Nat. Hazards Earth Syst. Sci. Discuss., 1, 471–508, 2013 www.nat-hazards-earth-syst-sci-discuss.net/1/471/2013/ doi:10.5194/nhessd-1-471-2013 © Author(s) 2013. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Natural Hazards and Earth System Sciences (NHESS). Please refer to the corresponding final paper in NHESS if available.

Integrating spatial and temporal probabilities for the annual landslide hazard maps in Shihmen watershed, Taiwan

C. Y. Wu and S. C. Chen

Department of Soil and Water Conservation, National Chung-Hsing University, Taichung 40227, Taiwan

Received: 23 February 2013 - Accepted: 5 March 2013 - Published: 19 March 2013

Correspondence to: S. C. Chen (scchen@nchu.edu.tw)

Published by Copernicus Publications on behalf of the European Geosciences Union.



Abstract

Landslide spatial probability, temporal probability, and landslide size probability were employed to perform landslide hazard assessment in this study. Following a screening process, landslide susceptibility-related factors included eleven intrinsic geomor-

- ⁵ phological factors and two extrinsic rainfall factors, which were evaluated as effective factors because of the higher correlation with the landslide distribution. Landslide area analysis was first employed to establish the power law relationship between landslide area and noncumulative number, and a probability density function was then used to convert this relationship to cumulative probability of landslide area. The exceedance
- probability of rainfall with different recurrence intervals was used to determine the temporal probability of those events. Finally, the landslide spatial probability, landslide area probability, and exceedance probability were integrated to estimate the annual probability of each slope-unit with a landslide area exceeding a certain threshold in a watershed. The slope-units with high landslide probability were concentrated in Taigang Diverse to the slope and which should be the landslide to react of future measurement efforts.
- ¹⁵ River watershed, which should be the leading target of future management efforts.

1 Introduction

Taiwan is often affected by landslides because of its steep topography, fragile geology, seismic activity, and rapid development in mountainous regions. After the Chichi earth-quake (*M*_L = 7.3), the susceptibility of the affected areas to landslides increased, as
²⁰ heavy rainfall during typhoons or storms causes large landslides of loosened soil (Wu and Chen, 2009). Furthermore, climate change increases the amount of bare land and numbers of landslides in Taiwan (Chen and Huang, 2010). Due to the uncertainties associated with natural disasters, risk management is necessary to minimize losses (Chen et al., 2010). In view of the growing emphasis on risk management in disaster
²⁵ prevention work, the quantitative assessment of landslide risk is becoming increasingly important. In particular, landslide hazard analysis is the most important step in risk



assessment. Thus, a landslide hazard model that can be used as a basis for landslide risk analysis was consequently established in this study.

The definition of landslide hazard was "the probability of occurrence within a specified period of time and within a given area of a potentially damaging phenomenon"

- ⁵ (Varnes and IAEG, 1984). Guzzetti et al. (1999) incorporated "magnitude of event" in this definition, obtaining the new definition of landslide hazard as "the probability of occurrence within a specified period of time and within a given area of a landslide event with a certain magnitude". Guzzetti et al. (2005) further established a landslide hazard probability model that could be used to predict landslide location, frequency, and size.
- ¹⁰ Thus, the landslide spatial probability, landslide temporal probability, and landslide size probability were combined as a landslide hazard probability in this study.

Landslide spatial probability is also known as landslide susceptibility, which can be estimated by qualitative or quantitative methods. The quantitative method can further employ either statistical analysis or artificial intelligence. The statistical analysis

- ¹⁵ methods used to determine landslide susceptibility chiefly include bivariate analysis (Chung and Fabbri, 1993; Zêzere et al., 2007) and multivariate analysis. The multivariate analysis method may use multivariate regression (Carrara, 1983; Baeza and Corominas, 2001), logistic regression (Lee et al., 2008; Rossi et al., 2010; Nefeslioglu and Gokceoglu, 2011), or discriminant analysis (Guzzetti et al., 2006; Carrara et al., 2008) to obtain a sot of linear equations distinguishing landslides.
- 20 2008) to obtain a set of linear equations distinguishing landslides and non-landslides, which can be used to derive landslide susceptibility indices for analysis units and to complete landslide susceptibility maps.

With regard to temporal probability, the Poisson probability model and binomial probability model are often used to analyze the recurrence probability of naturally occurring

²⁵ random events in time (Crovelli, 2000; Önöz and Bayazit, 2001). Furthermore, the Poisson probability model has been used to estimate the temporal recurrence probability of events in many studies, including flooding probability research (Önöz and Bayazit, 2001) and landslide recurrence probability research (Guzzetti et al., 2005; Ghosh et al., 2012b). Nevertheless, due to the limited lengths of time for which natural hazard



data is available or the limitations imposed by the assumption that past conditions will remain unchanged in the next few years, it is sometimes necessary to employ flexible methods. Taking landslide and debris flow hazard as examples, rainfall factors can be considered important triggering factors; rainfall intensity in different return periods lead to different scale of landslide and debris flow hazard. The use of the exceedance probability of different rainfall return periods to estimate the probability of landslide and debris flow events can also achieve the goal of estimating temporal probability to a certain degree (Bründl et al., 2009; Chen et al., 2010).

5

As for probability of landslide size, Bak et al. (1988) argued that the phenomenon of self-organized criticality (SOC) occurs in connection with natural landslides, which implies that landslide area and frequency follow a power law. After analyzing three landslide inventories from California, central Italy, and Guatemala respectively, Malamud et al. (2004) verified the power law relationship between landslide area and noncumulative frequency. They also fit probability density function of landslide area with common functions, and found that there is a good agreement with a truncated inverse gamma distribution. For their part, Stark and Hovius (2001) achieved a good agreement after

employing a double Pareto distribution to fit a probability density function of landslide area.

