Geological Society of Malaysia, Bulletin 46
May 2003; pp. 173–179

Probabilistic landslide susceptibility analysis and verification using GIS and remote sensing data at Penang, Malaysia

JASMI AB TALIB

Malaysian Centre For Remote Sensing (MACRES)
No. 13, Jalan Tun Ismail, 50480, Kuala Lumpur, Malaysia
E-mail: jasmi@macres.gov.my

Abstract: The aim of this study is to generate and evaluate the landslide hazard map for Penang Island using a Geographic Information System (GIS) and remote sensing techniques. Landslide locations were identified in the study area from imagery and aerial photograph interpretations followed by field surveys. The topographic and geologic data and satellite images were collected, processed and constructed into a spatial database using GIS and image processing. The factors that influence landslide occurrences such as topographic slope, topographic aspect, topographic curvature and distance from drainage were retrieved from topographic database whereas, geology and distance from lineament were retrieved from the geologic database; land use from TM satellite images and vegetation index value from SPOT satellite data. Landslide hazard areas were analysed and mapped using the landslide occurrence factors by probability – likelihood ratio method. Several areas are considered as hazardous, such as Paya Terubung, Bukit Relau, Bukit Gemuruh and Teluk Bahang. The results of the analysis then were verified using the landslide location data. The validation results showed satisfactory agreement between the hazard map and the existing data on landslide location.

Abstrak: Tujuan kajian dijalankan adalah untuk menjana dan menilai peta bencana tanah runtuh dengan menggunakan kaedah Sistem Maklumat Geografi (GIS) dan remote sensing. Data lokaliti tanah runtuh diperolehi dan dikenalpasti hasil interpretasi data-data satelit, foto-foto udara dan juga maklumat kerja lapangan. Data-data topografi, geologi dan juga imej-imej satelit pula dikumpul, diproses dan dijana dengan menggunakan GIS dan kaedah pemprosesan imej serta disimpan di storan *database.* Peta-peta parameter seperti peta cerun, peta *aspect,* peta bentuk cerun serta peta jarak daripada sungai dijanakan daripada *database* topografi manakala peta geologi dan peta jarak daripada lineamen dijanakan daripada *database* geologi. Disamping itu pula penggunaan data-data satelit menghasilkan parameter-parameter seperti peta nilai indeks tumbuhan yang diperolehi daripada data satelit SPOT dan peta guna tanah hasil interpretasi ke atas data satelit Landsat TM. Kesemua peta-peta parameter tersebut dianalisa dengan menggunakan kaedah kebarangkalian *likelihood.* Hasil pengkelasan jumlah kebarangkalian kesemua peta parameter tersebut menghasilkan peta bencana tanah runtuh. Di antara kawasan-kawasan yang dikenalpasti sebagai kawasan berpotensi tinggi bencana tanah runtuh ialah Paya Terubung, Bukit Relau, Bukit Gemuruh dan Teluk Bahang. Analisa selanjutnya mendapati bahawa kejituan peta bencana tanah runtuh yang dihasilkan ini adalah tinggi.

INTRODUCTION

Recently, there has been an increasing occurrence of landslides in our country. Most of these landslides were on cut slopes or embankments along the roads and highways in mountainous areas (Mahadzer, 2001). Some of these landslides occurred near high-rise apartments and residential areas, creating great anxiety to various groups of people. A few major and catastrophic landslides also occurred within the last ten years. The landslides include the tragic Highland Tower landslide, Genting Sempah landslide, Gua Tempurung landslide, Paya Terubung landslide and Bukit Antarabangsa landslides (Mahadzer, 2001). The frequent landslide tragedies resulted in significant damage to people and property. In Penang Island, much damage was caused on these occasions (Mahadzer, 2001). The reasons for the landslide occurrences were heavy rainfall (Mohd. Asbi, 2001) and, as there was little effort to assess or predict the event, damages were extensive. Through scientific analysis of landslides, we can assess and predict landslide-susceptible areas, and thus decrease landslide damage through proper preparation. In order to achieve this, landslide hazard analysis techniques were developed, applied, and verified in the study area.

There have been many studies of landslide hazard evaluation using GIS; Guzzetti *et al.* (1999) summarized many landslide hazard evaluation studies. Recently, there were studies for landslide hazard evaluation using GIS, among them are Chung and Fabbri (1999), Dhakal *et al.*, (1999), Gokceoglu *et al.* (2000), Lee and Min (2001), Clerici *et al.* (2002), Dai and Lee (2002), Donati and Turrini (2002), Lee *et al.* (2002a) and Lee *et al.* (2002b) have applied probabilistic and statistical method to landslide hazard mapping.

