GIS Application in Landslide Hazard Analysis – An Example from the Shihmen Reservoir Catchment Area in Northern Taiwan

Chyi-Tyi Lee, Institute of Applied Geology, National Central University,
No.300, Jungda Rd., Jungli City, Taoyuan County, 32001, Taiwan
Email: ct@ncu.edu.tw

Abstract: This study used GIS as a tool to map storm-induced landslides from SPOT5 images. Digital elevation model (DEM) of 10m x 10m resolution was used to extract geomorphic landslide causative factors: slope gradient, slope roughness, tangential curvature, relative slope height, total slope height and wetness index etc. Digital geological map was used to extract geologic causative factors: lithology and distance to fault trace. SPOT image taken prior to a typhoon was also used to calculate an environmental factor - NDVI (normalized differential vegetation index). Triggering factor was tested to be the maximum rainfall intensity. These causative factors and triggering factors were used to build a landslide susceptibility model via logistic regression. Validation result shows that this model could be used for the prediction of future landslides. GIS is a useful tool for the construction of landslide prediction model and for application in regional planning, hazard mitigation, and sediments yield estimation.

Introduction

Advances in Geographical Information Systems (GIS) technology and the mathematical/statistical tools for modeling and simulation, have led to the growing application of quantitative techniques in many areas of the earth sciences (Carrara & Pike, 2008). The study of landslide hazard also applied these basic tools frequently with intensively use of digital elevation models (DEMs) and SOPT images.

This study firstly used GIS to map storm-induced landslide distributions from SPOT5 images taken prior and after a typhoon storm, and an event-based landslide inventory was built. For the purpose of constructing a landslide hazard model, high resolution DEM was used to extract geomorphic landslide causative factors, such as: slope gradient, slope roughness, tangential curvature, relative slope height, total slope height and wetness index etc. These factors were processed in a raster GIS – Erdas Imagine. Digital geological map was used to extract geologic causative factors: lithology and distance to fault trace. They were processed in vector GIS – MapInfo, and were transferred to raster cells in the raster GIS. SPOT image taken prior to the typhoon was also used to calculate an environmental factor - NDVI (normalized differential vegetation index) in Erdas Imagine system. Hourly rainfall data from 26 rain gauge stations in and around the catchment area were used to process rainfall factors at each station, and then these point data were spatially interpolated to each raster cell in the study area.

Landslide causative factors and triggering factors were used to build a landslide susceptibility model via logistic regression. A validation work was also done to check the performance of using this model for prediction. Result shows that this model could be used for the prediction of future landslide occurrences during a scenario event. It may be useful for the decision making of slope remedial measure, regional planning and hazard mitigation policy.

Regional Setting

The Shihmen Reservoir is an important water resources reservoir in northern Taiwan (Fig. 1). It has a catchment area of 763.4 km^2. Elevations in the watershed range from 135 m in the northwest to 3,524 m in the southeast, with generally rugged topography. Slopes with gradient greater than 55% covers 60.5% of the catchment area, slopes gradient of 30~55% occupies 29.3% area, only 10.2% is gentle slopes with gradient less than 30%. Rocks are composed of folded and faulted Miocene and Paleogene indurate sandstone and mudrocks (Fig. 1). Terraces on the river sides are composed of sandy gravels and they may be covered by lateritic soils on high terraces. Slopes are commonly mantled by shallow slope washes or colluvium. Nearly 90% of the study area is forested. The climate is influenced by typhoons in summer and the northeast monsoon in winter. The mean annual temperature is 20°C, with a mean monthly temperature of 27.5 °C in July and 14.2 °C in January. The annual precipitation averages

2,370 mm. Because of the visiting of typhoons, large rainfall events usually happen from May to September.

From 23 to 25 August 2004, Typhoon Aere crossed the northern part of Taiwan. The passage of Typhoon Aere brought 1,604 mm of rainfall to the study area, with a maximum rainfall intensity of 51.7 mm/hr. During the Typhoon Aere, there occurred numerous landslides in the catchment area and caused the reservoir water become turbid, and the sediments even blocked the water intake and suspend water supply for 20 days. Millions of people and thousands of factories suffered this disaster.