In summary, a watershed was divided into a number of slope units; then the the-²⁰ matic variables of individual slope units were derived, screened, and entered in logistic regression to perform landslide susceptibility analysis. The exceedance probability of rainfall triggering factor and probability density function of landslide area were also employed to establish a probability model for rainfall-induced landslide hazard, which is applied to the Shihmen watershed. This watershed covers an area of 760 km², and

²⁵ is one of the main sources for northern Taiwan. In view of the fact that landslides induced by typhoons and torrential rain events may affect downstream tap water quality by causing surges in turbidity, the landslide hazard analysis of this area may serve as a reference for management of the watershed.



2 Methodology

The landslide hazard is defined as the probability of occurrence within a specified period of time and within a given area of a landslide event with a certain magnitude (Guzzetti et al., 2005; Ghosh et al., 2012a). Thus, the landslide hazard probability, (H_L),

⁵ within a given area can be obtained from the conditional probability of landslide spatial probability, $P(S_L)$, of the temporal probability of a landslide event, $P(N_L)$, and of the landslide size probability $P(A_L)$. The H_L can be calculated based on the independence assumption among the three probabilities using the following equation:

 $H_{\mathsf{L}} = P(S_{\mathsf{L}}) \ P(N_{\mathsf{L}}) \ P(A_{\mathsf{L}}) \,.$

20

25

Landslide inventory maps, thematic variables of landslide susceptibility factors, and rainfall data of landslide events were used to perform landslide hazard analysis, which included landslide susceptibility (spatial probability), occurrence probability of landslide event (temporal probability), and landslide size probability. In this study, because most landslides included in the inventory maps had been induced by typhoons or torrential rains, rainfall was chosen as the sole triggering factor.

2.1 Landslide spatial probability distribution

A watershed was divided into a number of slope units in this study; then the thematic variables of individual slope units were derived, screened, and entered in logistic regression to perform landslide susceptibility analysis. The landslide spatial probability was obtained after testing and validating model.

Due to their relatively unbroken geomorphological boundaries, slope units have more geomorphological and geological significance than grid units. The slope units were consequently employed as basic units of analysis in this study. Guided by the division method employed by Xie et al. (2004), the GIS hydrology module was used to divide the watershed into slope units. The smallest slope units had areas larger than the average landslide area (Van Den Eeckhaut et al., 2009), which reduced the chance



(1)

that a single landslide would be divided among different slope units, ensuring relatively optimal analytical results.

Over 50 types of landslide thematic variables have been considered or used in relevant studies (Lin, 2003). Based on the references, lithology, slope, aspect, elevation,
 normalized differential vegetation index (NDVI), terrain roughness, slope roughness, total slope height, distance from road, distance from fault, and distance from river were preliminarily selected as intrinsic causative factors in this study while various rainfall-related data were employed as extrinsic triggering factors. Referring to the quantitative landslide thematic variable screening procedures of the Central Geological Survey (2009), the landslide thematic variables were selected as effective variables using

- a success rate curve (SRC), landslide ratio plot, frequency distribution of landslide and non-landslide group, and probability-probability plot (P-P plot) for each variable. In addition, the model accuracy assessment included classification error matrix, SRC, and frequency distribution of landslide and non-landslide group. A classification error matrix was used to assess the model accuracy by comparing with the group the slope unit
- trix was used to assess the model accuracy by comparing with the group the slope unit actually belonged to.

After establishing a landslide susceptibility model and calculating landslide susceptibility index for each slope unit, ordinarily the slope units would be ranked as high susceptibility, medium susceptibility, and low susceptibility grades on the basis of their susceptibility indices, enabling the drawing of landslide susceptibility maps. However, the level of susceptibility index (0–1) cannot be directly treated as landslide spatial probability. The spatial probability in this study was therefore determined using the relationship between landslide ratio and landslide susceptibility index.

This was done by calculating the ratio of the landslide sample numbers to the num-²⁵ ber of slope units for each susceptibility index interval, then plotting the relationship between landslide ratio and different value intervals, and converting the different susceptibility indices to spatial probabilities. The relationship plots are also used to verify whether the actual landslide trends are consistent with the degrees of landslide susceptibility.



2.2 Temporal probability of landslides

With regard to analysis of landslide temporal probability, two categories of methods could be chosen based on the number of years of landslide data. The first category consisted only of landslide data before and after a single landslide event. The hourly

rainfall data were collected during the typhoon or torrential rain triggering landslides from rain gauge stations in the study area. Frequency analysis of the rainfall data was employed to derive the exceedance probability of each relevant rainfall event, and thus to obtain the temporal probability of event-based landslides.

The second category consisted of multi-year landslide inventory. In this case, the Poisson probability model was employed to calculate the recurrence intervals of historical landslide events and the temporal probability of landslides based on the assumption that the mean recurrence of events will remain the same in the future. The Poisson probability model of experiencing n landslides during time t is given by the following equation:

¹⁵
$$P[N(t) = n] = \exp(-\lambda t)^* (\lambda t)^n / n!,$$

where λ is the mean occurrence probability of landslides, and its reciprocal μ is mean recurrence interval between landslides in years. The probability that one or more landslides will occur during time *t* is given by the following equation:

$$P[N(t) \ge 1] = 1 - P[N(t) = 0] = 1 - \exp(-t/\mu).$$
(3)

20 2.3 Landslide size analysis

Bak et al. (1988) derived the distribution of landslide area and landslide noncumulative number, and found that the number of landslides increases with landslide area up to a highest value; then it decays following a power law:

 $N_{\rm L}=C'A_{\rm L}^{-\beta},$

(2)

(4)

where A_{L} is the landslide area, N_{L} is the noncumulative number of that landslide area, and β and C' are constants.

