

Incorporating the effects of topographic amplification in the analysis of earthquake-induced landslide hazards using logistic regression

S. T. Lee, T. T. Yu, W. F. Peng, and C. L. Wang

Department of Resources Engineering, National Cheng Kung University, 701 Tainan, Taiwan

Received: 2 November 2009 – Revised: 23 October 2010 – Accepted: 28 October 2010 – Published: 3 December 2010

Abstract. Seismic-induced landslide hazards are studied using seismic shaking intensity based on the topographic amplification effect. The estimation of the topographic effect includes the theoretical topographic amplification factors and the corresponding amplified ground motion. Digital elevation models (DEM) with a 5-m grid space are used. The logistic regression model and the geographic information system (GIS) are used to perform the seismic landslide hazard analysis. The 99 Peaks area, located 3 km away from the ruptured fault of the Chi-Chi earthquake, is used to test the proposed hypothesis. An inventory map of earthquake-triggered landslides is used to produce a dependent variable that takes a value of 0 (no landslides) or 1 (landslides). A set of independent parameters, including lithology, elevation, slope gradient, slope aspect, terrain roughness, land use, and Arias intensity (I_a) with the topographic effect. Subsequently, logistic regression is used to find the best fitting function to describe the relationship between the occurrence and absence of landslides within an individual grid cell. The results of seismic landslide hazard analysis that includes the topographic effect (AUROC = 0.890) are better than those of the analysis without it (AUROC = 0.874).

1 Introduction

Landslides triggered by earthquakes are one of the most destructive natural disasters. Seismic landslide hazard analysis is used to estimate the probability of slope failure due to an earthquake. It is a useful tool for land-use planners for preventing disasters. The highly susceptible slopes in conventional hazard maps represent areas of high potential failure triggered by earthquakes.

Many modeling approaches for landslide hazard prediction can be used to produce statistics-based susceptibility maps. Logistic regression and discriminant analysis are the most frequently used models (Brenning, 2005). Logistic regression and statistical models have been developed using the geographic information system (GIS) for landslide susceptibility mapping (Atkinson and Massari, 1998; Rowbotham and Dudycha, 1998; Lee and Min, 2001; Ayalew and Yamagishi, 2005; Mathew et al., 2007; Chang et al., 2008; Garcia-Rodriguez et al., 2008; Yilmaz, 2009). Superposing several layers of landslide-related factors using GIS can be used to create a composite model based on various weighting parameters. In order to predict the potential of landslides, landslide-related factors must first be selected. The main trigger factor of earthquake-induced landslides is the intensity of seismic shaking. Some studies have not included this factor, stating that seismicity was relatively uniform throughout the study area (Ayalew and Yamagishi, 2005; Garcia-Rodriguez et al., 2008). Other regression analyses have used trigger factors such as peak ground acceleration (PGA), peak ground velocity (PGV), and Arias intensity (I_a) with the interpolation method (Jibson et al., 2000; Luzi and Pergalani, 2000) or empirical equations (Carro et al., 2003; Pelaez et al., 2005; Lee and Evangelista, 2006; Wang et al., 2008). Most conventional analyses neglect the topographic amplification effect.

Many studies have found sufficient evidence that shows that topographic irregularity considerably affects the amplitude and frequency content of ground motion (Geli et al., 1988; Bard and Riepl-Thomas, 2000; Assimaki et al., 2005; Nguyen and Gatmiri, 2007). The occurrence of landslides at or near ridge crests during moderate to large earthquakes has also been attributed to the topography affect (Harp et al., 1981; Harp and Keefer, 1990; Densmore and Hovius, 2000; Sepulveda et al., 2005a, b). The topographic amplification effect refers to the effect of surface topography on the magnitude of ground acceleration. In general, seismic



Correspondence to: T. T. Yu
(yutt@mail.ncku.edu.tw)

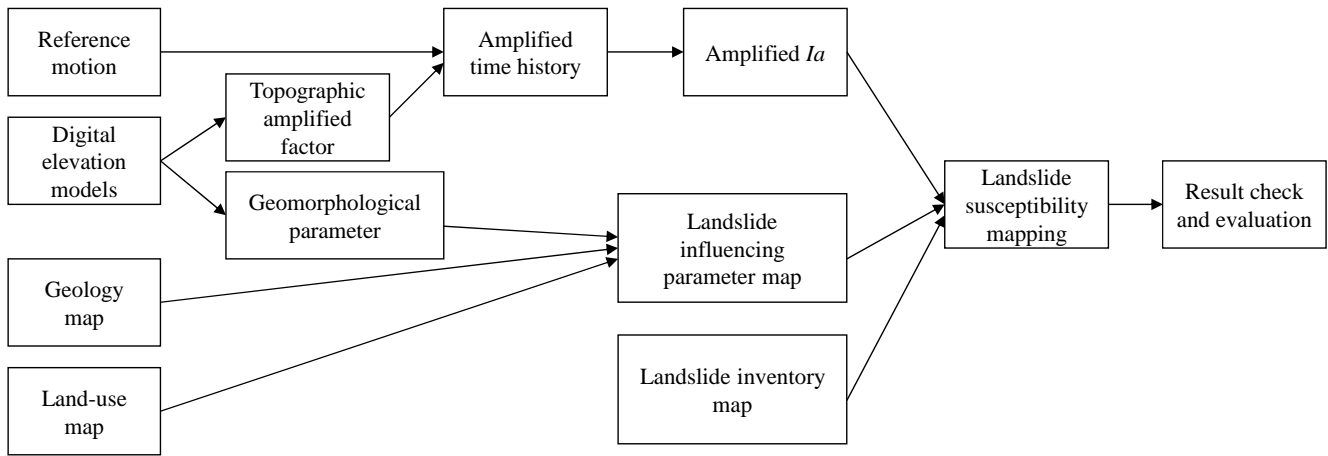


Fig. 1. Flowchart of the proposed method.

intensities are amplified at ridge crests and are de-amplified at canyons or hill toes. Consequently, conventional analysis that does not consider the topographic amplification effect underestimates the stability in the crest area. Although the natural topography is too complex for an accurate model of topographic amplification to be developed for it, Geli et al. (1988) indicated that it is possible to qualitatively predict the magnitude of ground amplification due to the topographic effect. Paolucci (2002) introduced a transfer function to obtain the topographic amplification factors.

In the present study, a landslide triggered by 21 September 1999, Chi-Chi earthquake is used to build a seismic landslide hazard model. Based on the amplified ground motions, the I_a values are extracted for each grid cell. With geomorphological, geological, land-cover, and seismic parameters, a logistic regression model is used to assess the hazard of earthquake-induced landslides in the 99 Peaks region. A flowchart of the proposed procedure is shown in Fig. 1.

2 Study area

The study area is located in the Juo-Juo-Fong (99 Peaks) region, central Taiwan, which lies within 24.000°–24.115° latitude and 120.761°–120.813° longitude. The epicentre of 21 September 1999, Chi-Chi earthquake ($M_w = 7.6$) was located about 17 km south of the study area. Surface fault-rupture occurred along the Chelungpu fault about 3 km west of the study area.

Geology is an important factor in landslide potential. The geological map of the 99 Peaks region is shown in Fig. 2. Chen and Wu (2006) surveyed the area and estimated that Pliocene and Pleistocene rocks dominate the 99 Peaks region. The Toukoshan formation in the 99 Peaks region can be divided into two areas based on lithology. The eastern part is Houyenshan member (Tkh), dominated by conglomerate

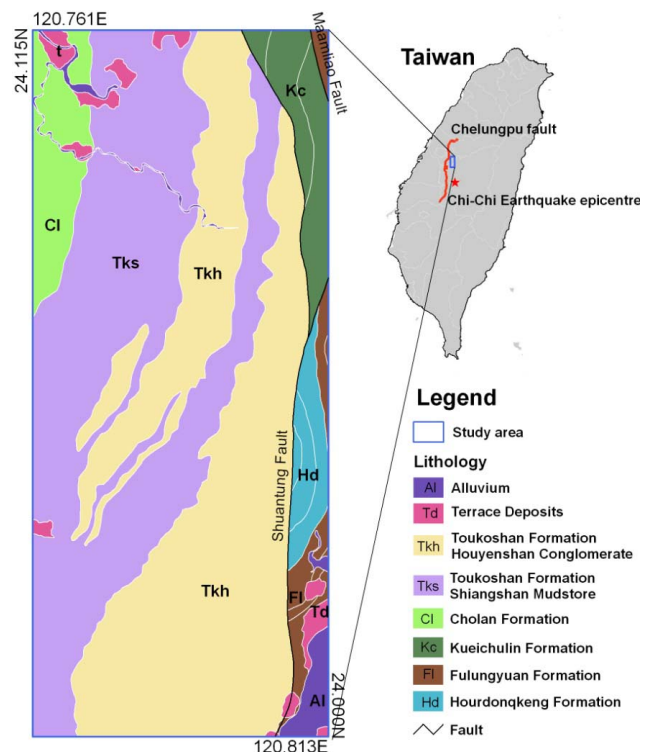


Fig. 2. Location and geological map of the study area.

with thin to thick sandstone. The conglomerate layer in the 99 Peaks region has extremely high permeability (Chen and Wu, 2006). The western part is Shiangshan member (Tks), dominated by alternating layers of sandstone-intercalated sandstone and shale and thin-bedded conglomerate. Other formations are found at the base (Lo et al., 1999):

1. *Alluvium* (AI): composed of unconsolidated mud, sand, and gravel;

2. *Terrace Deposits* (Td): composed of gravel;
3. *Cholan Formation* (Cl): composed of sandstone, siltstone, mudstone, and shale in a repeating alternating sequence;
4. *Kueichulin Formation* (Kc): mainly composed of gray to yellowish-brown massive sandstone which is partly whitish-gray;
5. *Fulungyuan Formation* (Fl): composed of sandstone and shale;
6. *Hourdonqkeng Formation* (Hd): composed of siltstone and shale.

