Hindawi
Advances in Civil Engineering
Volume 2024, Article ID 6554818, 11 pages
https://doi.org/10.1155/2024/6554818

Research Article

Copula-Based Probabilistic Hazard Assessment Model for Debris Flow Considering the Uncertainties of Multiple Influencing Factors

Mi Tian,[1,2] **Yuan Shen**,[1] **Long Fan**,[3] **and Xiao-Tao Sheng**[4]

[1]*School of Civil Engineering, Architecture and Environment, Hubei University of Technology, 28 Nanli Road, Wuhan 430068, China*
[2]*Innovation Demonstration Base of Ecological Environment Geotechnical and Ecological Restoration of Rivers and Lakes, Hubei University of Technology, 28 Nanli Road, Wuhan 430068, China*
[3]*Wuhan Maritime Communications Research Institute, 312 Luoyu Road, Wuhan 430070, China*
[4]*The Institute of Smart Water, Wuhan University, 8 Donghu South Road, Wuhan 430072, China*

Correspondence should be addressed to Mi Tian; mitianhbut@hbut.edu.cn

Received 17 July 2023; Revised 4 November 2023; Accepted 21 November 2023; Published 8 January 2024

Academic Editor: Guang-Liang Feng

This paper proposes a probabilistic hazard assessment model for debris flows considering the uncertainties of multiple influencing factors based on copula approaches. Fifty-nine rainfall-induced debris flows occurred between 2001 and 2009 in Taiwan are taken as an illustrative example to validate the proposed approaches. A copula-based probabilistic model is developed to model the joint probability distribution of debris-flow volume V and its influencing factors (e.g., rainfall intensity, RI and landslide area, A_L). The developed model is then used to make probabilistic prediction of debris-flow volume for a specific hazard level, and compared with the empirical approaches. The proposed probabilistic model is also used to develop the exceedance probability charts of quantities for a specific debris-flow basin. Results show that the developed V–RI–A_L probabilistic model can provide reasonable estimates of debris-flow volume in Taiwan for a specific probability level of 0.94, and show better predictive performance than the empirical relationships by using an independent debris-flow dataset in Taiwan for validation. The developed multivariate joint probabilistic model can also provide the exceedance probability of debris flows through considering the uncertainties of debris flow and its influencing factors, providing a preliminary reference for hazard assessment of the debris flows.

1. Introduction

Debris flows are characterized with large discharge, high velocity, and enormous destructive power that can usually cause devastating damages to the buildings, infrastructures, local residents in the downstream [1, 2]. To reduce the damages of debris flows, performing accurate hazard assessment of debris flows is necessary for risk management and remedial measure designs [3–5]. Hazard assessment of debris flows generally includes the quantitative estimation of the most important parameters (e.g., the debris-flow volume, mean flow velocity, and runout distance), and determination of the probability that a debris-flow event can occur in a specific debris-flow basin [6, 7]. Among the most important debris-flow parameters, the debris-flow volume, V, is one of the most important parameters for potential debris-flow hazard assessment, which is a prerequisite for predicting the peak discharge and runout distance of debris flow [8–10]. Meanwhile, the probability level of debris flows can be quantitatively determined by a probabilistic model of debris-flow volume [11, 12]. However, due to various uncertainties and variabilities in debris flows, it is difficult to accurately estimate the volume of debris flow and its corresponding probability of occurrence.

Various attempts have been made to estimate the debris-flow volume for a specific debris-flow basin, including theoretical and empirical approaches. Theoretical methods are physically based models to simulate the dynamic processes of debris flows [13–16]. Theoretical approaches generally have the limitation of selecting appropriate rheological parameters [17]. Due to the scarcity of rheological parameter data and size effects in laboratory tests, the rheological parameters are associated with great uncertainties. Empirical approaches usually provide a simple and useful tool for its convenience, which have been widely used in estimating the debris-flow volume. Most empirical equations are developed by relating the debris-flow volume to the basin area of a specific debris-flow basin [18–21]. Some empirical equations were also proposed to consider more factors (e.g., geological and rainfall parameters) to make more accurate estimation of debris-flow volume [22–24]. For example, Gartner et al. [20] developed a multivariate nonlinear empirical equation for predicting the debris-flow volume by considering the basin area and the total storm rainfall. Chang et al. [24] proposed an empirical relationship by including geological, topographic, rainfall parameters. However, there are various unavoidable uncertainties and variabilities that exist in debris flows [25–27]. For example, due to the sedimental production, erosion, and deposition, the debris-flow volume for a specific basin may change over the time. Moreover, the topography of a particular debris-flow basin also changes over time, affecting the propagation of debris flow [28]. These influencing factors are also associated with various uncertainties, which are usually ignored in estimating the debris-flow volume. Thus, it is necessary to take various uncertainties into account in the hazard assessment of debris flows.

This paper proposes probabilistic hazard assessment models for debris flow considering the uncertainties of multiple influencing factors based on the copula approaches. Probabilistic analyses are conducted on the 59 past rainfall debris-flow events in Taiwan. First, 59 datasets of past debris flows are divided in to 50 sets of data for model construction and 9 sets of data for validation. A multivariate copula model is developed based on the 50 sets of observation data to model the joint probability distribution of debris-flow volume *V* and its influencing factors (e.g., rainfall intensity, *RI* and landslide area, A_L). Then, the developed V–RI–A_L copula model in Taiwan is used to make probabilistic prediction of the debris-flow volume for a specific hazard level, and compared with the empirical approaches. Finally, the proposed probabilistic model is validated by the independent nine sets of data. The proposed probabilistic model is also used to develop the exceedance probability charts of quantities (e.g., the debris-flow volume, *V* and rainfall intensity, *RI*) considering a given landslide area, A_L for a specific debris flow. This paper successfully characterizes the high uncertainties and variabilities of debris-flow volume and its influencing factors, and provide a preliminary reference for debris-flow risk assessment and control measure design.

2. Methodology

2.1. Definition of Multivariate Copula Model. According to Sklar's [29] theorem, the basic idea of copula theory is to decompose a joint probability function into the modeling of dependence and the modeling of multiple marginal distributions. Let $X_1, X_2, \ldots, X_n$ $(n > 2)$ denote debris-flow parameters. The multivariate joint cumulative distribution function (CDF) of $X_1, X_2, \ldots, X_n$ is denoted as $F_{1,2,\cdots,n}(x_1, x_2, \cdots, x_n)$. Then, $F_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n)$ can be presented as follows [30, 31]:

$$\begin{aligned} F_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n) &= P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n) \\ &= C[F_1(x_1), F_2(x_2), \ldots, F_n(x_n); \theta] \\ &= C(u_1, u_2, \ldots, u_n; \theta) \end{aligned}, \tag{1}$$

where $P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n)$ is the cumulative probability given $X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n$; $C(u_1, u_2, \ldots, u_n; \theta)$ is the multivariate copula function with the copula parameter θ; $u_1 = F_1(x_1), u_2 = F_2(x_2), \ldots, u_n = F_n(x_n)$ are the marginal CDFs of $X_1, X_2, \ldots, X_n$, respectively. The most appropriate marginal distributions of $X_1, X_2, \ldots, X_n$ can be determined by Akaike information criterion (AIC) based on the observation data among various candidate distributions (e.g., Truncated Normal, Lognormal, Weibull, and Truncated Gumbel distributions). From Equation (1), the joint probability-density function (PDF) of $X_1, X_2, \ldots, X_n, f_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n)$ is written as follows:

