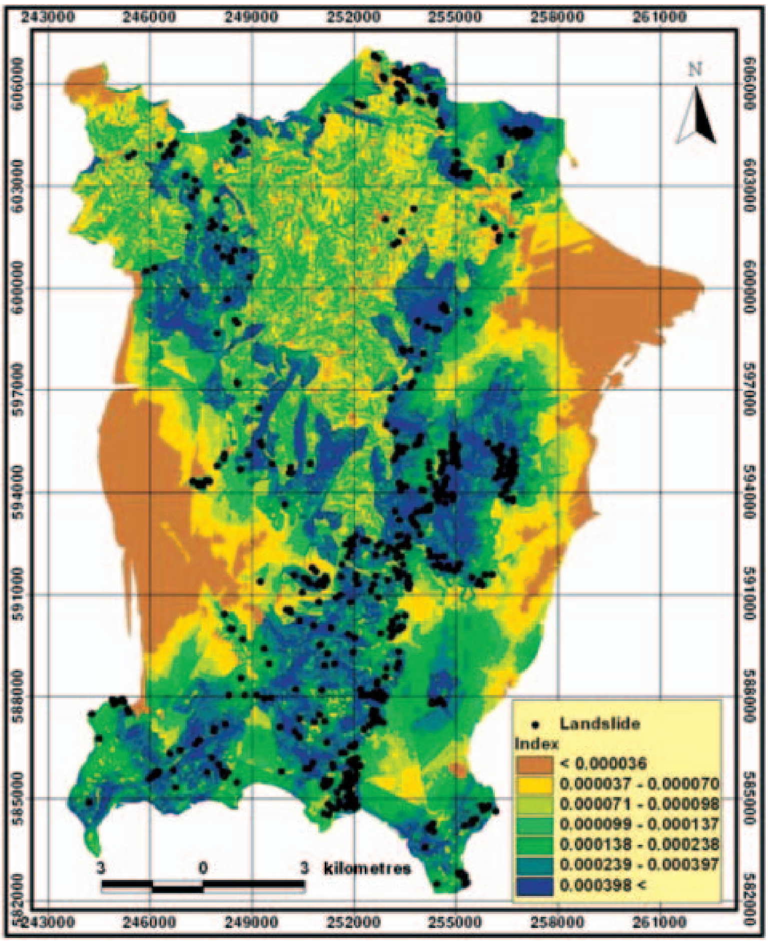

Figure 2. Landslide susceptibility map based on logistic regression.

factors. Therefore, the frequency ratios of each factor's type or range were calculated from their relationship with landslide events as shown in table 3. In the relation analysis, the ratio is that of the area where landslides occurred to the total area, so that a value of 1 is an average value. If the value is greater than 1, it means a higher correlation, and value lower than 1 means lower correlation.

To calculate the Landslide Susceptibility Index (LSI), each factor's frequency ratio values were summed to the training area as in equation (4). The landslide susceptibility value represents the relative susceptibility to landslide occurrence. So the greater the value, the higher the susceptibility to landslide occurrence and the lower the value, the lower the susceptibility to landslide occurrence.

$$\mathrm{LSI} = Fr_1 + Fr_2 + \cdots + Fr_n \tag{4}$$

where Fr is the rating of each factor's type or range.

The landslide susceptibility map was made using the LSI values and for interpretation, it is shown in figure 3.

Table 3. Frequency ratio to landslide occurrences.