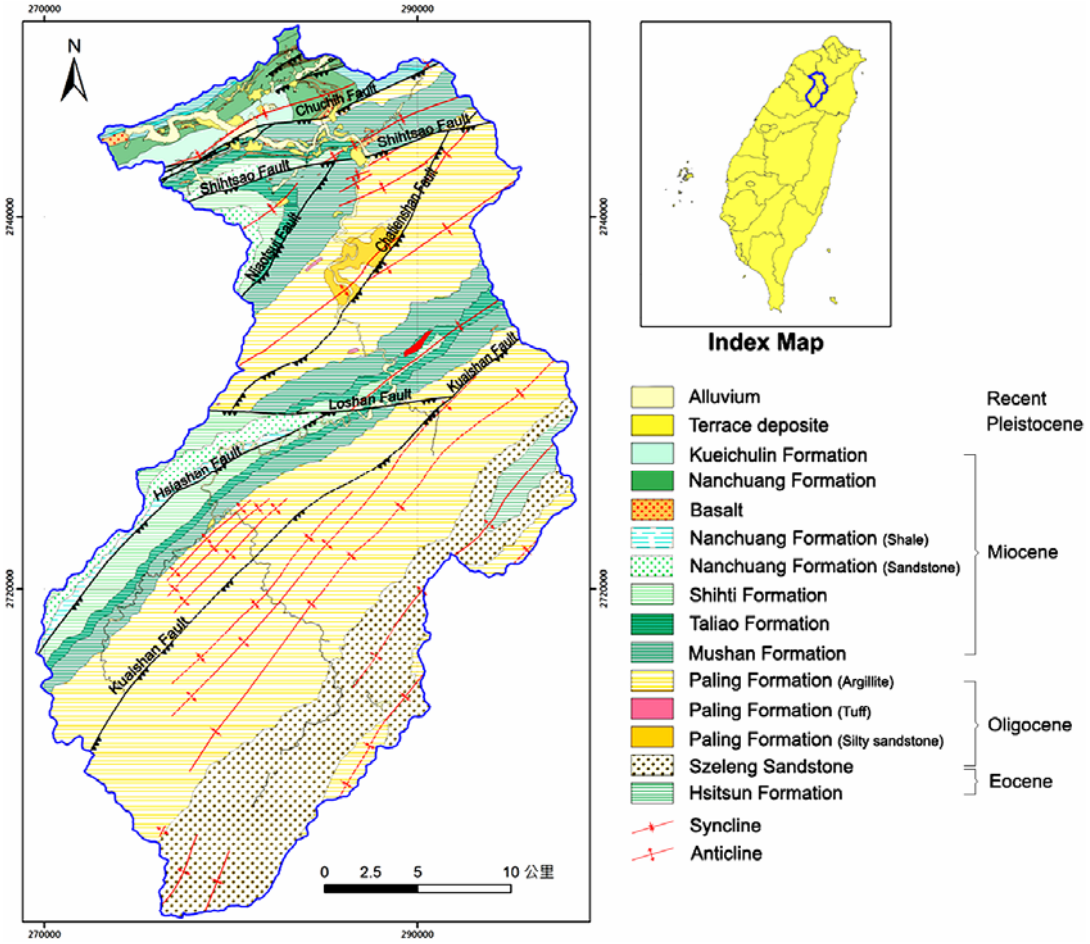

Fig. 1 Geological map of the Shihmen Reservoir catchment area

Methodology and Working Procedure

The methods and working procedure utilized in the present study generally follow those of Lee et al. (2008a, 2008b). The first step includes image and data collection. Then an event-based landslide inventory is established. In parallel with this, the causative factors of the landslides are processed and the triggering factors determined. These factors are then statistically tested, and the effective factors selected for susceptibility analysis. The working procedure is shown in Fig. 2.

Logistic regression allows us to determine a linear function of factors for interpreting the landslide distribution from a set of training data. The methodology has been well established in many previous studies (e.g., Carrara, 1983; Atkinson & Massari, 1998; Dai et al., 2001; Ayalew & Yamagishi, 2005; Eeckhaut et al., 2006; Greco et al., 2007). The linear function is used to calculate the landslide susceptibility index (LSI) for each cell. The LSIs are then used to establish a probability of failure to LSI curve and determine the spatial probability of landslide occurrence at each cell. The probability of failure used here is the ratio of landslide cells to total cells in an LSI bin (Lee et al., 2008a, 2008b). It is called the proportion of landslide cells in Jibson et al. (2000). The spatial probability

of landslides is then used for landslide hazard mapping.