$$\begin{aligned} f_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n) &= \frac{\partial^n C[F_1(x_1), F_2(x_2), \ldots, F_n(x_n); \theta]}{\partial F_1(x_1) \partial F_2(x_2) \ldots \partial F_n(x_n)} \prod_{i=1}^{n} \frac{\partial F_i(x_i)}{\partial x_i} \\ &= c[F_1(x_1), F_2(x_2), \ldots, F_n(x_n); \theta] \prod_{i=1}^{n} f_i(x_i) \\ &= c(u_1, u_2, \ldots, u_n; \theta) \prod_{i=1}^{n} f_i(x_i), \end{aligned} \tag{2}$$

where $f_1(x_1), f_2(x_2), \ldots, f_n(x_n)$ are the marginal PDFs of $X_1, X_2 \ldots, X_n$, respectively; $c(u_1, u_2, \ldots, u_n; \theta)$ is the multivariate copula density function associated with the copula function $C(u_1, u_2, \ldots, u_n; \theta)$, which is given by:

$$c(u_1, u_2, \ldots, u_n; \theta) = \partial^3 C(u_1, u_2, \ldots, u_n; \theta) / \partial u_1 \partial u_2 \ldots \partial u_n. \tag{3}$$

Multivariate copula functions, such as elliptical copulas (e.g., Gasussian copula) and Archimedean copulas (e.g., Frank, Clayton, and Clayton copulas) have *n*-dimensional generalizations, which can be used in Equation (3) to construct the joint distribution of multivariate parameters. The copula parameter θ can be estimated by the maximum likelihood estimation (MLE) based on the observation data $\{x_{1j}, x_{2j}, \ldots, x_{nj}, j = 1, 2, \ldots, N\}$ with a sample size of *N*, which is derived as follows:

$$L(\theta) = \sum_{j=1}^{N} \ln c(u_{1j}, u_{2j}, \ldots, u_{nj}; \theta), \tag{4}$$

where $(u_{1j}, u_{2j}, \ldots, u_{nj})$ are the empirical distributions of observation data $(x_{1j}, x_{2j}, \ldots, x_{nj})$, which are defined as follows [30, 31]:

$$\begin{cases} u_{1j} = \dfrac{\text{rank}(x_{1j})}{N+1}, j = 1, 2, \ldots, N \\ u_{2j} = \dfrac{\text{rank}(x_{2j})}{N+1}, j = 1, 2, \ldots, N \\ \vdots \\ u_{nj} = \dfrac{\text{rank}(x_{nj})}{N+1}, j = 1, 2, \ldots, N \end{cases}, \quad (5)$$

where $\text{rank}(\cdot)$ is the ranking function. For example, $\text{rank}(x_{1j})$ denotes the rank of x_{1j} among the list $\boldsymbol{x}_1 = (x_{11}, x_{12}, \ldots, x_{1N})$ in an ascending order. Then, the copula parameter can be obtained by maximizing the likelihood function $L(\theta)$, which can be expressed as follows:

$$\widehat{\theta} = \arg\max L(\theta) = \arg\max \sum_{j=1}^{N} \ln c(u_{1j}, u_{2j}, \ldots, u_{nj}; \theta). \quad (6)$$

The best-fit copula among the set of candidate copulas can be selected by the AIC score, which is given by [28]:

$$\text{AIC} = N \ln \left\{ \frac{1}{N-k} \sum_{j=1}^{N} \left[F_m(x_{1j}, x_{2j}, \ldots, x_{nj}) - C(u_{1j}, u_{2j}, \ldots, u_{nj}; \theta) \right]^2 \right\} + 2k, \quad (7)$$

where $F_m(x_{1j}, x_{2j}, \ldots, x_{nj}) = P(X_1 < x_{1j}, X_2 < x_{2j}, \ldots, X_n < x_{nj}) = \frac{1}{N+1} \sum_{l=1}^{N} \sum_{m=1}^{N} \sum_{n=1}^{N} N_{l,m,n}$ is the measured joint probability of multivariate parameters (i.e., $X_1, X_2, \ldots, X_n$); $C(u_{1j}, u_{2j}, \ldots, u_{nj}; \theta)$ is the candidate copula; k is the number of copula parameters for the candidate copulas.

2.2. Construction of Multivariate Copula Model. The n-dimensional multivariate copula model in Equation (1) can be established based on the multivariate observation data $\{x_{1j}, x_{2j}, \ldots, x_{nj}, j = 1, 2, \ldots, N\}$ for $X_1, X_2 \ldots, X_n$. The construction of multivariate copula model generally includes four steps:

(1) Determine the most appropriate marginal PDFs $f_1(x_1), f_2(x_2), \ldots, f_n(x_n)$ of $X_1, X_2 \ldots, X_n$ among various candidate distributions (e.g., Truncated Normal, Lognormal, Weibull, and Truncated Gumbel distributions) by AIC;
(2) Determine the copula parameter θ for a candidate copula function by MLE (i.e., Equation (6)). Note that the measured data $\{x_{1j}, x_{2j}, \ldots, x_{nj}, j = 1, 2, \ldots, N\}$ are transformed to the empirical distributions $(u_{1j}, u_{2j}, \ldots, u_{nj})$ using Equation (5);
(3) Identify the best-fit multivariate copula model for $X_1, X_2 \ldots, X_n$ by using Equation (7);
(4) Develop the joint PDF $f_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n)$ of $X_1, X_2 \ldots, X_n$ by using Equation (2), based on the marginal PDFs of $X_1, X_2 \ldots, X_n$ and the best-fit multivariate copula model.

3. Data Sources and Marginal Probability Models

3.1. Data Sources of Rainfall-Induced Debris Flows. This paper collected 59 datasets of rainfall debris-flow events occurred between 2001 and 2009 in Taiwan from Chang et al. [24]. These 59 debris-flow basins were affected by frequent earthquakes, which often induced a large number of landslides. A large amount of loose material deposited in the debris-flow basins, making it easy to form debris flows after intense rainfall. Chang et al. [24] collected physiographical parameters, geological index, and rainfall factors. Considering that the debris flows were mainly triggered by the intensive rainfall in the study area, thus, the rainfall intensity, *RI*, and landslide area, A_L are taken as the main influencing factors. Fifty-nine datasets of past debris flows are divided into 50 sets of data for model construction and 9 sets of data for validation, as shown in Table 1. More details of the geological conditions and debris flows in the study area are refered to the study by Chang et al. [24].

Table 1 summarizes the statistical characteristics of *RI*, A_L, and V of 50 datasets of debris flows in Taiwan. It shows that the mean value of A_L in the study area is 263,785.70 m^2, with a coefficient of variation (COV) of 2.37. Even with the same rainfall conditions in the study area, the landside area for different debris-flow basins vary by many orders of magnitude. The mean value of V is 97,905.920 m^3 with a COV of 1.37. Compared with the variability of *RI* with a COV of 0.28, A_L and V show great variability. Considering that there are the high uncertainties in the debris-flow volume and its influencing factors, *RI*, A_L, and V are considered as random variables, where $X_1 = RI$, $X_2 = A_L$, and $X_3 = V$.

The Kendall rank correlation coefficient is used to illustrate the cross-correlation between the debris-flow volume and its influencing factors (i.e., *RI*, A_L, and V), as shown in Table 2. It is found that these three variables are positively correlated. It can be seen that there is a strong positive correlation between A_L and V with Kendall rank correlation coefficient of 0.603. Such a strong correlation between A_L and V may implies that more loose material can be transported downstream to form larger debris flows. RI–A_L and RI–V are also positively correlated with the Kendall rank correlation coefficients of 0.180 and 0.181, respectively.