Factor	Class	Landslide occurrence points	Landslide occurrence points (%)	Pixels in domain	Pixel %	Frequency ratio
Slope (°)	0~5	63	11.65	1 324 669	45.24	0.26
	6~10	41	7.58	119 251	4.07	1.86
	11~15	71	13.12	200 779	6.86	1.91
	16~20	102	18.85	327 812	11.19	1.68
	21~25	108	19.96	366 266	12.51	1.60
	26~30	90	16.64	312 124	10.66	1.56
	3 ~87	66	12.20	277 477	9.48	1.29
Aspect	Flat	49	9.06	599 634	20.48	0.44
	N	39	7.21	194 419	6.64	1.09
	NE	65	12.01	211 666	7.23	1.66
	E	72	13.31	460 442	15.72	0.85
	SE	87	16.08	361 722	12.35	1.30
	S	66	12.20	184 387	6.30	1.94
	SW	53	9.80	235 551	8.04	1.22
	W	57	10.54	359 246	12.27	0.86
	NW	53	9.80	321 311	10.97	0.89
Curvature	−38.65~−1	165	30.50	681 534	23.27	1.31
	0	171	31.61	1 537 754	52.51	0.60
	1~32.26	205	37.89	709 090	24.21	1.56
Distance from drainage (m)	0~200	416	76.89	2 136 667	72.96	1.05
	201~400	109	20.15	496 867	16.97	1.19
	401~600	16	2.96	164 354	5.61	0.53
	601~80	0	0.00	69 278	2.37	0.00
	801~2200	0	0.00	61 212	2.09	0.00
Lithology	No data	0	0.00	379	0.01	0.00
	Alluvium	98	18.11	1 000 697	34.17	0.53
	Granite	443	81.89	1 927 302	65.81	1.24
Distance from lineament (m)	0~200	176	32.53	894 416	30.54	1.07
	201~400	163	30.13	492 910	16.83	1.79
	401~600	90	16.64	309 474	10.57	1.57
	601~800	47	8.69	229 217	7.83	1.11
	801~1000	31	5.73	183 714	6.27	0.91
	1001~1200	21	3.88	137 806	4.71	0.82
	1201~1400	6	1.11	99 222	3.39	0.33
	1401~1600	7	1.29	80 340	2.74	0.47
	1601~1800	0	0.00	67 134	2.29	0.00
	1801~2000	0	0.00	58 035	1.98	0.00
	2001~6800	0	0.00	37 6110	12.84	0.00
Land use	No data	4	0.74	51 178	1.75	0.42
	Urban	113	20.89	828 373	28.29	0.74
	Mixed	292	53.97	949 299	32.42	1.66
	Forest	65	12.01	732 320	25.01	0.48
	Scrub	7	1.29	6746	0.23	5.62
	Aquaculture	5	0.92	35 965	1.23	0.75
	Swamp	0	0.00	35 383	1.21	0.00
	Rubber	50	9.24	138 921	4.74	1.95
	Rice	5	0.92	136 099	4.65	0.20
	Coconut	0	0.00	9871	0.34	0.00
	Barren	0	0.00	2094	0.07	0.00
	Oil palm	0	0.00	2129	0.07	0.00

Table 3. (Continued).

Factor	Class	Landslide occurrence points	Landslide occurrence points (%)	Pixels in domain	Pixel %	Frequency ratio
Vegetation index	No data	0	0.00	2765	0.09	0.00
	−0.80∼−0.60	0	0.00	11 909	0.41	0.00
	−0.60∼−0.40	4	0.74	39 455	1.35	0.55
	−0.40∼−0.20	29	5.36	233 399	7.97	0.67
	−0.20∼0.00	42	7.76	317 483	10.84	0.72
	0.00∼0.20	54	9.98	295 647	10.10	0.99
	0.20∼0.40	189	34.94	1 016 029	34.70	1.01
	0.40∼−061	223	41.22	1 011 691	34.55	1.19

Number of total cells in study area: 2 928 378 (without no data).
Number of landslide occurrence points: 541.

5. Validation comparison of the models

For validation of landslide susceptibility calculation models, two basic assumptions are needed. One is that landslides are related to spatial information such as

