

Original Article

Evaluation of rainfall threshold models for debris flow initiation in the Jiangjia Gully, Yunnan Province, China

YANG Hongjuan¹  <https://orcid.org/0000-0003-0635-6764>; e-mail: yanghj@imde.ac.cnZHANG Shaojie^{1*}  <https://orcid.org/0000-0001-5908-9554>;  e-mail: sj-zhang@imde.ac.cnHU Kaiheng¹  <https://orcid.org/0000-0001-8114-5743>; e-mail: khhu@imde.ac.cnWEI Fangqiang^{2,3}  <https://orcid.org/0000-0001-8734-0881>; e-mail: fqwei@imde.ac.cnLIU Yanhui⁴  <https://orcid.org/0009-0003-9722-1618>; e-mail: lyanhui@mail.cgs.gov.cn

*Corresponding author

¹ State Key Laboratory of Mountain Hazards and Engineering Resilience, Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610299, China² Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China³ Chongqing School, University of Chinese Academy of Sciences, Chongqing 400714, China⁴ China Institute of Geo-Environment Monitoring, Beijing 100081, China**Citation:** Yang HJ, Zhang SJ, Hu KH, et al. (2024) Evaluation of rainfall threshold models for debris flow initiation in the Jiangjia Gully, Yunnan Province, China. Journal of Mountain Science 21(6). <https://doi.org/10.1007/s11629-023-8507-6>

© Science Press, Institute of Mountain Hazards and Environment, CAS and Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract: Systematically determining the discriminatory power of various rainfall properties and their combinations in identifying debris flow occurrence is crucial for early warning systems. In this study, we evaluated the discriminatory power of different univariate and multivariate rainfall threshold models in identifying triggering conditions of debris flow in the Jiangjia Gully, Yunnan Province, China. The univariate models used single rainfall properties as indicators, including total rainfall (R_{tot}), rainfall duration (D), mean intensity (I_{mean}), absolute energy (E_{abs}), storm kinetic energy (E_s), antecedent rainfall (R_a), and maximum rainfall intensity over various durations ($I_{\text{max_dur}}$). The evaluation reveals that the $I_{\text{max_dur}}$ and E_{abs} models have the best performance, followed by the E_s , R_{tot} , and I_{mean} models, while the D and R_a models have poor performances. Specifically, the $I_{\text{max_dur}}$ model has the highest performance metrics at a 40-min duration. We used logistic regression to

combine at least two rainfall properties to establish multivariate threshold models. The results show that adding D or R_a to the models dominated by E_{abs} , E_s , R_{tot} , or I_{mean} generally improve their performances, specifically when D is combined with I_{mean} or when R_a is combined with E_{abs} or E_s . Including R_a in the $I_{\text{max_dur}}$ model, it performs better than the univariate $I_{\text{max_dur}}$ model. A power-law relationship between $I_{\text{max_dur}}$ and R_a or between E_{abs} and R_a has better performance than the traditional $I_{\text{mean}}-D$ model, while the performance of the E_s-R_a model is moderate. Our evaluation reemphasizes the important role of the maximum intensity over short durations in debris flow occurrence. It also highlights the importance of systematically investigating the role of R_a in establishing rainfall thresholds for triggering debris flow. Given the regional variations in rainfall patterns worldwide, it is necessary to evaluate the findings of this study across diverse watersheds.

Received: 26-Nov-2023
1st Revision: 24-Mar-2024
2nd Revision: 26-Apr-2024
Accepted: 10-May-2024

Keywords: Rainfall threshold; Logistic regression; Maximum rainfall intensity; Absolute energy; Antecedent rainfall

1 Introduction

Debris flows are mixtures of poorly sorted sediment and water that are gravity-driven and exhibit behaviors intermediate between rock avalanches and water floods (Iverson 1997). They occur widely in mountainous regions and pose a threat to public and infrastructure safety due to their long run-out distances and strong destructive power, causing impact, erosion, and inundation (Dowling and Santi 2014; Kean et al. 2019; Nieto et al. 2021). In some cases, the significant amount of sediment carried by debris flow may block the main river and lead to a cascading disaster (Cui et al. 2013; An et al. 2022). To mitigate debris flow hazards, early warning systems for debris flows have been established at the local or regional level in areas prone to these hazards (Baum and Godt 2010; Osanai et al. 2010; Berenguer et al. 2015; Devoli et al. 2018). As rainfall is the primary trigger for debris flow, establishing rainfall thresholds is an effective way to issue early warnings for debris flows.