STUDY AREA

The Penang Island that had much landslide damages, was selected as a suitable case to evaluate the frequency and distribution of landslides (Fig. 1). The Island covering an area of 285 km^2 is bounded by latitudes, 5°15' N to 5°30'N and longitudes 100°10'E to 100°20'E. Rainfall in

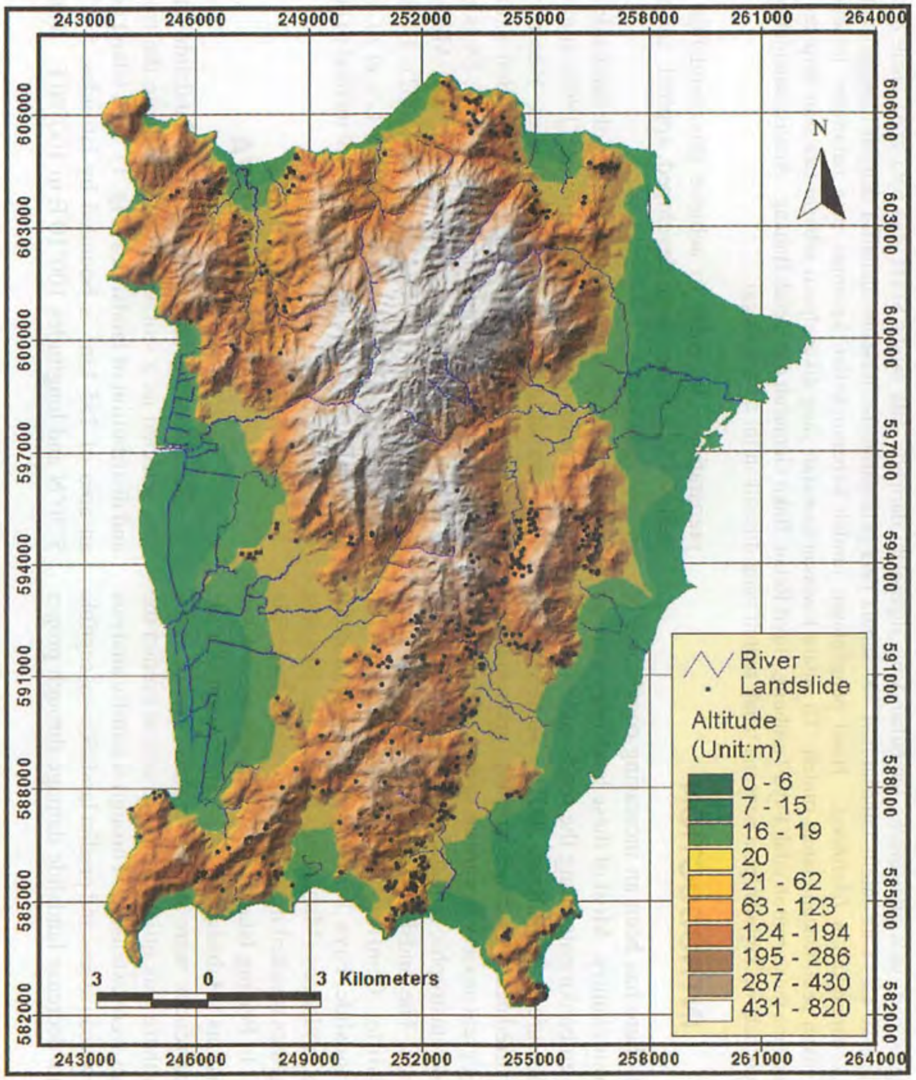

Figure 1. Hillshaded map of the study area showing location of landslides.