The Chi-Chi earthquake struck the 99 Peaks region and triggered a large number of landslides. The digital inventory of landslides was estimated by matching the ground surface variation from SPOT images taken before and after the events by the Central Geological Survey in Taiwan (CGST¹). For each landslide, the location, elevation, aspect, slope angle, and area were recorded. The inventory documents 1136 landslides of various types that were triggered in the study area. Most of the landslides were shallow slope failures (Chen and Wu, 2006). Figure 3a shows the spatial distribution of landslides induced by the Chi-Chi earthquake. A total collapsed area of 7.22 km² was caused by the Chi-Chi earthquake. Landslide size ranged from less than 1.95 × 10² m² to more than 1.86 × 10⁵ m². However, not all landslide inventories were necessarily correctly mapped. To ensure that landslide data were suitable for analysis, they were screened in accordance to the following rules.

1. **Landslide area:** Fig. 4 shows the landslide frequency/area graph (Stark and Hovius, 2001; Malamud et al., 2004; Catani et al., 2005) produced by plotting the slope of the cumulative number of landslides (y-axis) versus their area (x-axis). This distribution shows the scale behavior of natural event records (earthquakes, landslides, etc.) characterized by a linear relationship for large events and a rollover for smaller events with decreasing frequency for very small landslide areas. The rollover can partly be explained by the undercounting of small events (less accurate mapping) which can be taken into consideration by using particular distribution fits proposed by Stark and Hovius (2001). For our dataset, the proposed scaling law holds for areas greater than 10³ m². It can thus be concluded that there is a good degree of confidence in the accuracy of the inventory for landslides with an area greater than 10³ m². This is, landslides with areas smaller 10³ m² were not analyzed.

¹<http://www.moeacgs.gov.tw>, 2008

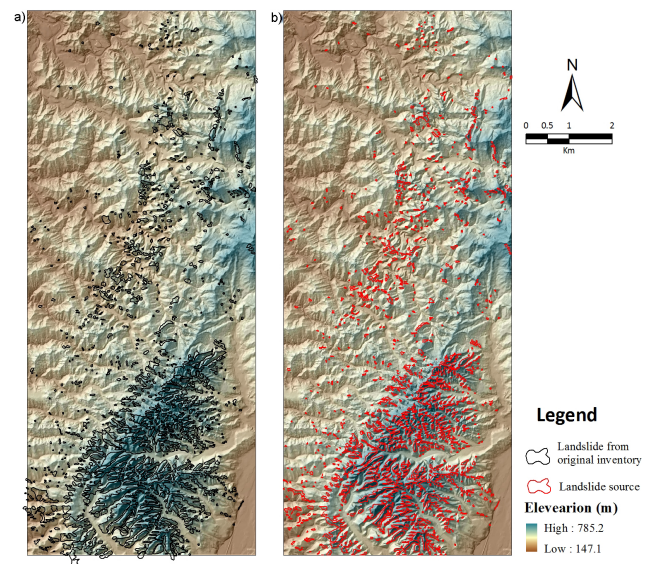


Fig. 3. Terrain map showing the distribution of landslides triggered by the Chi-Chi earthquake in 1999: (a) original inventory, (b) landslide source.

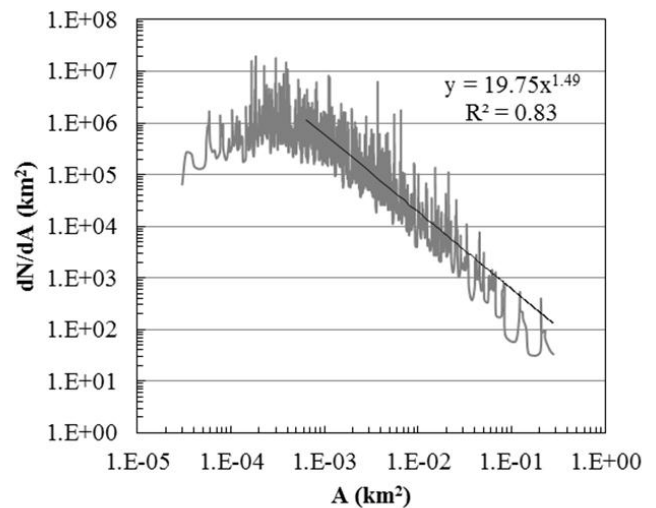


Fig. 4. Landslide area frequency plot. dn/dA is the derivative of landslide frequency with respect to the total area and A is the landslide area.

2. **Slope gradient:** the alluvium and cells of slopes of less than 10° are regarded as stable during shaking. We removed landslides with slope gradients of less than 10°.

Landslide source areas were defined to include grid cells with elevations above the median elevation for each landslide. That is, the upper half of each landslide was considered as the source area, following Jibson et al. (2000).

After the screening process, the number of landslides was reduced to 1077; their distribution is shown in Fig. 3b.

3 Amplified ground motion

The main trigger factor of earthquake-induced landslides is the seismic shaking intensity. Arias intensity (I_a) is an important measurement of ground motion since it simultaneously reflects multiple characteristics of the motion, whereas most common scalar ground-motion measures, such as peak ground acceleration (PGA), peak ground velocity (PGV), and individual ordinates of spectral acceleration, reflect a very specific aspect. I_a can capture and represent multiple attributes of the overall ground motion based on the record. The Arias intensity (Arias, 1970) is defined as:

$$I_a = \frac{\pi}{2g} \int_0^d a^2(t) dt \quad (1)$$

where $a(t)$ is the recorded ground acceleration, g is the acceleration due to gravity, d is the total duration of earthquake motion, and t is time. The Arias intensity was first used for analyzing the occurrence of landslides by Wilson and Keefer (1985); they found it to be the most efficient intensity measure of earthquake-induced landslide potential.

The free-field strong-motion stations around the 99 Peaks region are TCU052, TCU065, TCU067, TCU071, TCU072, TCU075, and TCU137. Station TCU052 is located at site C (very dense soil or soft rock) and all other seismic stations are located at site D (stiff soil). For the 99 Peaks region, a Kriging algorithm was used to interpolate the shaking intensities recorded at these stations to generate the contour of shaking intensities. The distribution of I_a (Fig. 5) in the study area was evaluated. It should be noted that the distribution of shaking intensities was derived from the seismic stations outside the study area. The distribution of shaking intensities cannot be used to determine the response of the topographic effect inside the study area.

This study considers the topographic effect in relation to seismic intensity. A three-dimensional (3-D) staggered-grid finite differences model (Ohminato and Chouet, 1997) is constructed to estimate the effect of topography on ground motion amplification. The scheme introduces stress-free boundary conditions into the finite difference method by including 3-D topographic characteristics. The displacements and stresses in the three directions (NS, EW, up, down) in each cell unit (within the medium or below the free surface) can be calculated during the wave propagation. The model is suitable for simulating seismic wave propagation. Details of the formulation and implementation can be found in Ohminato and Chouet (1997).