Considering that *RI*, A_L, and V are positively correlated variables, three-dimensional copula density functions are used to construct the joint probability model of *RI*, A_L, and V. Four three-dimensional Archimedean copula functions, namely

Table 1: Fifty-nine datasets of debris flows in Taiwan and the statistics of *RI*, A_L, and *V*.

No.	*RI* (mm)	A_L (m^2)	*V* (m^3)	No.	*RI* (mm)	A_L (m^2)	*V* (m^3)
50 Sets of data for model construction							
1	25.85	1,518,000	455,500	26	19.13	19,440	15,000
2	26.77	1,326,000	342,300	27	14.15	38,472	22,400
3	30.85	595,000	236,000	28	15.09	90,720	50,000
4	30.97	588,000	155,500	29	19.25	6,370	52,800
5	25.75	144,000	99,400	30	28.84	30,000	5,250
6	31.09	84,000	157,700	31	26.67	2,530	12,600
7	27.6	614,000	604,606	32	34.14	10,000	6,000
8	36.56	3,884,000	402,585	33	29.28	700	2,600
9	12.75	16,100	8,400	34	22.62	13,400	9,200
10	11.81	1680	5,200	35	24.5	8,570	50,000
11	16.11	1624	4,830	36	10.56	11,000	75,000
12	13.4	3,520	2,700	37	23.45	1,005,000	326,000
13	23.04	600	2,700	38	27.3	20,800	45,000
14	18.71	104,593	79,400	39	27.42	28,400	15,000
15	19.68	9,410	15,369	40	27.42	236,900	150,000
16	19.7	75,620	87,000	41	26.8	25,700	30,000
17	19.67	231,600	180,000	42	16.66	485,400	70,000
18	21.98	30,561	20,900	43	20.17	26,000	6,000
19	19.17	9,490	8,000	44	21.2	660,450	50,000
20	10.12	472,872	269,500	45	18.75	18,520	30,000
21	11.92	8,000	7,000	46	21.47	20,900	30,000
22	17.86	493,923	256,000	47	21.43	116,700	15,000
23	19.22	5,700	19,428	48	21.89	31,400	68,000
24	31.65	50,400	200,000	49	22.25	6,500	100,000
25	18.82	1,120	19,428	50	22.74	5,600	20,000
Mean	22.08	263,785.70	97905.920	Max	111	3,884,000	604,606
COV	0.28	2.37	1.37	Min	15	600	2,600
9 Sets of data for validation							
51	10.47	112,200	49,140	56	22.76	2,080	3,500
52	18.26	199,100	46,800	57	26.88	94,390	120,000
53	16.23	7,320	9,000	58	19.84	96,600	9,000
54	18.69	11,410	40,000	59	14.44	21,700	10,000
55	23.06	7,960	9,000	—	—	—	—

Table 2: Kendall rank correlation coefficients of *RI*, A_L, and *V*.

Statistics	X_1	X_2	X_3
$X_1 = RI$	1.000	0.180	0.181
$X_2 = A_L$	0.180	1.000	0.603
$X_3 = V$	0.181	0.603	1.000

Table 3: Three-dimensional Archimedean copula functions.

Copula	$C(u_1, u_2, u_3; \theta)$	Range of θ
Clayton	$\max[(u_1^{-\theta} + u_2^{-\theta} + u_3^{-\theta} - 2)^{-\frac{1}{\theta}}, 0]$	$[-1, \infty)\backslash\{0\}$
Frank	$-\frac{1}{\theta}\ln\left[1 + \frac{(e^{-\theta u_1}-1)(e^{-\theta u_2}-1)(e^{-\theta u_3}-1)}{(e^{-\theta}-1)^2}\right]$	$(-\infty, \infty)\backslash\{0\}$
Ali–Mikhail–Haq	$\frac{u_1 u_2 u_3}{1-\theta(1-u_1)(1-u_2)(1-u_3)}$	$[-1, 1)$
Gumbel–Hougaard	$e^{-[(-\ln u_1)^{\theta}+(-\ln u_2)^{\theta}+(-\ln u_3)^{\theta}]^{\frac{1}{\theta}}}$	$[-1, \infty)$

Clayton, Frank, Ali–Mikhail–Haq, and Gumbel–Houggard copula, are considered as the possible copulas to characterize the dependence structure of *RI*, A_L, and *V*. These four commonly used copula functions and the copula parameters are presented in Table 3. Because, the copula functions in Table 3 are single-parameter copulas. Therefore, the number of copula parameters for the four candidate copulas *k* equals to 1.

3.2. Best-Fit Marginal Distributions. To best-fit the marginal distributions of *RI*, A_L, and *V*, four commonly used marginal distributions, namely truncated Normal (TruncNormal), Lognormal, Weibull, and truncated Gumbel (TruncGumbel) distributions in the geotechnical literatures [26–28], are used in this paper. Table 4 summarizes the CDFs and PDFs of these four distributions.

Based on the 50 sets of debris-flow data, the optimal marginal distributions of *RI*, A_L, and *V* are identified by AIC, respectively. Table 5 presents the AIC scores for the

Table 4: Four commonly used marginal distributions.

Marginal distribution	CDF $F(x;\,p,\,q)$	PDF $f(x;\,p,\,q)$	μ and σ
TruncNormal	$\frac{\Phi(\frac{x-p}{q})-\Phi(-\frac{p}{q})}{1-\Phi(-\frac{p}{q})}$	$\frac{\frac{1}{\sqrt{2\pi}q}\exp[-\frac{(x-p)^2}{2q^2}]}{1-\Phi(-\frac{p}{q})}$	$\mu=p, \sigma=q$
Lognormal	$\Phi(\frac{\ln x-p}{q})$	$\frac{1}{\sqrt{2\pi}xq}\exp[-\frac{(\ln x-p)^2}{2q^2}]$	$\mu=\exp(p+0.5q^2)$ $\sigma^2=[\exp(q^2)-1]\exp(2p+q^2)$
Weibull	$1-\exp[-(\frac{x}{p})^q]$	$\frac{q}{p}(\frac{x}{p})^{q-1}\exp[-(\frac{x}{p})^q]$	$\mu=p\Gamma(1+1/q)$ $\sigma^2=p^2[\Gamma(1+2/q)-\Gamma^2(1+1/q)]$
TruncGumbel	$\frac{\exp\{-\exp[-q(x-p)]\}-\exp[-\exp(pq)]}{1-\exp[-\exp(pq)]}$	$\frac{q\exp\{-q(x-p)-\exp[-q(x-p)]\}}{1-\exp[-\exp(pq)]}$	$\mu=p+0.5772/q, \sigma^2=\pi^2/6q^2$

Note. Φ denotes the CDF of standard normal distribution; p and q are the parameters of the marginal probability distribution function; μ is the mean value; σ is the standard deviation.

Table 5: Summary of the calibration for the best-fit marginal distributions.

