

Article

A Landslide Probability Model Based on a Long-Term Landslide Inventory and Rainfall Factors

Chun-Yi Wu *  and Yen-Chu Yeh

Department of Soil and Water Conservation, National Chung Hsing University, Taichung 402, Taiwan;
cxz7.997.9@gmail.com

* Correspondence: cywu@nchu.edu.tw; Tel.: +886-4-2284-0381 (ext. 605)

Received: 11 February 2020; Accepted: 24 March 2020; Published: 26 March 2020



Abstract: The prediction and advanced warning of landslide hazards in large-scale areas must deal with a large amount of uncertainty, therefore a growing number of studies are using stochastic models to analyze the probability of landslide occurrences. In this study, we used a modified Thiessen's polygon method to divide the research area into several rain gauge control areas, and divided the control areas into slope units reflecting the topographic characteristics to enhance the spatial resolution of a landslide probability model. We used a 2000–2015 long-term landslide inventory, daily rainfall, and effective accumulated rainfall to estimate the rainfall threshold that can trigger landslides. We then employed a Poisson probability model and historical rainfall data from 1987 to 2016 to calculate the exceedance probability that rainfall events will exceed the threshold value. We calculated the number of landslides occurring from the events when rainfall exceeds the threshold value in the slope units to estimate the probability that a landslide will occur in this situation. Lastly, we employed the concept of conditional probability by multiplying this probability with the exceedance probability of rainfall events exceeding the threshold value, which yielded the probability that a landslide will occur in each slope unit for one year. The results indicated the slope units with high probability that at least one rainfall event will exceed the threshold value at the same time that one landslide will occur within any one year are largely located in the southwestern part of the Taipei Water Source Domain, and the highest probability is 0.26. These slope units are located in parts of the study area with relatively weak lithology, high elevations, and steep slopes. Compared with probability models based solely on landslide inventories, our proposed landslide probability model, combined with a long-term landslide inventory and rainfall factors, can avoid problems resulting from an incomplete landslide inventory, and can also be used to estimate landslide occurrence probability based on future potential changes in rainfall.

Keywords: landslide; rainfall threshold; landslide probability model; Taiwan

1. Introduction

Taiwan is a relatively new island formed by plate movements. Due to its high mountains, steep slopes, and relatively unstable geological conditions, as well as frequent typhoons and torrential rains, slope and disasters are common in mountainous areas. Thus, slope hazard prevention and mitigation projects are necessary. In slope hazard prevention work, landslides have a high level of unpredictability. In particular, estimating the likelihood of landslides in large watersheds using deterministic models is difficult when no detailed geomorphological and hydrological data have been collected for the whole area. Therefore, the use of a stochastic model to assess landslide probability is more feasible. According to the definition, landslide hazard involves both spatial and temporal probability [1]. The analysis of landslide spatial probability is generally seen as a landslide

susceptibility analysis in the research [2–12]. The landslide temporal probability, normally expressed in terms of frequency, return period, or exceedance probability [13], was analyzed in the research [14–21].

Methods of performing landslide temporal probability analysis can be classified as hydrological models and approaches based on exceedance probability [22]. The hydrological models employ infiltration models to determine the critical rainfall triggering landslides, which requires the estimation and validation of soil parameters over large areas, and therefore makes these models impractical for regional-scale applications. The approaches based on exceedance probability can be further subdivided into two types, where the first type employs a landslide inventory induced by a single rainfall event and rainfall data for that event to analyze the return period of the landslide event [8,23,24], and the second type employs a long-term landslide inventory to calculate the exceedance probability for the occurrence of landslides. Concerning the latter type, the Poisson probability model [17,25–29], binomial probability model, and empirical model [20] are commonly used to analyze the recurrence probability of landslide events. As a consequence, when a research area has a long-term landslide inventory, the Poisson probability model can be employed to estimate the temporal probability of landslides under the assumption that the frequency of future landslides occurring is the same as in the past. However, due to the constraints of this assumption, the Poisson probability model cannot separate the effect of geomorphological and hydrological factors on landslides, and therefore cannot be used to infer how landslide probability will change when climate change causes changes in the frequency of torrential rain and in the rainfall patterns. If the effects of geomorphological and hydrological factors can be considered separately and the occurrence probability of torrential rain events can be estimated independently, then the landslide temporal probability can be estimated correctly based on the change trends of the estimated torrential rain occurrence probability [30]. One approach to separate the effects of geomorphological and hydrological factors in landslide probability models is to employ the concept of conditional probability to separately estimate rainfall probability and landslide probability under these rainfall conditions. In this approach, a Poisson probability model is first used to calculate the exceedance probability of rainfall events that may trigger landslides, the landslide probability under these rainfall conditions is then calculated, and the two are multiplied to obtain the temporal probability of landslides [26,28,29].

Before calculating the probability of rainfall events that may trigger landslides, the scale of rainfall events that trigger landslides or the threshold rainfall events that must be exceeded to trigger landslides must first be understood. The minimum amount of rainfall needed to trigger landslides was first considered by Endo [31], and the rainfall thresholds triggering landslide events were quantified by Onodera et al. [32]. Campbell [33] suggested that the combined effect of both antecedent rainfall and rainfall intensity on the landslides needed to be considered, and a warning system could be based on the relationship between antecedent rainfall and critical rainfall [34]. Caine [35] used rainfall intensity and rainfall duration to establish global shallow landslide rainfall thresholds. Methods of establishing rainfall thresholds were classified as physical models and empirical models [36], where physical models employ detailed spatial information on hydrological, lithological, morphological, and soil characteristics as a basis for modeling the relationship between rainfall, infiltration, and landslide events. However, the information is hard to collect accurately over large areas. Empirical models can be grouped as thresholds combining rainfall duration, total event rainfall, or rainfall intensity parameters, thresholds considering antecedent rainfall, and thresholds combining other parameters, where the first two groups can be further subdivided into the following three categories based on the parameters used for determining rainfall thresholds [21]: the first category consists of intensity and duration parameters [18,20,29,34,35,37–39], the second category consists of antecedent rainfall conditions [26,28,29,40], and the third category consists of accumulative event rainfall and duration parameters [41]. Although rainfall intensity–duration models have been most commonly used in recent years [21], thresholds for rainfall-induced landslides may define the rainfall, soil moisture, or hydrological conditions that, when reached or exceeded, are likely to trigger landslides [36]. Some research has also suggested that groundwater and soil moisture are factors influencing the initiation of

landslides [42,43], and antecedent rainfall can affect both of these factors. Accordingly, antecedent rainfall can be used to determine when landslides may occur [36]. In research on rainfall thresholds incorporating the antecedent rainfall conditions, large differences exist in the number of days of antecedent rainfall that were employed in each study. For example, daily rainfall was employed in conjunction with 15-day antecedent rainfall [39], both daily and 3-day cumulative rainfall were used [44], and 3-day and 30-day antecedent rainfall were employed [45]. Guzzetti et al. [36] suggested that the large variability in the antecedent rainfall may be attributed to three types of factors concerning the research area: diversity in lithological, morphological, vegetation, and soil conditions; differences in climatic regimes and meteorological circumstances; and the heterogeneity and incompleteness in the rainfall and landslide data used to establish the rainfall thresholds. As a consequence, the local conditions and availability of data in the research area must be assessed when choosing the number of days of antecedent rainfall.