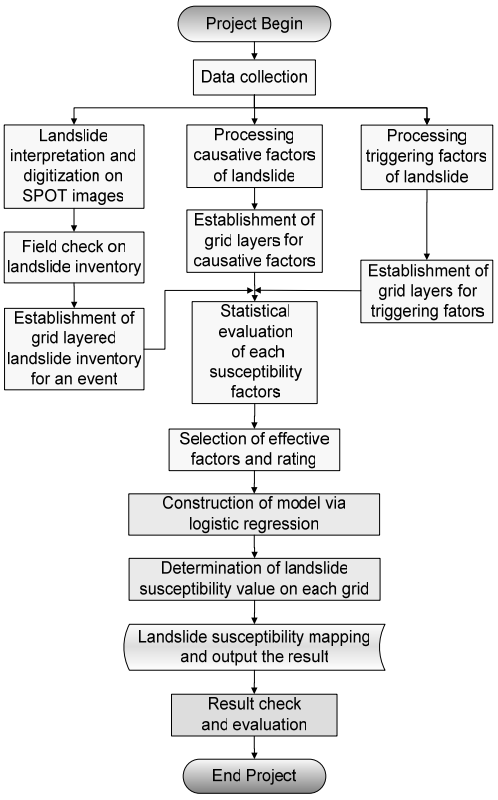

Fig. 2 Working procedure for event-based landslide susceptibility analysis in this study.

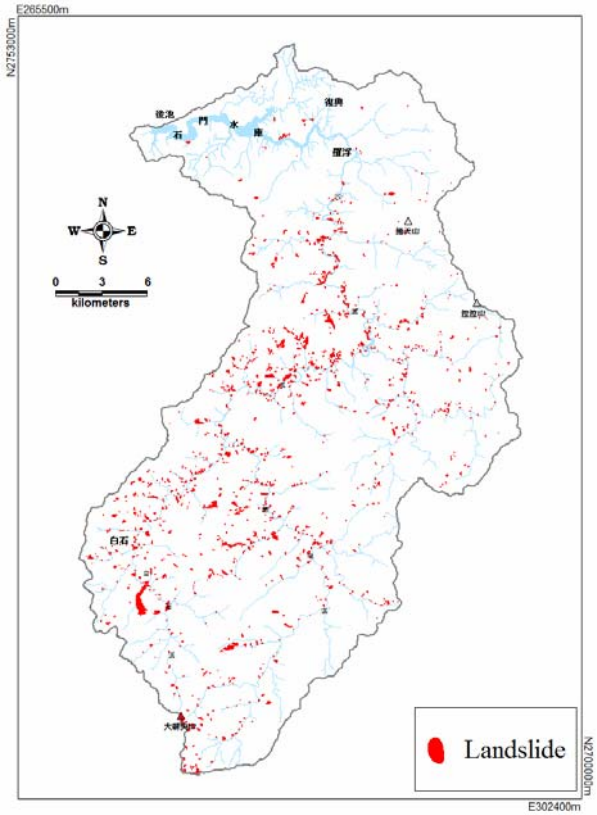

Fig. 3 Spatial distribution of landslides induced by Typhoon Aere in the study region.

Data Acquisition and Processing

The basic data utilized in this study included a 5m x 5m grid DEM, SPOT5 images, 1/5,000 photo-based contour maps, 1/50,000 geologic maps, and hourly rainfall data. The DEMs were collected from Central Geological Survey (CGS), Taiwan. They were transferred to a color-shaded image and were visually checked. Defects were replaced by re-digitizing from a 1/5,000 scale photo-based contour map. Other abnormal points were corrected using a median filter. Finally the DEMs were reduced to 10m x 10m grid-cells for use.

SPOT5 images taken before and after the typhoon event were received, processed and rectified by the Center for Space and Remote Sensing Research, National Central University, Taiwan. Both multi-spectral (XS) and panchromatic (PAN) images were used. A fusing technique was utilized to produce a higher resolution (2.5 m) false-color composite image to facilitate landslide recognition.

1/50,000 geological maps were collected from the CGS. Each map was overlaid with a shaded DEM and visually inspected in a GIS. Some abnormal boundaries, mostly associated with alluvial and terrace deposits, were corrected. The ERDAS IMAGINE system was used to transform the geologic vector map to a raster image of 10m x 10m grid-cells.

The hourly rainfall data in and around the study region were collected from the Central Weather Bureau and the Water Resources Agency, Taiwan. These data were first plotted and visually inspected to compare individual records for consistency with neighboring gauge stations; abnormal data was deleted. Rainfall data were finally interpolated into 10m x 10m grid-cells data. All later processing and analysis for each susceptibility factor and for logistic analysis are based on the 10m x 10m grid-cells unit.