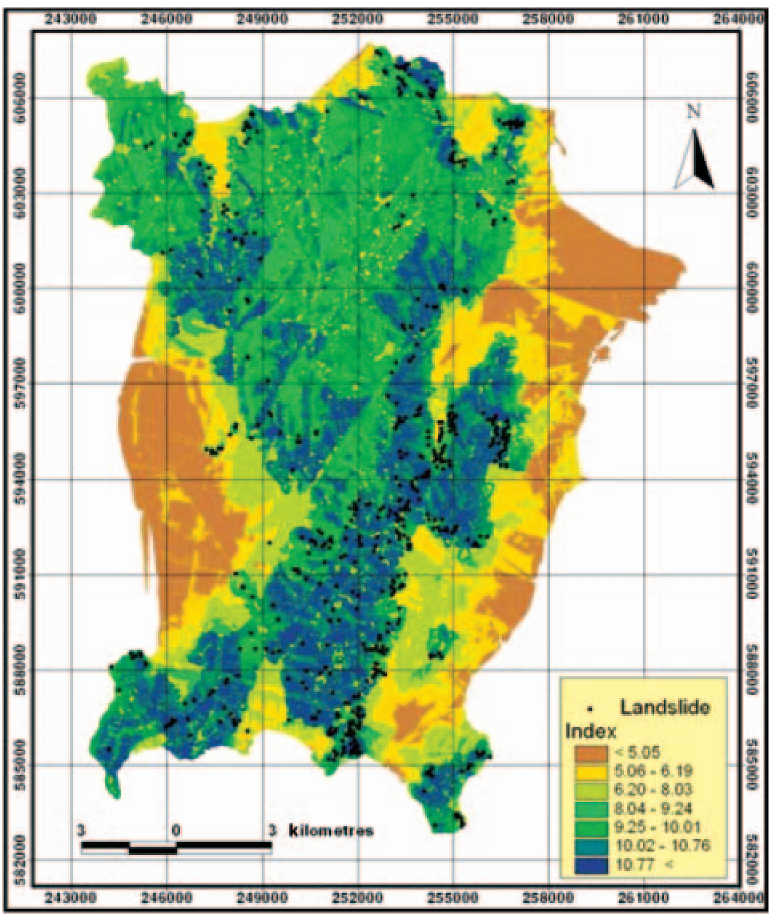

Figure 3. Landslide susceptibility map based on frequency ratio.

topography, soil, forest and land use, and the other is that future landslides will be triggered by a specific factor such as rainfall or earthquake. In this study, the two assumptions are satisfied because the landslides were related to the spatial information and the landslides were triggered by one cause—heavy rainfall in the study area.

The landslide susceptibility analysis result was validated using known landslide locations. Validation was performed by comparing the known landslide location data with the landslide susceptibility map. Each factor used and frequency ratio was compared. The rate curves were created and their areas of the under curve were calculated for all cases. The rate explains how well the model and factor predict the landslide. So, the area under curve can assess the prediction accuracy qualitatively. To obtain the relative ranks for each prediction pattern, the calculated index values of all cells in the study area were sorted in descending order. Then the ordered cell values were divided into 100 classes, with accumulated 1% intervals. The rate verification results appear as a line in figure 4. For example, in the case of all factors used, 90–100% (10%) class of the study area where the landslide susceptibility index had a higher rank could explain 41% of all the landslides. In addition, the 80–100% (20%) class of the study area where the landslide susceptibility index had a higher rank could explain 58% of the landslides using the logistic regression model. To compare the result quantitatively, the areas under the curve were re-calculated as the total area is 1 which means perfect prediction accuracy.

So, the area under a curve can be used to assess the prediction accuracy qualitatively. The area under the curve is shown in table 4. In the case of all factors

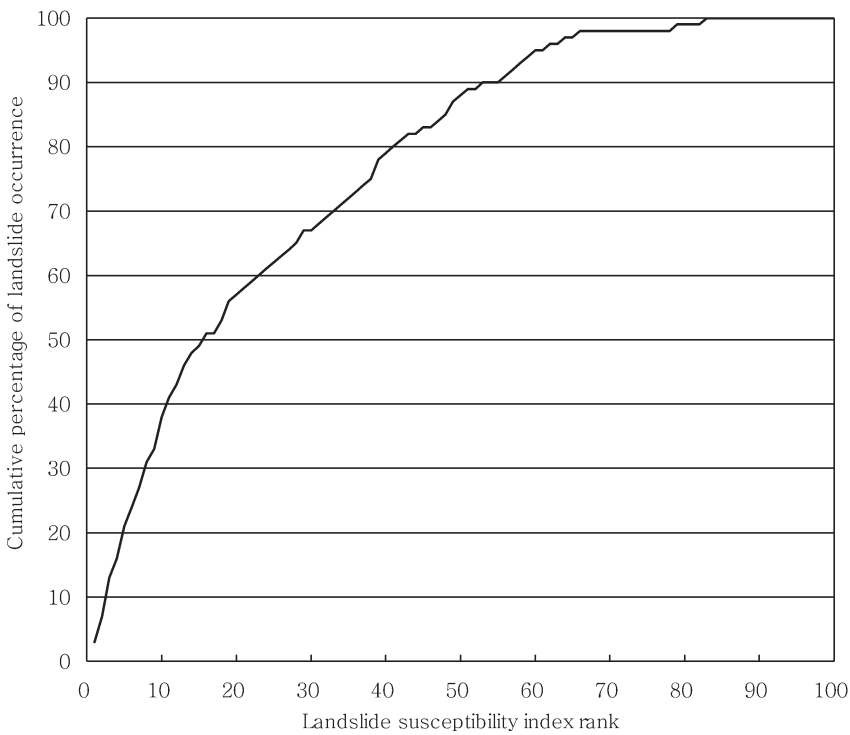

Figure 4. Cumulative frequency diagram showing landslide susceptibility index rank occurring in cumulative percentage of landslide occurrence in the case all factors used.

Table 4. Validation result.