Since the rainfall threshold determined using a single rain gauge for a large area constitutes one value for the entire area, as soon as rainfall reaches or exceeds the threshold, landslides may occur anywhere in that area, and knowing their precise locations is impossible. As a consequence, a denser array of rain gauges can be employed to acquire rainfall data with finer spatial resolution [21], and the research areas can be subdivided into analytical units with a smaller area, which can better account for the spatial variability of rainfall patterns in the analytical units and the spatial resolution of landslide prediction. However, 19.1% of recent studies on this subject failed to subdivide their research areas, and those studies that did subdivide their research areas had resulting analytical units with an average area of 302.0 km² [21]. For instance, a research area of 4660 km² was subdivided into 12 analytical units with an average area of 388.3 km² [39], but excessively large analytical units make it impossible to identify the precise possible locations of landslides. In addition, the subdivision approaches employed in some studies run into the problem of incomplete coverage. For instance, although a 25 km² research area was subdivided into eight analytical units, the landslide prediction results only represented the paths of roads in the subdivisions and not the entire subdivisions because most landslides (94%) in the study occurred on roadside slopes [26]. Althuwaynee et al. [28] divided the research area into six circular analytical units with their centers at rain gauges, but the analytical units did not cover the entire research area and also overlapped. Although these studies subdivided their research areas into different analytical units, the units could not provide a landslide probability distribution with a finer spatial resolution because they were excessively large, or experienced problems such as incomplete coverage and overlap. If the method of subdividing a research area into analytical units is improved so that the units are smaller in area, the spatial resolution of the landslide probability estimation results could be improved. There are seven types of analytical units subdivided in research areas: grid cell, terrain unit, unique condition unit, slope unit, geo-hydrological unit, topographical unit, and administrative unit [46,47]. The slope units are suitable for landslide probability analysis because they express topographic features and slope characteristics. In this study, we consequently selected slope units as our analytical unit.

2. Research Area and Materials

2.1. Environmental Setting of Taipei Water Source Domain

Taipei Water Source Domain is located in the northeast part of Taiwan and supplies tap water for five million people in the greater Taipei area. The area is characterized by hilly and mountainous topography, as well as the Xueshan Range extending to the northeast and a subrange of Mt. Qilan extending to the northwest, both of which account for the area's high terrain in the south and low terrain in the north. Elevations in the area range from 12 to 2130 m (Figure 1).

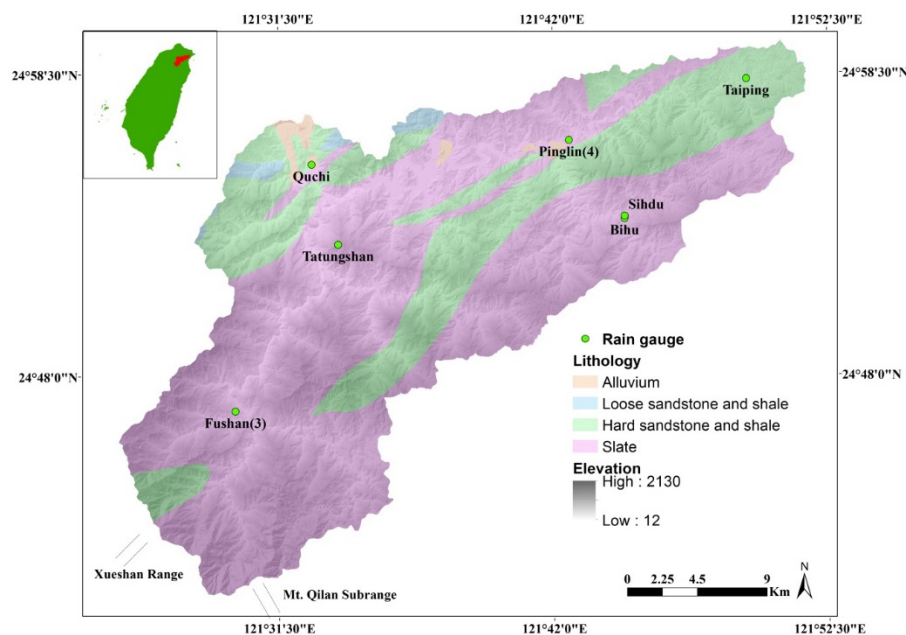


Figure 1. Distribution of elevation, lithology, and rain gauges in the Taipei Water Source Domain.

Concerning the distribution of lithology, we followed the classification approach proposed by Lin et al. [48] by dividing the Taipei Water Source Domain into areas underlain by alluvium, loose sandstone and shale, hard sandstone and shale, and slate. Whereas alluvium found at the confluence of rivers and in downstream areas covers only a small part of the research area, hard sandstone and shale as well as slate underlay most of the research area. Of these types of lithology, areas underlain by slate had the highest number of landslides and the greatest landslide area. Wu et al. [49] indicated that the areas underlain by hard sandstone and shale as well as slate in the Kaoping River Watershed had the highest landslide ratios in 2008 and 2009. This indicates that the lithology condition of most areas is fragile. Typhoons and torrential rain events can readily wash away unconsolidated sand and gravel as well as trigger landslides, which deposit large loads of sediment in rivers and reservoirs.

2.2. Rainfall Data

The rain gauges employed in this study were located as shown in Figure 1, and rainfall data between 1987 to 2016 from these rain gauges were used. Average daily rainfall for the entire area during the same period as the 2000–2015 landslide inventory is shown in Figure 2. Figure 2 shows that apart from the eight typhoon events causing the corresponding landslide inventory, other events of high daily rainfall occurred without a significant increase in landslides. As a consequence, apart from calculating the exceedance probability that rainfall events will exceed the rainfall threshold, we also calculated the probability of landslides when rainfall events exceed the threshold. In addition, Figure 3 shows the average daily rainfall and standard deviation of the eight typhoon events during the 2000–2015 period in each control area of a rain gauge divided by a modified Thiessen polygon method, considering the morphology of the area, proposed by Salvaticic et al. [19].

2.3. Landslide Inventory

After selecting eight major typhoon events occurring in the research area during the 2000–2015 period—typhoons Xangsane (2000), Nari (2001), Aere (2004), Sinlaku (2008), Morakot (2009), Parma (2009), Megi (2010), and Soudelor (2015)—we collected satellite images before and after each typhoon event, calculated and classified the normalized difference vegetation index (NDVI) to find the possible locations of landslide sites, and eliminated and revised unlikely landslide sites according to the slope, drainage, and land use maps in the study area. In the process of mapping the source areas of landslides

from the satellite images, we often found that human mapping errors affected interpretation quality. We followed the recommended procedures proposed by Liu et al. [50] to map landslides in the research area. Table 1 shows the dates of the eight landslide events and landslide statistical data. The size of landslides ranged from 16 to 118,108 m² and the average area was 2474 m². The resulting distribution of landslides caused by the eight typhoon events was shown in Figure 4, which revealed that landslide sites were concentrated in the southwestern portion of the research area.

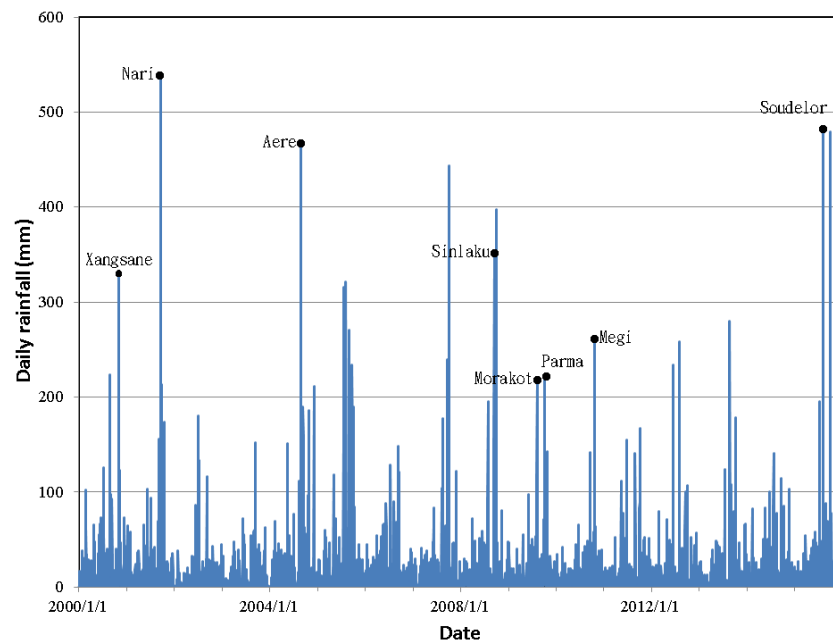


Figure 2. Average daily rainfall within the Taipei Water Source Domain, 2000–2015. The dots represent the eight typhoon events causing the corresponding landslide inventory.