Many studies have verified the power law relationship between landside area and noncumulative frequency, including studies of rainfall-induced landslides (Fujii, 1969; Hovius et al., 2000; Weng, 2009; Jaiswal et al., 2011; Ghosh et al., 2012b) and earthquake-induced landslides (Guzzetti et al., 2002).

The probability density function of landslide area was fitted with a Pearson type 5 distribution (i.e. inverse gamma distribution). After ranking landslide area from small to large, various parameters of this distribution function, estimated by fitting, were used to calculate the corresponding cumulative probability of different landslide areas. Thus, the probability of one specific landslide area could be predicted when a landslide occurs

3 Data acquisition and processing

5

10

in slope units.

3.1 Environmental setting of the Shihmen watershed

The Shihmen watershed straddles Taoyuan, Hsinchu, and Yilan counties, and the reservoir is chiefly fed by the Dahan River. This watershed has an area of approximately 760 km², and the Shihmen Reservoir is the third largest reservoir in Taiwan and one of the chief sources of water for northern Taiwan. The geographical extent and river system of the watershed are shown in Fig. 1. The area, higher in the south than in the north, consists chiefly of mountains; the elevation ranges from 236 m to 3526 m, and has an average of approximately 1409 m. The average slope is approximately 34°, and the slope decreases progressively from the southeast to the northwest. With regard to the regional geology, outcrops in the area chiefly consist of the Oligocene Baling stratum, which occupies roughly 35.07% of the total area, Eocene Siling sandstone, which occupies roughly 16.20% of the area, and the Miocene Wenshui stratum, which occupies 12.43% of the area. As far as land use is concerned, most land within the



area consists of undeveloped forest, which occupies 92.44 % of the total area, followed by farmland, which occupies 2.71 % of the overall area. New and enlarged landslides occupied 579 ha, or 0.76 % of the watershed, following Typhoon Aere in 2004, and the greatest number of landslides are found in the upstream basin of Baishi River.

5 3.2 Selection of intrinsic causative variables

The lithology, slope, aspect, elevation, normalized differential vegetation index (NDVI), terrain roughness, slope roughness, total slope height, distance from road, distance from fault, and distance from river were preliminarily selected as intrinsic causative factors in this study. Lithology was chiefly classified as argillite, quartzitic sandstone, hard sandstone and shale, sandstone and shale, terrace deposits, and alluvium on the basis of the 1:50 000 scale geological maps from Central Geological Survey.

Slope, aspect, and elevation data were acquired from a digital elevation model (DEM) using the ArcGIS program. The standard deviation of the elevation calculated in each slope unit, which expresses the degree of terrain irregularity, was taken as the terrain roughness (Wilson and Gallant, 2000). Similarly, the standard deviation of the slope calculated in each slope unit, which expresses the degree of slope variation, was taken as the slope roughness (Wilson and Gallant, 2000).

In addition, the height differential from the crest to the toe of the slope in each slope unit was taken as the total slope height. The NDVI, ranged from -1 to 1, was determined by taking advantage of the absorption of red light and reflection of near infrared

²⁰ mined by taking advantage of the absorption of red light and reflection of near infrared by green plants. The horizontal distance of each slope unit from roads, faults, or perennial rivers can reflect the effect of roads, faults, and rivers on landslides.

As for screening variables, the success rate curve (SRC), landslide ratio plot, frequency distribution of landslide and non-landslide group, and probability-probability plot

(P-P plot) for each variable were used to select the effective variables. Since the area under the curve (AUC) can be used as a basis for determining the effectiveness of a model (Chung and Fabbri, 1999), the SRCs were used to determine the ability of the model to explain training data. The AUC value can range from 0 to 1; the closer the



480

value is to 1, the better the result. This AUC value of SRC was used to assess the ability of thematic variables to predict landslides.

After calculating the ratio of landslide sample numbers to total number of slope units in each value interval for each variable, landslide ratio plots showing the relationship between landslide ratios and different value intervals were drawn to check whether landslide trends are consistent with the physical meanings of the variables. The goal of these frequency distribution plots was to determine whether the frequency distribution of landslide and non-landslide groups could be differentiated, and hence whether the variable could be used to distinguish landslide and non-landslide group. In addition, P-P plot was used to inspect the relationship between a certain variable and a specific

¹⁰ P-P plot was used to inspect the relationship between a certain variable and a specific distribution.

The analysis results of the success rate curve (SRC), landslide ratio plot, frequency distribution of landslide and non-landslide group, and probability-probability plot (P-P plot) are used for subsequent variable selection. Only results of two variables are shown in Fig. 2; terrain roughness is a representative variable which could be used to distinguish landslide and non-landslide group while average NDVI is a representative variable which could not be used. Furthermore, in the frequency distribution of landslide and non-landslide group, the discriminant D_j was also used to judge the variables' ability to distinguish between the landslide and non-landslide groups. In $D_j = (\overline{A_j} - \overline{B_j})/S_j$, $\overline{A_j}$, is the mean value for the landslide group, $\overline{B_j}$ is the mean value for the non-landslide group, and S_j is the pooled standard deviation of the two groups. The AUC value and D_j for each variable are shown in Table 1.

