Influence of Rainfall on Landslide Susceptibility along 50 to 110k Section of the Southern Cross Island Highway, Taiwan

Hsun-Chuan Chan,[1] Jhih-Syong Peng[1] and Chia-Chi Chang[1,*]

[1] Dept. of Soil and Water Conservation, National Chung Hsing Uni. (250 Kuo Kuang Rd., Taichung 402, Taiwan)
*Corresponding author. E-mail: d100042004@mail.nchu.edu.tw

The inventories of landslide during typhoon Mindulle, Morakot, and 0719 rainfall event by Central Geological Survey were selected as the landslide data. The elevation, slope, slope aspect, slope high, lithology, terrain roughness, slope roughness, plan curvature, profile curvature, total curvature, distance of road, and distance of river were first chosen as the landslide causative factors, according to the previous studies. Secondly, the calibration and selection procedure were performed to select the factors efficiently. Logistic regression method was used for establishing the landslide susceptibility model. Furthermore, the rainfall intensities of different rainfall duration were used as a landslide triggering factor in different rainfall events. All landslide prediction accuracies for the above event models are not far from 70%. The model of event with long rainfall duration and high rainfall intensity shows the largest accuracy increase after using rainfall as triggering factor.

Key words: Logistic regression, the Southern Cross Island Highway, rainfall intensity

1. INTRODUCTION

The Southern Cross Island Highway of Taiwan runs through hills, valleys and mountain peaks on the southern ridges of Central Mountain Range, Taiwan. It plays a major role in transportation between east and west part of southern Taiwan and has colorful landscapes along the way which are important national nature properties. Typhoons and rainfalls in recent years cause many slope failures along the Southern Cross Island Highway. Slope failures along the road not only block traffic but also endanger the drivers (**Fig. 1**). Hence, this study uses slopes along section between 50k and 110k of the Southern Cross Island Highway with road corridors range of 1400m as research objects to investigate the influence of different rainfall events on landslide susceptibility analysis.

The previous methods for landslide susceptibility analysis can be divided into two major categories: qualitative analysis and quantitative analysis. Early landslide susceptibility analysis is often performed with qualitative analysis; that is, direct or indirect landslide susceptibility analysis performed with specialized knowledge and experience of experts [*Kienholz*, 1978; *Ives and Bovis*, 1978]. However, the results of qualitative methods are easily influenced by personal subjectivity. Hence, qualitative analysis is rarely used these days. In recent years, artificial neural network [*Lin et al.*, 2009] and statistical analysis are mainstream methods for landslide susceptibility. Artificial neural networks connect a large number of simple artificial neurons to simulate the capability of biological neural networks. With trained artificial neural network, the analysis accuracy is comparatively better than other methods. Because artificial neural network which finished learning has recall capability, the regression fit results of artificial neural network are better than ordinary linear equations. However, the time consuming try and error process is disadvantageous for wide area analysis. Statistical analysis method is categorized into bivariate analysis and multivariate analysis. Statistical analysis method establishes the inventory of existing landslide location data including landslide causative factor characteristics like topography, geology and location. The established inventory is analyzed to choose a proper factor set for discriminating between landslides and non-landslides in study area and effectively

predicting landslide susceptibility for locations at which landslide still not occurs or area of similar geological characteristics. The deterministic analysis method based on dynamic principles is more exact. However, geotechnical coefficients of wide area are difficult to obtain. Hence, the predictive capability of deterministic analysis on wide area is limited. This study chooses wide-application statistical method for landslide susceptibility analysis. Considering bivariate method may produce model prediction error in the case of analyzing two highly correlating factors. The analysis of this study also includes factors of non-quantifiable categories. Hence, the logistic regression of multivariate analysis is used for analysis in this study.

Most researches on the relationship between rainfall and slope failure focus on rainfall intensities and total rainfall [*Chen et al.*, 2003]. However, the time correlation of rainfall causes larger difficulty and uncertainty for rainfall measurement. [*Yang*,1989] investigated the relationship between rainfall and landslide timing with landslide cases in 1965 to 1987 and concluded that most landslides occur at times with largest rainfall intensities, not at times with largest total rainfall. However, studying a single factor of certain rainfall event cannot ensure that the factor under study is sufficient to represent the characteristics of every rainfall events. Hence, this study attempt to choose factors which represent the characteristics of each rainfall event and has the capability to discriminate between landslide and non-landslide data samples by performing analysis with the average rainfall intensities in different rainfall durations. Then, discuss the variable of the analysis model and hope the analysis process can improve the accuracy of landslide potential analysis.