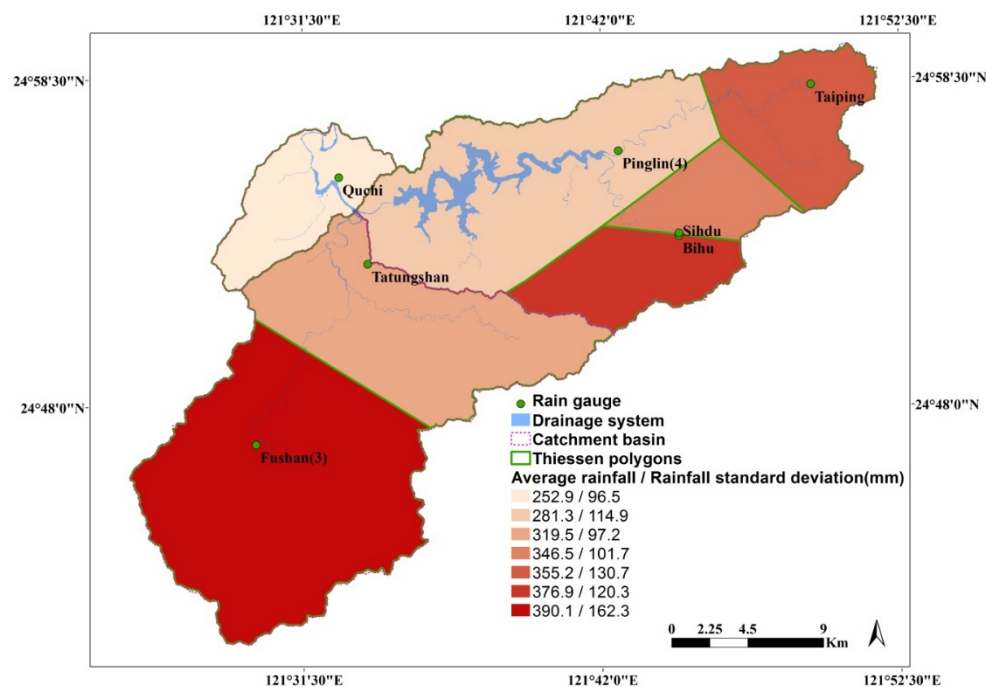


Figure 3. Average daily rainfall and standard deviation of the eight typhoon events in each control area of the rain gauge divided by the modified Thiessen polygon method.

Table 1. Landslide inventory for the eight typhoon events.

Typhoon Event	Date (MM/DD/YYYY)	Average Rainfall at the Date (mm)	Number of New Landslide Sites	Smallest Landslide Area (m²)	Largest Landslide Area (m²)	Total Area of Landslides (m²)	Average Area of Landslides (m²)
Xangsane	11/01/2000	326.67	42	326	19,619	131,148	3123
Nari	09/16/2001	538.05	92	107	68,032	261,650	2844
Aere	08/24/2004	465.57	97	140	27,270	239,856	2473
Sinlaku	09/13/2008	348.18	32	475	21,101	71,111	2222
Morakot	08/07/2009	219.83	173	16	118,108	1,016,448	5875
Parma	10/05/2009	221.79	302	47	49,369	484,785	1605
Megi	10/21/2010	262.35	47	407	27,318	118,874	2529
Soudelor	08/08/2015	478.99	589	257	48,041	1,075,263	1826

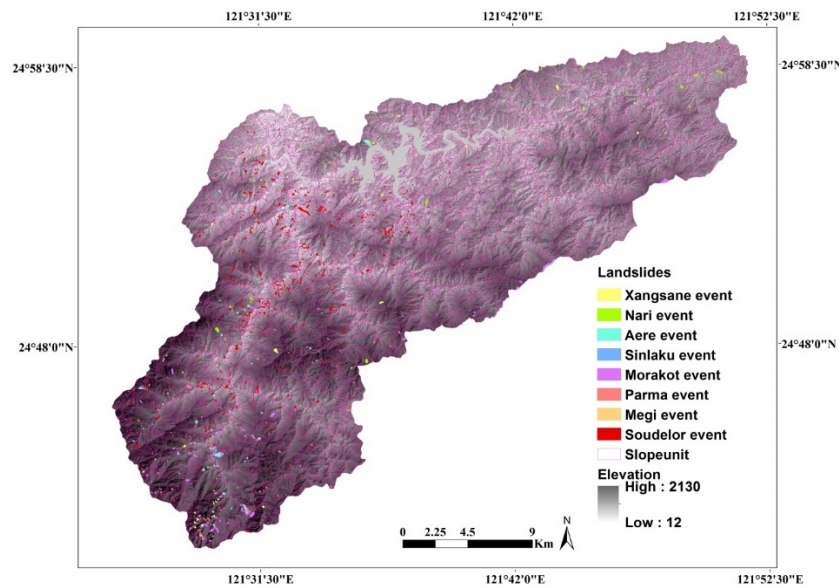


Figure 4. Distribution of slope units and landslide sites caused by the eight typhoon events in the Taipei Water Source Domain.

2.4. Analytical Units and Rain Gauge Control Areas

We employed slope units as analytical units due to their relatively well-defined topographic boundaries, as well as topographic and geological meaning. We employed the subdivision method used by Xie et al. [51] to divide the watershed into slope units. The original topography could be divided into sub-watersheds, and the combination of sub-watershed units before and after reversal yielded the slope units. We ensured that the smallest area of slope units was larger than the average area of landslides [47], which minimized the chance that any specific landslide site would be a part of different slope units, and thereby confuse the analysis results. We also divided the research area into rain gauge control areas (Figure 3) based on rain gauge locations and using the modified Thiessen polygon method. The rainfall measured by each rain gauge was taken as representative of the control area in which that gauge was located, and we expected this approach to reflect the different rainfall distribution characteristics within the research area.

3. Methods

3.1. Analysis of Discrete Rainfall Groups

The two rainfall parameters considered in this study consisted of daily rainfall (I) and effective accumulated rainfall (R_t). After selecting rain gauges near the research area with rainfall data for recent years, we obtained daily rainfall data for the 1987–2016 period from the Water Resources Agency and Central Weather Bureau. This study calculated the effective accumulated rainfall based on rainfall for that day and rainfall during the previous 7 days using the method proposed by Jan [52]; this calculation was performed using Equation (1):

$$R_t = R_0 + \sum_{i=1}^7 \alpha^i R_i = \sum_{i=0}^7 \alpha^i R_i \quad (1)$$

where R_0 is the rainfall amount on that day, R_1 is the rainfall amount on the day before that day, and so on, and the weighting coefficient $\alpha = 0.7$ proposed by Jan [52].

Adopting the concept proposed by Tsai [53], after using daily rainfall data to calculate effective accumulated rainfall (R_t), we obtained a group of daily rainfall and effective accumulated rainfall (I , R_t) for each day. The daily rainfall and effective accumulated rainfall were continuous variables and

would not facilitate subsequent calculation of a joint cumulative distribution function, therefore we rounded off the daily rainfall and effective accumulated rainfall values to the 10th place and made them discrete variables. The group of daily rainfall and effective accumulated rainfall (I, R_t) for each day was termed as “discrete rainfall group” in this study.

We defined different rainfall events by the continuity of daily rainfall. Consecutive days of non-zero daily rainfall were considered to be the same rainfall event, and the number of the consecutive days varied from event to event. We then calculated the distance (d) from each discrete rainfall group to the origin (0, 0), and assumed that the greater the value of d , the greater the likelihood of landslides. The discrete rainfall group with the greatest d in each rainfall event was used to represent that rainfall event in subsequent analysis.