With regard to the selection standard used in variable selection, the first step was to check whether the AUC value was greater than 0.5; if less than 0.5, the factor was considered a random variable in the model, and was assumed to increase model error (Dahal et al., 2008). Furthermore, the landslide ratio plot had to be consistent with the physical meaning of each variable. For instance, the greater the distance from road, the smaller the landslide ratio. According to the analysis results, the variable eliminated in the first step was average NDVI. In the second step, the absolute value of the



discriminant D_j must exceed 0.1 (a D_j value greater than 0 indicates that the mean value of the landslide group is relatively large, while a value less than 0 indicates that the non-landslide group has a larger mean value), or the P-P plot indicates that the values have a normal distribution. As the analysis results, the average elevation and distance from road were eliminated in the second step. Finally, maximum slope, average slope, slope roughness, highest elevation, total slope height, terrain roughness, average aspect, minimum NDVI, distance from fault, distance from river, and lithology were selected as intrinsic thematic factors.

3.3 Selection of extrinsic triggering factors

- Because Typhoon Aere was selected as a research subject, the 96 h of rainfall data from 12 a.m. on 22 August to 12 a.m. on 26 August 2004 were used to perform the analysis. After calculating the maximum 1-, 2-, 3-, 6-, 12-, 24-, 48-, 72-, and 96-h rainfall at each rainfall station during Typhoon Aere, the geostatistics method was employed to estimate the rainfall distribution throughout the entire research area.
- With regard to geostatistics analytical methods, ordinary kriging was first used in analysis. Cokriging was also used in analysis with rainfall as the primary variable and elevation as a secondary variable; another cokriging method was used with rainfall as the primary variable and elevation, slope, and aspect as secondary variables. When performing analysis using these three methods, spherical and Gaussian models were chosen as semivariogram models, which therefore yielded six combinations.

Several indicators of prediction error could be inspected to compare different models. A model complying with the following conditions is optimal: a mean value close to 0, a mean standardized value close to 0, the smallest root-mean-square, the average standard error closest to the root-mean-square, and the root-mean-square standard-

²⁵ ized value closest to 1. In addition, the rainfall distribution estimated using geostatistics methods was compared with the distribution of landslides triggered by Typhoon Aere. Then, SRCs of various rainfall distributions were drawn for calculating the AUC values.



In accordance with the results of the foregoing comparison, maximum 1-h rainfall (cokriging with Gaussian semivariogram model and elevation variable) and maximum 24-h rainfall (ordinary kriging with Gaussian semivariogram model) were employed as extrinsic triggering factors in the landslide susceptibility model. The distributions of these factors are shown in Figs. 3 and 4. The purpose of selecting two different sets of rainfall data – namely maximum 1-h rainfall and a longer duration rainfall – was to reflect the form of rainfall during this typhoon.

3.4 Power law of landslide area

5

All new landslides for past years were classified among three groups according to the slope height of slope unit. The first group consisted of slope units with slope height of less than 379 m, the second of units with slope height of 379–514 m, and the third of units with slope height of more than 514 m. The groups' criteria were determined in order to ensure similar numbers of slope units in each group. There were 4016 landslides in the first group with a β value of 2.1724; there were 3985 landslides in the second group with a β value of 2.2546; there were 4007 landslides in the third group with a β value of 2.1265. The higher the β value, the lower the ratio of landslides with large areas.

4 Results of landslide probabilities

4.1 Spatial probability analysis

The DEM (5 m) of Shihmen watershed was used to divide the watershed into slope units. The original topography could be divided into 659 sub-watersheds, and the combination of sub-watershed units before and after reversal yielded the slope units. A total of 9181 slope units were obtained, and the average size of a slope unit was approximately 8.28 ha.



From the results of screening variables, maximum slope, average slope, slope roughness, highest elevation, total slope height, terrain roughness, average aspect subgroup, minimum NDVI, distance from fault, distance from river and lithology type were selected as intrinsic causative factors while maximum 1-h rainfall and maximum 24-h rainfall were selected as extrinsic triggering factors. The coefficients of variables used 5 in Logistic regression equation are shown in Table 2; landslide susceptibility map for Typhoon Aere based on logistic regression model is shown in Fig. 5. The classification error matrix is shown in Table 3; the SRC and frequency distribution are shown in Fig. 6. The applicability of the model can be seen from an accuracy rate of 77.8% for the landslide group, an accuracy rate of 72.8% for the non-landslide group, the AUC 10 value of 0.788, and the separation of the two groups in the frequency distribution plot. After establishing a landslide susceptibility model for Typhoon Aere and calculating landslide susceptibility index for each slope unit, the relationship between landslide ratio and landslide susceptibility index was employed to determine spatial probability of landslides. As shown in Fig. 7, the landslide ratio increases with landslide susceptibility 15 index, which is consistent with the expected results. The landslide ratio was therefore used to determine the spatial probability of landslides in any particular susceptibility interval. In other words, the relationship equation could be used to convert landslide susceptibility indices to landslide ratios. The results will express the probability of land-

²⁰ slides in those slope units when identical rainfall conditions occur in the future.

25

4.2 Temporal probability of multi-year landslide inventory

The multi-year landslide inventory was used to calculate the temporal probability of landslide occurrence in each slope unit using the Poisson probability model. After assuming that landslides will occur with the same rate during the coming 20 yr as during the past 20 yr, the probability of landslides during 1, 2, 5, 10, and 20 yr was calculated for each slope unit. Figure 8 shows probability of landslide occurrence during the next 1-yr period.



Under the assumed conditions, the areas with the highest landslide probability in the Shihmen watershed are those areas with many landslides during the past 20 yr. The slope units with the highest landslide probability clustered in the southwest part of the watershed.