Marginal distribution	AIC value		
	RI	A_L	V
TruncNormal	328.52	1438.31	1299.18
Lognormal	331.68	**1273.85**	**1239.22**
Weibull	**328.24**	1281.22	1243.13
TruncGumbel	336.80	1397.95	1275.52

Note. Bold indicates the minimum AIC value among the four candidate marginal distributions.

four candidate marginal distributions of RI, A_L, and V. It can be seen that as for RI, Weibull distribution has the minimum AIC value (i.e., 328.24) among the candidate distributions. Thus, Weibull distribution is taken as the most suitable marginal distribution for RI in the study area. As for A_L, the AIC values for TruncNormal, Lognormal, Weibull, and TruncGumbel distributions are 1,438.31, 1,273.85, 1,281.22, and 1,397.95, respectively. It is obvious that the Lognormal distribution for A_L has the lowest AIC value (i.e., 1,273.85), which is identified as the most suitable marginal distribution for A_L in the study area. Similarly, the most suitable marginal distribution for V in the study area is Lognormal distribution (Table 5).

Figures 1–3 show the PDFs and CDFs of the measured data and four candidate marginal distributions for RI, A_L, and V, respectively. It is clear that as for RI, the Weibull distribution fits well with the observation data of RI. Lognormal distribution is in good agreement with the measured data of A_L and V. This validates the accuracy of the identification of marginal distributions of RI, A_L, and V.

4. Probabilistic Prediction Model of Debris-Flow Volume Based on Copula Approaches

Based on the 50 sets of observation data of RI, A_L, and V in Table 1, copula approaches are used to construct their joint probabilistic model. Table 6 presents the results of calibration of the most suitable copula function and its corresponding copula parameter. The AIC values for Clayton, Frank, Ali–Mikhail–Haq, and Gumbel–Hougaard copulas are −235.346, −244.087, −197.52, and −100.01, respectively. It is obvious that Frank copula has the minimum AIC value. Therefore, it is identified as the most suitable copula to characterize the dependence structure of RI, A_L, and V in the study area. The corresponding copula parameter θ of Frank copula is calculated as 2.808.

By substituting the best-fit marginal distributions for RI, A_L, and V (i.e., Weibull, Lognormal, and Lognormal distributions), and the most suitable copula function (i.e., Frank copula) into Equation (1), the three-dimensional joint distribution for RI, A_L, and V can be expressed as follows:

$$\begin{aligned} F(RI, A_L, V) &= P(RI \le RI^*, A_L \le A_L^*, V \le V^*) = C(u_{RI}, u_{A_L}, u_V; \theta) \\ &- \frac{1}{2.808}\ln\left[1+\frac{(e^{-2.808u_{RI}}-1)(e^{-2.808u_{A_L}}-1)(e^{-2.808u_V}-1)}{(e^{-2.808}-1)^2}\right], \end{aligned} \quad (8)$$

where RI^*, $A_L{}^*$, and V^* are the arguments for RI, A_L, and V, respectively; u_{RI}, u_{A_L}, and u_V are the marginal CDFs for RI, A_L, and V, respectively. Then, based on the probability theory, the conditional probability α that the debris-flow volume V for a specific basin equals to V_0 considering $RI = RI_0$ and $A_L = A_{L0}$ can be calculated as follows:

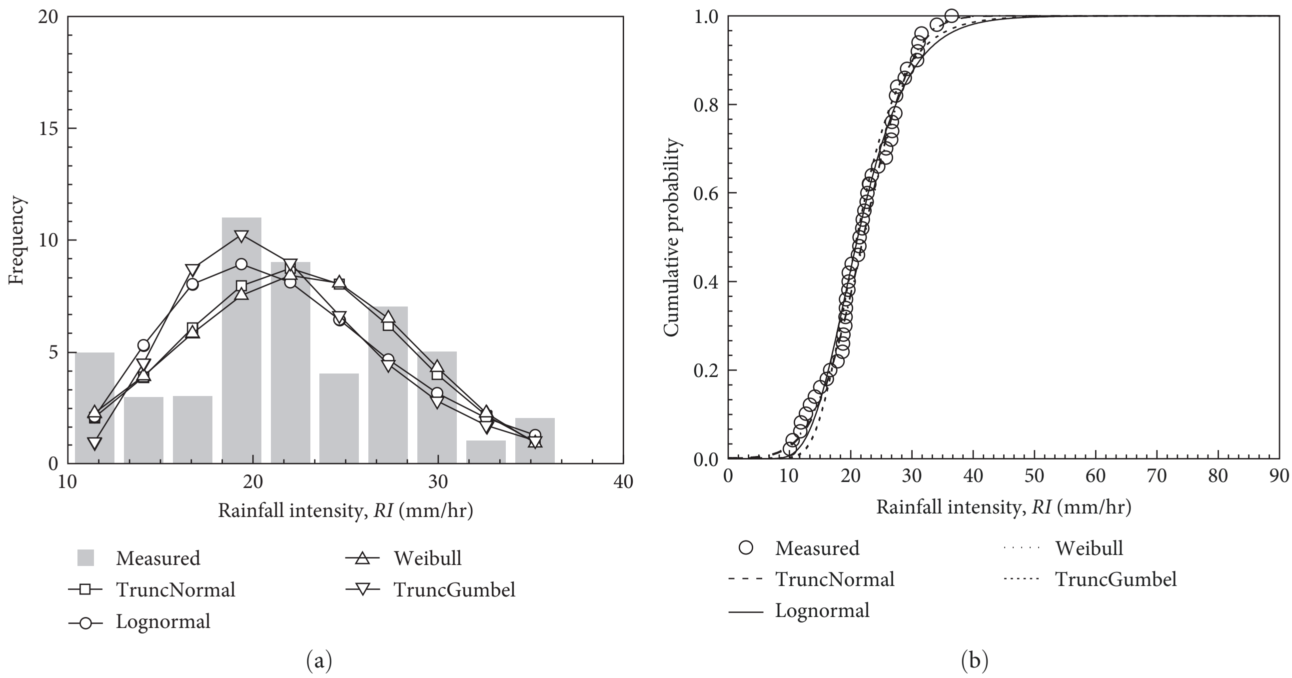


FIGURE 1: Best-fit marginal distribution of *RI*: (a) PDF and (b) CDF.