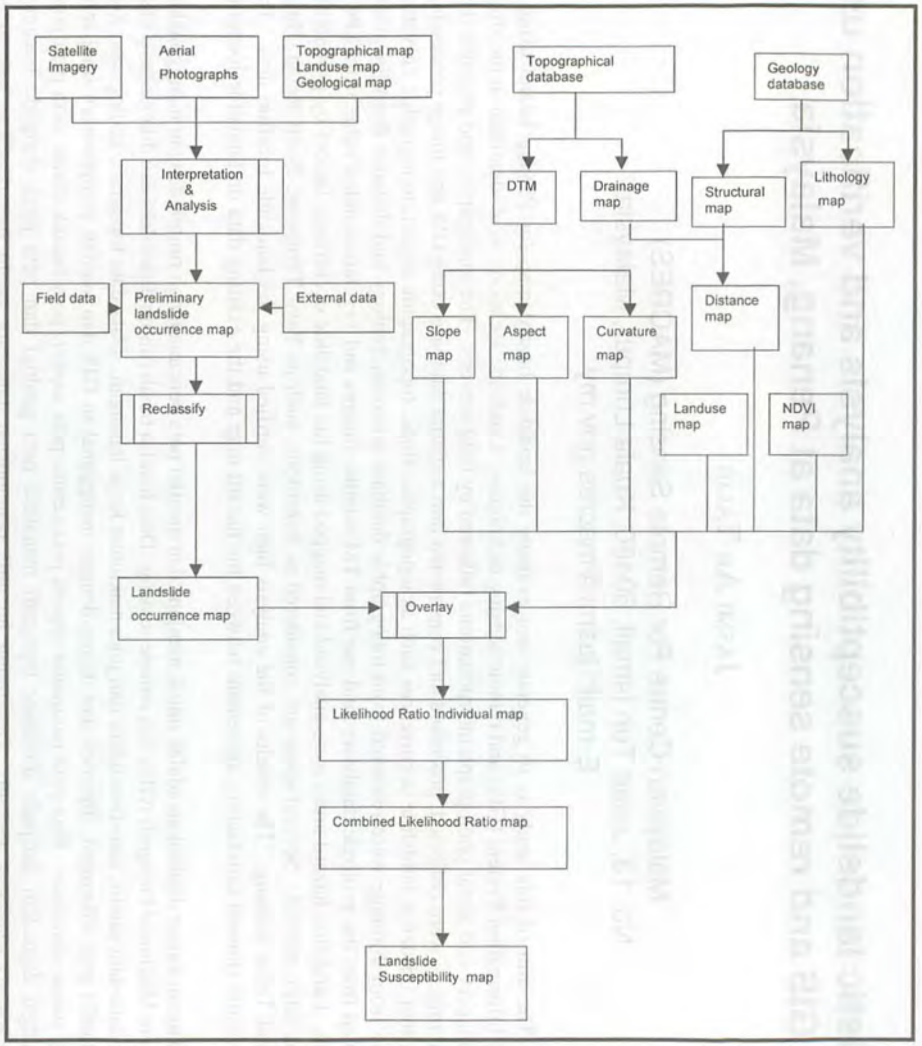

Figure 2. Flowchart of methodology.

Table 1. Average rainfalls recorded at the Air Itam Reservoir Meteorological Station for the year 1975 to 2000.

	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
Min (mm)	0.0	0.0	1.0	19.0	71.0	14.0	0.5	28.5	13.0	86.0	19.5	24.5
Mean (mm)	49.3	79.8	106.1	201.0	209.0	154.2	189.2	232.1	338.9	352.9	258.8	102.1
Max (mm)	198.7	370.0	372.0	681.5	392.8	424.5	445.0	601.9	601.5	581.0	684.0	257.0

Table 2. Data layer of the 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 / Raster	10m∞10m
	Vegetation Index (NDVI)	GRID / Raster	10m∞10m

this area is considered quite evenly distributed throughout the year, with more rain from September to November. Mean rainfall recorded for the past 20 years from Air Itam Reservoir Meteorological Station as shown in Table 1, showed that the amounts for September to November exceeds annual average. The lithology of this area on the other hand consists mainly of granite.

MATERIALS AND METHODOLOGY

In landslide-hazard analysis, data were collected and stored into a spatial database. All landslides related factors then were extracted and their likelihood ratio calculated. Each likelihood ratio was summed together and reclassified to generate a landslide susceptibility map. Lastly this susceptibility map needed to be verified.

A key assumption using probability, likelihood ratio approach, is that the potential (occurrence possibility) of landslides will be comparable to the actual frequency of landslides. Landslide occurrence areas were detected in Penang area by aerial photograph interpretations and field surveys. For this study area, 1:6,000–1:40,000 scale aerial photographs taken from 1981 to 2000 were used to delineate landslide locations, which later were verified by fieldwork. Recent landslides were observed in aerial photographs as a break 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 the scars. In total, 541 landslides were mapped and about one tenth of that was verified in the field. A map of landslides was then developed in combination with GIS and later was used to evaluate the frequency and distribution of shallow landslides in the area.