Fig. 1 Landslide in the Southern Cross Island Highway

2. STUDY AREA

The area under study is located at southern Taiwan and encompasses part of Tainan and Kaohsiung City, as shown in **Fig.2**. The length of chosen section is about 60km; the area around chosen section is about 147km^2; the elevation of chosen section ranges from 102m to 1607m. The slope ranging from 55% to 100% in chosen section accounts for 38.53% of the area around chosen section. The area under study is mostly located in mountain country. It is expected that the study area with rugged topography and steep slope is extremely easy for serious slope failure to occur after typhoon, heavy rainfall or earthquake strikes. The distribution of slope aspect is homogeneous. The proportion of southeast aspect (15.78%) and south aspect (14.41%) are slightly higher than the other aspect. The proportion of other aspect shows little difference. In recent 11 years, the annual rainfall the study area is more than 3000mm due to monsoon climate and local topography, which is much more than the average annual rainfall of Taiwan (2510mm). Moreover, 80% to over 90% of abundant annual rainfall concentrates between May and September. A small portion of rainfall is from plum rain period, but most rainfall is from typhoon events.

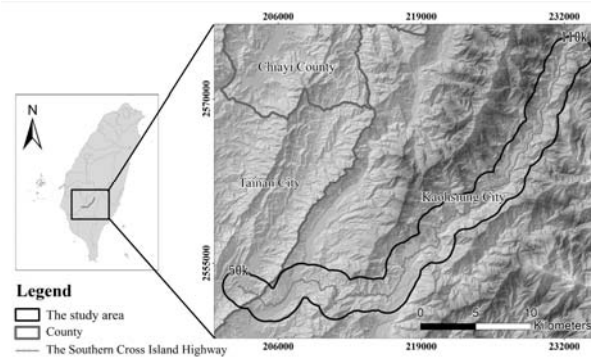

Fig. 2 The study area

3. RESEARCH METHOD

3.1 Slope unit

Slope unit is the analysis cell used in this study. The following analysis uses the slope unit produced by the method of reverse DEM [*Xie et al.*, 2004].

3.2 Logistic regression

Linear regression model is the most common statistical method, but it is not useful for the case that variables are classified variables. This study uses logistic regression model for using classified variable to the analysis. Logistic regression has S-shaped distribution curve whose value range is between 0 and 1. The distribution curve is similar to an accumulated distribution curve of random

variable. Logistic regression model is represented as the following equations [*Gregory and John*, 2003]:

$$P(y_i = 1|x_i) = \frac{1}{1 + e^{-(Z)}} \quad (1)$$

$$Z = \alpha + \beta x_i \quad (2)$$

where $P(y_i = 1)$ is the condition probability of landslide event; x_i is the independent variable column vector; α and β are regression intercept range and row vector of regression coefficient. In the case of $P(y_i = 1|x_i)$, the above equations can be shown as:

$$\ln\left(\frac{P_i}{1 - P_i}\right) = \alpha + \sum_{i=1}^{k} \beta_k x_{ki} \quad (3)$$

where $P_i = P(y_i = 1 \mid x_{1i}, x_{2i}, ..., x_{ki})$ denotes the probability of an event as a function of independent variables $x_{1i}, x_{2i}, \ldots, x_{ki}$.

3.3 Model capability evaluation

This study validates models with classification error matrix table and receiver operation characteristic curve to examine the analysis performance of models.

4. LANDSLIDE SUSCEPTIBILITY MODEL

4.1 Basic Data

(1) High accuracy digital elevation model(DEM) of 5×5m size.
(2)The 1:250000 geography map made by Central Geological Survey.
(3)The 2011 digital road network map published by Institute of Transportation, MOTC and including highway and river networks of Taiwan
(4)The hourly rainfall during typhoon Mindulle, Morakot, and 0719 rainfall events.
(5)The landslide event inventories, established by Central Geological Survey, Taiwan, for typhoon Mindulle, Morakot, and 0719 rainfall events.

4.2 Analysis Samples

The slope unit is defined as landslide and non-landslide group. The classification standard is 100m2 per grid area; that is, any slope unit is classified as landslide group only in the case that the landslide area in the slope unit is larger than 100m^2; otherwise, the slope unit is non-landslide group. By the classification standard mentioned above, the numbers of slope unit as landslide group in typhoon Mindulle, Morakot, and 0719 rainfall events are 206, 1223 and 1400, respectively. The numbers of slope unit as non-landslide group are in typhoon Mindulle, Morakot, and 0719 rainfall events 1892, 1875 and 1698, respectively.