⁵ Furthermore, the landslide probability during the next 1-yr period obtained using the Poisson probability model can be validated using subsequent estimated annual probability; the results of validation are discussed in Sect. 5.

4.3 Cumulative probability of landslide area

Landslide inventory for the research area was used to perform landslide size analysis. Pearson type 5 and Pearson type 5 (3P) probability density functions were employed to convert power law to cumulative probability of landslide area. Thereby the probability of landslides exceeding a certain size threshold could be derived. The results obtained using Pearson type 5 and Pearson type 5 (3P) probability density functions were extremely similar for all groups. As a consequence, the Pearson type 5 (3P) probability density function was used to show the result for all data and the three groups (see Fig. 9). When a landslide occurs in any slope unit within the research area, the probability that the landslide area will exceed 1000 m² is approximately 58.3 %, and the probability the area will exceed 1000 m² is approximately 6.8 %.

4.4 Validations of spatial and size probabilities

Maximum 1-h rainfall and maximum 24-h rainfall during Typhoon Krosa in 2007 were employed as extrinsic triggering factors, in conjunction with the established intrinsic causative factors, to calculate landslide probability maps. The landslide probability in each slope unit was estimated based on the rainfall conditions occurring during Typhoon Krosa. The resulting landslide spatial probability map was compared with the actual distribution of new landslides to assess the model's prediction ability.



The landslide susceptibility model had an overall prediction accuracy rate of 82.3 % during Typhoon Krosa, which is a very good result. However, the accuracy rate for the landslide group was only 62.1 %. This can be attributed to the fact that rainfall during Typhoon Aere, which formed a basis for the model, was especially heavy. Neverthe-⁵ less, rainfall was significantly lighter during Typhoon Krosa, and this increased the error rate. Furthermore, the AUC value of success rate curve was 0.796, which indicates that landslides uniformly took occurred in areas with relatively high susceptibility. The separation of the landslide group and non-landslide group in the frequency distribution plot reveals that the model's parameters possess the ability to distinguish landslides and non-landslides. This result indicates that this model also retains fairly good accuracy under relatively light rainfall conditions such as those during Typhoon Krosa.

Furthermore, there were 611 landslides, caused by Typhoon Krosa, with its new area greater than 100 m^2 . The cumulative percentages of landslide areas are shown in Fig. 10. A total of 296 landslides had areas greater than 1000 m^2 , which constituted 48.4% of all landslides and was less than the predicted 58.3%. The probability

- ¹⁵ tuted 48.4% of all landslides and was less than the predicted 58.3%. The probability of landslides had areas greater than 1000 m², caused by relatively little rainfall event such as Typhoon Krosa, was less than the 58.3% predicted using the probability density function. The result indicated that there was a higher than expected proportion of landslides with areas less than 1000 m². As a consequence, a hazard model derived
- ²⁰ using a Pearson type 5 (3P) probability density function will overestimate the probability of landslides with a certain size under conditions of relatively light rainfall. However, with regard to hazard assessment, this situation can be considered a conservation result, and the model can still facilitate determination of problem areas.



5 Annual landslide probability

5

25

Rainfall data were collected and used to perform frequency analysis; then landside probabilities corresponding to rainfall events with different recurrence intervals were derived. Exceedance probabilities for particular rainfall events were used as event-based landslide temporal probabilities in the calculation of annual landside probabilities.

5.1 Landslide ratio in rainfall events with different recurrence intervals

More than ten years of data from 30 rainfall stations in the vicinity of the study area were used to analyze the amount of rainfall from rainfall events with different durations and recurrence intervals. Then the maximum 1-h rainfall and maximum 24-h rainfall with different recurrence intervals were obtained. Geostatistics methods were sequentially used to estimate the maximum 1-h rainfall and maximum 24-h rainfall throughout the entire area with different recurrence intervals. These results were used in conjunction with the already-determined intrinsic causative factors to calculate landslide probability maps for the entire area. These maps showed the probability of landslide events in each slope unit at different recurrence intervals. It must be noted, however, that because the

maximum values at individual rainfall stations were employed in estimating the spatial distribution of maximum 1-h rainfall and maximum 24-h rainfall, the resulting landslide susceptibilities reflect simultaneously maximum rainfall values for all stations. These can be seen as "worst case" prediction results.

20 5.2 Annual landslide ratio

The landslide probabilities derived from rainfall events with different recurrence intervals represent the landslide probability of each slope unit following such an event. However, it is difficult to derive the occurrence probabilities of such events. In these circumstances, analysis of the occurrence probability of triggering factors can serve as an alternative method. Therefore, the exceedance probability of rainfall events with



different recurrence intervals was used as the occurrence probability of such an event. Additionally, this probability was used in conjunction with the results of landslide probability to calculate annual landslide probability in the study area (see Fig. 11). The highest landslide probability, annual landslide probability of approximately 40%, is distributed over the south of this watershed, whereas the probability is less than 10% in

5 tributed over the south of this watershed, whereas the probability is less than 10% in most areas.

5.3 Validation of temporal probability

20

The multi-year landslide inventory was used to calculate the temporal probability of landslide occurrence in each slope unit using the Poisson probability model. After assuming that landslides will occur with the same rate during the coming 20 yr as during the past 20 yr, the probability of landslide occurrence for each slope unit during the next one-year period are obtained (see Fig. 8). In addition, annual landslide probabilities obtained using the exceedance probabilities of rainfall events with different recurrence intervals are shown in Fig. 11. The difference between these two was calculated by the annual landslide probability minus the Poisson landslide probability for each slope unit (Fig. 12).