Topography and lithology databases were constructed as shown in Table 2, whereas lineament, land use and vegetation index value were extracted from Landsat TM and SPOT XS satellite images as for the analysis. Furthermore maps relevant to landslide occurrences were constructed to a vector type spatial database using the GIS software. First, using the topographic database, the digital elevation model (DEM) with 10 m resolutions was created and used for generating slope, aspect and curvature map. Besides that by using the topographic database also, the distance from drainage was calculated. The lithology map was extracted and distance from lineament was calculated using the geology database. The buffer interval used for distance calculation was in 100 m range. Land use map was extracted from Landsat TM satellite images and field checks whereas vegetation index value map was calculated from SPOT satellite image. This Normalized Difference Vegetation Index (NDVI) value was calculated using the formula of (IR − R)/(IR + R), where IR stands for the infrared portion of the electromagnetic spectrum, and R stands for the red portion of the electromagnetic spectrum. NDVI defines vegetated areas in the imagery.

All the calculated and extracted factors were converted to raster maps of 10 m x 10 m grid for the analysis. Univariant probability analysis; using likelihood ratio method were used to analyse the spatial relationships between the landslide locations and each landslide-related factor. Furthermore the factor's ratings were summed to produce landslide hazard index and hazard maps. Finally, the hazard map was verified using existing landslide locations. Flowchart of the methodology is shown in Figure 2.

RESULTS AND DISCUSSION

Application of probabilistic method and its interpretation

Generally, for the prediction of landslide, the following assumption is necessary. Landslide occurrence is determined from landslide related factor and the future landslide can occur in the same condition with past landslide (Chung and Fabbri, 1999). Based on the assumption, the relationship between areas with landslide occurrences and landslide related factors could be distinguished from the relationship between areas without occurrences of landslide and landslide related factors. To represent the distinction quantitatively, the likelihood ratio was used for this study. The likelihood ratio is a ratio between probability of occurrence to probability of not-occurrence in a certain attribute (Bonham-Carter, 1994). Therefore, if the ratio is

Table 3. Likelihood ratio of slope to landslide occurrences.

Class	Landslide occurrence Points	Landslide occurrence Points %	Pixels in domain	Pixel %	Likelihood ratio
0°~5°	63	11.65	1324669	45.24	0.26
6°~10°	41	7.58	119251	4.07	1.86
11°~15°	71	13.12	200779	6.86	1.91
16°~20°	102	18.85	327812	11.19	1.68
21°~25°	108	19.96	366266	12.51	1.60
26°~30°	90	16.64	312124	10.66	1.56
31° ~87°	66	12.20	277477	9.48	1.29

Table 4. Likelihood ratio of aspect to landslide occurrences.

Class	Landslide Occurrence Points	Landslide Occurrence Points %	Pixels in domain	Pixel %	Likelihood ratio
Flat	49	9.06	599634	20.48	0.44
N	39	7.21	194419	6.64	1.09
NE	65	12.01	211666	7.23	1.66
E	72	13.31	460442	15.72	0.85
SE	87	16.08	361722	12.35	1.30
S	66	12.20	184387	6.30	1.94
SW	53	9.80	235551	8.04	1.22
W	57	10.54	359246	12.27	0.86
NW	53	9.80	321311	10.97	0.89

Table 5. Likelihood ratio of curvature to landslide occurrences.

Class	Landslide Occurrence Points	Landslide Occurrence Points %	Pixels in domain	Pixel %	Likelihood ratio
-38.65 ~ -1	165	30.50	681534	23.27	1.31
0	171	31.61	1537754	52.51	0.60
1 ~ 32.26	205	37.89	708885	24.21	1.56
negative curvatures:concave; zero curvature: flat; positive curvatures:convex					

Table 6. Likelihood ratio of distance from drainage to landslide occurrences.

Class	Landslide Occurrence Points	Landslide Occurrence Points %	Pixels in domain	Pixel %	Likelihood ratio
0~200m	416	76.89	2136251	72.96	1.05
201~400m	109	20.15	496758	16.97	1.19
401~600m	16	2.96	164338	5.61	0.53
601~800m	0	0.00	69278	2.37	0.00
801~2200m	0	0.00	61212	2.09	0.00

Table 7. Likelihood ratio of lithology generated with respect to landslide occurrences.