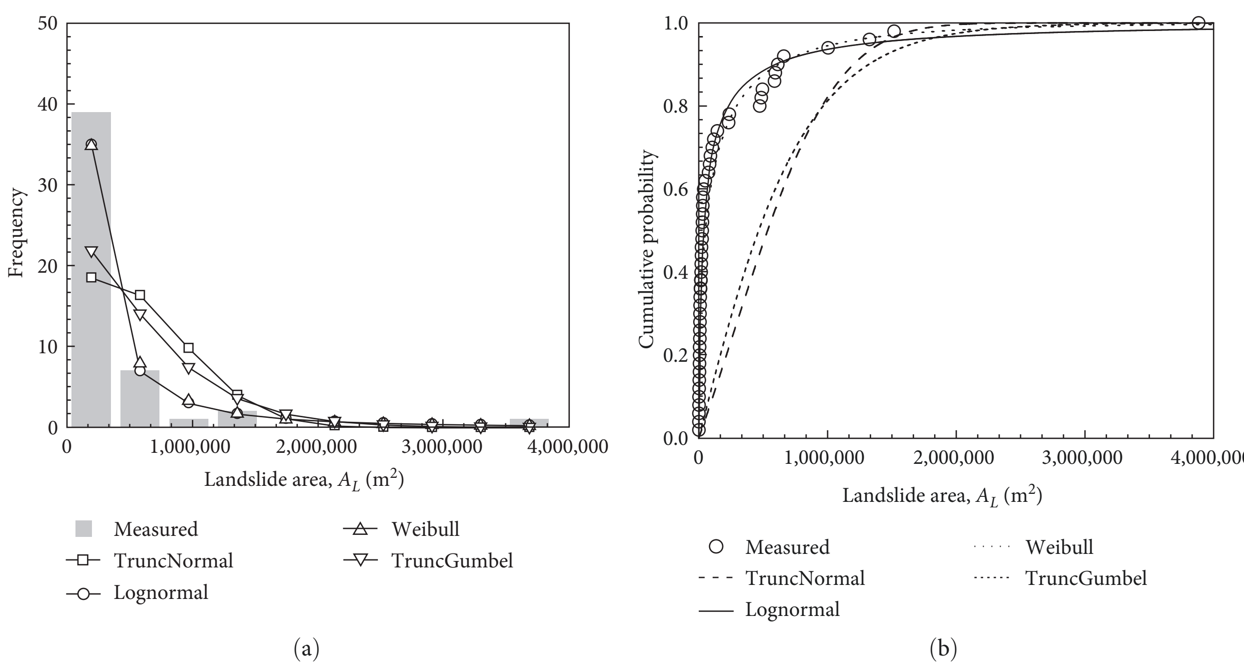


FIGURE 2: Best-fit marginal distribution of A_L: (a) PDF and (b) CDF.

$$\begin{aligned}\alpha &= P(V \leq V_0 | RI = RI_0, A_L = A_{L0}) \\ &= C(V \leq V_0 | RI = RI_0, A_L = A_{L0}) = \frac{\frac{\partial^2}{\partial u_{RI} \partial u_{A_L}} C\left(u_{RI_0}, u_{A_{L0}}, u_{V_0}\right)}{\frac{\partial^2}{\partial u_{RI} \partial u_{A_L}} C\left(u_{RI_0}, u_{A_{L0}}\right)},\end{aligned} \tag{9}$$

where $C(u_{RI_0}, u_{A_{L0}})$ is the bivariate joint distribution of RI and A_L. From Equation (9), it can be seen that given $RI = RI_0$ and $A_L = A_{L0}$, the risk level of a potential debris-flow event that the debris volume equals to V_0 is quantitatively

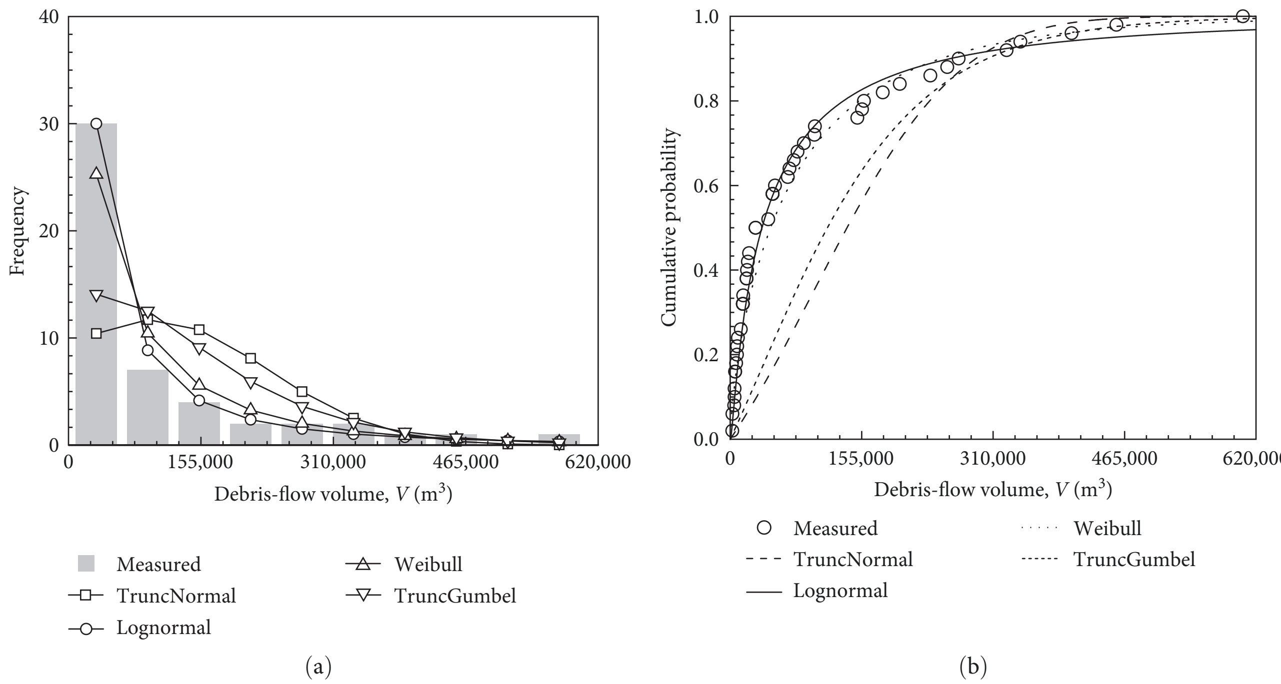


Figure 3: Best-fit marginal distribution of V: (a) PDF and (b) CDF.

Table 6: Identification of best-fit copula function and copula parameters.

Copula	AIC value	θ
Clayton	−235.346	0.797
Frank	**−244.087**	**2.808**
Ali–Mikhail–Haq	−197.52	0.990
Gumbel–Houggard	−100.01	0.000

Note. Bold indicates the minimum AIC value among the four candidate copulas.

characterized as α. Equation (9) also shows that the debris-flow volume can be estimated with a given probability level α considering $RI = RI_0$ and $A_L = A_{L0}$. Obviously, there exists a value of α that can derive the best-fit estimates of debris-flow volume. Mean-square error (MSE) is a measure of the difference between the measured and estimated value, which is given by:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(V_i' - V_i\right)^2, \tag{10}$$

where N is the sample size of the debris-flow observation data; V_i' is the estimated value of debris-flow volume with a given probability level α; V_i is the measured value of debris-flow volume. The probability level α with minimum MSE is considered as the best-fit risk level, which is subsequently used to estimate the debris-flow volume by using Equation (9).

Figure 4 shows the relationship between MSE and the probability level α. It is clear that MSE has the minimum value when the probability level α equals to 0.94. Given the conditional probability $\alpha = 0.94$ and 50 sets of observation data in Table 1, the volumes for the 50 past debris-flow events are estimated by using Equation (9). Figure 5 shows the results of estimated debris-flow volumes by copula approaches. The determination coefficient is about 0.743. It is obvious that estimated debris-flow volumes are generally around the 1 : 1 line. About 98% estimated values are within the 95% confidence interval.

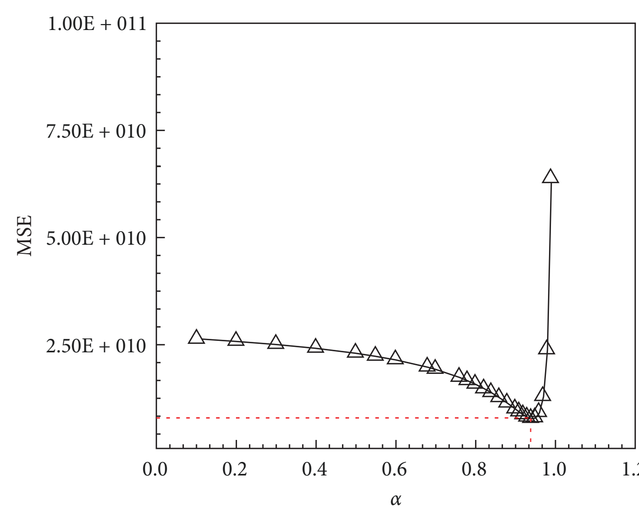


Figure 4: Relationship between MSE and the probability level α.