3.2. Joint Cumulative Distribution Function

The joint cumulative distribution function was obtained from the joint probability mass function of the foregoing discrete rainfall groups. The probability ($P_{I,R_t}(I_i, R_{tj})$) of each discrete rainfall group (I_i, R_{tj}) was defined [54] as shown in Equation (2):

$$P_{I, R_t}(I_i, R_{tj}) = P(I = I_i \cap R_t = R_{tj}) \quad (2)$$

where $i = 0, 10, 20, 30, \dots$; $j = 0, 10, 20, 30, \dots$; the joint probability mass function has a probability value only when I and R_t are multiples of 10 and the probability values in other places are 0.

The foregoing joint probability mass function yielded a joint cumulative distribution function using:

$$F_{I, R_t}(I_i, R_{tj}) = \sum_{0}^i \sum_{0}^j P_{I, R_t}(I_i, R_{tj}). \quad (3)$$

The joint cumulative distribution function was a monotonic increasing function with a range between 0 and 1, and had the form of a three-dimensional curved surface when plotted on coordinate axes. The farther the point (I_i, R_{tj}) from the origin, the greater its probability value. The probability of a discrete rainfall group on the curved surface expressed the cumulative probability of all discrete rainfall groups, which were nearer to the origin than this discrete rainfall group (I_i, R_{tj}).

3.3. Selection of a Rainfall Probability Threshold

After establishing a joint cumulative distribution function, taking each 0.05 as an interval, we set 20 rainfall probability thresholds ranging from 0.05 to 1.00, and employed the error matrix concept to calculate the true positive rate (TPR), true negative rate (TNR), and positive predictive value (PPV) for each rainfall probability threshold at each rain gauge control area. The rainfall probability threshold was treated as the threshold of cumulative probability of the discrete rainfall groups which was used to predict whether rainfall events could trigger landslides. Here, TPR expresses the ratio of discrete rainfall groups that correctly predicted landslide occurrence to discrete rainfall groups triggering landslides actually, TNR expresses the ratio of discrete rainfall groups that correctly predicted no landslide occurrence to discrete rainfall groups triggering no landslides actually, and PPV expresses the ratio of discrete rainfall groups that correctly predicted landslide occurrence to discrete rainfall groups predicting landslides. To capture the performance of each threshold, PPV and Youden's index were used for comprehensive consideration. The higher the PPV and Youden's index values, the more accurate the rainfall probability threshold at classifying landslide occurrence and landslide occurrence for discrete rainfall groups. The TPR, TNR, PPV, and Youden's index calculations were performed employing Equations (4)–(7).

$$\text{TPR (\%)} = \frac{\text{Number of discrete rainfall groups predicting landslides when landslides actually occurred}}{\text{Number of all discrete rainfall groups triggering landslides actually}} \quad (4)$$

$$\text{TRN (\%)} = \frac{\text{Number of discrete rainfall groups predicting no landslides when no landslides occurred}}{\text{Number of all discrete rainfall groups triggering no landslides actually}} \quad (5)$$

$$\text{PPV (\%)} = \frac{\text{Number of discrete rainfall groups predicting landslides when landslides actually occurred}}{\text{Number of all discrete rainfall groups predicting landslides}} \quad (6)$$

$$\text{Youden's index} = \text{TPR} + \text{TNR} - 1 \quad (7)$$

3.4. Poisson Probability Model

A Poisson probability model relies on the past frequency of events to predict their occurrence probability in the future. The basic assumption underlying this type of model is that future events will occur with the same frequency as past events. In this model, the probability of at least one event occurring in the time interval (t) is given by Equation (8):

$$P(N(t) \geq 1) = 1 - e^{-\lambda t} \quad (8)$$

where $P(N(t) \geq 1)$ indicates the probability of at least one event occurring within a period of t years; this probability is known as the exceedance probability.

We calculated the number of discrete rainfall groups exceeding the threshold at each rain gauge in the past using the optimal rainfall probability thresholds and then divided by the years of the rainfall data to obtain the occurrence frequency (λ), which was used to calculate the exceedance probability. The exceedance probability indicated the probability of at least one rainfall event exceeding the threshold of discrete rainfall groups within any one year.

3.5. Conditional Probability

We employed the concept of conditional probability in the analysis. We first used the Poisson probability model to calculate the exceedance probability of at least one rainfall event exceeding the threshold of discrete rainfall groups within any one year at each rain gauge control area. We divided the number of landslides occurring in each slope unit by the number of rainfall events exceeding the threshold of discrete rainfall groups to estimate the probability that a landslide would occur in that slope unit when the rainfall exceeded the threshold. Lastly, we multiplied the two probabilities together to obtain the probability that a rainfall event would exceed the threshold of discrete rainfall groups and at least one landslide would also occur in each slope unit within any one year, as shown in Equation (9):

$$P(R \geq RT \cap L) = P(R \geq RT) \times P(L|R \geq RT) \quad (9)$$

where $R \geq RT$ indicates rainfall events exceed the threshold of the discrete rainfall group and L indicates the occurrence of a landslide.

4. Results and Discussion

4.1. Joint Cumulative Distribution Functions of the Rain Gauges

In this study, we collected multi-year daily rainfall data from each rain gauge and calculated the effective accumulated rainfall (R_t) by employing Equation (1), which yielded rainfall and effective accumulated rainfall for each day. We then rounded off the daily rainfall and effective accumulated rainfall values to the 10th place, which yielded discrete rainfall groups including both daily rainfall and effective accumulated rainfall. The next step was establishing frequency tables for different discrete rainfall groups, which we used to show the frequency of the discrete rainfall groups. Figure 5 shows the frequency of discrete rainfall groups at the BiHu rain gauge with daily rainfall and effective accumulated rainfall (R_t) ranging from 0 to 100 mm. The depth axis represents daily rainfall, the horizontal axis represents the effective accumulated rainfall (R_t), and the vertical axis represents the frequency of a discrete rainfall group. We then calculated the cumulative frequency of each discrete rainfall group on this basis, and this represented the frequency of all discrete rainfall groups with

values lower than that of any designated discrete rainfall group. The cumulative frequency was then divided by the total frequency of all discrete rainfall groups, which yielded the cumulative probability of each discrete rainfall group.

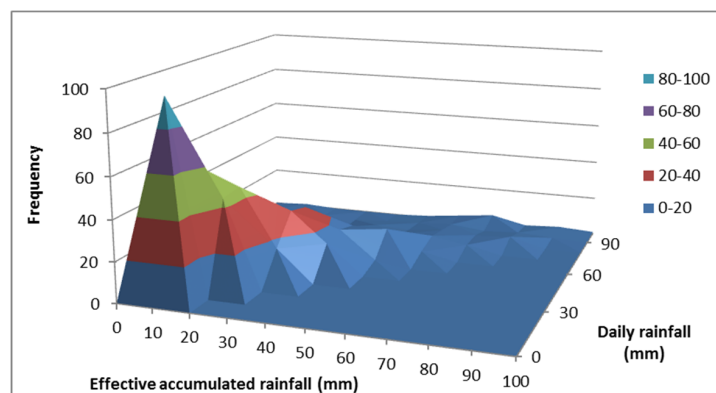


Figure 5. Frequency distribution of discrete rainfall groups at the BiHu rain gauge.

The joint cumulative distribution function of each rain gauge was then obtained from the cumulative probability of the discrete rainfall groups, and this function was used to plot a joint cumulative distribution chart. Figures 6 and 7 are joint cumulative distribution functions for the BiHu and Fushan (3) rain gauges, and daily rainfall and effective accumulated rainfall (R_t) are shown within a 0–300 mm range. The joint cumulative distribution functions have areas with gentler slopes indicating fewer and more dispersed discrete rainfall groups within a certain interval, and areas with steeper slopes indicating more and more concentrated discrete rainfall groups within a certain interval.

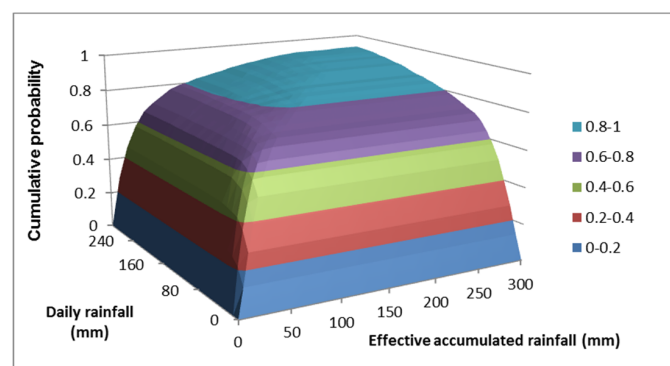


Figure 6. The joint cumulative distribution function for the BiHu rain gauge.