Class	Landslide Occurrence Points	Landslide Occurrence Points %	Pixels in domain	Pixel %	Likelihood ratio
Alluvium-Quaternary	98	18.11	997934	34.08	0.53
Granite	443	81.89	1924800	65.73	1.25

Table 8. Likelihood ratio of distance from lineament to landslide occurrences.

Class	Landslide Occurrence Points	Landslide Occurrence Points %	Pixels in domain	Pixel %	Likelihood ratio
0~200m	176	32.53	894416	30.54	1.07
201~400m	163	30.13	492910	16.83	1.79
401~600m	90	16.64	309474	10.57	1.57
601~800m	47	8.69	229217	7.83	1.11
801~1000m	31	5.73	183714	6.27	0.91
1001~1200m	21	3.88	137806	4.71	0.82
1201~1400m	6	1.11	99222	3.39	0.33
1401~1600m	7	1.29	80340	2.74	0.47
1601~1800m	0	0.00	67134	2.29	0.00
1801~2000m	0	0.00	58035	1.98	0.00
2001~6800m	0	0.00	376110	12.84	0.00

higher than 1, there is a relationship between landslide with certain factors' attribute.

The relationship between landslide and slope angle is as shown in Table 3, the steeper the slope, the greater the probability landslide occurrences. Below 5°, the ratio is lower than 1.00, indicating a very low probability of 0.26 and above 6°, the ratio is greater than 1.00, indicating a higher probability. As the slope angle increases, shear stress in soil or other unconsolidated material generally increases as well (Varnes, 1984). Steep natural slopes resulting from outcropping bedrock, however, may not be susceptible to shallow landslides. In the case of the aspect (Table 4), landslides are most abundant on south-facing and northeast-facing slopes. The frequency of landslides is lowest on east-facing, west-facing and northwest-facing slopes except flat area. Curvature values represent the morphology of the topography. A positive curvature indicates that the surface is upwardly convex at that cell. A negative curvature indicates that the surface is upwardly concave at that cell. A value of zero indicates that the surface is flat. According to curvature (Table 5), positive value has higher probability than the negative value although both were having higher probability of landslide occurrences whereas flat area has a low value of 0.60. The reason is that a convex or concave slope has more water and retains it longer during or after heavy rainfall (Lee, 2002a).

An analysis has been carried out to assess the influences of drainage on landslide occurrences. As shown in Table 6, it was found that as the distance from drainage increases, landslide frequency generally decreases. Below distance of 400 m, the ratio is higher than 1.00, indicating a high probability whereas distance above than 600 m, the ratio is 0.00, indicating zero probability. This may be due to terrain modification caused by gully erosion that may influence the initiation of landslides.

For geological factors such as lithology (Table 7), shows that the likelihood ratio is higher in granite areas; 1.25, and is lower in alluvium areas, 0.53. In case of the distance from lineament (Table 8), the closer to the lineament the greater the probability of landslide occurrences. Distance below 800 m shows a ratio of above 1.00, indicating a high probability. As the distance from lineament decreases, the fracture of the rock increases and degree of weathering generally increases as well.

In the case of land use (Table 9), landslide-occurrence values were higher for scrub, rubber and mixed areas but lower for rice, swamp, coconut, barren and oil palm areas. The reason was that the landslides occurred mainly in inclined and disturbed mountainous areas. A vegetation index (Table 10), value of below 0.20, the ratio is lower than 1.00, indicating a low probability and value above 0.20, the ratio is higher than 1.00, indicating a high probability. The result indicates that the landslide probability increases with density of vegetation.

Landslide susceptibility mapping and verification

The correlation ratings were calculated from relation analysis between landslides and the relevant factors. Therefore, the rating of each factor's type or range was assigned as the relationship between landslide and each factor's type or range. That relationship was represented as the ratio of the number of cells where landslides were occurred to the number of cells where landslides not occurred as shown in Table 3 to Table 10. The landslide hazard index (LHI) is calculated by summation of each factor's ratio value.

$$\mathbf{LHI} = \Sigma \mathrm{Fr} \qquad (1)$$

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

The calculated LHI have a minimum value of 1.93 and a maximum value of 15.82. The average value is 8.02 whereas standard deviation value is 2.32. The relation analysis is the ratio of the area where landslides occurred to the total area, and the average value of 8 were used. A value greater than 8, indicates a higher correlation, and a value lower than 8 indicates lower correlation. The landslide-hazard map was made using the LHI value index and then was classified using equal areas and grouped into six classes as shown in Figure 3. The indicated hazardous areas are Paya Terubung, Bukit Relau, Bukit Gemuruh and Teluk Bahang.