Event-based Landslide Inventory

Landslides triggered by Typhoon Aere were interpreted and delineated by comparing SPOT5 images taken before and after the typhoon. Landslides found in both inventories were examined very carefully for changes in tone and/or enlargement of extent. Typhoon Aere triggered 703 landslides, of which 50 were enlarged or reactivated old landslides. Most observed slope failures were shallow landslides on soil mantled slopes with depths less than 2 m. To develop our susceptibility model, we only considered new landslides triggered by typhoons. Fig. 3 shows the spatial distribution of landslides triggered by Typhoon Aere.

A landslide area is composed of source area and deposit area. Landslide deposits were identified by comparing the GIS landslide layer with the 1/5,000 scale photo-based contour map. The slope angle or concentration of contour lines was used to differentiate deposits from sources. Only landslide source area was used in building the susceptibility model.

Selection of Factors for Modeling

There are more than fifty different landslide-related factors commonly used (both in Taiwan and worldwide) for LSA (Lin, 2003). In the present storm event-induced landslide study, we first selected sixteen of the most frequently used, based on data abundance and availability. These factors were further tested, including the normality of each factor, standardized differences (Davis, 2002) between the landslide group and non-landslide group for each factor, probability of failure curve, success rate curve (Fig. 4) and correlation coefficient between any two factors. A final selection of effective factors was decided based upon the evaluation and test results. In this study, 8 causative factors: lithology, slope gradient, NDVI, slope roughness, profile curvature, relative slope height, total slope height, topographic wetness index, distance to a fault, and a triggering factor - maximum rainfall intensity, were selected for building the susceptibility model.

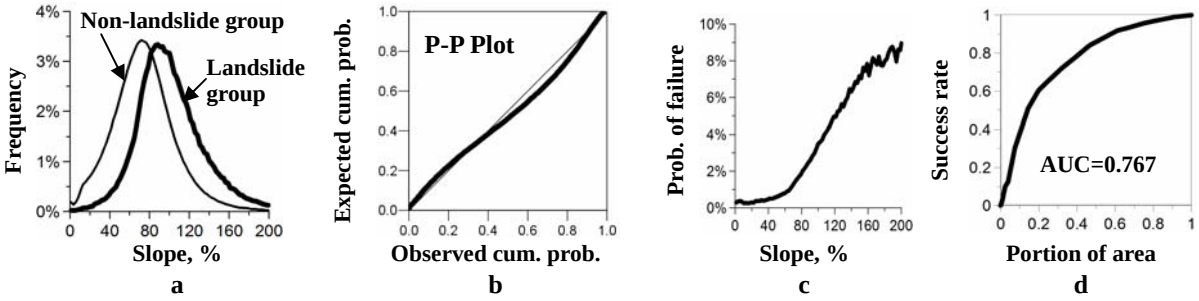

Fig. 4 Processing and selection of a factor, using slope factor as an example. (a) Frequency distribution curves, (b) P-P plot for testing of normal distribution, (c) Probability of failure curve, (d) Success rate curve.

The 10 selected factors are described as follows:

1. Lithology: Lithologic units were reduced from the 1 to 50,000 scale geological map. Each unit was used as an independent factor. These include: sandstone and shale unit, indurate sandstone and shale unit, argillite unit, and quartzite and argillite unit. Alluvium and diluvium are present at flat area and were classified into stable area, which was not included in the analysis.

2. Slope gradient: Slope is the first derivative of topographic surface. Following Wilson & Gallant (2000), we used 3×3 subgrid to calculate slope gradient from grid DEM in Erdas Imagine system.

3. NDVI: A SPOT5 image taken before Typhoon Aere was used to calculate NDVI according to,

$$\mathrm{NDVI} = \frac{\mathrm{IR} - \mathrm{R}}{\mathrm{IR} + \mathrm{R}}, \qquad (2)$$

where IR is reflectivity of near infrared band, R is reflectivity of red band. NDVI is ranging from -1 to 1; higher value indicates denser vegetation; low value indicates bare land.

4. Slope roughness: Referring to Wilson & Gallant (2000), we used 13×13 subgrid to calculate the standard deviation of slope gradient and used this standard deviation as slope roughness.