To further validate the proposed copula approaches, nine sets of independent observation data in Table 1 (Nos. 51–59) are employed to forecast the debris-flow volume using Equation (9), where the conditional probability α is set as 0.94. Meanwhile, an empirical relationship is developed with the 50 sets of “training” data in Table 1, as shown in Figure 6. The determination coefficient of empirical relationship is about

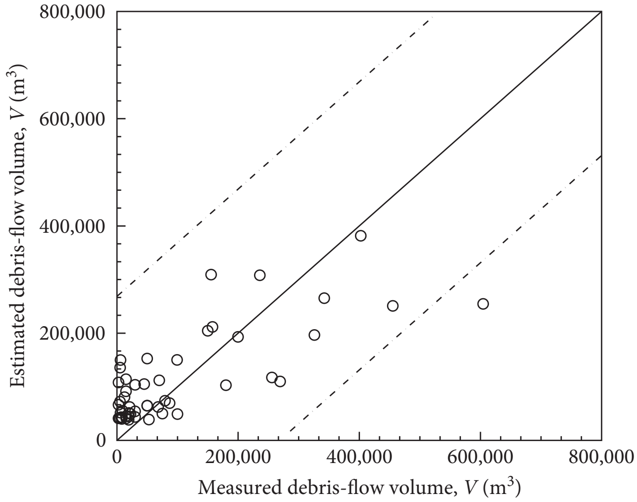


FIGURE 5: Comparison of measured debris-flow volume with the estimates obtained from the copula approaches.

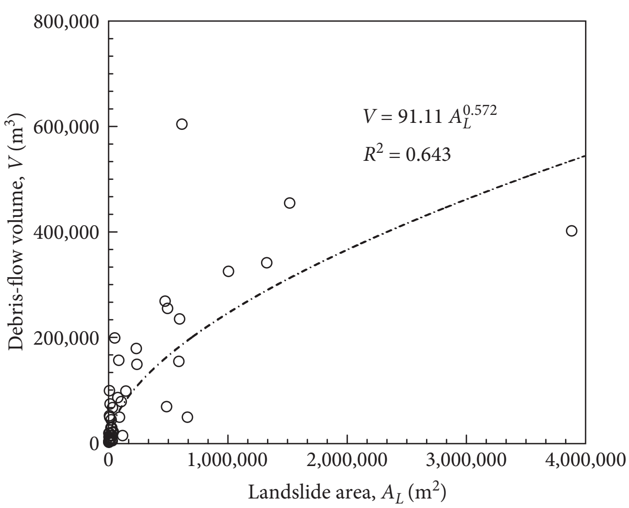


FIGURE 6: Empirical relationship for predicting the debris-flow volume in the study area.

0.643, which is then used for comparison. Figure 7 shows the measured values and predicted values obtained from the proposed copula approaches and the empirical relationship. It is clear that the V values predicted by the proposed copula approaches are closer to the measured values. The V values estimated by the proposed copula approaches show smaller scatters than that obtained from the empirical relationship. This indicates that the proposed copula approaches properly characterizes the high uncertainties and variabilities of the

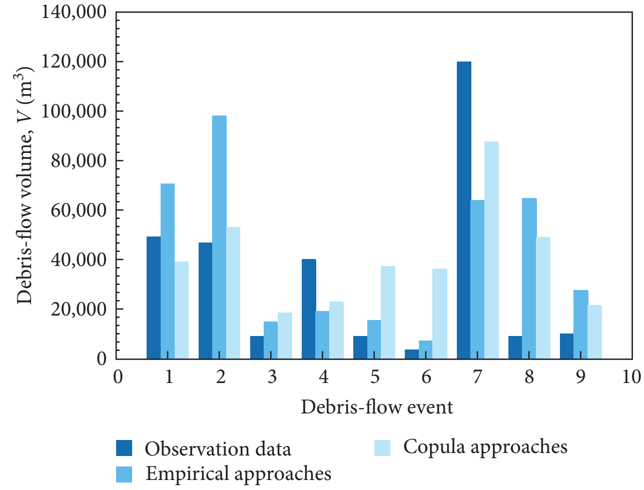


FIGURE 7: Validation of the proposed copula approaches based on the nine sets of independent data.

debris-flow volume and its influencing factors, and can provide reasonable forecasting of the debris-flow volume.

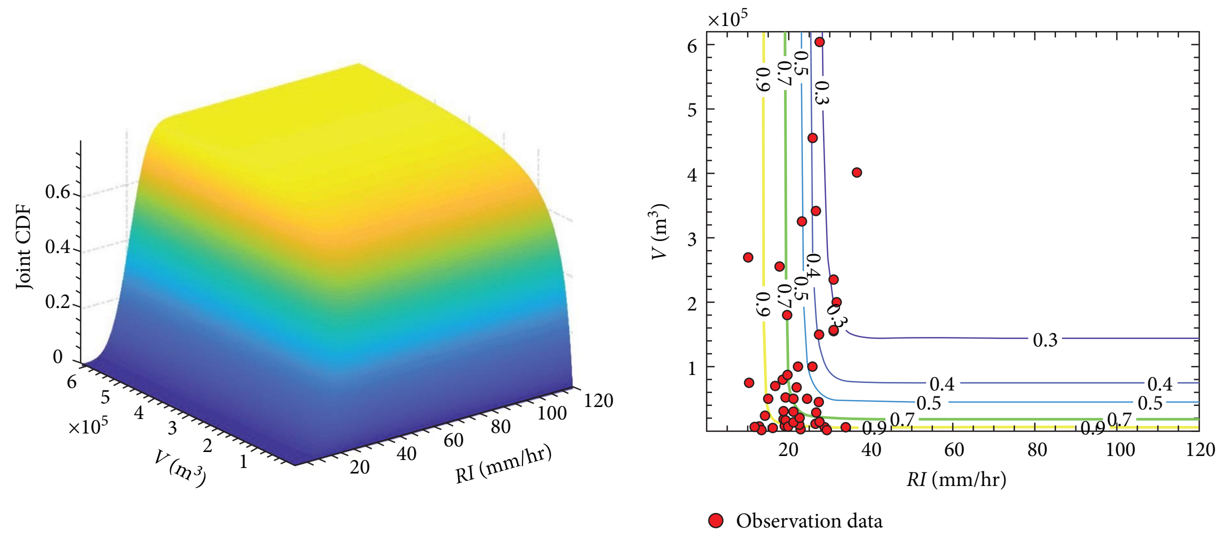
In addition, it should be noted that the debris-flow volume also depends on the debris-flow parameters (e.g., flow velocity) and physiographical parameters. These factors are not considered in this paper and should be considered for more rigorous and more accurate estimates of debris-flow volume. Although the proposed method has the above limits, these limits do not impact the proposed method itself but are the common problems of most physical and empirical methods. The probabilistic model is developed based on a limited number of observation data from debris-flow events in the study area. If more debris-flow event data are available, the joint probability model can be recalibrated to substantially improve the prediction accuracy of debris-flow volume.