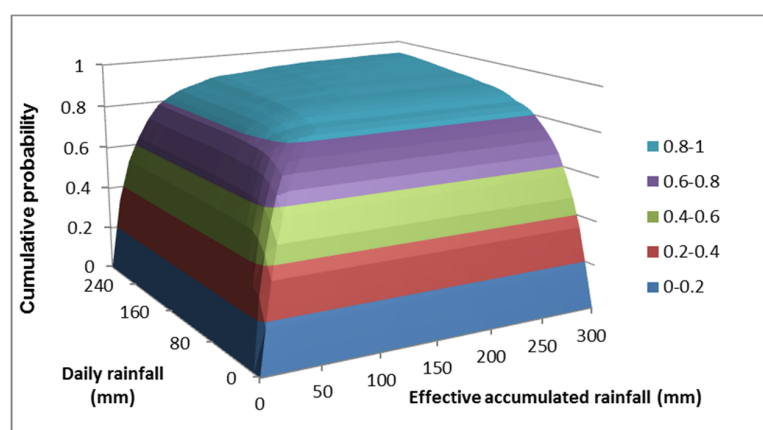


Figure 7. The joint cumulative distribution function for the Fushan (3) rain gauge.

4.2. Selection of Rainfall Probability Thresholds of the Rain Gauges

Following the analysis results of the joint cumulative distribution functions of the rain gauges, we used rainfall data from the rain gauges during the eight rainfall events triggering landslides to select rainfall probability thresholds. The rainfall probability threshold was treated as the threshold of cumulative probability of the discrete rainfall groups which was used to predict whether rainfall events could trigger landslides. Starting with a rainfall probability threshold value of 0.05, we set a rainfall probability threshold at each interval of 0.05 until a value of 1.00 was reached, and then calculated the TPR, TNR, PPV, and Youden's index of each rainfall probability threshold. Here, the number of landslide events predicted correctly divided by the number of rainfall events triggering landslides actually equaled TPR, the number of no landslide events predicted correctly divided by the number of rainfall events triggering no landslides actually equaled TNR, and the number of landslide events predicted correctly divided by the number of rainfall events predicting landslides equaled PPV. Table 2 shows the results of these calculations for the Bihu rain gauge. In the analysis results for the individual rain gauges, the rainfall probability thresholds with the highest Youden's index were within the probability interval of 0.85–0.95, and the rainfall probability thresholds with the highest PPV were at the probability of 0.95 in all cases. We consequently opted to use a rainfall probability threshold value of 0.95 for the whole area. The TPR, TNR, PPV, and Youden's index for all rain gauges when the rainfall probability threshold was 0.95 are shown in Table 3.

Table 2. True positive rate (TPR), true negative rate (TNR), positive predictive value (PPV), and Youden's index for Bihu rain gauge at different rainfall probability thresholds.

Rainfall Probability Threshold	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
Number of landslide events predicted correctly	8	8	8	8	8	8	8	8	8	8
Number of rainfall events triggering landslides actually	8	8	8	8	8	8	8	8	8	8
TPR	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Number of no landslide events predicted correctly	0	0	92	92	173	237	237	292	332	361
Number of rainfall events triggering no landslides actually	608	608	608	608	608	608	608	608	608	608
TNR	0%	0%	15%	15%	28%	39%	39%	48%	55%	59%
PPV	1.3%	1.3%	1.5%	1.5%	1.8%	2.1%	2.1%	2.5%	2.8%	3.1%
Youden's index	0%	0%	15%	15%	28%	39%	39%	48%	55%	59%
Rainfall Probability Threshold	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
Number of landslide events predicted correctly	8	8	8	8	8	8	8	8	7	0
Number of rainfall events triggering landslides actually	8	8	8	8	8	8	8	8	8	8
TPR	100%	100%	100%	100%	100%	100%	100%	100%	88%	0%
Number of no landslide events predicted correctly	407	431	469	494	526	549	575	588	602	608
Number of rainfall events triggering no landslides actually	608	608	608	608	608	608	608	608	608	608
TNR	67%	71%	77%	81%	87%	90%	95%	97%	99%	100%
PPV	3.8%	4.3%	5.4%	6.6%	8.9%	11.9%	19.5%	28.6%	53.8%	-
Youden's index	67%	71%	77%	81%	87%	90%	95%	97%	87%	0%

Table 3. TPR, TNR, PPV, and Youden's index for all rain gauges at a rainfall probability threshold of 0.95.

Rainfall Gauge	Bihu	Fushan (3)	Tatungshan	Pinglin (4)	Sihdu	Taiping	Quchi
Number of landslide events predicted correctly	7	4	8	5	8	7	7
Number of rainfall events triggering landslides actually	8	8	8	8	8	8	8
TPR	88%	50%	100%	63%	100%	88%	88%
Number of no landslide events predicted correctly	602	586	652	598	563	549	610
Number of rainfall events triggering no landslides actually	608	599	662	605	583	565	629
TNR	99%	98%	98%	99%	97%	97%	97%
PPV	53.8%	23.5%	44.4%	41.7%	28.6%	30.4%	26.9%
Youden's index	87%	48%	98%	61%	97%	85%	84%

4.3. Landslide Probability Analysis Employing a Rainfall Probability Threshold and a Long-Term Landslide Inventory

After determining a rainfall probability threshold for the rain gauges, we calculated the number of discrete rainfall groups exceeding this threshold at each rain gauge during the 1987–2016 period. We then divided these values by the years of statistics at each gauge, which yielded the λ values in Equation (8). Substituting $t = 1$ year into Equation (8) allowed us to calculate the probability of at least one rainfall event exceeding the threshold of discrete rainfall group within any one year (i.e., $P(R$

$\geq RT$) in Equation (9)) under the assumption that future rainfall conditions will be the same as past conditions. Figure 8 shows the exceedance probability value calculated for each rain gauge overlaid on each rain gauge control area. The Quchi rain gauge control area had the highest probability of 0.76502 that at least one rainfall event will exceed the threshold of the discrete rainfall group within any one year, whereas the Bihu rain gauge control area had the lowest probability of 0.43886.

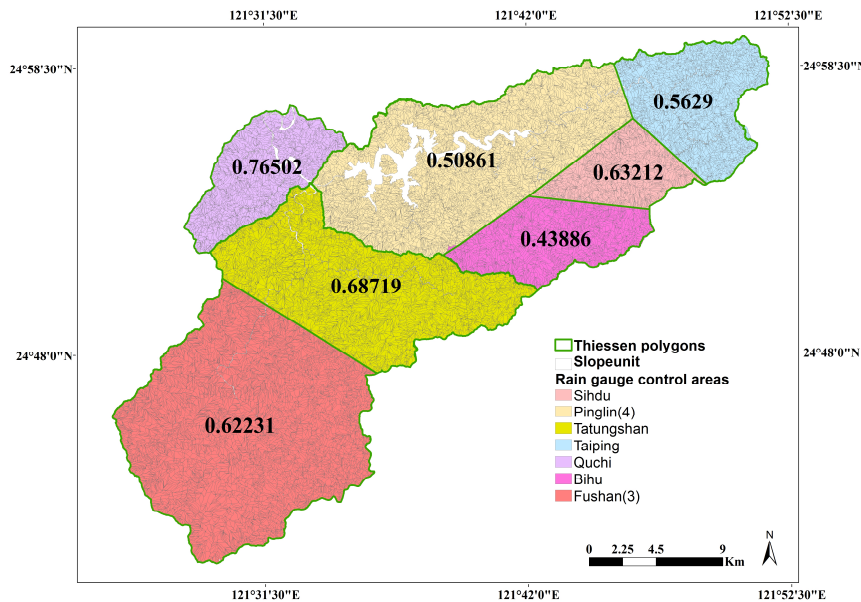


Figure 8. The exceedance probability that at least one rainfall event will exceed the threshold of discrete rainfall group within any one year in each rain gauge control area.