	Slope	Aspect	Curvature	Distance from lineament	Distance from drainage	NDVI	Land use	Lithology	All factors used	Frequency ratio
Area	6569	6163	5363	4086	5028	5603	6616	5850	7862	7106
Area ratio	0.657	0.616	0.536	0.409	0.503	0.560	0.662	0.585	0.786	0.711
Area ratio (%)	65.7	61.6	53.6	40.9	50.3	56.0	66.2	58.5	78.6	71.1

and logistic regression model used, the area ratio was 0.786 and we could say the prediction accuracy is 78.6%. In the case of all factors and frequency ratio model used, the area ratio was 0.711 and we could say the prediction accuracy is 71.1%. In the case of slope factor and logistic regression model used, the prediction accuracy is 65.7%. Overall the case of all factors and logistic regression model used showed a higher accuracy than cases of each factor and logistic regression used and all factors and frequency ratio model used.

6. Conclusions and discussion

Landslides are among the most hazardous natural disasters. Government and research institutions worldwide have attempted for years to assess the landslide hazard and risk and to show its spatial distribution. In this study, a probabilistic approach to estimating the susceptible area of landslides using GIS and remote sensing is presented.

The result of validation of logistic regression and frequency ratio model, the logistic regression model showed the better prediction accuracy, more than 7.5%. In the case of each factor used with logistic regression, the slope used case showed the best prediction accuracy (65.7%). But, there is some difference (12.9%) from the case all used (78.7%). So, it could be concluded that the case of all factors and logistic regression model used had best prediction accuracy in landslide susceptibility mapping.

The area ratio value from the effect analysis can be used to weight the relative importance of these factors, and can improve the prediction accuracy of the landslide susceptibility map. The frequency ratio model is simple; the process of input, calculation and output can be readily understood. The large amount of data can be processed in the GIS environment quickly and easily. The logistic regression model requires conversion of the data to ASCII or other formats for use in the statistical package, and later re-conversion to incorporate them into the GIS database. Moreover, it is hard to process the large amount of data in the statistical package. However, correlation of landslide and other factors can be analysed qualitatively in statistical package. The logistic regression model showed better accuracy than frequency ratio model in this study and the use of all factors showed the better results. In the case of a similar statistical model (discriminant analysis), the factors must have a normal distribution, and in the case of multi-regression analysis, the factors must be numerical. However, for logistical regression, the dependent variable must be input as 0 or 1, therefore the model applies well to landslide occurrence analysis.

Landslide hazard 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, but the models used in the study are valid for generalized planning and assessment purposes, although they may be less useful at the site-specific scale where local geological and geographical heterogeneities may prevail. For the model to be more generally applied, more landslide data are needed, as well as application to more regions.

References

Atkinson, P.M. and Massari, R., 1998, Generalized linear modeling of susceptibility to landsliding in the central Apennines, Italy. *Computer and Geosciences*, **24**, pp. 373–385.

BAEZA, C. and COROMINAS, J., 2001, Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surface Processes and Landforms*, **26**, pp. 1251–1263.

CARRO, M., DE AMICIS, M., LUZI, L. and MARZORATI, S., 2003, The application of predictive modeling techniques to landslides induced by earthquakes, the case study of the 26 September 1997 Umbria–Marche earthquake (Italy). *Engineering Geology*, **69**, pp. 139–159.

CLERICI, A., PEREGO, S., TELLINI, C. and VESCOVI, P., 2002, A procedure for landslide susceptibility zonation by the conditional analysis method. *Geomorphology*, **48**, pp. 349–364.

DAI, F.C. and LEE, C.F., 2002, Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology*, **42**, pp. 213–228.

DAI, F.C., LEE, C.F., LI, J. and XU, Z.W., 2001, Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environmental Geology*, **40**, pp. 381–391.

DONATI, L. and TURRINI, M.C., 2002, An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology, application to an area of the Apennines (Valnerina, Perugia, Italy). *Engineering Geology*, **63**, pp. 277–289.

ERCANOGLU, M. and GOKCEOGLU, C., 2002, Assessment of landslide susceptibility for a landslide-prone area (north of Yenice, NW Turkey) by fuzzy approach. *Environmental Geology*, **41**, pp. 720–730.

GOKCEOGLU, C., SONMEZ, H. and ERCANOGLU, M., 2000, Discontinuity controlled probabilistic slope failure risk maps of the Altindag (settlement) region in Turkey. *Engineering Geology*, **55**, pp. 277–296.

GUZZETTI, F., CARRARRA, A., CARDINALI, M. and REICHENBACH, P., 1999, Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study. Central Italy. *Geomorphology*, **31**, pp. 181–216.

