Engineering Geology 182 (2014) 201–212

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo

Statistical seismic landslide hazard analysis: An example from Taiwan

Chyi-Tyi Lee *

Graduate Institute of Applied Geology, National Central University, No. 300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan

ARTICLE INFO

Article history:
Accepted 28 July 2014
Available online 13 August 2014

Keywords:
Landslides
Landslide inventory
Landslide susceptibility
Landslide hazard
Earthquake-induced landslides
Arias intensity

ABSTRACT

Following the work of Lee et al. (2008a), a statistical approach is applied to seismic landslide hazard analysis for the whole of Taiwan. All the work is done using new data sets, which include a new and carefully mapped Chi-Chi earthquake-induced landslide inventory, a 5-m DEM, and a new version of the 1:50,000-scale geologic map of Taiwan. Landslide causative factors used in the susceptibility analysis include the slope gradient, slope aspect, terrain roughness, slope roughness, total curvature, total slope height, and lithology. The corrected Arias intensity taking topographic amplification into consideration is used as a triggering factor.

Firstly, a susceptibility model is built using the 1999 Chi-Chi shallow landslides as a training data set and multivariate logistic regression as the analytical tool. This model is validated by using the 1998 Jueili earthquake-induced landslide data. Then, a probability-of-failure curve is established by comparing the Chi-Chi landslide data and the susceptibility values, after which the spatial probability of landslide occurrence is drawn. The temporal probability may be accounted for with the triggering factor (the hazard level of the Arias intensity), which was obtained through regular probabilistic seismic hazard analysis. Finally, the susceptibility model and the probability-of-failure curve are applied to the whole of Taiwan using the topographically corrected 475-year Arias intensity as a triggering factor to complete a seismic shallow-landslide probability map for ground-motions having a 475-year return period.

1. Introduction

Seismic landslide hazard maps have commonly been prepared based on a deterministic approach using the Newmark displacement method (Wilson and Keefer, 1985; Jibson and Keefer, 1993; Harp and Wilson, 1995; Jibson et al., 1998, 2000; Liao, 2004). However, in recent years, a statistical approach for interpreting seismic landslide distribution has been proposed (Lee et al., 2008a). It is based on an event-based landslide susceptibility model with the earthquake intensity as a triggering factor. The advantage of the statistical approach is that it does not require failure depth, material strength, or groundwater data, and it may produce a better prediction rate (Lee, 2006; Lee et al., 2008a, 2008b). This study basically follows the methodology proposed by Lee et al. (2008a), but the data set is updated and the temporal probability of earthquake intensity is added to complete the hazard analysis.

Rapid mapping of landslides from SPOT images for the whole region of Taiwan was carried out by Liao and Lee (2000). They documented 9272 larger landslides of various types (having areas greater than 625 m^2) covering a total area of 127.8 km^2. Factors controlling the earthquake-induced landslides were evaluated by Liao et al. (2002) based on this landslide inventory, and by Khazai and Sitar (2003) based on another preliminary landslide inventory.The satellite images used in our work, from the year 2000, did not have a high resolution and were analyzed rapidly, and so we re-mapped the earthquake-induced landslides from high-resolution SPOT images and confirmed the landslides by using aerial photo-pairs. Landslide types were recognized, source and deposit areas were separated, and a GIS database was built. This work was carried out from 2003 to 2008. The characteristics of these landslides and evaluation of their controlling factors were introduced in Lee (2013).

This study uses shallow landslides to train a susceptibility model in the vicinity of the mesoseismal region of the Chi-Chi earthquake using a multivariate analysis of landslides and their controlling factors. This model is then applied to the whole of Taiwan using the topographically corrected 475-year Arias intensity as a triggering factor.

2. Regional setting and the Chi-Chi earthquake

The island of Taiwan has an area of 36,188 km^2. The highest peak is Yushan, which rises 3952 m above sea level, but there are numerous other peaks of over 3000 m. Taiwan is tectonically active, being in the collision zone between the Asiatic continent and the Luzon Arc (Teng, 1990). Active crustal deformation (Bonilla, 1975; Yu et al., 1997; Lee, 1999), frequent earthquakes (Tsai et al., 1977; Wu, 1978), numerous typhoons, and a high erosion rate (Dadson et al., 2003) currently affect the region.

The backbone of the island is formed by a mountain range extending north-south that is bordered by foothills and coastal plains on the western flank. Geologically, the range has a metamorphic core surrounded by slate formations and fold-and-thrust Neogene sedimentary strata. The Chelungpu fault, which ruptured during the 1999 Chi-Chi earthquake,

* Corresponding author.
E-mail address: ct@ncu.edu.tw.

http://dx.doi.org/10.1016/j.enggeo.2014.07.023
0013-7952/

is a thrust fault in the fold-and-thrust belt. A coastal plain and gentle slopes lie to the west of the fault, hills and high mountains to the east.

At 1:47 a.m. local time on September 21, 1999, a shallow M_w7.6 earthquake struck central Taiwan rupturing the Chelungpu fault. The hypocenter was 8 km below the village of Chi-Chi. The main shock severely shook the entire island of Taiwan (Ma et al., 1999; Kao and Chen, 2000) and caused considerable structural damage as well as triggering thousands of landslides. The majority of the earthquake-induced landslides were located east of the Chelungpu fault on hilly and mountainous lands.

The climate in the Taiwan area is subtropical, with an average annual precipitation of about 2500 mm. Although it is humid and wet, no significant precipitation was observed within one month before and for a half month after the Chi-Chi earthquake, which simplifies the study of earthquake-induced landslides.