In this study, we also divided the number of landslides occurring in each slope unit during the 2000–2015 period by the number of rainfall events exceeding the threshold of discrete rainfall group at the rain gauges to which the slope units were assigned during the same period to estimate the landslide probability in the slope units when the rainfall exceeded the threshold, which is $P(L|R \geq RT)$ in Equation (9). The resulting probability distribution is shown in Figure 9. Figure 9 shows that the different slope units within a single rain gauge control area have different landslide probabilities, and these differences should be attributed to different geomorphological conditions in the slope units.

Lastly, employing Equation (9), we multiplied the probability $P(R \geq RT)$ that at least one rainfall event will exceed the threshold of discrete rainfall group within any one year in each rain gauge control area by the landslide probability $P(L|R \geq RT)$ in each slope unit when rainfall exceeds the threshold, which yielded the probability that at least one rainfall event exceeds the threshold of discrete rainfall group at the same time that one landslide will occur in each slope unit during the future one-year period (Figure 10). The two probability maps shown in Figures 9 and 10 were validated by the landslide inventory data respectively. The landslides were mainly distributed in the slope units where the landslide probability values were greater than 0.01. The top 2% of slope units ranked with landslide probabilities included 50.40% of slope units where landslides occurred while the top 6% of slope units ranked with landslide probabilities included 100.00% of slope units where landslides occurred in Figure 10. The results indicated these maps had reasonable landslide probability distributions. Figure 10 reveals that the Fushan (3) rain gauge control area, which is located in the southwest part of the research area, contained relatively many slope units with high landslide probability, and the highest probability value was 0.26. Apart from having fragile lithology consisting of hard sandstone and shale as well as slate, this area has a higher elevation and steeper slopes than other control areas, which suggests that elevation and slope have a definite correlation with landslide occurrence.

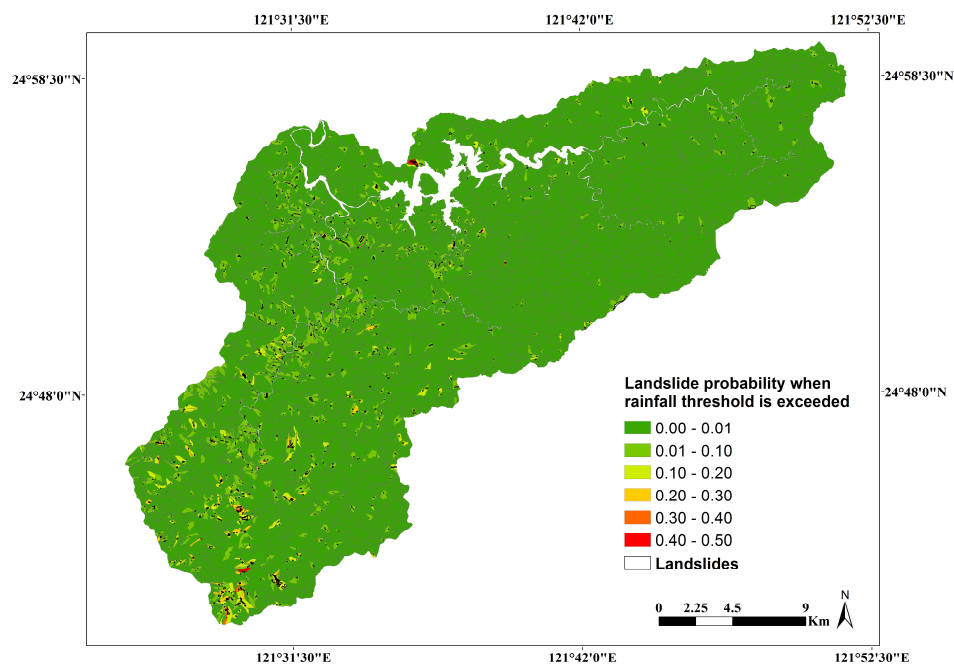


Figure 9. The landslide probability of slope units in the Taipei Water Source Domain when rainfall exceeds the threshold of discrete rainfall group.

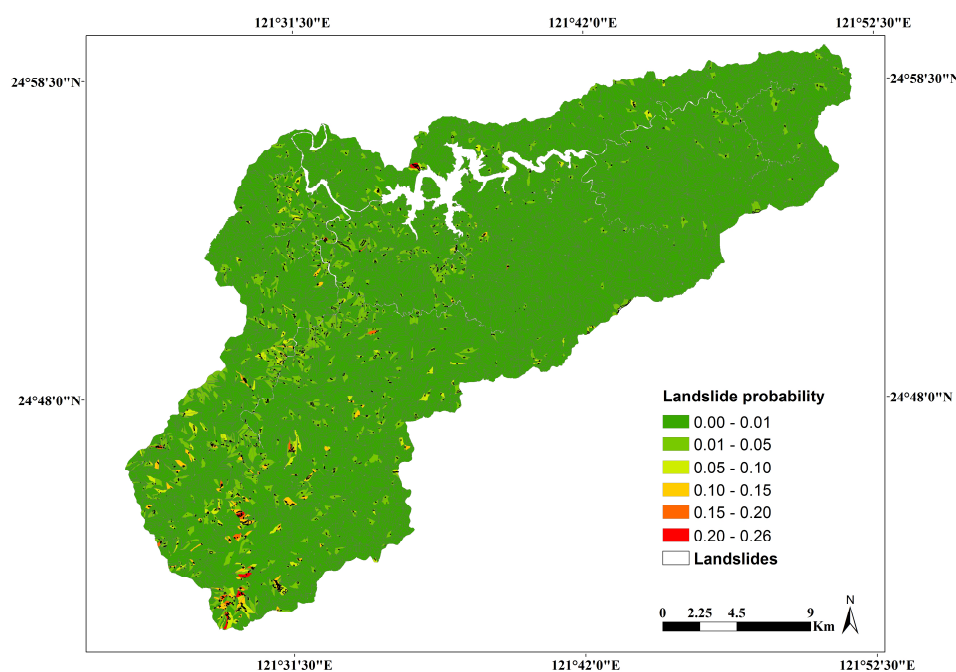


Figure 10. The probability that at least one rainfall event exceeds the threshold of discrete rainfall group and one landslide will also occur during the future one-year period within the Taipei Water Source Domain.

4.4. Discussion

In comparison with a landslide probability model based solely on the use of landslide inventories, our landslide probability model based on the use of landslide inventories and rainfall factors reflect different basic assumed conditions. The assumption of the landslide probability model incorporating rainfall factors is that the frequency of future rainfall events exceeding the threshold and the frequency of landslides occurring when the threshold has been exceeded are the same as in the past. In contrast,

the assumption of a landslide probability model based solely on the use of landslide inventories is that the frequency of future landslides occurring is the same as in the past. As a consequence, landslide probability models incorporating rainfall factors possess the following advantages: (1) This model can reflect the differences in landslide probability between the rain gauge control areas that have different rainfall conditions. (2) When rainfall data were added in the analysis, the probability model we obtained yielded more reliable results because the rainfall data were collected from a longer period (29–45 years) than the landslide inventory (16 years). (3) If we know how the probability of at least one rainfall event exceeding the threshold will change in the future, the incorporation of rainfall factors in the landslide probability model will allow the effect of possible rainfall changes on the landslide probability to be assessed.