5. Profile curvature: Curvature is the second derivative of topographic surface. It may be classified as profile curvature, plane curvature, tangential curvature, and total curvature. We used 3×3 subgrid to calculate profile curvature from a 13×13 subgrid smoothed DEM in Erdas Imagine system. Using a smoothed DEM is to reduce local topographic effect.

6. Total slope height: We took a topographic profile through a point of interest. Upper turning point is taken as crest; lower turning point is taken as toe. Height difference from crest to toe is defined as "total slope height".

7. Relative slope height: Height difference from point of interest to toe is defined as "slope height". Slope height divided by total slope height is defined as "relative slope height".

8. Topographic wetness index: Topographic wetness index (Wilson & Gallant, 2000; Kirkby, 1975) ω is defined as:

$$\omega = \ln(\frac{A_S}{\tan\theta}) \quad . \tag{3}$$

Where, A_S is upslope contribution area, θ is slope angle.

9. Distance to a fault: Using buffer tool in GIS to select 0-500, 500-1000, 1000-1500, 1500-2000, 2000,2500, 2500-3000, 3000-3500, 3500-4000, and 4000-4500 meters as 9 different distance zones. A probability of failure value was calculated at each zone and this value was used as rating of the zone.

10. Maximum rainfall intensity: There are 9 rain gauge stations in the catchment area. We first processed these 9 stations and more stations around the catchment area to obtain the maximum rainfall intensity and total rainfalls for each station. Then we used these parameters and incorporate topographic heights as auxiliary variables to perform spatial interpolation at each grid-cell using kriging with varying local means. Results are shown in Fig. 5 and Fig. 6. Maximum rainfall intensity was finally selected as an effective triggering factor for building the susceptibility model.

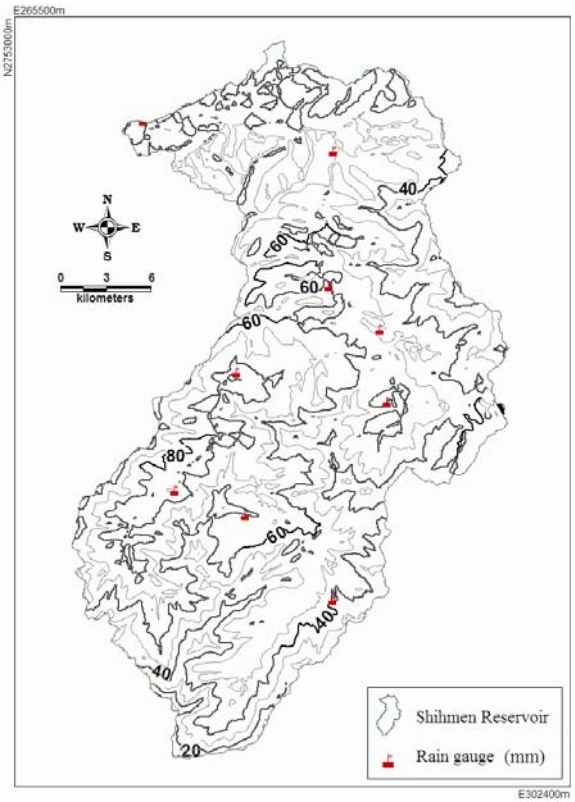

Fig. 5 Maximum hourly rainfalls in Typhoon Aere event.

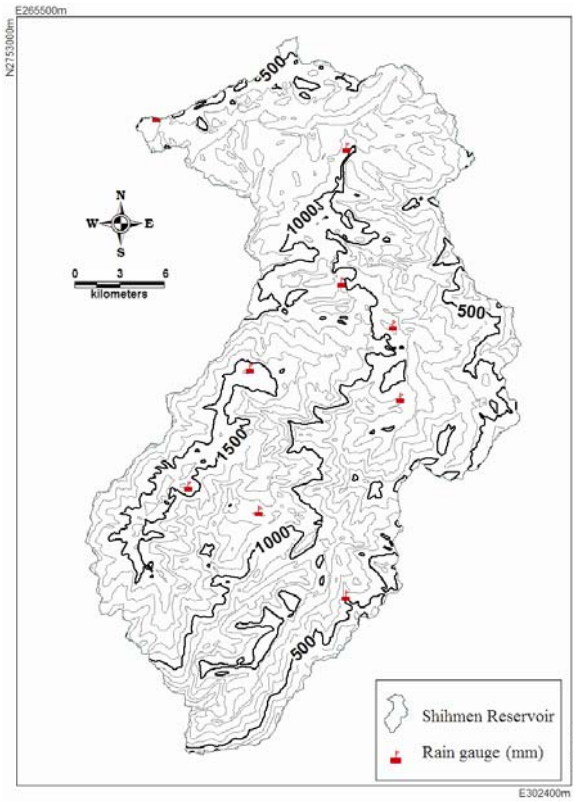

Fig. 6 Total rainfalls in Typhoon Aere event.