3. Methodology

We adopted the statistical seismic landslide susceptibility analysis method proposed by Lee et al. (2008a) to train a susceptibility model and to develop a probability-of-failure curve for the transformation of susceptibility values into landslide spatial probability. Temporal probability in this study is accounted for using specific return-period earthquake intensity in this study. With the addition of the temporal probability, the susceptibility model is upgraded to a hazard model, and the spatial probability is used to determine the hazard level for certain return periods or exceedance probabilities of earthquakes.

A geographic information system (GIS) was used to digitize the landslides and to build the landslide inventory. Spatial functions in GIS were used to analyze the relationships between the landslide distribution and factors associated with the landslides. The Erdas Imagine system was used to process the digital elevation model (DEM) and derive topographic factors for the statistics. Strong-motion data were processed using a standardized method and FORTRAN code.

3.1. Methods for evaluating the effectiveness of a landslide susceptibility factor

If a factor can be used to interpret the landslide spatial distribution to some extent, it is deemed effective. We propose three different methods to evaluate the effectiveness of a factor. For any given factor, the data set is divided into a *landslide* group and a *non-landslide* group for purposes of analysis. Theoretically, if the two groups have almost no intersection and can be easily separated, then this factor should be a perfectly effective factor. If the percentage of landslides increases or decreases with the factor score, then this factor is also considered an effective factor.

Flat areas and gentle slopes, where the slope gradient is less than 10% but has an area greater than 1 ha, are regarded as stable and were not included either in the analysis or in the validation.

3.1.1. Difference between landslide and non-landslide groups

The difference between two groups can be visually inspected by plotting the frequency distribution of the two groups and then quantified by computing a standardized difference D (Davis, 2002), from which the effectiveness of a factor as a discriminator can be determined:

$$D_j = \frac{\overline{A}_j - \overline{B}_j}{S_{Pj}}, \tag{1}$$

where $\overline{A}_j$ is the mean of factor j for group A (*landslide*); $\overline{B}_j$ is the mean of factor j for group B (*non-landslide*); S_{Pj} is the pooled standard deviation of factor j; and D_j is the standardized difference of factor j. The larger the standardized difference, the more effective the factor.

Before calculating the standardized difference of a factor, a test of normality is required. A standardized difference value for evaluation is valid only for a normally distributed data set.

3.1.2. Probability of failure curve

An effective factor should be correlated with the proportion of landslide cells (Jibson et al., 2000) or the landslide ratio (landslide pixels to total pixels ratio in a factor interval) (Lee et al., 2005, 2008a, 2008b). In the present study, it is called the probability of failure. The correlation can be visually inspected by plotting a probability-of-failure curve.

A correlation coefficient between the landslide ratio and the factor score can be calculated. Either a positive correlation or a negative correlation is good for an effective factor. A threshold can exist in the probability-of-failure curve for landslides. When the factor score is less than the threshold, the probability of failure may be zero. Only factor scores greater than the threshold are utilized in the calculation of the correlation coefficient.

3.1.3. Success rate curve

The success-rate curve (Chung and Fabbri, 1999) is a cumulative percentage of a landslide area against the percentage of the total area, calculated starting from the highest susceptibility of a model. The success rate of a factor is calculated starting from the highest factor score if there is positive correlation between landslide ratio and factor score or from the lowest factor score if there is a negative correlation. If this success-rate method is used in validating a model that uses a different data set, then this is called the prediction rate (Chung and Fabbri, 2003), and a prediction-rate curve is built.

The success-rate curve indicates how well a model (or a factor) interprets the data (landslides). The prediction-rate curve indicates how well a model predicts future landslides. The success-rate curve and prediction-rate curve have computed areas under the curve (AUC) that range between 0 and 1; higher values indicate a higher success rate, and values near or less than 0.5 means that the factor is not effective. In the model evaluation, we classify AUC > 0.9 as excellent, 0.9 > AUC > 0.8 as good, 0.8 > AUC > 0.7 as fair, 0.7 > AUC > 0.6 as poor, AUC < 0.6 as very poor (Lee et al., 2008a, 2008b). In the evaluation of a specific factor, we classify AUC > 0.8 as excellent, 0.8 > AUC > 0.7 as good, 0.7 > AUC > 0.6 as fair, 0.6 > AUC > 0.55 as poor, and AUC < 0.55 as very poor.

3.2. Construction of a susceptibility model

The construction of a landslide susceptibility model can be performed using a multivariate statistical method (Carrara, 1983; Atkinson and Massari, 1998; Guzzetti et al., 1999; Dai et al., 2001; Ayalew and Yamagishi, 2005; Eeckhaut et al., 2006; Greco et al., 2007) or other methods (Varnes, 1984; Hutchinson, 1995; Mantovani et al., 1996; Aleotti and Chowdury, 1999; Chung and Fabbri, 1999; Chung, 2006) using selected effective factors and a landslide inventory. Discriminant analysis was selected in Lee et al. (2008a) to build a susceptibility model with causative factors, a triggering factor, and an event-based landslide inventory. However, in routine analysis of landslide susceptibility for many different drainage basins in Taiwan in recent years, we find that logistic regression is better for three reasons: (1) the shape and concentration of data points in the probability-of-failure curve are always better when the logistic regression method is used, (2) logistic variables can be used for categorical data, such as lithology and slope aspect, in the logistic regression, and (3) although normality is also required in logistic regression, it is not as important as in discriminant analysis. Therefore, the logistic regression method is adopted in the present study.

3.3. Spatial probability of landslides

After a susceptibility model is built, the landslide inventory can be used again to construct a probability-of-failure curve for the model by comparing actual landslide densities and the susceptibility values. This represents the spatial probability of landslide occurrences during a triggering event. A landslide spatial-probability map can be produced by