It can be seen from Fig. 12 that absolute values of the probability differentials are almost less than 0.15, which indicate a ratio exceeding 91.9%. The result indicates that the estimated annual landslide probabilities are very close to the estimated one-year landslide probabilities. Additionally, the feasibility of the use of exceedance probability as a basis for determining the temporal probability of event-based landslides was verified.

5.4 Annual probability of landslides exceeding a certain area

The annual probability of landslides exceeding a certain area in any slope unit can ²⁵ be derived by further analysis involving annual landslide probability and landslide area probability in the Shihmen watershed. For instance, Fig. 13 shows the annual



probability of landslides with areas exceeding 3000 m² in each slope unit. The resulting landslide probability model can be used as a basis for future landslide risk analysis. Furthermore, the annual risk can be estimated based on the annual landslide probability instead of a scenario-based probability. Thus, the annual benefit of a risk reduction
⁵ program can be evaluated as the reduced annual risk, and the benefit-cost analysis of the program can be successively achieved. In addition, possible future research analyzing the sediment delivery ratio of each slope unit can be used to estimate the volume of sediment transported downstream during landslide events.

6 Conclusions

A watershed was divided into a number of slope units; then the thematic variables of individual slope units were derived, screened, and entered in logistic regression to perform landslide susceptibility analysis. The exceedance probability of rainfall triggering factor and probability density function of landslide area were also employed to establish a probability model for rainfall-induced landslide hazard. Lastly, results for rainfall events with different recurrence intervals are obtained and used to estimate the annual

probabilities of landslides in each slope unit with areas exceeding a certain threshold.

The applicability of the model can be seen from an overall accuracy rate of 75.3% and the AUC value of 0.788. This AUC value is far greater than the AUC values of the various landslide thematic variables, which indicates that the ability of this landslide susceptibility model to predict landslides is better than that of a model based on only

²⁰ susceptibility model to predict landslides is better than that of a model based on only one variable. According to the validation result, this landslide susceptibility model can be used to predict the spatial probability distribution of landslides caused by rainfall events with different recurrence intervals.

The probability of landslides with areas exceeding 3000 m² in each slope unit was derived based on the landslide hazard analysis for the Shihmen watershed. Additionally, the south of this watershed is especially prone to landslides and should therefore be a main target of future soil conservation efforts. Possible future research analyzing



489

the sediment delivery ratio of each slope unit can be used to estimate the volume of sediment transported downstream during landslide events.

The difference between the annual landslide probability and the Poisson landslide probability for each slope unit was compared to verify the feasibility of the use of ex-

- ceedance probability as a basis for determining the temporal probability of event-based landslides. The resulting landslide probability model can be used as a basis for future landslide risk analysis. This method can avoid the inconsistency between the Poisson probability model assumption and the real situation. The assumption is that landslides will occur with the same rate during the coming few years as during the past, which
 can lead a fallacy that slope units where landslides have not occurred in the past have
 - future landslide probabilities of 0.

Acknowledgements. The authors would like to thank the National Science Council, Taiwan, for financially supporting this research (NSC 101-2811-M-005-033). The Soil and Water Conservation Bureau and involved counties are commended for their full cooperation.

15 References

20

25

Baeza, C. and Corominas, J.: Assessment of shallow landslide susceptibility by means of multivariate statistical techniques, Earth Surf. Proc. Land., 26, 1251–1263, doi:10.1002/esp.263, 2001.

Bak, P., Tang, C., and Wiesenfeld, K.: Self-organized criticality, Phys. Rev., A38, 364–374, doi:10.1103/PhysRevA.38.364, 1988.

Bründl, M., Romang, H. E., Bischof, N., and Rheinberger, C. M.: The risk concept and its application in natural hazard risk management in Switzerland, Nat. Hazards Earth Syst. Sci., 9, 801–813, doi:10.5194/nhess-9-801-2009, 2009.

Carrara, A.: A multivariate model for landslide hazard evaluation, Math. Geol., 15, 403–426, doi:10.1007/BF01031290, 1983.

Carrara, A., Crosta, G., and Frattini, P.: Comparing models of debris-flow susceptibility in the alpine environment, Geomorphology, 94, 353–378, doi:10.1016/j.geomorph.2006.10.033, 2008.



- Discussion Pape Central Geological Survey: Landslide susceptibility zonation in slopeland of urban area, Sinotech Engineering Consultants Inc., 2009 (in Chinese). Chen, S. C. and Huang, B. T.: Non-structural mitigation programs for sediment-related disasters
- after the Chichi Earthquake in Taiwan, J. Mt. Sci., 7, 291-300, 2010.
- ⁵ Chen, S. C., Wu, C. Y., and Huang, B. T.: The efficiency of a risk reduction program for debrisflow disasters – a case study of the Songhe community in Taiwan, Nat. Hazards Earth Syst. Sci., 10, 1591–1603, doi:10.5194/nhess-10-1591-2010, 2010.
 - Chung, C. F. and Fabbri, A.: The representation of geoscience information for data integration, Nonrenewable Resources, 2, 122-138, doi:10.1007/BF02272809. 1993.
- Chung, C. F. and Fabbri, A. G.: Probabilistic prediction models for landslide hazard mapping, 10 Photogramm. Eng. Rem. S., 65, 1389–1399, 1999.
 - Crovelli, R. A.: Probability models for estimation of number and costs of landslides, United States Geological Survey Open File Report 00-249, 2000.