However, several aspects connected to the application of this landslide probability model still require further investigation: (1) The method of analyzing landslide probability proposed in this study requires a long-term landslide inventory and rainfall data, therefore attention must be paid to the completeness of rainfall data for the research area and handling methods when data are incomplete. (2) Whereas the rainfall factors used in this study reflect daily rainfall and effective accumulated rainfall, the use of different rainfall factors will yield different analysis results, which may be explored further in future research. (3) Apart from the modified Thiessen polygon method, the division of rain gauge control areas can be performed using other methods, such as the height–balance polygon method. A finer division method should yield more precise results of a landslide probability distribution, therefore future research can also compare the applicability of different methods of division into rain gauge control areas. (4) We obtained a long-term landslide inventory consisting of only eight events, therefore all events collected were used in the process of building the model. The landslide inventory covering the period of other events may be collected to verify the predictive ability of this landslide probability model.

5. Conclusions

In this study, we employed joint cumulative distribution functions to calculate the TPR, TNR, PPV, and Youden's index for different rainfall probability thresholds, selected a threshold of 0.95 as suitable for the research area, and used this rainfall probability threshold to calculate the Poisson probability of at least one rainfall event exceeding the threshold of discrete rainfall groups at each rain gauge within the future one-year period. We then combined this probability with the landslide probability in individual slope units when rainfall exceeded the threshold value, which allowed us to estimate the probability that a landslide will occur in individual slope units during the future one-year period. Many of the slope units with a high landslide probability are located in the Fushan (3) rain gauge control area, and the highest probability is 0.26. Apart from fragile lithology, this area is characterized by high elevations and steep slopes, which indicates that the elevation and slope have a significant influence on the occurrence of landslides. This finding suggests that this area should be a focal area for landslide prevention and mitigation efforts.

The landslide probability model established based on the use of a long-term landslide inventory and rainfall factor had a finer spatial resolution and data for a longer period, which yielded more reliable results and enabled the effect of possible rainfall changes on the landslide probability to be assessed. The effects of the completeness of rainfall data for the research area, the use of different rainfall factors, as well as the different methods of division into rain gauge control areas on the landslide probability analysis results can be significant and still require further investigation.

Author Contributions: Conceptualization, C.-Y.W.; data curation, Y.-C.Y.; formal analysis, Y.-C.Y.; methodology, C.-Y.W. and Y.-C.Y.; supervision, C.-Y.W.; visualization, Y.-C.Y.; writing—original draft, C.-Y.W. and Y.-C.Y.; writing—review & editing, C.-Y.W.; funding acquisition, C.-Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and Technology, Taiwan (MOST 107-2311-B-005-001).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Varnes, D.J. *IAEG Landslide Hazard Zonation: A Review of Principles and Practice (No.3)*; United Nations: New York, NY, USA, 1984.
2. Carrara, A. Multivariate models for landslide hazard evaluation. *J. Int. Assoc. Math. Geol.* **1983**, *15*, 403–426. [[CrossRef](#)]
3. Chung, C.J.F.; Fabbri, A.G. The representation of geoscience information for data integration. *Nonrenew. Resour.* **1993**, *2*, 122–139. [[CrossRef](#)]
4. Baeza, C.; Corominas, J. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surf. Process. Landf.* **2001**, *26*, 1251–1263. [[CrossRef](#)]
5. Guzzetti, F.; Galli, M.; Reichenbach, P.; Ardizzone, F.; Cardinali, M. Landslide hazard assessment in the Collazzone area, Umbria, Central Italy. *Nat. Hazards Earth Syst. Sci.* **2006**, *6*, 115–131. [[CrossRef](#)]
6. Zêzere, J.L.; Oliveira, S.C.; Garcia, R.A.C.; 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* **2007**, *4*, 123–136. [[CrossRef](#)]
7. Carrara, A.; Crosta, G.; Frattini, P. Comparing models of debris-flow susceptibility in the alpine environment. *Geomorphology* **2008**, *94*, 353–378. [[CrossRef](#)]
8. Lee, C.T.; Huang, C.C.; Lee, J.F.; Pan, K.L.; Lin, M.L.; Dong, J.J. Statistical approach to storm event-induced landslides susceptibility. *Nat. Hazards Earth Syst. Sci.* **2008**, *8*, 941–960. [[CrossRef](#)]
9. Rossi, M.; Guzzetti, F.; Reichenbach, P.; Mondini, A.C.; Peruccacci, S. Optimal landslide susceptibility zonation based on multiple forecasts. *Geomorphology* **2010**, *114*, 129–142. [[CrossRef](#)]
10. Nefeslioglu, H.A.; Gokceoglu, C. Probabilistic risk assessment in medium scale for rainfall-induced earthflows: Catakli catchment area (Cayeli, Rize, Turkey). *Math. Probl. Eng.* **2011**, *2011*, 1–21. [[CrossRef](#)]
11. Chan, H.C.; Chen, P.A.; Lee, J.T. Rainfall-induced landslide susceptibility using a rainfall–runoff model and logistic regression. *Water* **2018**, *10*, 1354. [[CrossRef](#)]
12. Roccati, A.; Faccini, F.; Luino, F.; Ciampalini, A.; Turconi, L. Heavy rainfall triggering shallow landslides: A susceptibility assessment by a GIS-approach in a Ligurian Apennine catchment (Italy). *Water* **2019**, *11*, 605. [[CrossRef](#)]
13. Corominas, J.; van Westen, C.; Frattini, P.; Cascini, L.; Malet, J.P.; Fotopoulou, S.; Catani, F.; Van Den Eeckhaut, M.; Mavrouli, O.; Agliardi, F.; et al. Recommendations for the quantitative analysis of landslide risk. *Bull. Eng. Geol. Environ.* **2014**, *73*, 209–263. [[CrossRef](#)]
14. Guzzetti, F.; Reichenbach, P.; Cardinali, M.; Galli, M.; Ardizzone, F. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology* **2005**, *72*, 272–299. [[CrossRef](#)]
15. Jaiswal, P.; van Westen, C.J.; Jetten, V. Quantitative landslide hazard assessment along a transportation corridor in southern India. *Eng. Geol.* **2010**, *116*, 236–250. [[CrossRef](#)]
16. Das, I.; Stein, A.; Kerle, N.; Dadhwal, V.K. Probabilistic landslide hazard assessment using homogeneous susceptible units (HSU) along a national highway corridor in the northern Himalayas, India. *Landslides* **2011**, *8*, 293–308. [[CrossRef](#)]
17. Ghosh, S.; van Westen, C.J.; Carranza, E.J.M.; Jetten, V.G.; Cardinali, M.; Rossi, M.; Guzzetti, F. Generating event-based landslide maps in a data-scarce Himalayan environment for estimating temporal and magnitude probabilities. *Eng. Geol.* **2012**, *128*, 49–62. [[CrossRef](#)]
18. Segoni, S.; Tofani, V.; Rosi, A.; Catani, F.; Casagli, N. Combination of Rainfall Thresholds and Susceptibility Maps for Dynamic Landslide Hazard Assessment at Regional Scale. *Front. Earth Sci.* **2018**, *6*, 85. [[CrossRef](#)]
19. Salvatici, T.; Tofani, V.; Rossi, G.; D’Ambrosio, M.; Stefanelli, C.T.; Masi, E.B.; Rosi, A.; Pazzi, V.; Vannocci, P.; Petrolo, M.; et al. Application of a physically based model to forecast shallow landslides at a regional scale. *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 1919–1935. [[CrossRef](#)]
20. Rosi, A.; Canavesi, V.; Segoni, S.; Nery, T.D.; Catani, F.; Casagli, N. Landslides in the Mountain Region of Rio de Janeiro: A Proposal for the Semi-Automated Definition of Multiple Rainfall Thresholds. *Geosciences* **2019**, *9*, 203. [[CrossRef](#)]
21. Segoni, S.; Piciullo, L.; Gariano, S.L. A review of the recent literature on rainfall thresholds for landslide occurrence. *Landslides* **2018**, *15*, 1483–1501. [[CrossRef](#)]