After the staggered-grid finite difference model is constructed with the real digital elevation model, the topographic amplification factors $H_f^k(f)$ are calculated by subjecting the model to seismic motion with incident plane shear waves, with two cases of polarization, in the EW and NS directions,

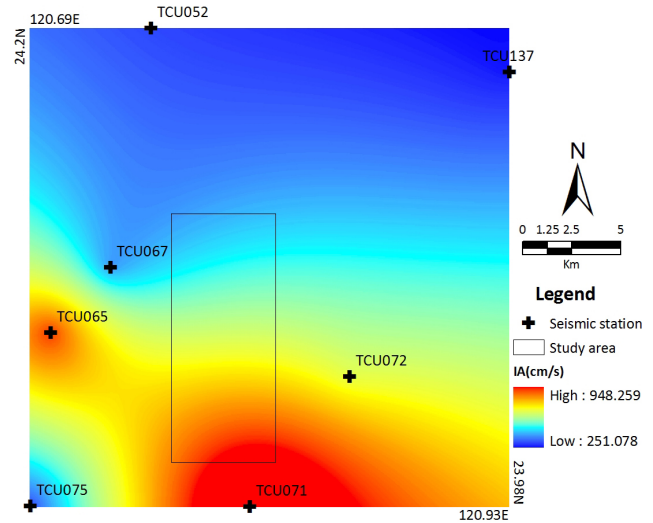


Fig. 5. Arias intensity generated by interpolating the shaking intensity recorded at a seismic station.

respectively. The time dependence of the incident wave is described in terms of a Ricker wavelet:

$$f(t) = (a - 0.5) \exp(-a) \quad (2)$$

$$a = |\pi(t - t_0)/f_p|^2 \quad (3)$$

where $f(t)$ is the Ricker wavelet, f_p is the peak frequency of the Fourier spectrum, and t_0 is an appropriate time shift parameter.

The topographic amplification factors $H_{NS}^k(f)$ at site k in the NS direction are calculated using the following steps (Paolucci, 2002): (1) input the values of previous incident waves with the polarization in the NS direction; (2) calculate the responses at site k along the NS direction based on the digital elevation model and transform them into the frequency domain; (3) calculate the responses of the flat topography model at the free surface and transform them into the frequency domain $U_{mod\ el}(f)$; (4) calculate the topographic amplification factors for each frequency using:

$$H_{NS}^k(f) = Y_{NS_mod\ el}^k(f) / U_{mod\ el}(f) \quad (4)$$

Similarly, $H_{EW}^k(f)$ is calculated with polarization of the input incident wave in the EW direction.

The frequency response at the surface of irregular topography can be expressed as (Paolucci, 2002):

$$Y_i^k(f) = H_i^k(f) \cdot U_i(f) \quad (5)$$

where $Y_i^k(f)$ is the i -th component of the seismic response at site k , and $U_i^k(f)$ is the corresponding reference motion at the reference site (located at the flat surface) in the i -th direction. In this paper, the ground motion record at a nearby seismic station is regarded as the reference motion. The transfer function, $H_i^k(f)$, for the i -th direction at site k is regarded as the topographic amplification factor in this study.

Table 1. Data sources and data types used in this work.

Data	Data source	Data type	Resolution/Scale
DEM	Satellite Geoinformatics Research Centre, NCKU, Taiwan	Grid	5 m × 5 m
Geologic map	Central Geological Survey, Taiwan	Polygon coverage	1/25 000
Land cover	Soil and Water Conservation Bureau, Taiwan	Polygon coverage	1/5000

The steps of the procedure are summarized as follows:

1. Compute the topographic amplification factor $H_i^k(f)$:
 - i. The seismic wave propagation model is constructed with the digital elevation model (DEM) of the study area.
 - ii. The response motions $Y_{i_mod\ el}^k(f)$ of the mountain topography model are calculated by subjecting the model to an incident plane wave.
 - iii. The model reference motion $U_{mod\ el}(f)$ of the flat surface model is derived by subjecting the model to the same incident wave.
 - iv. The topographic amplification factors $H_i^k(f)$ are calculated by combining the response motions of the mountain model and model reference motion $U_{mod\ el}(f)$ of the flat model with Eq. (6). The amplification factors at each cell are in the NS and EW directions.
2. Calculate the topographic amplified motion $a^k(t)$, I_a :
 - i. The reference motion $U_i(f)$ is defined as the ground motion without considering topographic effects. The ground motion record at the seismic station near the study area is chosen as the reference motion $U_i(f)$.
 - ii. The amplified motions $Y_i^k(f)$ are obtained by combining the topographic amplification factors $H_i^k(f)$ and reference motion $U_i(f)$ with Eq. (4). The amplified motions are then transformed into the time domain to generate the amplified time history $a^k(t)$ for each cell in the NS and EW directions.
 - iii. Based on the time history $a_i^k(t)$ for each cell, an I_a map is generated.

4 Landslide-influencing parameters

The factors that affect landslides are very complex. Various factors are used for landslide hazard assessment. The present study investigates the effects of topographic amplification in the analysis of earthquake-induced landslide hazards. Hence, we selected seven of the most frequently used factors based

on the general consensus that each variable must be independent, operational, complete, non-uniform, measurable, and non-redundant. Lithologic, geomorphological (elevation, slope gradient, slope aspect, terrain roughness), land-cover, and triggering parameters (I_a) were selected. The datasets used in the landslide hazard assessment include: (1) a geological map, (2) high-resolution digital elevation models of the topography, and (3) a land-cover map. Table 1 shows the sources and resolutions of these variables. All factors were constructed using the ArcGIS software package.

A 5-m resolution DEM was produced by the Satellite Geoinformatics Research Centre in the National Cheng Kung University (NCKU) based on aerial photogrammetry. All the geomorphological variables were derived from 5-m resolution DEM data instead of being cited from the standard 40-m DEM to preserve the integrity of the database. The study area was divided into a grid of 5 m × 5 m cells. The area contained 2534 rows and 1064 columns, for a total of 2 696 176 grid cells. The Chi-Chi earthquake induced landslides within 210 832 of the cells.

4.1 Lithologic parameters

Information on lithology classification (Fig. 6a) was derived from the 1:25 000 geological database of the Central Geological Survey of Taiwan (CGST¹). The following lithologies were used in the regression model: alluvium, terrace deposits, Cl, Kc, Fl, Hd, Tkh, and Tks. The first class (alluvium) was used as reference category. We calculated the landslide density for each class of variable (Fig. 7). Figure 7a shows the landslide densities computed for the eight classes of lithologic parameter. The class “Tkh” had the highest landslide density (11.3%).

4.2 Geomorphological parameters

Many previous landslide studies used logistic regression, elevation, and slope angle as the best predictor variables for estimating the probability of landslide occurrence (Garcia-Rodriguez et al., 2008; Ohlmacher and Davis, 2003; Ayalew and Yamagishi, 2005). Each 5-m space cell in the DEM had an interpolated elevation value (Fig. 6b). The landslide density tended to increase with elevation until the maximum density (26.74%) was reached in the range of 573.4–785.2 m (Fig. 7b).