JIBSON, W.R., EDWIN, L.H. and JOHN, A.M., 2000, A method for producing digital probabilistic seismic landslide hazard maps. *Engineering Geology*, **58**, pp. 271–289.

LEE, S. and CHOI, U., 2003, Development of GIS-based geological hazard information system and its application for landslide analysis in Korea. *Geoscience Journal*, **7**, pp. 243–252.

LEE, S. and MIN, K., 2001, Statistical analysis of landslide susceptibility at Yongin, Korea. *Environmental Geology*, **40**, pp. 1095–1113.

LEE, S., CHWAE, U. and MIN, K., 2002a, Landslide susceptibility mapping by correlation between topography and geological structure: the Janghung area, Korea. *Geomorphology*, **46**, pp. 149–162.

LEE, S., CHOI, J. and MIN, K., 2002b, Landslide susceptibility analysis and verification using the Bayesian probability model. *Environmental Geology*, **43**, pp. 120–131.

LEE, S., RYU, J.H., MIN, K. and WON, J.S., 2003a, Landslide susceptibility analysis using GIS and artificial neural network. *Earth Surface Processes and Landforms*, **27**, pp. 1361–1376.

LEE, S., RYU, J.H., LEE, M.J. and WON, J.S., 2003b, Landslide susceptibility analysis using artificial neural network at Boun, Korea. *Environmental Geology*, **44**, pp. 820–833.

LEE, S., CHOI, J. and MIN, K., 2004a, Probabilistic landslide hazard mapping using GIS and remote sensing data at Boun, Korea. *International Journal of Remote Sensing*, **25**, pp. 2037–2052.

LEE, S., RYU, J.H., WON, J.S. and PARK, H.J., 2004b, Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. *Engineering Geology*, **71**, pp. 289–302.

LUZI, L., PERGALANI, F. and TERLIEN, M.T.J., 2000, Slope vulnerability to earthquakes at subregional scale, using probabilistic techniques and geographic information systems. *Engineering Geology*, **58**, pp. 313–336.

OHLMACHER, G.C. and DAVIS, J.C., 2003, Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. *Engineering Geology*, **69**, pp. 331–343.

PARISE, M. and JIBSON, W.R., 2000, A seismic landslide susceptibility rating of geologic units based on analysis of characteristics of landslides triggered by the 17 January, 1994 Northridge, California earthquake. *Engineering Geology*, **58**, pp. 251–270.

PISTOCCHI, A., LUZI, L. and NAPOLITANO, P., 2002, The use of predictive modeling techniques for optimal exploitation of spatial databases: a case study in landslide hazard mapping with expert system-like methods. *Environmental Geology*, **41**, pp. 765–775.

RAUTELAL, P. and LAKHERA, R.C., 2000, Landslide risk analysis between Giri and Tons Rivers in Himachal Himalaya (India). *International Journal of Applied Earth Observation and Geoinformation*, **2**, pp. 153–160.

REFICE, A. and CAPOLONGO, D., 2002, Probabilistic modeling of uncertainties in earthquake-induced landslide hazard assessment. *Computers & Geosciences*, **28**, pp. 735–749.

ROMEO, R., 2000, Seismically induced landslide displacements: a predictive model. *Engineering Geology*, **58**, pp. 337–351.

ROWBOTHAM, D. and DUDYCHA, D.N., 1998, GIS modeling of slope stability in Phewa Tal watershed, Nepal. *Geomorphology*, **26**, pp. 151–170.

SHOU, K.J. and WANG, C.F., 2003, Analysis of the Chiufengershan landslide triggered by the 1999 Chi-Chi earthquake in Taiwan. *Engineering Geology*, **68**, pp. 237–250.

TEMESGEN, B., MOHAMMED, M.U. and KORME, T., 2001, Natural hazard assessment using GIS and remote sensing methods, with particular reference to the landslides in the Wondogenet area, Ethiopia. *Physics and Chemistry of the Earth, Part C: Solar, Terrestrial & Planetary Science*, **26**, pp. 665–675.

ZHOU, C.H., LEE, C.F., LI, J. and XU, Z.W., 2002, On the spatial relationship between landslides and causative factors on Lantau Island, Hong Kong. *Geomorphology*, **43**, pp. 197–207.

ZHOU, G., ESAKI, T., MITANI, Y., XIE, M. and MORI, J., 2003, Spatial probabilistic modeling of slope failure using an integrated GIS Monte Carlo simulation approach. *Engineering Geology*, **68**, pp. 373–386.