Dahal, R. K., Hasegawa, S., Nonomura, A., Yamanaka, M., Dhakal, S., and Paudyal, P.: Predic-

- tive modelling of rainfall-induced landslide hazard in the Lesser Himalava of Nepal based on 15 weights-of-evidence, Geomorphology, 102, 496-510, doi:10.1016/j.geomorph.2008.05.041, 2008.
 - Fujii, Y.: Frequency distribution of landslides caused by heavy rainfall, Journal Seismological Society Japan, 22, 244-247, 1969.
- Ghosh, S., van Westen, C. J., Carranza, E. J. M., and Jetten, V. G.: Integrating spatial, temporal, 20 and magnitude probabilities for medium-scale landslide risk analysis in Darjeeling Himalayas, India, Landslides, 9, 371-384, doi:10.1007/s10346-011-0304-6, 2012a.
 - Ghosh, S., van Westen, C. J., Carranza, E. J. M., Jetten, V. G., Cardinali, M., Rossi, M., and Guzzetti, F.: Generating event-based landslide maps in a data-scarce Himalayan en-
- vironment for estimating temporal and magnitude probabilities, Eng. Geol., 128, 49-62, 25 doi:10.1016/j.enggeo.2011.03.016, 2012b.
 - Guzzetti, F., Carrara, A., Cardinali, M., and Reichenbach, P.: Landslide hazard evaluation: an aid to a sustainable development, Geomorphology, 31, 181–216, 1999.
 - Guzzetti, F., Malamud, B. D., Turcotte, D. L., and Reichenbach, P.: Power-law correlations of landslide areas in central Italy, Earth Planet. Sc. Lett., 195, 169-183, doi:10.1016/S0012-
- 30 821X(01)00589-1, 2002.





Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., and Ardizzone, F.: Probabilistic landslide hazard assessment at the basin scale, Geomorphology, 72, 272–299, doi:10.1016/j.geomorph.2005.06.002, 2005.

Guzzetti, F., Galli, M., Reichenbach, P., Ardizzone, F., and Cardinali, M.: Landslide hazard as-

sessment in the Collazzone area, Umbria, Central Italy, Nat. Hazards Earth Syst. Sci., 6, 115–131, doi:10.5194/nhess-6-115-2006, 2006.

Hovius, N., Stark, C. P., Chu, H. T., and Lin, J. C.: Supply and removal of sediment in a landslidedominated mountain belt: Central Range, Taiwan, J. Geol., 108, 73–89, doi:10.1086/314387, 2000.

- Jaiswal, P., van Westen, C. J., and Jetten, V.: Quantitative assessment of landslide hazard along transportation lines using historical records, Landslides, 8, 279–291, doi:10.1007/s10346-011-0252-1, 2011.
 - Lee, C.-T., Huang, C.-C., Lee, J.-F., Pan, K. -L., Lin, M.-L., and Dong, J.-J.: Statistical approach to storm event-induced landslides susceptibility, Nat. Hazards Earth Syst. Sci., 8, 941–960, doi:10.5104/phase.8.041.2008.2008

doi:10.5194/nhess-8-941-2008, 2008.

- Lin, Y. H.: Application of neural network to landslide susceptibility analysis. Master thesis, National Central University, Taiwan, 2003 (in Chinese).
- 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.
- Nefeslioglu, H. A. and Gokceoglu, C.: Probabilistic risk assessment in medium scale for rainfall-induced earthflows: Catakli Catchment Area, Math. Probl. Eng., 2011, 1–21, doi:10.1155/2011/280431, 2011.
 - Önöz, B. and Bayazit, M.: Effect of the occurrence process of the peaks over threshold on the flood estimates, J. Hydrol., 244, 86–96, doi:10.1016/S0022-1694(01)00330-4, 2001.
- Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A. C., and Peruccacci, S.: Optimal landslide susceptibility zonation based on multiple forecasts, Geomorphology, 114, 129–142, doi:10.1016/j.geomorph.2009.06.020, 2010.
 - Stark, C. P. and Hovius, N.: The characterization of landslide size distributions, Geophys. Res. Lett., 28, 1091–1094, doi:10.1029/2000GL008527, 2001.
- ³⁰ Van Den Eeckhaut, M., Reichenbach, P., Guzzetti, F., Rossi, M., and Poesen, J.: Combined landslide inventory and susceptibility assessment based on different mapping units: an example from the Flemish Ardennes, Belgium, Nat. Hazards Earth Syst. Sci., 9, 507–521, doi:10.5194/nhess-9-507-2009, 2009.

- Varnes, D. J. and IAEG Commission on Landslides and other Mass-Movements: Landslide hazard zonation: a review of principles and practice, NESCO Press, Paris, 1984.
- Weng, K. L.: Types and distribution of landslides in the Yufong River watershed, M.S. thesis, National Chung-Hsing University, Taiwan, 2009 (in Chinese).
- ⁵ Wilson, J. P. and Gallant, J. C.: Terrain analysis principles and applications, John Wiley & Sons, New York, 2000.
 - Wu, C. H. and Chen, S. C.: Determining landslide susceptibility in Central Taiwan from rainfall and six site factors using the analytical hierarchy process method, Geomorphology, 112, doi:10.1016/j.geomorph.2009.06.002, 190–204, 2009.
- 10 Xie, M., Esaki, T., and Zhou, G.: Gis-based probabilistic mapping of landslide hazard using a three-dimensional deterministic model, Nat. Hazards, 33, 265–282, doi:10.1023/B:NHAZ.0000037036.01850.0d, 2004.
 - Zêzere, J. L., Oliveira, S. C., Garcia, R. A. C., and Reis, E.: Landslide risk analysis in the area North of Lisbon (Portugal): evaluation of direct and indirect costs resulting from a motorway
- disruption by slope movements, Landslides, 4, 123–136, doi:10.1007/s10346-006-0070-z, 2007.