Results and Evaluation

Data in the Shihmen Reservoir catchment area for shallow landslides triggered by Typhoon Aere were used for the analysis. The data set from landslide group contains 38,403 cells, 20,000 cells were randomly selected for building the model and the other 18,403 cells were used for validation. A randomly selected non-landslide data set of similar size was used in the logistic regression and validation. Gentle slopes, i.e., a slope gradient of less than 10%, were regarded as stable, and were not included in the analysis. These areas were also not considered in the evaluation of the results or the calculation of the success rate or prediction rate. The results of the logistic regression included a coefficient (apparent weight) for each factor as follows,

$$\ln\left(\frac{p}{1-p}\right) = \begin{aligned} &-1.148F_1 + 0.805F_2 + 0.897F_3 + 0.631F_4 + 0.127F_5 - 0.306F_6 + 0.062F_7 \\ &+0.277F_8 + 0.126F_9 - 0.685F_{10} + 0.074F_{11} - 0.747F_{12} + 1.304F_{13} - 2.617 . \end{aligned} \quad (4)$$

Where, occurrence probability p is taken as LSI in this study; it ranges from 0 to 1. F_1 is sandstone and shale unit, F_2 is indurate sandstone and shale unit, F_3 is argillite unit, F_4 is quartzite and argillite unit, F_5 is slope gradient, F_6 is NDVI，F_7 is slope roughness, F_8 is profile curvature, F_9 is total slope height，F_{10} is relative slope height, F_{11} is topographic wetness index, F_{12} is distance to a fault, F_{13} is maximum rainfall intensity.

All the factors were standardized before analysis except lithology factor and distance to a fault factor; the former was assign 1 if present and 0 if not present, the later was normalized between 0 and 1. Among the 10 factors used, the maximum rainfall intensity factor has the highest coefficient and a large percentage of the weighting. NDVI, relative slope height, and distance to a fault factors show negative coefficients, because actual landslides are negatively correlated with these factors. Sandstone and shale lithologic unit shows negative coefficient, for its relative young geologic age and of soft rock in nature.

Equation (4) was used to calculate the LSI for each grid-cell. Then LSIs were used to calculate the probability of failure for each LSI bins and a probability of failure curve was determined (Fig. 7). The probability of failure curve does show a strong trend of increasing probability of landslide occurrence with increasing susceptibility values. This relation was fitted with a function as follow,

$$P_{ls} = 0.007587\left(\frac{\lambda}{1-\lambda}\right)^{0.6403} . \quad (5)$$

Where P_{ls} is probability of failure, or spatial probability of shallow landslide occurrence. λ is LSI. The spatial probability of landslide can then be used to map the susceptibility classes, as shown in Fig. 8.

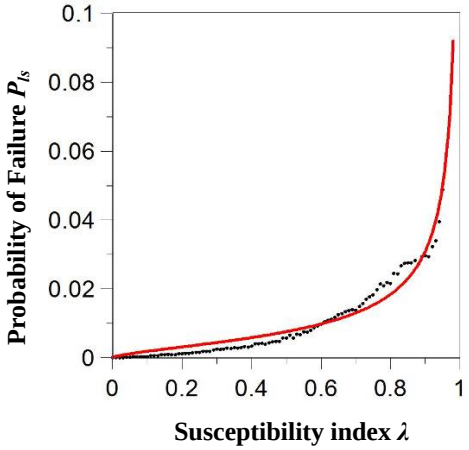

Fig. 7 Probability of failure with respect to landslide susceptibility index.

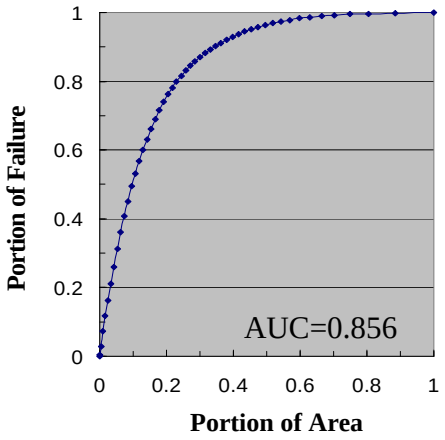

Fig. 9 Prediction rate curve for the model.