For the verification of the landslide hazard calculation methods, two basic assumptions were needed. Firstly, landslides were related to factors such as slope, aspect, curvature, distance from drainage, geology, distance from lineament, land use and vegetation index, and secondly, future landslides can be predicted by a specific impact factor such as rainfall or earthquake (Chung and Fabbri, 1999). In this study, the two assumptions as mentioned by Chung and Fabbri, 1999, are satisfied because the landslides are related to the spatial information and the caused by heavy rainfall for the study area.

The success rate of verification results from comparing the hazard calculation results and landslide occurrence location using likelihood method is shown as a line graph in Figure 4. The success rate illustrates how well the estimators performs (Chung and Fabbri, 1999). 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 accumulation of 1% intervals. An index value above of 11.33 indicates that 10% of the study area where landslide susceptibility index is higher in rank and comprises 41% of all the landslides. Furthermore, an index value above than 10.60 comprises 30% of the area and 68% of the landslides.

Table 9. Likelihood ratio of landuse generated with respect to landslide occurrences.

Class	Landslide Occurrence Points	Landslide Occurrence Points %	Pixels in domain	Pixel %	Likelihood ratio
Urban	113	21.04	34406	1.18	0.74
Mixed	292	54.38	828260	28.29	1.66
Forest	65	12.10	949007	32.42	0.48
Scrub	7	1.30	732255	25.01	5.62
Aquaculture	5	0.93	6739	0.23	0.75
Swamp	0	0.00	35960	1.23	0.00
Rubber	50	9.31	35333	1.21	1.95
Rice	5	0.93	138916	4.74	0.20
Coconut	0	0.00	136099	4.65	0.00
Barren	0	0.00	9871	0.34	0.00
Oil Palm	0	0.00	2094	0.07	0.00

Table 10. Likelihood ratio of vegetation index to landslide occurrences.

Class	Landslide Occurrence Points	Landslide Occurrence Points %	Pixels in domain	Pixel %	Likelihood ratio
-0.80 ~ -0.60	0	0.00	11909	0.41	0.00
-0.60 ~ -0.40	4	0.74	39451	1.35	0.55
-0.40 ~ -0.20	29	5.36	233370	7.97	0.67
-0.20 ~ 0.00	42	7.76	317441	10.84	0.72
0.00 ~ 0.20	54	9.98	295593	10.10	0.99
0.20 ~ 0.40	189	34.94	1015840	34.70	1.01
0.40 ~ 0.61	223	41.22	1011468	34.55	1.19
Number of total cells in study area: 2,928,378					
Number of landslide occurrence points: 541					

CONCLUSION

Landslide hazard maps are useful to planners and engineers for choosing suitable locations to implement developments. Although the results can be used as a basic data to assist slope management and land-use planning, the methods used in the study are only valid for generalized planning and assessment purposes, and may be less useful at the site-specific scale where local geological and geographic heterogeneities prevail.