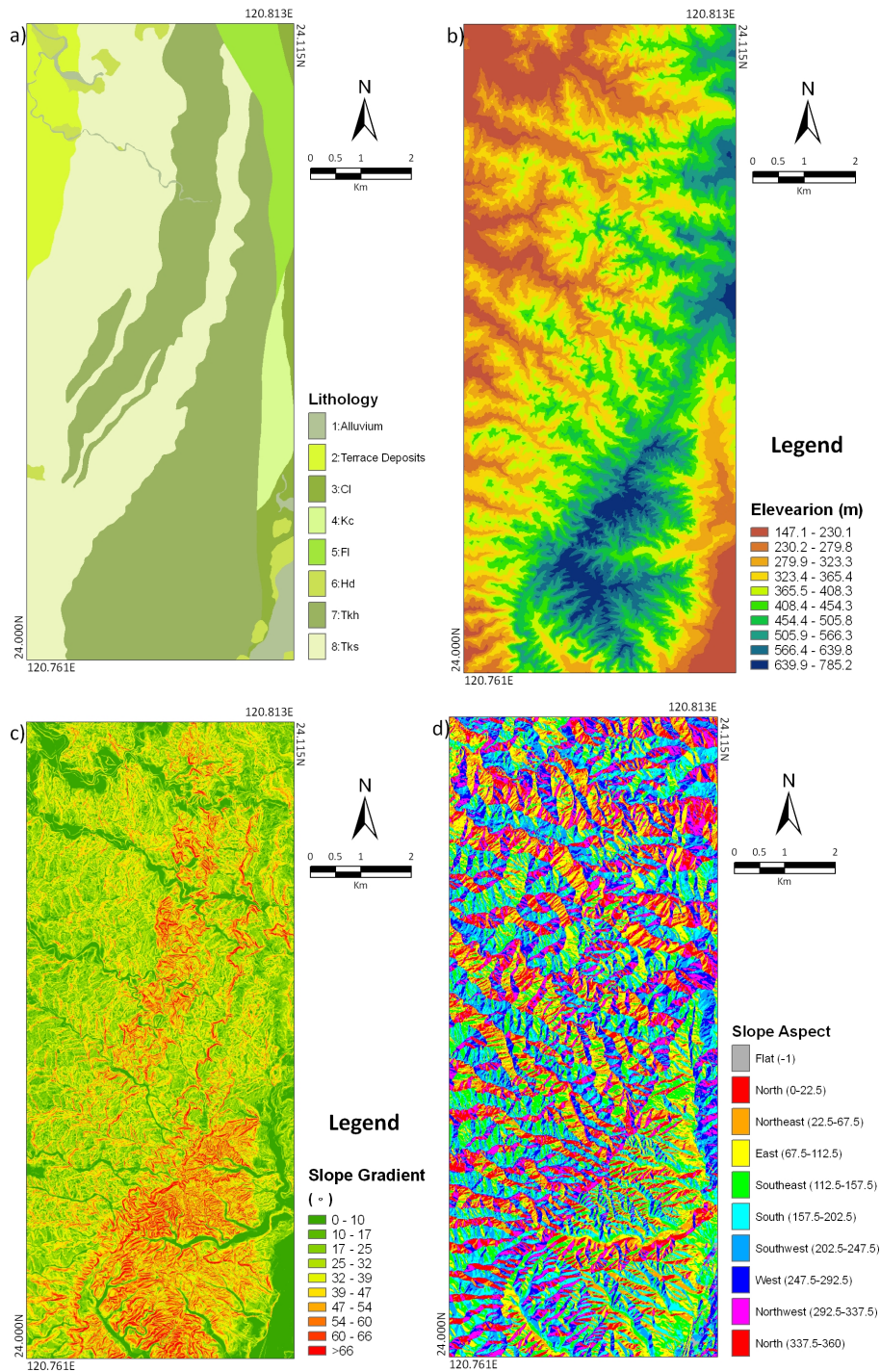


Fig. 6. Spatial distribution of factors: (a) lithologic unit, (b) elevation, (c) slope gradient map, (d) slope aspect map, (e) terrain roughness, (f) land-cover map.

Some geomorphological factors, such as slope gradient, slope aspect, and terrain roughness, were derived from the DEM. Conceptually, the slope gradient was calculated using 3×3 cell neighborhoods around the processing cell based on the average maximum technique (Burrough and McDonell, 1998). Landslides usually occur on steep slopes.

The cells of slopes less than 10° were not analyzed in the study area. The obtained slope gradient map is shown in Fig. 6c. Landslide density tended to increase with slope until the maximum density (21.9%) was reached in the range $63.94\text{--}83.6^\circ$ (Fig. 7c).

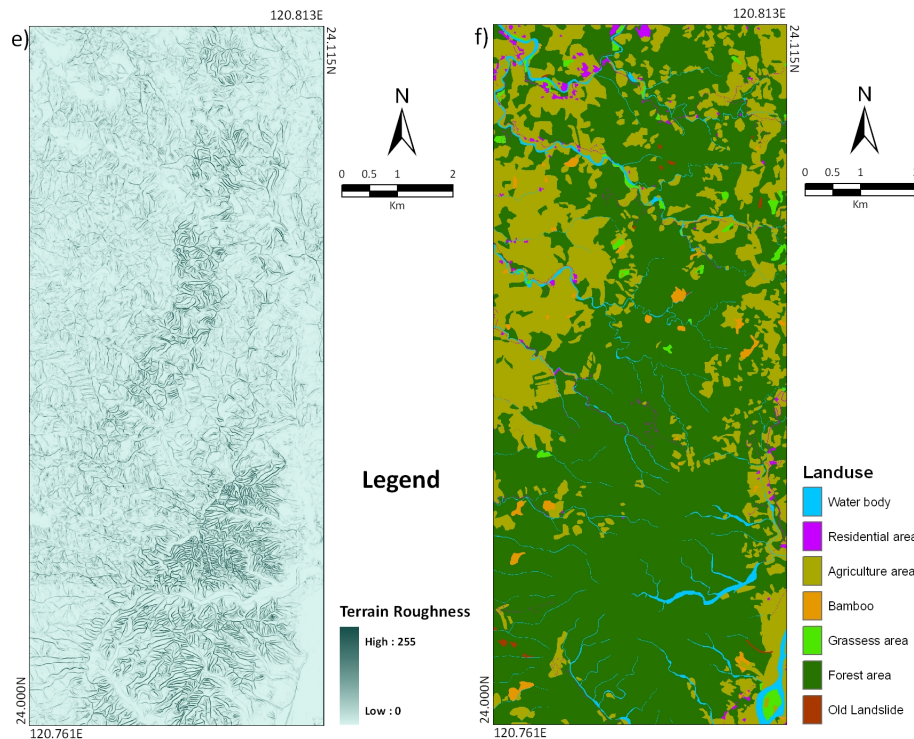


Fig. 6. Continued.

The slope aspect can influence soil strength and susceptibility to landslides because it affects moisture retention and vegetation cover. For earthquake-induced landslides, slope aspect can also affect the distribution of landslide density (Tibaldi et al., 1995; Chang et al., 2007; Garcia-Rodriguez et al., 2008; Lee and Evangelista, 2006; Lee et al., 2008). We used eight (North, Northeast, East, Southeast, South, Southwest, West, and Northwest) aspects (Fig. 6d) for logit modeling, with North as the reference aspect. Figure 7d shows that the majority of landslides occurred on South (9.7%) and Southeast (9.1%) facing slopes.

Terrain roughness measurements from the DEM provide a better description if the undulation of the topographic surface is considered. The spatial variability function of the DEM measures the dispersion of the vector perpendicular to the surface, following Garcia-Rodriguez et al. (2008). The obtained roughness values were standardized and a roughness map with 0–255 levels was created (Fig. 6e). The landslides density correlates with terrain roughness (Fig. 7e).

4.3 Land-cover parameters

Information of land cover was compiled by the Soil and Water Conservation Bureau of Taiwan (SWCB²), and directly used as one of the GIS layers (SWCB, 1999²) without any additional processing.

²<http://www.swcb.gov.tw/>, 1999

The land-cover map (Fig. 6f) was classified into eight classes, namely water bodies, residential, agriculture areas, bamboo areas, grass, forests, and old landslides. These classes were used for the logit model, with water bodies used as the reference category. The correlation of landslide density with land cover shows that the highest density occurs for old landslides (Fig. 7f).

4.4 Trigger parameters

The trigger factor is associated with seismic intensity. The parameter I_a is an ideal index for representing seismic intensity and is thus used for logistic regression. The recorded ground accelerations of the Chi-Chi earthquake at station TCU052 are shown in Fig. 8. Station TCU052 is on the Pliocene stratum at site C (very dense soil or soft rock). The elevation of the station is 215 m. The station is located on a flat surface. The variation of the topography of nearby hills is very slow. The seismic station might have been affected by a surface wave generated at a nearby hill, but the effect would have been small and could be ignored. Hence, the record from the station can be regarded as the reference motion (Eq. 5). For the Chi-Chi earthquake, the recorded PGA and I_a values at station TCU052 were 439 gal and 293 cm s⁻¹, respectively. The location of station TCU052 is about 2 km away from the Chelungpu fault, and the epicentral distance is about 39 km. The duration of the seismic record is about 25 s. The Chi-Chi earthquake produced a rupture trace more than 100 km long following