Table 1. The AUC value and D_j for each variable.

Variable	AUC	D_j	Variable	AUC	D_j
Maximum slope	0.678	0.736	Average aspect	0.526	-0.114
Average slope	0.629	0.526	Average NDVI	0.481	0.087
Slope roughness	0.603	0.398	Minimum NDVI	0.651	-0.650
Highest elevation	0.527	0.155	Distance from fault	0.521	-0.117
Average elevation	0.506	0.050	Distance from road	0.512	-0.062
Total slope height	0.683	0.779	Distance from river	0.609	-0.556
Terrain roughness	0.685	0.789	Lithology	0.523	-0.136

Discussion Pa	NHESSD 1, 471–508, 2013				
per Disc	The annua hazard i Shihmen v	The annual landslide hazard maps in Shihmen watershed			
ussion Pape	Title	Page			
	Abstract	Introduction			
	Conclusions	References			
iscussi	Tables	Figures			
on P	14	►I.			
aper	•	Þ			
	Back	Close			
Discus	Full Screen / Esc Screes Printer-friendly Version				
ssion					
Paper	Interactive Discussion				
		BY			

NHESSD 1, 471–508, 2013					
The annua hazard Shihmen	The annual landslide hazard maps in Shihmen watershed				
C. Y. Wu and	C. Y. Wu and S. C. Chen				
Title	Title Page				
Abstract	Introduction				
Conclusions	References				
Tables	Tables Figures				
14	۶I				
•	Þ				
Back	Close				
Full Scre	Full Screen / Esc				
Printer-frier	Printer-friendly Version				
Interactive Discussion					

Discussion Paper

Discussion Paper

Discussion Paper

Discussion Paper

Table 2. Coefficients of variables used in Logistic regression equation.

Variable	Coefficient	Variable	Coefficient	Variable	Coefficient
maximum slope	0.036	maximum 1-h rainfall	0.035	average aspect	
average slope	0.004	maximum 24-h rainfall	0.003	average aspect subgroup (1)	0.646
slope roughness	0.019	intercept	-8.053	average aspect subgroup (2)	1.294
highest elevation	0.001	lithology		average aspect subgroup (3)	1.340
total slope height	0.003	lithology type (1)	1.128	average aspect subgroup (4)	1.091
terrain roughness	0.002	lithology type (2)	-0.317	average aspect subgroup (5)	0.699
minimum NDVI	-2.602	lithology type (3)	-19.364	average aspect subgroup (6)	0.429
distance from fault	-0.033	lithology type (4)	-19.856	average aspect subgroup (7)	0.316
distance from river	-0.046				

Table 3. Classification error matrix for Typhoon Aere.

		Prediction results		Accuracy rate (%)
		Landslide group	Non-landslide group	
Observed data	Landslide group	311	89	77.8
	Non-landslide group	109	291	72.8
Overall				75.3

NHE 1, 471–5	NHESSD 1, 471–508, 2013				
The annua hazard r Shihmen v	The annual landslide hazard maps in Shihmen watershed				
C. Y. Wu and	C. Y. Wu and S. C. Chen				
Title I	Title Page				
Abstract	Introduction				
Conclusions	References				
Tables	Tables Figures				
14	۶I				
•	F				
Back	Close				
Full Scre	Full Screen / Esc				
Printer-frien	Printer-friendly Version				
Interactive Discussion					

Discussion Paper

Discussion Paper

Discussion Paper

Discussion Paper







Fig. 2. Success rate curves (SRCs), landslide ratio distributions (LRDs), frequency distributions of landslide and non-landslide group (FDs), and probability-probability plots (P-P plots) of representative variables.



Fig. 3. Maximum 1-h rainfall during Typhoon Aere.

Discussion Pa	NHESSD 1, 471–508, 2013				
per Discussior	The annual landslide hazard maps in Shihmen watershed C. Y. Wu and S. C. Chen				
1 Pape	Title	Page			
er -	Abstract	Introduction			
	Conclusions	References			
iscuss	Tables	Figures			
ion P	14	►I			
aper	•	•			
	Back	Close			
Discus	E Full Screen / Es				
sion	Printer-friendly Version				
Paper		Interactive Discussion			
	(cc)	\bullet			



Fig. 4. Maximum cumulative 24-h rainfall during Typhoon Aere.

Discussion Pa	NHESSD 1, 471–508, 2013				
per Discussior	The annual landslide hazard maps in Shihmen watershed C. Y. Wu and S. C. Chen				
n Pap	Title	Title Page			
er	Abstract	Introduction			
	Conclusions	References			
iscussi	Tables	Figures			
on P	I •	►I			
aper	•	Þ			
	Back	Close			
Discus	Full Scre	en / Esc			
sion	Printer-frien	Printer-friendly Version			
Paper	Interactive Discussion				



Fig. 5. Landslide susceptibility map for Typhoon Aere based on logistic regression model.





Fig. 6. (a) Success rate curve and (b) frequency distribution of landslide and non-landslide group during Typhoon Aere.





Fig. 7. Landslide spatial probability during Typhoon Aere.





Fig. 8. Probability of landslide events in the Shihmen watershed within a 1-yr period.



Fig. 9. Cumulative probability of landslide area based on Pearson Type 5 (3P) distribution.





Fig. 10. Cumulative percentage of landslide areas caused by Typhoon Krosa in 2007.





Fig. 11. Annual landslide probability in the Shihmen watershed.

506