5. Exceedance Probability Charts for Debris-Flow Hazard Assessment Based on Multivariate Joint Probabilistic Model

Considering that the debris-flow volume is a key parameter in the hazard assessment of debris flow, it is worthwhile developing exceedance probability charts for mitigation strategies design based on the previous analyses of this study. The developed probabilistic model (i.e., Equation (8)) can be used to develop exceedance probability design charts for debris-flow hazard assessment. From Equation (8), the exceedance probability of RI and V given a specific landslide area $A_L = A_{L0}$ is defined as follows [26]:

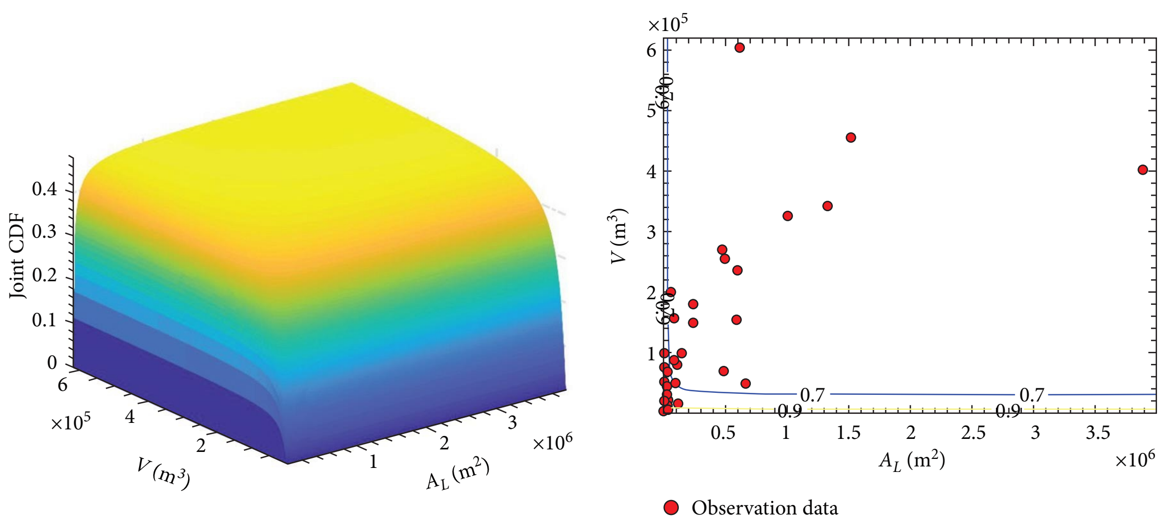
$$\begin{aligned} P[(V \geq V_0) \cup (RI \geq RI_0) | A_L = A_{L0}] &= 1 - P[(V \leq V_0) \cap (RI \leq RI_0) | A_L = A_{L0}] \\ &= 1 - C(V \leq V_0, RI = RI_0 | A_L = A_{L0}) \end{aligned} . \quad (11)$$

Using Equation (8), the respective exceedance probability values of RI and V can be calculated at different threshold values. Figure 8 shows the conditional probability distribution and bivariate exceedance probability chart considering



(a) (b)

Figure 8: Conditional probability distribution and bivariate exceedance probability chart considering $A_L = 263{,}785.70\ \mathrm{m}^2$ (i.e., the mean value in Table 1). (a) Conditional probability distribution of V and RI and (b) bivariate exceedance probability chart.



(a) (b)

Figure 9: Conditional probability distribution and bivariate exceedance probability chart considering $RI = 22.08$ mm/hr (i.e., the mean value in Table 1). (a) Conditional probability distribution of V and A_L and (b) bivariate exceedance probability chart.

$A_L = 263{,}785.70\ \mathrm{m}^2$ (i.e., the mean value in Table 1). Each line in Figure 8(b) implies an equal exceedance probability line. The bivariate exceedance probability chart of RI and V shows that the exceedance probability increases with the decreasing threshold values of RI and V. The exceedance probability chart of RI and V can also provide a means to determine the magnitude of a debris flow. For example, if the rainfall intensity $RI = 60$ mm/hr, the corresponding debris-flow volume can be estimated with different exceedance probability. In addition, the intensity of a debris flow can be characterized in a probability-based manner. For example, the exceedance probability of $RI = 60$ mm/hr and $V = 10{,}000\ \mathrm{m}^3$ equals to 0.7 from the chart, which can provide a preliminary reference for debris-flow risk assessment and design of control measures. Similarly, the exceedance probability chart of A_L and V given a specific rainfall intensity $RI = 22.08$ mm/hr (i.e., the mean value in Table 1) has the same results, as shown in Figure 9. The bivariate exceedance probability of A_L and V also increases with the decreasing threshold values of A_L and V. The exceedance probability chart of A_L and V can also provide a means to determine the magnitude of a debris flow.

6. Summary and Conclusions

Hazard assessment is crucial for debris-flow risk assessment and design of the control measures. This paper proposed probabilistic models for debris-flow hazard assessment considering the uncertainties of multiple influencing factors based on the copula approaches. The proposed probabilistic models not only can provide probabilistic estimation of the debris-flow volume, but also determine the probability of a potential debris-flow event. The proposed copula approaches were illustrated by using the 59 past rainfall debris-flow events in Taiwan. First, 59 datasets of past debris flows were divided in to 50 sets of data for model construction and 9 sets of data for validation. Then, a three-dimensional copula model incorporating the debris-flow volume V and its influencing factors (e.g., rainfall intensity, RI and landslide area, A_L) was developed based on the 50 sets of observation data. Finally, the developed V-RI-A_L joint probabilistic model in Taiwan was used to make probabilistic prediction of the debris-flow volume for a specific hazard level. The proposed approaches were validated and compared with the empirical approach by using nine sets of independent observation data in the study area. The proposed probabilistic model was also used to develop the exceedance probability charts of quantities (e.g., the debris-flow volume, V and rainfall intensity, RI) considering a given landslide area, A_L for a specific debris flow. The findings are summarized as follows:

(1) The statistical goodness-of-fit tests show that the Weibull distribution is the most appropriate marginal distribution for rainfall intensity, RI in Taiwan. Lognormal distribution is the most appropriate marginal distribution for debris-flow volume V and landslide area, A_L in the study area.

(2) Among the Clayton, Frank, Ali–Mikhail–Haq, and Gumbel–Houggard copula, the Frank copula is the best-fit copula for characterizing the dependence structure between RI, A_L, and V in the study area. A three-dimensional joint probabilistic model that incorporates Weibull–Lognormal–Lognormal distribution, and the Frank copula can be used to characterize the joint probability distribution of RI, A_L, and V in Taiwan.

(3) The joint probabilistic model of RI, A_L, and V can be used to provide reasonable prediction of debris-flow volume with a specific conditional probability $\alpha = 0.94$. Compared with the empirical relationship, the estimated debris-flow volume by using the proposed copula approaches are closer to the measured values. The proposed approaches can provide an alternative method for forecasting the magnitude of a potential debris-flow event in Taiwan.

(4) The developed probabilistic model of RI, A_L, and V can provide exceedance probability-design charts for debris-flow hazard assessment. The exceedance probability increases with the decreasing threshold values of RI and V given a specific landslide area $A_L = A_{L0}$. The exceedance probability chart can provide a preliminary reference for debris-flow risk assessment and design of the control measures.

Data Availability

All data in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Project No. 52009037), the Natural Science Foundation of Hubei Province of China (Project No. 2020CFB291), and the Wuhan Knowledge Innovation Special Project (No. 2022020801020268).