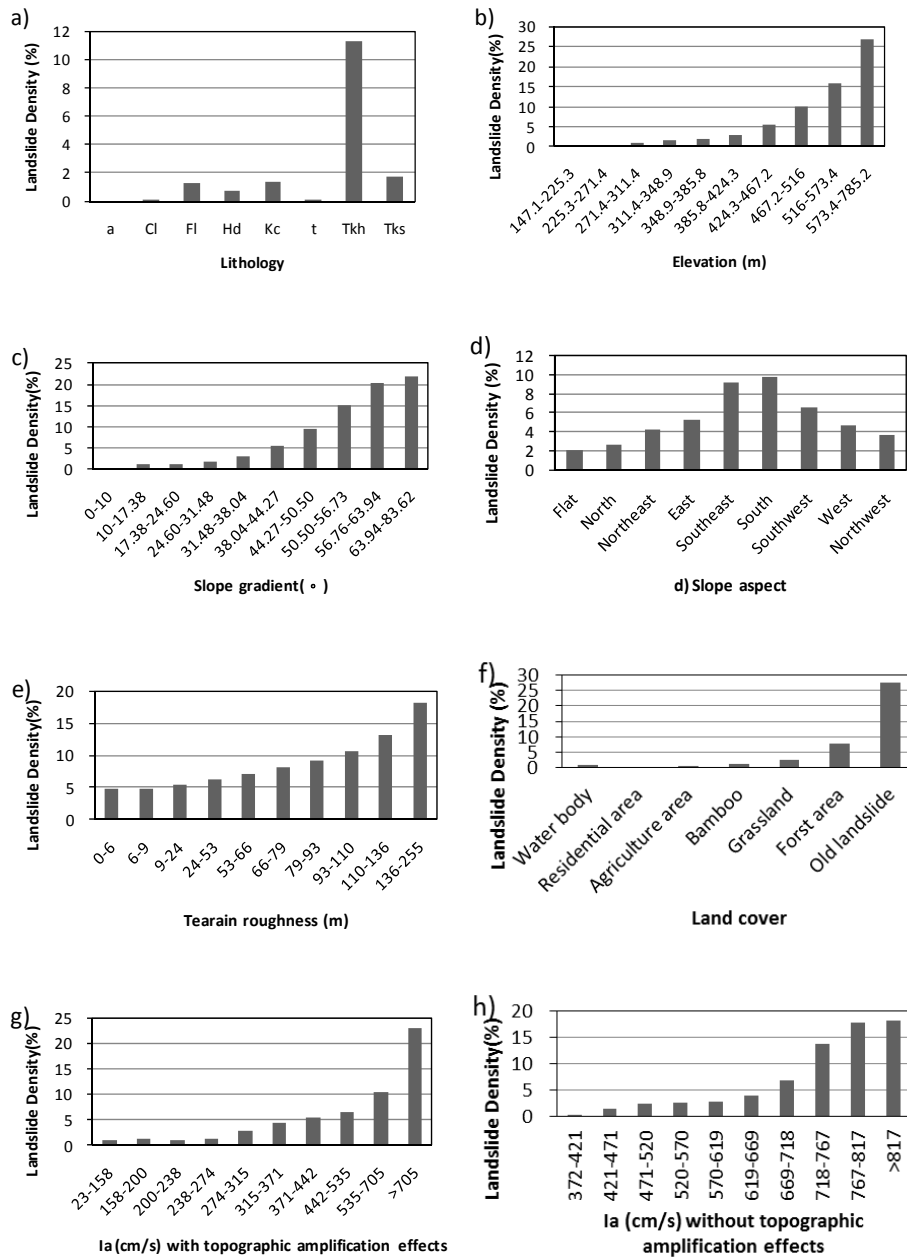


Fig. 7. Landslide density for each class of a variable: (a) lithology, (b) elevation, (c) slope gradient, (d) slope aspect, (e) terrain roughness, (f) land cover, (g) I_a with amplification effects, and (h) I_a without amplification effects.

the Chelungpu fault. The 99 Peaks region is just above the Chelungpu fault. The section of the Chelungpu fault under the 99 Peaks region has S strike and dips of N 6° W and 28° E, respectively. These values of the strike and dip were used to set up the incidence angle of incident shear waves in the wave propagation model.

The amplified time history $a_i^k(t)$ at each grid cell in the study area was generated using the previously mentioned procedure. I_a for the amplified time history was then determined at each grid cell, as shown in Fig. 9a. The I_a distribution includes the topographic effect during the

Chi-Chi earthquake. In conventional landslide hazard analysis, the distribution of shaking intensity is derived by interpolating the recorded shaking intensities around the study area. The interpolated shaking intensity distribution for the Chi-Chi earthquake in the study area is shown in Fig. 9b for comparison. The map of shaking intensity distribution shown in Fig. 9b was generated using the recorded intensity outside the study area, and thus does not indicate the response of topographic amplification effects inside the study area. Landslide density tended to increase with I_a until the maximum density (Fig. 7g and h).

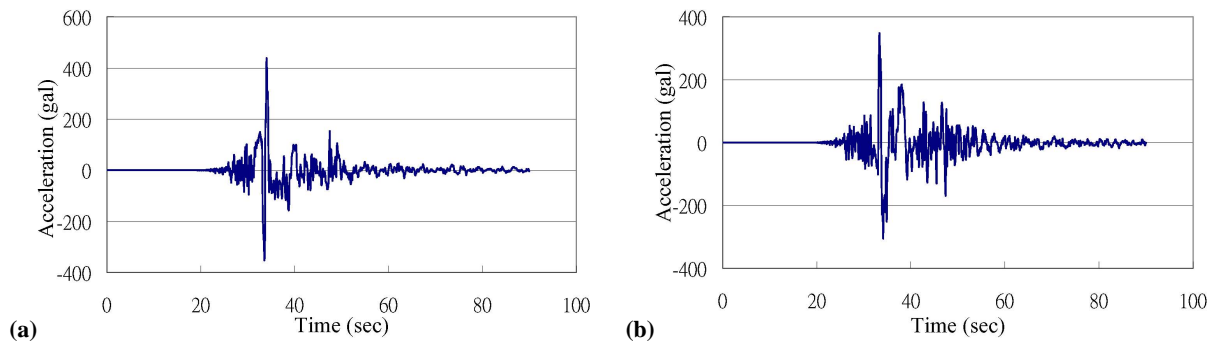


Fig. 8. Ground acceleration of Chi-Chi earthquake recorded at station TCU052 (a) in the NS direction and (b) in the EW direction.

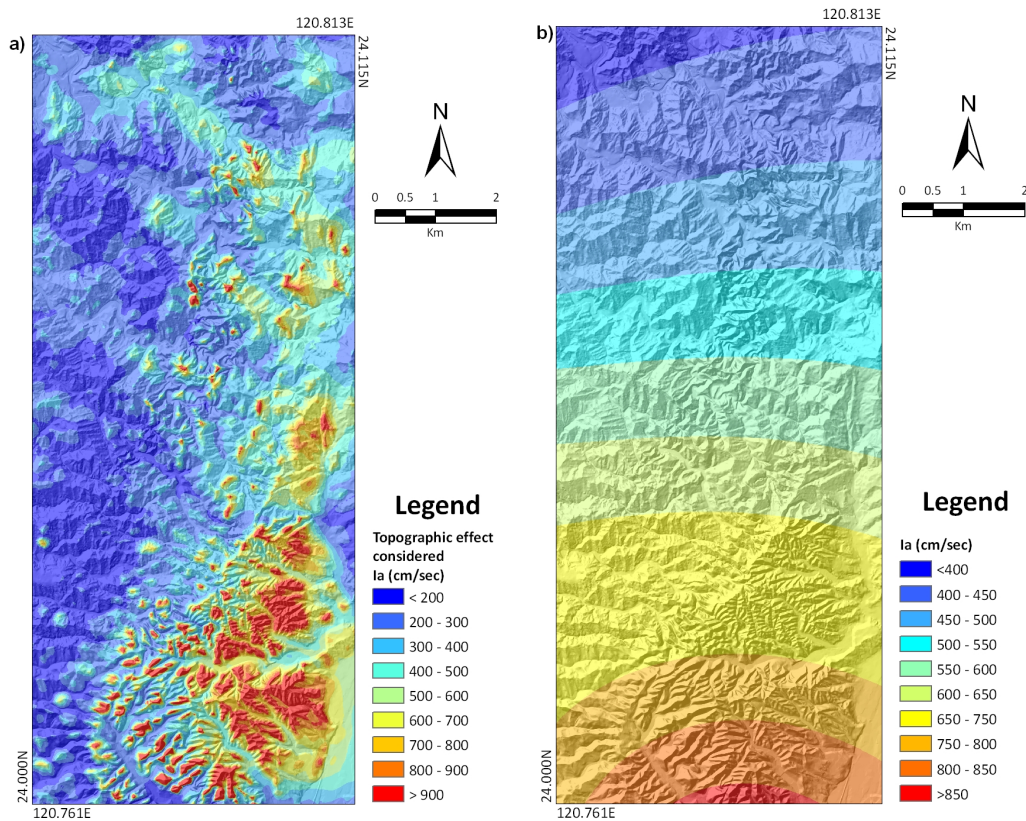


Fig. 9. Arias intensity generated using (a) the proposed procedure with the topographic effect and (b) interpolated from shaking intensity recorded in the seismic stations.

5 Logistic regression

Logistic regression is a mathematical modeling approach that describes the relationship between one or several independent variables and a dichotomous dependent variable (Hosmer and Lemeshow, 1989). For example, the values 1 and 0 can represent the presence and absence of a landslide, respectively. In logistic regression, the independent variables can be categorical, numerical, or both (Menard, 2002). For a binary response variable, it is convenient to model

probability in terms of the log odds of improvement, called the logit, denoted as (Atkinson et al., 1998):

$$\text{logit}(P_i) = \frac{e^{f(X_i)}}{1 + e^{f(X_i)}} \tag{6}$$

where P_i is the probability of an event occurring associated with a given observation i . The probability varies from 0 to 1 on an S-shaped curve and $f(x)$ is the linear combination. Logistic regression involves fitting an equation

of the following form to the data:

$$f(X_i) = \beta_0 + \sum_{i=1}^n \beta_i X_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (7)$$

where β_i ($i = 0, 1, 2, \dots, n$) are n coefficients to be estimated in the model, and X_i ($i = 0, 1, 2, \dots, n$) are the n values for each of the explanatory variables.

For landslide susceptibility mapping, a commonly used method is to create layers of binary values (dummy variables) for an each class of independent parameter (Guzzetti et al., 1999; Ayalew and Yamagishi, 2005; Dai and Lee, 2002; Ohlmacher and Davis, 2003). The spatial databases of each factor were used to determine the relationship between landslides and each calculated factor using the Statistical Package of Social Sciences (SPSS).

5.1 Analytical approaches

The following factors were used in this study: geomorphological, geological, land-cover, and seismic parameters. Four variables were numerical: elevation, slope gradient, terrain roughness, and I_a . Three variables were categorical: slope aspect, lithology, and land cover.