4.3 Landslide Factors

4.3.1 Potential factors

There are numerous landslide susceptibility factors in the previous literatures [*Atkinson and Massari*, 1998; *Lee and Min*, 2001; *Dai et al.*, 2003, 2004; *Ayalew and Yamagishi* , 2005; *Chang et al.*, 2007; *Lee and Pradhan*, 2007; *Chauhan et al.*, 2010]. The characteristics of operational, complete, non-uniform, measurable and non-redundant are first consideration of factor selection. The initially chosen factors can be used for landslide susceptibility analysis include elevation, slope, slope aspect, slope high, lithology, terrain roughness, slope roughness, plan curvature, profile curvature, total curvature, distance of road, and distance of river. In second selection, landslide and non-landslide frequency distribution map, proportion of landslide cell, discriminator, success rate curve, P-P plot (probability plot) and correlation coefficient are used to examine landslide discrimination capability of factors. The proper factors for model establishing of typhoon Mindulle, Morakot, and 0719 rainfall event are chosen with the selection procedure mentioned above. The chosen factors for typhoon Mindulle are maximum elevation, average slope, terrain roughness and profile curvature; the chosen factors for typhoon Morakot are average elevation, average slope, terrain roughness and profile curvature; the chosen factors for 0719 rainfall event are average elevation, average slope, slope high and plain curvature. The aspect and lithology are in classified form. Past literatures have acknowledged the influence of aspect and lithology on landslide research. Hence, these two factors are used for model establishment without further selection.

4.3.2Triggering factors

In researches on landslide susceptibility analysis, triggering factor means outside interference. Rainfall plays an important role in investigating the susceptibility analysis of rainfall triggered landslide. The effects of rainfall on slope face are also changed by the difference of topography, geology and environment.

This study uses average rainfall for representing the rainfall characteristics in study area (**Table 1**). In comparison of the three rainfall events, the Morakot event rainfall duration of 101.39 hour is much larger than those of other two events and even twice as long as that of 0719 rainfall event. The maximum hourly rainfall intensity of Morakot event shows the same characteristic. Hence, this study defines Morakot event as long duration and high intensity

rainfall event. The Mindulle event has rainfall duration of 71.63 hour and rainfall intensity of 67.00mm/hr. Although the Mindulle event rainfall intensity is smaller than that of Morakot event, it is still large enough to make the Mindulle event defined as middle duration and high intensity. The rainfall duration and intensity of 0719 rainfall event are only half of those of Morakot event. Hence, 0719 rainfall event is defined as short duration and middle intensity rainfall event. With the maximum average rainfall intensities of 1, 3, 6, 9, 12, 24 hours of three different rainfall events, this study uses kriging method for estimating rainfall space distributions. The estimation results provide advices for selecting proper time scale of rainfall factors in the study area.

Table 1 Average rainfall factor in different events

Event	Mindulle	Morakot	0719 rainfall
	2004	2009	2011
	June 28 to July 3	Aug. 5 to Aug.10	July 19 to July 20
Duration (hr)	71.63	101.39	52.69
Total rainfall (mm)	1182.25	1973.83	384.19
Average rainfall intensity (mm/hr)	15.71	19.23	7.29
1-h max rainfall intensity (mm/hr)	67.00	87.47	36.47
3-h max rainfall intensity (mm/hr)	45.05	73.59	25.71
6-h max rainfall intensity (mm/hr)	36.88	63.67	18.90
9-h max rainfall intensity (mm/hr)	32.18	60.81	16.80
12-h max rainfall intensity (mm/hr)	29.36	56.56	14.97
24-h max rainfall intensity (mm/hr)	24.69	43.32	10.64

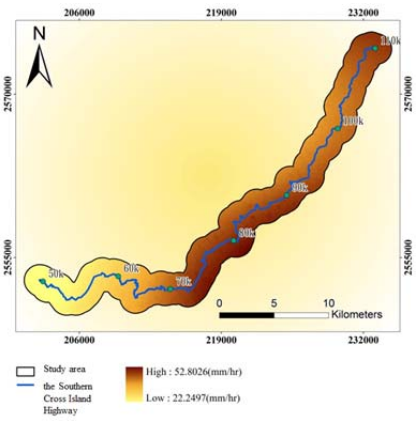

Fig. 3 The spatial distribution of the rainfall during typhoon Mindulle.