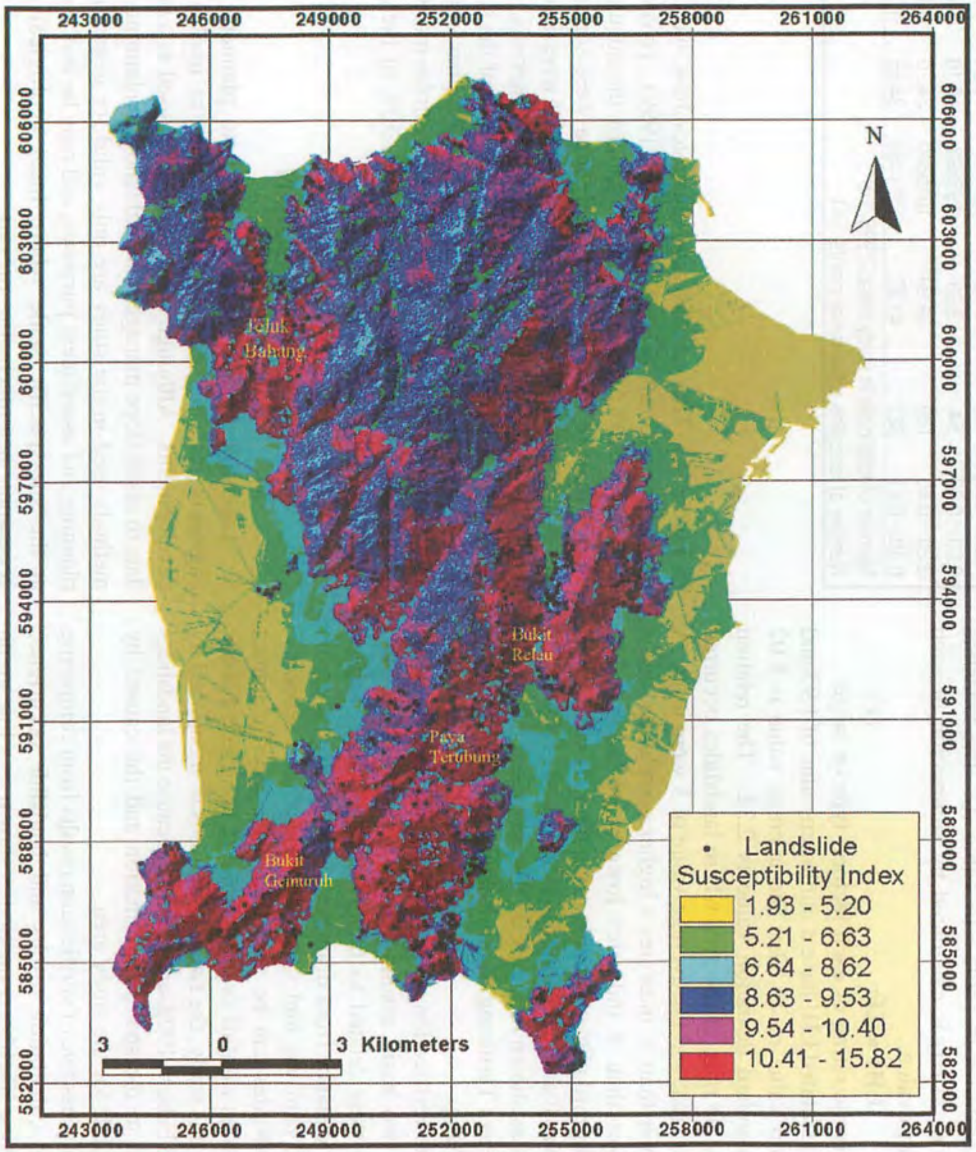

Figure 4. Cumulative frequency diagram showing landslide susceptibility index rank occurring in cumulative percentage of landslide occurrence.

Figure 3. Landslide susceptibility map based on likelihood ratio.

REFERENCES

BONHAM-CARTER, G.F., 1994. *Geographic Information Systems for geoscientists, modeling with GIS*. Oxford, Pergamon Press.

CHUNG, C.F. AND FABBRI, A.G., 1999. Probabilistic prediction models for landslide hazard mapping, *Photogrammetric Engineering & Remote Sensing 65*, 1389-1399.

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

DAI, F. AND LEE, C.F., 2002. Landslides on Natural Terrain: Physical Characteristics and Susceptibility Mapping in Hong Kong. *Mountain Research and Development, 22*, 40-47

DHAKAL, A.S., AMADA, T. AND ANIYA, M., 1999. Landslide Hazard Mapping and the Application of GIS in the Kulekhani Watershed, Nepal. *Mountain Research and Development, 19*, 3-16.

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*, 277-289.

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*, 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*, 181-216.

LEE, S. AND MIN, K., 2001. Statistical analysis of landslide susceptibility at Yongin, Korea. *Environmental Geology, 40*, 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*, 49-162.

LEE, S., CHOI, J. AND MIN, K., 2002b. Landslide susceptibility analysis and verification using the Bayesian probability model, *Environmental Geology 46*, 120-131.

MAHADZER, M., 2001. Landslide investigation and slope risk management. *National Slope Seminar 2001*, Cameron Highlands, Pahang Darul Makmur.

MOHD. ASBI, O. AND LLOYD, D.M., 2001. Slope instability problems of roads in mountainous terrain: A geotechnical perspective. *National Slope Seminar 2001*, Cameron Highlands, Pahang Darul Makmur.

VARNES, D.J., 1984. Landslide Hazard Zonation: a review of principles and practice. Commission on landslides of the IAEG, UNESCO. *Natural Hazards, 3*, 61p.

Manuscript received 21 February 2003