In order to prepare the database for the logistic regression, data with different measuring scales were combined to obtain better results. Before the main statistical analyses, the data were normalized for logistic regression. If this procedure is not carried out, it will be difficult to understand statistical results and evaluate the role of each independent variable in the final results. Many landslide susceptibility assessment studies using logistic regression or linear regression used landslide density to normalize each variable. In this study, we used a different approach. Our approach was to reclassify each of the seven variables into a range of 0–255 using Jenks natural breaks method (Jenks and Coulson, 1963). This study found that classification methods, such as equal interval, quantile, Jenks natural breaks, and geometrical interval, affect the accuracy of the final results. The Jenks natural breaks method improved landslide prediction.

The sample size used to create the dependent variable is important. It is generally recommended that similar proportions of 1 (“landslide”) and 0 (“non-landslide”) cells be used in the logistic regression. Hence, we took 2696 random samples, of which 1298 were landslide and 1398 were non-landslide cells (Fig. 10).

Considering the independent variables, we selected the stepwise method *Forward:Wald* (SPSS software) to avoid the multicollinearity problem between explanatory variables. We then performed entry testing based on the significance of the score statistics and removal testing based on the probability of the Wald statistics. The mode was adjusted using 7 steps: (a) elevation entered on step 1; (b) slope gradient entered on step 2; (c) terrain roughness entered on step 3; (d) slope aspect entered on step 4; (e) I_a entered on step 5; (f) land cover entered on step 6; (g) lithology entered

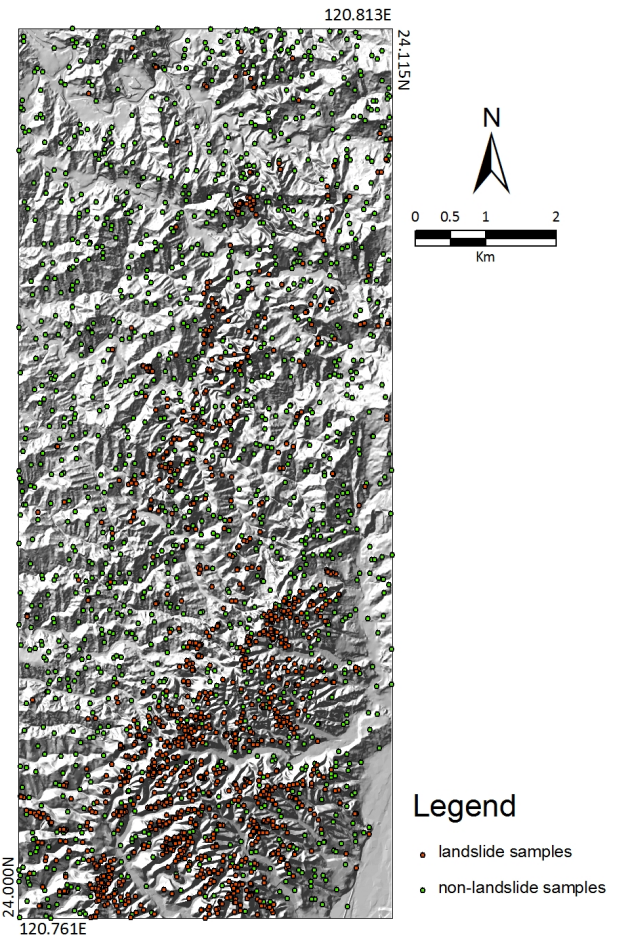


Fig. 10. Distribution of random samples in the logistic regression.

on step 7. All factors were used and logistic regression mathematical equations were formulated as shown in Eq. (7). Then, we obtained the following equation:

$$f(x) = 0.010 \cdot \text{Elevation} + 0.014 \cdot \text{Slope} + 0.009 \cdot \text{Roughness} + 0.018 \cdot I_a + \text{Aspect}_c + \text{Lithology}_c + \text{Landcover}_c - 5.142 \quad (8)$$

where Elevation is the elevation value, Slope is the slope gradient, Roughness is the terrain roughness value, I_a is the Arias intensity value, and Aspect_c , Lithology_c , and Landcover_c are the logistic regression coefficient values. The statistical significance of each coefficient in the model is listed in Table 2.

5.2 Results and verification

The proposed earthquake-induced landslide model using the logistic model is significant at the 0.01 level. The overall statistics of this model are shown in Table 3. Among the entire explanatory variable, are significant at 0.01 confidence level. The receiver operating characteristic (ROC) analysis summarizes the performance of a logistic regression model.

Table 2. Logistic regression results and coefficient values used for this study.

Step 7					
Variable	B	S.E.	Wald	Exp(B)	Sig.
Elevation	0.010	0.001	73.360	1.010	< 0.01
Slope gradient	0.014	0.001	161.757	1.014	< 0.01
Terrain roughness	0.009	0.001	65.966	1.009	< 0.01
I_a	0.018	0.003	34.874	1.018	< 0.01
Aspect			106.514		< 0.01
– Northeast	0.529	0.228	5.376	1.697	0.020
– East	1.102	0.220	25.028	3.011	< 0.01
– Southeast	1.498	0.209	51.401	4.474	< 0.01
– South	1.378	0.207	44.441	3.968	< 0.01
– Southwest	0.852	0.211	16.321	2.344	< 0.01
– West	0.341	0.238	2.057	1.406	0.152
– Northwest	–0.019	0.234	0.007	0.981	0.934
Lithology			17.242		< 0.01
– Cl	–0.995	1.601	0.386	0.370	0.534
– Fl	–0.893	1.353	0.435	0.410	0.509
– Hd	–0.229	1.327	0.030	0.795	0.863
– Kc	–0.985	1.301	0.573	0.373	0.449
– Td	–18.675	10.459.537	0.000	0.000	0.999
– Tkh	0.031	1.244	0.001	1.031	0.980
– Tks	0.323	1.245	0.067	1.381	0.795
Land use			29.389		< 0.01
– Residential area	–1.323	1.438	0.846	0.266	0.358
– Agriculture area	–1.867	0.881	4.489	0.155	0.034
– Bamboo	–1.924	1.381	1.943	0.146	0.163
– Grassland	–1.523	1.502	1.027	0.218	0.311
– Forest area	–0.819	0.855	0.918	0.441	0.338
– Old landslides	2.277	1.376	2.739	9.750	0.098
Constant	–5.142	1.375	13.982	0.006	> 0.01

B represents the estimated regression coefficients for the explanatory variables, with the standard error (S.E.) given. The Wald statistics is the ratio of B to S.E. of the regression coefficient squared. The significance of each explanatory variable (Sig.) is given by the p-value. Exp (B) is the predicted change in odds for a unit increase in the explanatory variable.

ROC curves can be used to provide predictions of landslide probability (Zweig and Campbell, 1993). The curves are obtained by plotting all combinations of sensitivities (on the y-axis) and the proportion of false negatives (1-specificity; on the x-axis) that can be obtained by varying the decision threshold. The area under the ROC curve (AUROC) is an estimate of the model accuracy (Hosmer and Lemeshow, 2000). The area measures discrimination, which is the ability of the test to correctly classify pixels with and without landslide risk. This measure of discrimination is threshold-independent and takes values from 0.5 (no discrimination) to 1 (perfect discrimination). Therefore, an ROC plot closer to the upper left corner indicates a higher overall accuracy of the test. The area corresponding to our study from the ROC curve is shown in Fig. 12. The AUROC is 0.890, which is considered excellent discrimination. The results of the analysis that considers the topography effect are better than those for the analysis without the topography effect (AUROC = 0.874).

Table 3. Overall statistics of the logistic regression model.

Step	–2 Log likelihood	Cox and Snell R Square	Nagelkerke R Square	Hosmer-Lemeshow Chi-square
1	2882.483	0.271	0.361	13.765 n.s
2	2669.977	0.326	0.435	12.507 n.s
3	2559.234	0.353	0.471	20.663
4	2428.880	0.384	0.512	15.746 n.s
5	2391.706	0.392	0.523	11.744 n.s
6	2344.843	0.403	0.537	14.389 n.s
7	2323.067	0.407	0.543	15.081 n.s.

n.s. = not significant at the 0.05 level.