References

[1] M. Jakob and O. Hungr, *Debris-Flow Hazards and Related Phenomena*, Springer, Berlin, 2005.

[2] R. Y. Ngadisih, N. P. Bhandary, and R. K. Dahal, "Integration of statistical and heuristic approaches for landslide risk analysis: a case of volcanic mountains in West Java Province, Indonesia," *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, vol. 8, no. 1, pp. 29–47, 2013.

[3] D. Rickenmann, "Empirical relationships for debris flows," *Natural Hazards*, vol. 19, no. 1, pp. 47–77, 1999.

[4] S. McDougall, "2014 Canadian Geotechnical Colloquium: landslide runout analysis—current practice and challenges," *Canadian Geotechnical Journal*, vol. 54, no. 5, pp. 605–620, 2017.

[5] Q.-X. Deng, J. He, Z.-J. Cao, I. Papaioannou, D.-Q. Li, and K.-K. Phoon, "Bayesian learning of Gaussian mixture model for calculating debris flow exceedance probability," *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, vol. 16, no. 1, pp. 154–177, 2022.

[6] M. Hürlimann, R. Copons, and J. Altimir, "Detailed debris flow hazard assessment in Andorra: a multidisciplinary approach," *Geomorphology*, vol. 78, no. 3-4, pp. 359–372, 2006.

[7] O. Hungr, S. McDougall, M. Wise, and M. Cullen, "Magnitude–frequency relationships of debris flows and debris avalanches in relation to slope relief," *Geomorphology*, vol. 96, no. 3-4, pp. 355–365, 2008.

[8] S. P. Schilling and R. M. Iverson, "Automated, reproducible delineation of zones at risk from inundation by large volcanic debris flows," in *Proceedings of the 1997 1st International Conference on Debris-Flow Hazards Mitigation: Mechanics, Prediction, and Assessment*, pp. 176–186, ASCE, San Francisco, CA, 7-9 August 1997.

[9] G. Liu, E. Dai, Q. Ge, W. Wu, and X. Xu, "A similarity-based quantitative model for assessing regional debris-flow hazard," *Natural Hazards*, vol. 69, no. 1, pp. 295–310, 2013.

[10] T. de Haas and A. L. Densmore, "Debris-flow volume quantile prediction from catchment morphometry," *Geology*, vol. 47, no. 8, pp. 791–794, 2019.

[11] D. Rickenmann, "Runout prediction methods," in *Debris-Flow Hazards and Related Phenomena*, M. Jakob and O. Hungr, Eds., Springer Praxis Books book series (GEOPHYS), pp. 305–324, Springer, Berlin, Heidelberg, 2005.

[12] J. He, D.-Q. Li, Z.-J. Cao, M. Tian, and Y. Hong, "Bayesian Estimation of Exceedance Probability of Debris Flows," in

Proc. of the 6th Intl. Symposium on Reliability Engineering and Risk Management (6ISRERM), Singapore, 31 May - 1 June 2018.

[13] V. Medina, M. Hürlimann, and A. Bateman, "Application of FLATModel, a 2D finite volume code, to debris flows in the northeastern part of the Iberian Peninsula," *Landslides*, vol. 5, no. 1, pp. 127–142, 2008.

[14] J. Wang, S. N. Ward, and L. Xiao, "Numerical simulation of the December 4, 2007 landslide-generated tsunami in Chehalis Lake, Canada," *Geophysical Journal International*, vol. 201, no. 1, pp. 372–376, 2015.

[15] R. M. Iverson and D. L. George, "Modelling landslide liquefaction, mobility bifurcation and the dynamics of the 2014 Oso disaster," *Géotechnique*, vol. 66, no. 3, pp. 175–187, 2016.

[16] D. Felix, X. Li, and J. Zhao, "A unified CFD-DEM approach for modeling of debris flow impacts on flexible barriers," *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 42, no. 14, pp. 1643–1670, 2018.

[17] M. Pirulli, "On the use of the calibration-based approach for debris-flow forward-analyses," *Natural Hazards and Earth System Sciences*, vol. 10, no. 5, pp. 1009–1019, 2010.

[18] O. Hungr, G. C. Morgan, and R. Kellerhals, "Quantitative analysis of debris torrent hazards for design of remedial measures," *Canadian Geotechnical Journal*, vol. 21, no. 4, pp. 663–677, 1984.

[19] P.-S. Lin, J.-Y. Lin, J.-C. Hung, and M.-D. Yang, "Assessing debris-flow hazard in a watershed in Taiwan," *Engineering Geology*, vol. 66, no. 3-4, pp. 295–313, 2002.

[20] J. E. Gartner, S. H. Cannon, P. M. Santi, and V. G. Dewolfe, "Empirical models to predict the volumes of debris flows generated by recently burned basins in the western U.S." *Geomorphology*, vol. 96, no. 3-4, pp. 339–354, 2008.

[21] J. E. Gartner, S. H. Cannon, and P. M. Santi, "Empirical models for predicting volumes of sediment deposited by debris flows and sediment-laden floods in the transverse ranges of southern California," *Engineering Geology*, vol. 176, pp. 45–56, 2014.

[22] T.-C. Chang and Y.-H. Chien, "The application of genetic algorithm in debris flows prediction," *Environmental Geology*, vol. 53, no. 2, pp. 339–347, 2007.

[23] D. Rickenmann and A. Koschni, "Sediment loads due to fluvial transport and debris flows during the 2005 flood events in Switzerland," *Hydrological Processes*, vol. 24, no. 8, pp. 993–1007, 2010.

[24] C.-W. Chang, P.-S. Lin, and C.-L. Tsai, "Estimation of sediment volume of debris flow caused by extreme rainfall in Taiwan," *Engineering Geology*, vol. 123, no. 1-2, pp. 83–90, 2011.

[25] L. Franzi and G. Bianco, "A statistical method to predict debris flow deposited volumes on a debris fan," *Physics and Chemistry of the Earth, Part C: Solar, Terrestrial & Planetary Science*, vol. 26, no. 9, pp. 683–688, 2001.

[26] Y. Hong, J. P. Wang, D. Q. Li, Z. J. Cao, C. W. W. Ng, and P. Cui, "Statistical and probabilistic analyses of impact pressure and discharge of debris flow from 139 events during 1961 and 2000 at Jiangjia Ravine, China," *Engineering Geology*, vol. 187, pp. 122–134, 2015.

[27] X.-S. Tang, J. P. Wang, W. Yang, and D.-Q. Li, "Joint probability modeling for two debris-flow variables: copula approach," *Natural Hazards Review*, vol. 19, no. 2, Article ID 05018004, 2018.

[28] M. Tian and X.-T. Sheng, "Copula-based probabilistic approaches for predicting debris-flow runout distances in the Wenchuan earthquake zone," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, vol. 8, no. 1, Article ID 04021070, 2022.

[29] A. Sklar, "Fonctions de répartition àn dimensions et leurs marges," *Publications de l'Institut de Statistique de l'Université de Paris*, vol. 8, pp. 229–231, 1959.

[30] R.-B. Nelsen, *An Introduction to Copulas*, Springer, New York, 2006.

[31] X.-S. Tang, M.-X. Wang, and D.-Q. Li, "Modeling multivariate cross-correlated geotechnical random fields using vine copulas for slope reliability analysis," *Computers and Geotechnics*, vol. 127, Article ID 103784, 2020.