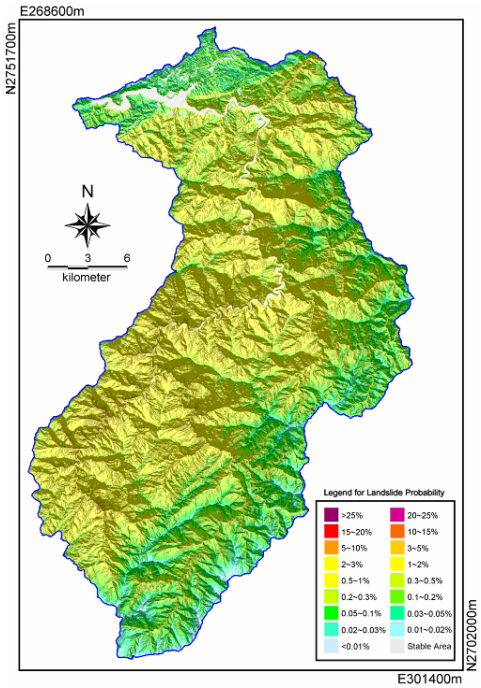

Fig. 8 Spatial landslide probability map of Typhoon Aere event.

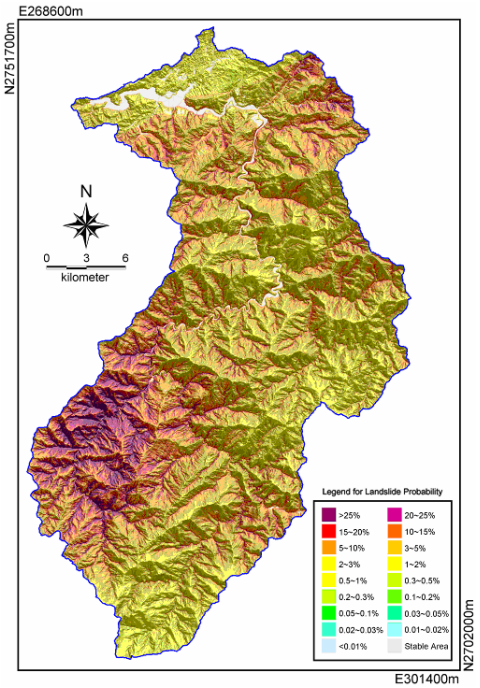

Fig. 10 Spatial landslide probability map of 100-year return-period rainfall intensity.

In validation, the accurate rate is 83.1% for landslide group and 75.1% for non-landslide group using error matrix method (Stehman, 1997). The total accurate rate is 79.1%. When using prediction rate curve method (Chung & Fabbri, 2003) for whole data set, the model predicts 50% landslides with in 10% of high susceptibility areas; the area under curve (AUC) is 0.856.

In hazard mapping, we used rainfall intensity of 100-year return period for the rainfall factor and calculate LSIs using equation (4). These LSIs were then transferred to probability of failures according to equation (5). The spatial probability of landslide under of 100-year return period rainfall can then be mapped, as shown in Fig. 10.

Model Applications

The landslide hazard model could be used for the prediction of future landslides providing a scenario rainfall distribution is given. The 100-year return period landslide hazard map could be used for general purposes, like for the decision making of slope remedial measure, regional planning and hazard mitigation policy. The landslide hazard model could also be used for sediment prediction of a catchment area.

From probability of landslide occurrence map, we map estimate shallow landslide area as A_{sl} follow,

$$A_{sl} = \sum_i a P_{sl_i} \,. \tag{8}$$

Where, P_{sl_i} is probability of failure at grid-cell i, a is area of a cell; it is 100 m^2 in this study. From spatial landslide probability map of Typhoon Aere event (Fig. 8), we can estimate landslide area to be 28,928,800m^2 according to equation (8). Actual landslide from inventory is 2,872 landslides with a total area of 21,053,000m^2. The difference indicates that mapping of the landslide inventory may have missed some landslides in shallow of a SPOT image, deleting of repeated landslides in the event-based landslide inventory may also reduce the amount of landslide as compared to the prediction one.