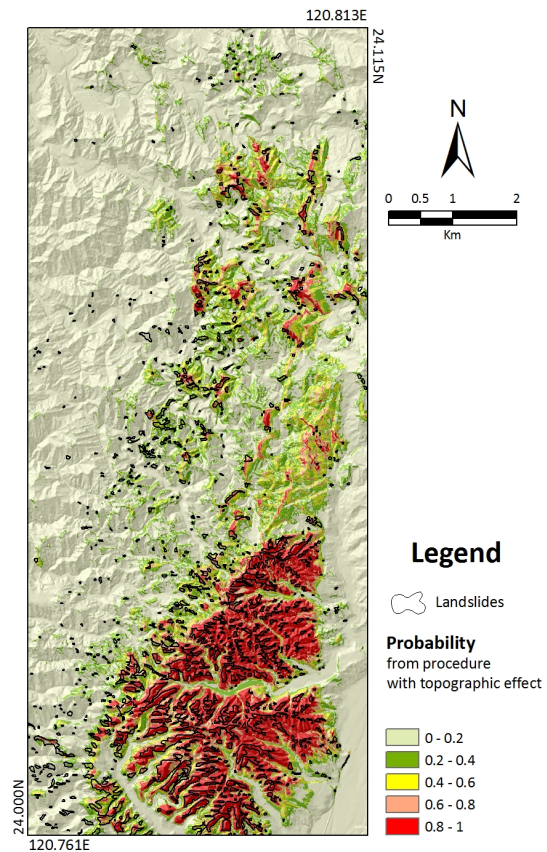


Fig. 11. Landslide susceptibility map based on logistic regression.

The probabilities of landslide occurrence were calculated using Eqs. (6) and (8). The resultant landslide susceptibility map is shown in Fig. 11. In order to evaluate the performance of the proposed method, the calculated landslide susceptibility maps were compared with the inventory of landslides that were actually triggered by the Chi-Chi earthquake. The accuracy curves (Chung and Fabbr, 1997) are introduced to achieve this. To obtain the accuracy curves, the landslide

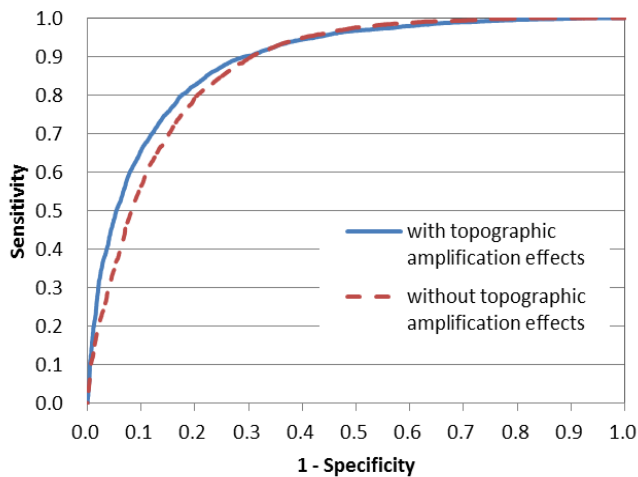


Fig. 12. Receiver operating characteristic (ROC) curve.

hazard results are sorted in descending order. The values of ordered cells were divided into 100 classes with accumulated 1% intervals. The number of landslide cells within the period for each hazard class was counted. Relationships between the cumulative hazard index ranking and the cumulative percent of landslide occurrence are shown in Fig. 13. The cumulative hazard index rank (x-axis) goes from high to low susceptibility (i.e., from dangerous to safe areas). Therefore a lower hazard index rank indicates a higher hazard and higher potential for landslide occurrence. For example, 10% class of the study area where the landslide hazard index hand a higher can account for 63% of all landslides caused by the Chi-Chi earthquake. In addition, the 20% class of the study area where here the landslide hazard index hand a higher can account for 81% of all landslides.

6 Discussion

A model that introduces the topographic effect into landslide hazard analysis, as suggested by Peng et al. (2009a, b), simply uses the vertically incident shear waves as the incident wave for the wave propagation model. Although this is a reasonable engineering simplification, it is unlikely to represent the real situation during earthquakes. In the present study, the strike and dip angles of the Chelungpu fault were adopted to set up the incidence angle of shear waves in the wave propagation model. This makes the incidence angle in the wave propagation model close to that in real situations. The reference motion (recorded in TCU052) is also used to preserve information, making the modified ground motion ideal for landslide hazard analysis.

Although the model includes the topographic effect in the landslide hazard analysis, it ignores the distance effect (path effect) for far-field earthquakes and the highly variable

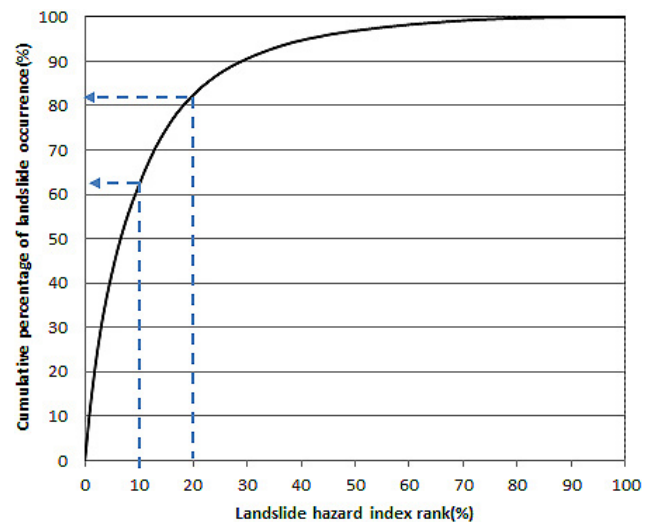


Fig. 13. Cumulative frequency diagram showing cumulative landslide hazard index rank occurring in cumulative percentage of landslide occurrence.

near-field ground motion because the reference motion was recorded at a single seismic station. In the small study area, this problem may be insignificant because the distances from cells to the fault surface are approximately equal for a far-field earthquake. However, for a large study area or for an area controlled by near-field ground motion, the accuracy of the prediction may be reduced. For example, Fig. 5 shows that there is an obvious variation between records from stations TCU052 and TCU071; therefore, there may be an inconsistency between the predicted and actual seismic intensity in the southern part of the study area because the reference motion used to simulate the wave propagation was obtained from seismic station TCU052. This drawback reduces the accuracy that predicts the hazard between the south and north parts of study area in this research, this is the main reason that the accuracy (AUROC) of this modified model is not obviously increased relative to the conventional model that estimates shaking intensity by interpolating. Further research is thus required to combine the effects of highly variable near-field ground motion (or the path effect of far-field earthquakes) with the topographic effect in hazard analysis.

7 Conclusions

The topographic effect was included in the analysis of earthquake-induced landslides in the 99 Peaks region, central Taiwan. The topographic effect was estimated using back analysis. The 3-D staggered-grid finite difference model proposed by Ohminato and Chouet (1997) was constructed to calculate the theoretical topographic amplification factors in the 99 Peaks region. The ground motion record

of the Chi-Chi earthquake at a seismic station near the 99 Peaks region was chosen as the reference motion. By combining the topographic amplification factors with the reference motion, the amplified acceleration time history and amplified shaking intensity parameters (I_a) of the Chi-Chi earthquake were obtained. A logistic regression model was adopted to perform the seismic landslide hazard analysis. The regression model indices, including the topography effect, were then used to generate the landslide susceptibility maps. In order to evaluate the performance of this procedure, a conventional method that ignores the topographic amplification effect was also used. The performance was tested by comparing the known landslide location data with those calculated using landslide susceptibility maps. The conventional method excludes the factor of seismicity when calculating the hazard potential for landslides as it lacks distinct variation throughout the study area. This work shows that the topography amplification effect is too significant to be ignored.

Acknowledgements. The authors would like to thank the Satellite Geoinformatics Research Centre, NCKU, for providing the digital elevation model and the Central Geological Survey for providing the digital inventory of landslides triggered by earthquakes. This work was partially supported by the National Science Council of Taiwan under grant NSC 95-2116-M-006-004.