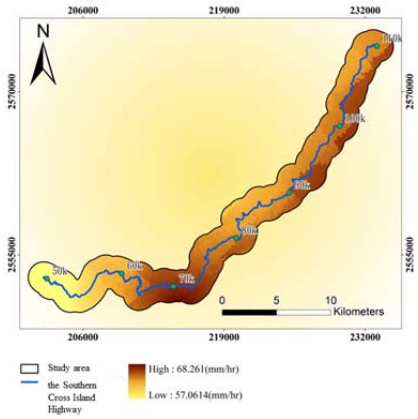

Fig. 4 The spatial distribution of the rainfall during typhoon Morakot.

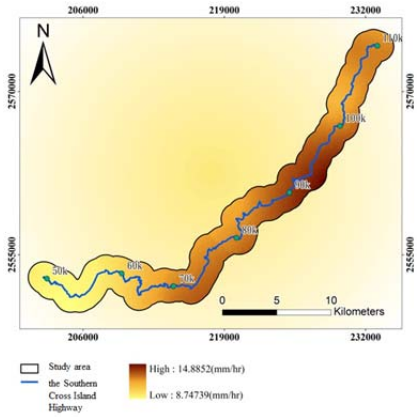

Fig. 5 The spatial distribution of the rainfall during 0719 rainfall.

The spatial distribution of rainfall data during different rainfall events are in Fig.3- Fig.5.

5. RESULTS AND DISCUSSION

Table 2 displays the regression coefficients of the results of establishing model with causative factors in the three events. Positive regression coefficient value means the landslide probability increases with factor increase; otherwise. Negative regression coefficient value means negative correlation between factor value and landslide probability; that is, the landslide probability decreases with factor increasing. Hence, the factor weight is represented by the absolute value of factor regression coefficient. Factor with large factor weight has large influence on landslide or non-landslide occurrence, so it is also more important than factor with smaller factor weight. **Table 3** shows the accuracies and ROC curves of the established models of the three events. The model established with landslide event inventory of Mindulle has total accuracy of 75%, area under ROC curve of 0.816 and about 80% accuracy in

non-landslide event prediction. Although both the total accuracies of Morakot and 0719 rainfall event model are below 70%, the accuracies of landslide groups are still acceptable. The major flaw of Morakot and 0719 rainfall event model is their relatively low predictive capability of landslide event. The possible cause of model capability difference is the difference between the landslide data inventory distributions of the events. Mindulle typhoon mainly affects middle Taiwan, so most landslide events in the study area occur at locations along the highway with steep topography. Only a few landslide events occur at locations with placid topography. The landslide event distributions of Morakot and 0719 rainfall are obviously wider; landslide events occur regardless of topography steepness. However, all of the three event models have slope factors positively correlating with landslide probability. Hence, the larger influence of slope on discrimination between landslide and non-landslide grid-cells causes Mindulle typhoon event model to have higher landslide prediction accuracy than other models.

Table 2 The results of building model in different events by causative factor

		The regression coefficients				
		Mindulle		Morakot	0719 rainfall	
	maximum elevation	1.832	average elevation	-1.595	average elevation	-1.405
	average slope	5.746	average slope	3.924	average slope	4.825
	terrain roughness	1.645	slope high	1.317	slope high	1.334
	profile curvature	-0.077	terrain roughness	0.234	plan curvature	0.291
			slope roughness	0.858		
			profile curvature	0.397		
lithology	1 (Lushan Fm)	—		—		—
	2(juifang group)	-0.241		-2.062		-1.896
	3(Sanhsia group)	-1.106		0.249		-0.113
	4(Cholan formation)	-0.232		1.156		0.327
	5(accumulation platform)	1.682		-2.818		-20.353
	6(alluvium)	-17.666		-19.104		-18.591
slope aspect	north	—		—		—
	northeast	-0.644		-0.015		0.137
	east	0.298		-0.009		0.502
	southeast	0.934		0.056		0.470
	south	1.231		0.142		0.372
	southwest	0.961		0.057		0.199
	west	0.418		-0.074		-0.004
	northwest	-0.252		-0.571		-0.470
	constant	-4.998		-2.442		-2.857

Table 3 The results of evaluating the accuracy of landslide predicted by the causative factor model

	The accuracy on classification error matrix (%)			Area under ROC curve
	Landslide	Non-landslide	Overall correct	
Mindulle	69.50	80.50	75.00	0.816
Morakot	70.02	62.93	66.82	0.734
0719 rainfall	78.83	53.64	68.88	0.756