22. Vasu, N.N.; Lee, S.R.; Pradhan, A.M.S.; Kim, Y.T.; Kang, S.H.; Lee, D.H. A new approach to temporal modelling for landslide hazard assessment using an extreme rainfall induced-landslide index. *Eng. Geol.* **2016**, *215*, 36–49. [\[CrossRef\]](#)
23. Lee, C.T. Multi-Stage Statistical Landslide Hazard Analysis: Earthquake-Induced Landslides. In *Landslide Science for a Safer Geoenvironment, Proceedings of the International Consortium on Landslides (ICL) Third Landslide Forum, Beijing, China, 2–6 June 2014*; Springer: Cham, Switzerland, 2014.
24. Lee, C.T.; Chung, C.C. Common Patterns Among Different Landslide Susceptibility Models of the Same Region. In *Advancing Culture of Living with Landslides, Proceedings of the 4th World Landslide Forum, Ljubljana, Slovenia, 29 May–2 June 2017*; Springer: Cham, Switzerland, 2017.
25. Önöz, B.; Bayazit, M. Effect of the occurrence process of the peaks over threshold on the flood estimates. *J. Hydrol.* **2001**, *244*, 86–96. [\[CrossRef\]](#)
26. Jaiswal, P.; van Westen, C.J. Estimating temporal probability for landslide initiation along transportation routes based on rainfall thresholds. *Geomorphology* **2009**, *112*, 96–105. [\[CrossRef\]](#)
27. Wu, C.Y.; Chen, S.C. Integrating spatial, temporal, and size probabilities for the annual landslide hazard maps in the Shihmen watershed, Taiwan. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 2353–2367. [\[CrossRef\]](#)
28. Althuwaynee, O.F.; Pradhan, B.; Ahmad, N. Estimation of rainfall threshold and its use in landslide hazard mapping of Kuala Lumpur metropolitan and surrounding areas. *Landslides* **2015**, *12*, 861–875. [\[CrossRef\]](#)
29. Afungang, R.N.; Bateira, C.V. Temporal probability analysis of landslides triggered by intense rainfall in the Bamenda Mountain Region, Cameroon. *Environ. Earth Sci.* **2016**, *75*, 1032. [\[CrossRef\]](#)
30. Sangelantoni, L.; Gioia, E.; Marincioni, F. Impact of climate change on landslides frequency: The Esino river basin case study (Central Italy). *Nat. Hazards* **2018**, *93*, 849–884. [\[CrossRef\]](#)
31. Endo, T. *Probable Distribution of the Amount of Rainfall Causing Landslides*; Annual Report 1968; Hokkaido Branch: Sapporo, Japan, 1969; pp. 122–136.
32. Onodera, T.; Yoshinaka, R.; Kazama, H. Slope failures caused by heavy rainfall in Japan. *J. Jpn. Soc. Eng. Geol.* **1974**, *15*, 191–200. [\[CrossRef\]](#)
33. Campbell, R.H. *Soil Slip, Debris Flows, and Rainstorms in the Santa Monica Mountains and Vicinity, Southern California*; US Government Printing Office: Washington, DC, USA, 1975; Volume 851, p. 51.
34. Aleotti, P. A warning system for rainfall-induced shallow failures. *Eng. Geol.* **2004**, *73*, 247–265. [\[CrossRef\]](#)
35. Caine, N. The rainfall intensity-duration control of shallow landslides and debris flows. *Geogr. Ann. Ser. A Phys. Geogr.* **1980**, *62*, 23–27.
36. Guzzetti, F.; Peruccacci, S.; Rossi, M.; Stark, C.P. Rainfall thresholds for the initiation of landslides in central and southern Europe. *Meteorol. Atmos. Phys.* **2007**, *98*, 239–267. [\[CrossRef\]](#)
37. Cannon, S.H.; Gartner, J.E.; Wilson, R.C.; Bowers, J.C.; Laber, J.L. Storm rainfall conditions for floods and debris flows from recently burned areas in southwestern Colorado and southern California. *Geomorphology* **2008**, *96*, 250–269. [\[CrossRef\]](#)
38. Brunetti, M.T.; Peruccacci, S.; Rossi, M.; Luciani, S.; Valigi, D.; Guzzetti, F. Rainfall thresholds for the possible occurrence of landslides in Italy. *Nat. Hazards Earth Syst. Sci.* **2010**, *10*, 447–458. [\[CrossRef\]](#)
39. Ciervo, F.; Rianna, G.; Mercogliano, P.; Papa, M.N. Effects of climate change on shallow landslides in a small coastal catchment in southern Italy. *Landslides* **2017**, *14*, 1043–1055. [\[CrossRef\]](#)
40. Tien Bui, D.; Pradhan, B.; Lofman, O.; Revhaug, I.; Dick, Ø.B. Regional prediction of landslide hazard using probability analysis of intense rainfall in the Hoa Binh province, Vietnam. *Nat. Hazards* **2013**, *66*, 707–730. [\[CrossRef\]](#)
41. Peruccacci, S.; Brunetti, M.T.; Luciani, S.; Vennari, C.; Guzzetti, F. Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in central Italy. *Geomorphology* **2012**, *139*, 79–90. [\[CrossRef\]](#)
42. Crozier, M.J. *Landslides: Causes, Consequences & Environment*; Taylor & Francis: Oxfordshire, UK, 1986.
43. Wieczorek, G.F. *Landslides: Investigation and Mitigation. Chapter 4-Landslide Triggering Mechanisms*; Special Report 247; Turner, A.K., Schuster, R.L., Eds.; Transportation Research Board: Washington, DC, USA, 1996.
44. Lee, S.; Won, J.S.; Jeon, S.W.; Park, I.; Lee, M.J. Spatial landslide hazard prediction using rainfall probability and a logistic regression model. *Math. Geosci.* **2015**, *47*, 565–589. [\[CrossRef\]](#)
45. Saadatkhah, N.; Kassim, A.; Lee, L.M. Hulu Kelang, Malaysia regional mapping of rainfall-induced landslides using TRIGRS model. *Arab. J. Geosci.* **2015**, *8*, 3183–3194. [\[CrossRef\]](#)
46. Guzzetti, F. *Landslide Hazard and Risk Assessment*. Ph.D. Thesis, Universitäts und Landesbibliothek Bonn, Bonn, Germany, 2006.

47. Van Den Eeckhaut, M.; Reichenbach, P.; Guzzetti, F.; Rossi, M.; 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.* **2009**, *9*, 507–521. [[CrossRef](#)]
48. Lin, C.W.; Lin, M.L.; Chang, C.P.; Wu, M.C.; Wang, T.T.; Chen, T.C. *Geohazards Susceptibility Analysis of Damaged Areas in Typhoon Morakot (1/3)*; Central Geological Survey MOEA: Taipei, Taiwan, 2010. (In Chinese)
49. Wu, C.Y.; Tsai, C.W.; Chen, S.C. Topographic characteristic analysis of landslides in Kaoping River Watershed. *J. Chin. Soil Water Conserv.* **2016**, *47*, 156–164. (In Chinese)
50. Liu, J.K.; Weng, T.C.; Hung, C.H.; Yang, M.T. Remote Sensing Analysis of Heavy Rainfall Induced Landslide. In Proceedings of the 21st Century Civil Engineering Technology and Management Conference, Hsinchu, Taiwan, 28 December 2001; Minghsin University of Science and Technology: Xinfeng Township, Taiwan, 2001; pp. C21–C31. (In Chinese).
51. Xie, M.; Esaki, T.; Zhou, G. GIS-based probabilistic mapping of landslide hazard using a three-dimensional deterministic model. *Nat. Hazards* **2004**, *33*, 265–282. [[CrossRef](#)]
52. Jan, C.D. *Using Rainfall Factors to Determine Debris-Flow Warning Criteria*; Soil and Water Conservation Bureau, Council of Agriculture, Executive Yuan: Tainan, Taiwan, 2003. (In Chinese)
53. Tsai, M.C. Establishment of Critical Line for Early Warning of Debris Flow Based on Two-Factors Survival Analysis—A Case Study in Sheng-Mu Area. Ph.D. Thesis, Feng Chia University, Taichung, Taiwan, 2 June 2016. (In Chinese).
54. Song, W.M. *Principle of Probability and Inferential Statistics*; McGraw-Hill: Taipei, Taiwan, 2007. (In Chinese)



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).