We further adopted a soil-thickness prediction model from Chung (2008) as follow,

$$h = 4.9429 - 1.0939 \ln S \,. \tag{9}$$

Where, h is soil thickness (m)，S is slope (°). Inserting slope data into equation (9), we can get soil thickness at each

grid-cell in the study area, or we can estimate landslide volume V_{sl} as follow,

$$V_{sl} = \sum_i (494.29 P_{sl_i} - 109.39 P_{sl_i} \ln S_i) \quad . \tag{10}$$

Using equation (10), we can estimate landslide volume induced by the Typhoon Aere to be 25,510,700m^3. Actual landslide volume using landslide area from inventory is 19,164,240m^3.

To this step, we can successfully predict landslide location, area and volume in a drainage basin or catchment area using GIS. However, the amount of sediment year in a stream still requires estimation of soil erosion on the slope and sediment transportation in the river. A distributed hydrological and sediment transport modeling should be further carried out so that sediment problem like that in the Shihmen Reservoir catchment area could be interpreted.

References

Atkinson, P. M., & Massari, R. (1998). Generalized linear modelling of susceptibility to landsliding in the central Apennines, Italy. *Computers & Geosciences*, 24, 373-385.

Ayalew, L., & Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, central Japan. *Geomorphology*, 65, 15-31.

Carrara, A. (1983) Multivariate models for landslide hazard evaluation. *Mathmatical Geology*, 15(3), 403-427.

Carrara, A., Guzzetti, F., Cardinali, M., & Reichenbach, P. (1999). Use of GIS technology in the prediction and monitoring of landslide hazard. *Natural Hazards*, 20, 117-135.

Carrara, A., & Pike, R. J. (2008). GIS technology and models for assessing landslide hazard and risk. *Geomorphology*, 94, 257-260.

Central Geological Survey (CGS) (2007). General report on geological investigation and database construction for the upstream watershed in flood-prone areas (first phase), geological investigation, landslide-debris flow investigation and their susceptibility evaluation on watershed (2007). CGS Report 095-08-12-2-001-01-0.

Chung, C. F., & Fabbri, A. G. (2003). Validation of spatial prediction models for landslide hazard mapping. *Natural Hazards*, 30, 451-472.

Chung, H. H. (2008). Terrain stability analysis using hydrologic model for predicting shallow landslides – a study on Piya Creek watershed. M.S. Thesis of Institute of Applied Geology, National Central University. (in Chinese with English abstract)

Dai, F. C., Lee, C. F., Li, J., & Xu, Z. W. (2001). Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environmental Geology*, 40, 381-391.

Davis, J. C. (2002). *Statistics and data analysis in geology*, 3rd ed. John Wiley and Sons.

Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G., & Vandekerckhove, L. (2006). Prediction of landslide susceptibility using rare events logistic regression, a case-study in the Flemish Ardennes. *Geomorphology*, 76, 392-410.

Greco, R., Sorriso-Valvo, M., & Catalano, E. (2007). Logistic regression analysis in the evaluation of mass movements susceptibility: the Aspromonte case study, Calabria, Italy. *Engineering Geology*, 89, 47-66.

Jibson, R. W., Harp, E. L., & Michael, J. A. (2000). A method for producing digital probabilistic seismic landslide hazard maps. *Engineering Geology*, 58, 271-289.

Kirkby, M. J. (1975). Hydrograph modelling strategies. In: Peel, R., Chisholm, M., & Haggett, P. (Eds.): *Process in physical and human geography*. Heinemann, London.

Lee, C. T., Huang, C. C., Lee, J. F., Pan, K. L., Lin, M. L., & Dong, J. J. (2008a). Statistical Approach to Earthquake-Induced Landslide Susceptibility. *Engineering Geology*, 100(1-2), 43-58.

Lee, C. T., Huang, C. C., Lee, J. F., Pan, K. L., Lin, M. L., & Dong, J. J. (2008b). Statistical approach to storm event-induced landslide susceptibility. *Natural Hazard and Earth System Sciences*, 8, 941-960.

Lin, Y. H. (2003). Application of neural networks to landslide susceptibility analysis. M.S. Thesis of Institute of Applied Geology, National Central University. (in Chinese with English abstract)

Stehman, S. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62, 77-89.

Wilson, J.P., & Gallant, J.C. (2000). *Terrain analysis, principles and applications*. John Wiley & Sons.

Acknowledgments

This research was supported by the Central Geological Survey (CGS), Ministry of Economic Affairs, Taiwan, under the Landslide Research Project.