The accuracies of every event model with the maximum average rainfall intensities in different durations are also derived. The representative triggering factors of the models are chosen as the durations of maximum average rainfall intensity with the highest model accuracies. The triggering factors of Mindulle, Morakot and 0719 rainfall events are chosen as 6 hour, 9 hour and 24 hour maximum average rainfall intensity respectively. The factor regression coefficients are sorted in **Table 4**. The model capability evaluations are shown in **Table 5**. In the case of Mindulle event, the model with rainfall factor shows little accuracy increase; in the case of Morakot event, the rainfall factor increases the total accuracy from 66.82% to 71.34%. The rainfall factor obviously plays an extremely important part in Morakot model establishment; in the case of 0719 rainfall event, the rainfall factor only increases the total accuracy from 68.88% to 70.72%; the total accuracy only increases less than 2%. Regardless of the total accuracy increase of model with rainfall triggering factor, the model establishment procedure formulated by this study builds models keeping landslide distribution prediction accuracies above 70% for different rainfall events. Furthermore, all of the areas under ROC curves shown in **Table 5** are above 0.75; the area under ROC curve of Mindulle event is even up to 0.817. It is clear that the present model achieve good results for all three events.

Furthermore, a probability value of 0.5 was selected as the cut-off value (Dai and Lee, 2002), which distinguishes the cells into 2 cases: places where landslide occurred and did not occur. **Fig.6** shows a comparison between predicted landslide which without rainfall data with actual landslide by logistic regression model. **Fig.7** shows a comparison between predicted landslides which join rainfall data with actual landslide by logistic regression model. The logistic regression model generally shows good classifications for the presence and absence of landslides.

Table 4 The results of building model in different events by triggering factor

The regression coefficients						
Mindulle			Morakot		0719 rainfall	
maximum elevation		1.771	average elevation	-3.676	average elevation	-2.430
average slope		5.731	average slope	4.108	average slope	4.195
terrain roughness		1.657	slope high	2.433	slope high	1.956
profile curvature		-0.070	terrain roughness	0.760	plan curvature	0.209
			slope roughness	0.628		
			profile curvature	0.302		
lithology	1 (Lushan Fm)	—		—		—
	2(juifang group)	0.095		-1.104		-0.467
	3(Sanhsia group)	-1.001		0.648		0.608
	4(Cholan formation)	-0.185		1.717		0.841
	5(accumulation platform)	2.132		-0.433		-18.935
	6(alluvium)	-17.334		-17.348		-17.152
slope aspect	north	—		—		—
	northeast	-0.655		0.037		0.065
	east	0.316		0.044		0.433
	southeast	0.938		0.141		0.401
	south	1.251		0.243		0.333
	southwest	0.965		0.103		0.165
	west	0.450		0.006		0.038
	northwest	-0.226		-0.523		-0.413
triggering factor		1.082		26.016		7.099
constant		-5.921		-26.964		-8.365

Table 5 The results of evaluating the accuracy of landslide predicted by the triggering factor model.

	The accuracy on classification error matrix (%)			Area under ROC curve
	Landslide	Non-landslide	Overall correct	
Mindulle	69.50	80.50	75.00	0.817
Morakot	73.14	69.14	71.34	0.778
0719 rainfall	78.67	58.54	70.72	0.780

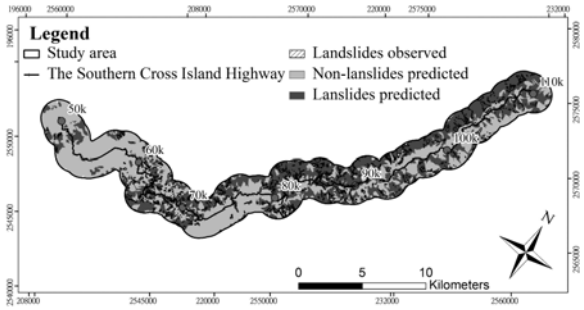

Fig. 6 Compared predicted landslide (without rainfall data) with actual landslide by logistic regression

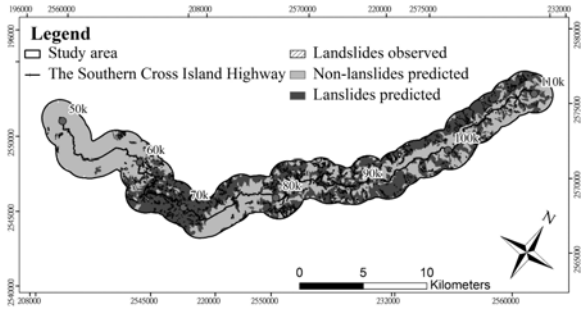

Fig. 7 Compared predicted landslide (join rainfall data) with actual landslide by logistic regression model