Edited by: T. Glade

Reviewed by: W. Murphy and two other anonymous referees

References

- Arias, A.: A measure of earthquake intensity, in: *Seismic design for nuclear power plants*, edited by: Hansen, R. J., Massachusetts Institute of Technology Press, Cambridge, 1970.
- Assimaki, D., Gazetas, G., and Kausel, E.: Effects of local soil conditions on the topographic aggravation of seismic motion: Parametric investigation and recorded field evidence from the 1999 Athens earthquake, *B. Seismol. Soc. Am.*, 95, 1059–1089, 2005.
- Atkinson, P., Jiskoot, H., Massari, R., and Murray, T.: Generalized linear modelling in geomorphology, *Earth Surf. Proc. Land.*, 23, 1185–1195, 1998.
- Atkinson, P. M. and Massari, R.: Generalised linear modelling of susceptibility to landsliding in the central Apennines, Italy, *Comput. Geosci.*, 24, 373–385, 1998.
- Ayalew, L. and Yamagishi, H.: The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko mountains, central Japan, *Geomorphology*, 65, 15–31, 2005.
- Bard, P. Y. and Riepl-Thomas, J.: Wave propagation in complex geological structures and their effects on strong ground motion, *Wave motion in earthquake engineering*, edited by: Kausel, E. and Manolis, G., WIT Press, Southampton, UK, 2000.
- Brenning, A.: Spatial prediction models for landslide hazards: review, comparison and evaluation, *Nat. Hazards Earth Syst. Sci.*, 5, 853–862, doi:10.5194/nhess-5-853-2005, 2005.
- Burrough, P. A. and McDonell, R. A.: *Principles of geographical information systems*, Oxford University Press, New York, 1998.
- Carro, M., De Amicis, M., Luzi, L., and Marzorati, S.: The application of predictive modeling techniques to landslides induced by earthquakes: The case study of the 26 September 1997 Umbria-Marche earthquake (Italy), *Eng. Geol.*, 69, 139–159, 2003.
- Catani, F., Casagli, N., Ermini, L., Righini, G., and Menduni, G.: Landslide hazard and risk mapping at catchment scale Arno river basin, *Landslides*, 2, 329–342, 2005.
- Chang, K. T., Chiang, S. H., and Hsu, M. L.: Modeling typhoon- and earthquake-induced landslides in a mountainous watershed using logistic regression, *Geomorphology*, 89, 335–347, 2007.
- Chang, K. T., Chiang, S. H., and Lei, F.: Analysing the relationship between typhoon-triggered landslides and critical rainfall conditions, *Earth Surf. Proc. Land.*, 33, 1261–1271, 2008.
- Chen, S. C. and Wu, C. H.: Slope stabilization and landslide size on mt. 99 peaks after chichi earthquake in Taiwan, *Environ. Geol.*, 50, 623–636, 2006.
- Chung, C. J. and Fabbri, A. F.: Sensitivity analysis of quantitative prediction models based on map overlays: an application to landslide hazard zonation, *Proceedings of the Int. Congr. Geomorphology*, Bologna, Italy, 28 August–3 September 1997.
- Dai, F. C. and Lee, C. F.: Landslides on natural terrain - physical characteristics and susceptibility mapping in Hong Kong, *Mt. Res. Dev.*, 22, 40–47, 2002.
- Densmore, A. L. and Hovius, N.: Topographic fingerprints of bedrock landslides, *Geology*, 28, 371–374, 2000.
- Garcia-Rodriguez, M. J., Malpica, J. A., Benito, B., and Diaz, M.: Susceptibility assessment of earthquake-triggered landslides in El Salvador using logistic regression, *Geomorphology*, 95, 172–191, 2008.
- Geli, L., Bard, P. Y., and Jullien, B.: The effect of topography on earthquake ground motion – a review and new results, *B. Seismol. Soc. Am.*, 78, 42–63, 1988.
- Guzzetti, F., Carrara, A., Cardinali, M., and Reichenbach, P.: Landslide hazard evaluation: A review of current techniques and their application in a multi-scale study, central Italy, *Geomorphology*, 31, 181–216, 1999.
- Harp, E. L. and Keefer, D. K.: Landslides triggered by the earthquake, The Coalinga, California, earthquake of May 2, 1983, edited by: Rymer, M. J. and Ellsworth, W. L., U.S. Geological Survey Professional Paper, 1487, 335–347, 1990.
- Harp, E. L., Wilson, R. C., and Wiczorek, G. F.: Landslides from the February 4, 1976, Guatemala earthquake, U.S. Geological Survey Professional Paper, 1204-A, 35 pp., 1981.
- Hosmer, D. W. and Lemeshow, S.: *Applied logistic regression*, Wiley, New York, 1989.
- Hosmer, D. W. and Lemeshow, S.: *Applied logistic regression*, 2nd ed, edited by: Cressie, N. A. C., Fisher, N. I., Johnstone, I. M., et al., Wiley, New York, 2000.
- Jenks, G. F. and Coulson, M. R.: Class intervals for statistical maps, *International Yearbook of Cartography*, 3, 119–134, 1963.
- Jibson, R. W., Harp, E. L., and Michael, J. A.: A method for producing digital probabilistic seismic landslide hazard maps, *Eng. Geol.*, 58, 271–289, 2000.
- Lee, C. T., Huang, C. C., Lee, J. F., Pan, K. L., Lin, M. L., and Dong, J. J.: Statistical approach to earthquake-induced landslide

- susceptibility, *Eng. Geol.*, 100, 43–58, 2008.
- Lee, S. and Evangelista, D. G.: Earthquake-induced landslide-susceptibility mapping using an artificial neural network, *Nat. Hazards Earth Syst. Sci.*, 6, 687–695, doi:10.5194/nhess-6-687-2006, 2006.
- Lee, S. and Min, K.: Statistical analysis of landslide susceptibility at Yongin, Korea, *Environ. Geol.*, 40, 1095–1113, 2001.
- Lo, W., Wu, L.-C., and Chen, H. W.: Geologic map and explanatory text of Taiwan, sheet 25: Kouhsing Central Geological Survey, Taipei, Taiwan, 1999.
- Luzi, L. and Pergalani, F.: A correlation between slope failures and accelerometric parameters: The 26 September 1997 earthquake (Umbria-Marche, Italy), *Soil Dyn. Earthq. Eng.*, 20, 301–313, 2000.
- Malamud, B. D., Turcotte, D. L., Guzzetti, F., and Reichenbach, P.: Landslide inventories and their statistical properties, *Earth Surf. Proc. Land.*, 29, 687–711, 2004.
- Mathew, J., Jha, V. K., and Rawat, G. S.: Application of binary logistic regression analysis and its validation for landslide susceptibility mapping in part of Garhwal Himalaya, India, *Int. J. Remote Sens.*, 28, 2257–2275, 2007.
- Menard, S.: *Applied logistic regression analysis*, 2nd edn., edited by: Lewis-Beck, M. S., Sage, Thousand Oaks, CA, USA, 2002.
- Nguyen, K. V. and Gatmiri, B.: Evaluation of seismic ground motion induced by topographic irregularity, *Soil Dyn. Earthq. Eng.*, 27, 183–188, 2007.
- Ohlmacher, G. C. and Davis, J. C.: Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA, *Eng. Geol.*, 69, 331–343, 2003.
- Ohminato, T. and Chouet, B. A.: A free-surface boundary condition for including 3d topography in the finite-difference method, *B. Seismol. Soc. Am.*, 87, 494–515, 1997.
- Paolucci, R.: Amplification of earthquake ground motion by steep topographic irregularities, *Earthq. Eng. Struct. D.*, 31, 1831–1853, 2002.
- Pelaez, J. A., Delgado, J., and Casado, C. L.: A preliminary probabilistic seismic hazard assessment in terms of Arias intensity in southeastern Spain, *Eng. Geol.*, 77, 139–151, 2005.
- Peng, W. F., Wang, C. L., Chen, S. T., and Lee, S. T.: A seismic landslide hazard analysis with topographic effect, a case study in the 99 peaks region, central Taiwan, *Environ. Geol.*, 57, 537–549, 2009a.
- Peng, W. F., Wang, C. L., Chen, S. T., and Lee, S. T.: Incorporating the effects of topographic amplification and sliding areas in the modeling of earthquake-induced landslide hazards, using the cumulative displacement method, *Comput. Geosci.*, 35, 946–966, 2009b.
- Rowbotham, D. N. and Dudycha, D.: Gis modelling of slope stability in Phewa Tal watershed, Nepal, *Geomorphology*, 26, 151–170, 1998.
- Sepulveda, S. A., Murphy, W., Jibson, R. W., and Petley, D. N.: Seismically induced rock slope failures resulting from topographic amplification of strong ground motions: The case of Pacoima canyon California, *Eng. Geol.*, 80, 336–348, 2005a.
- Sepulveda, S. A., Murphy, W., and Petley, D. N.: Topographic controls on coseismic rock slides during the 1999 chi-chi earthquake, Taiwan, *Q. J. Eng. Geol. Hydrogeol.*, 38, 189–196, 2005b.
- Stark, C. P. and Hovius, N.: The characterization of landslide size distributions, *Geophys. Res. Lett.*, 28, 1091–1094, 2001.
- Tibaldi, A., Ferrari, L., and Pasquare, G.: Landslides triggered by earthquakes and their relations with faults and mountain slope geometry – an example from Ecuador, *Geomorphology*, 11, 215–226, 1995.
- Wang, H. B., Wang, G., Wang, F., Sassa, K., and Chen, Y. P.: Probabilistic modeling of seismically triggered landslides using Monte Carlo simulations, *Landslides*, 5, 387–395, 2008.
- Wilson, R. C. and Keefer, D. K.: Predicting Areal Limits of Earthquake-induced Landsliding, in *Earthquake Hazards in the Los Angeles Region – An Earth-science Perspective*, edited by: Ziony, J. I., U.S. Geological Survey Professional Paper, 1360, 317–345, 1985.
- Yilmaz, I.: Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat-Turkey), *Comput. Geosci.*, 35, 1125–1138, 2009.
- Zweig, M. H. and Campbell, G.: Receiver-operating characteristic (ROC) plots – a fundamental evaluation tool in clinical medicine, *Clin. Chem.*, 39, 561–577, 1993.