6. CONCLUSION

From the available landslide inventory maps, this study chooses causative factors useful for establishing landslide susceptibility analysis models of Mindulle, Morakot and 0719 rainfall event. Models of different event with the chosen causative factors and average rainfall intensities in different rainfall durations as triggering factors are established and investigated for the model predictive capability variation correlating to rainfall events. The investigation results are described below:

(1)The accuracies of models established with causative factors are 75.00% for Mindulle event, 66.82% for Morakot event and 68.88% for 0719 rainfall event. All landslide prediction accuracies for the above event models are not far from 70%, so the landslide susceptibility analysis procedure established by this study is useful for predicting landslides in the study area.

(2)In the cases of models using triggering factors, the total accuracy variation of Mindulle event is nearly zero; the total accuracy of Morakot event increases 4.52%; the total accuracy of 0719 rainfall event increases 1.84%. In this study, the model of event with long rainfall duration and high rainfall intensity shows the largest accuracy increase after using rainfall as triggering factor. The models of other two events are effective enough without triggering factor.

(3)The results can support planning of the Southern Cross Island highway's surrounding area and location, and be a reference for disaster warning. For instance, the potential landslide zones are between 80k and 110k, when the typhoon comes, the relevant government agency can issue the warning to reduce the economic losses, deaths, and injuries.

(4)This study only uses three different types of rainfall events to establish the analysis process. Therefore, we suggest that add more different rainfall events in the future might prove the availability of the analysis model.

REFERENCES

Atkinson, P. M. and R. Massari. (1998): Generalized linear modelling of susceptibility to landsliding in the central Apennines, Italy, Comput. Geosci. Vol. 24, pp.373-385.

Ayalew, L. and H. Yamagishi. (2005): The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan, Geomorphology. Vol.65, pp.15-31.

Chang, K. T., S. H. Chiang and M. L. Hsu. (2007): Modeling typhoon- and earthquake-induced landslides in a mountainous watershed using logistic regression, Geomorphology. Vol.89, pp.335-347.

Chauhan, S., M. Sharma and M. K. Arora. (2010): Landslide susceptibility zonation of the Chamoli region, Garhwal Himalayas, using logistic regression model, Landslides. Vol.7, pp.411-423.

Chen, S.C., Chang, C.C., Chan, H.C., * Huang, L.M., and Lin, L.L. (2013): Modeling typhoon event-induced landslides using GIS-based logistic regression: A case study of Alisan Forestry Railway, Taiwan, Mathematical Problems in Engineering, Vol. 2013, Article ID 728304,

doi:10.1155/2013/728304.

Dai, F. C. and C. F. Lee. (2003): A spatiotemporal probabilistic modeling of storm-induced shallow landsliding using aerial photographs and logistic regression, Earth Surf. Process. Landf. Vol.28, pp.527-545.

Ives, J. D. and M. J. Bovis.(1978): Natural hazards maps for land-use planning, San Juan Mountains, Colorado, U.S.A, Arct. Antarct. Alp. Res. 10:185-212.

Kienholz, H. (1978),"Maps of geomorphology and natural hazards of Grindelwald, Switzerland, scale 1:10,000", Arct. Antarct. Alp. Res. 10:169-184.

Lee, S. and B. Pradhan. (2007): Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models, Landslides. Vol.4, pp.31-44.

Lee, S. and K. Min. (2001): Statistical analysis of landslide susceptibility at Yongin, Korea, Environmental Geology. Vol.40, pp.1095-1113.

Lin, H.M., Chang, S. K., Wu, J.H., Juang, C. H.(2008): Neural network-based model for assessing failure potential of highway slopes in the Alishan, Taiwan Area: Pre- and post-earthquake investigation", Engineering Geology,Vol. 42, pp.51–63.

Xie M., T. Esaki, and G. Zhou, (2004): GIS-Based Probabilistic Mapping of Landslide Hazard Using a Three-Dimensional Deterministic Model, Natural Hazards, Vol .33, pp.265-282.

Yang, J.S.(1989): Study on mechanical properties of representative rocks at Mt. So-San and In-Situ measurements of slope movements. Master thesis, National Cheng Kung University, Taiwan, ROC. (in Chinese)