Rainfall thresholds are defined as critical rainfall conditions that, when reached or exceeded, are likely to result in debris flows (Nikolopoulos et al. 2017). There are two methods for determining rainfall thresholds for debris flow: physical and empirical methods. Debris flows are mainly generated in two ways. They either originate from shallow landslides (Iverson et al. 1997) or initiate from erosion of in-channel sediment by runoff (Kean et al. 2013). Accordingly, physically based methods define critical rainfall conditions using hydrological methodologies to derive critical runoff and/or slope stability analysis (Berti et al. 2020; Pastorello et al. 2020; Li et al. 2021; Martinengo et al. 2023). However, because physically based methods require a large amount of high-quality input data, empirical methods that determine rainfall thresholds using historical rainfall and debris flow data are more commonly used (Hirschberg et al. 2021).

Empirical rainfall threshold models commonly use the power-law relationship between mean rainfall intensity (or total rainfall) and rainfall duration (Coe et al. 2008; Guzzetti et al. 2008; Chang et al. 2021; Liu 2023). Additionally, the maximum intensity over short durations (usually ≤ 60 min) is a good indicator of debris flow triggering (Staley et al. 2013; Abancó et al. 2016; Bel et al. 2017; Tsunetaka et al. 2021; Thomas et al. 2023). Recent research from the Goulunping catchment in central China demonstrated that the absolute energy, defined as the sum of squared values

of rainfall depth in each sampling period in a rainfall event, had strong predictive power for debris flow (Zhao et al. 2022). Moreover, the antecedent soil moisture conditions impact debris flow triggering in some watersheds (Bel et al. 2017; Oorthuis et al. 2023; Siman-Tov and Marra 2023). Therefore, a systematic study on the power of different rainfall properties and their potential combinations in distinguishing triggering from non-triggering conditions of debris flow (discriminatory power) is important for establishing rainfall thresholds.

When determining the rainfall threshold at the local or regional scale, uncertainty in rainfall data location and the record of hazard occurrence time can limit the accuracy of the threshold (Marra et al. 2016; Leonarduzzi et al. 2017). In situ debris flow monitoring has been conducted in many regions worldwide in recent years, providing valuable, high-quality data for evaluating empirical rainfall threshold models (Hürlimann et al. 2019 and references therein). At these monitored sites, rainfall data are recorded at 5–10 min intervals, and debris flow occurrence is recorded manually or detected by cameras or sensors (stage gages, geophones, infrasonic wave sensors, etc.). Although systematic studies on the evaluation of empirical rainfall threshold models have been performed at some monitored sites (Staley et al. 2013; Bel et al. 2017; Hirschberg et al. 2021; Zhao et al. 2022; Oorthuis et al. 2023), considering that environmental settings for debris flow are different in different regions, such work needs to be conducted at more monitored sites.

Therefore, the purpose of this study was to evaluate the discriminatory power of different rainfall threshold models for debris flow triggering in Jiangjia Gully, a monitored site in southern China. Using statistically-based skill scores, we first evaluated the discriminatory power of single rainfall properties and then evaluated the discriminatory power of multiple properties. Finally, limitations of the study and some uncertainties in evaluating the threshold models were discussed, including the duration considered in calculating antecedent rainfall, the rainfall temporal resolution, and the minimum inter-event time selected in separating continuous rainfall time series into individual events.

2 General Settings for the Study Area

Jiangjia Gully is located in northeastern Yunnan

Province, with a geographical location of $103^{\circ}05'46''$ – $103^{\circ}13'01''$ E and $26^{\circ}13'16''$ – $26^{\circ}17'13''$ N. It has a drainage area of 48.6 km² and faces west, with elevations ranging from a minimum of 1040 m a.s.l. to a maximum of 3260 m a.s.l. The area's climate is mainly affected by the Indian and East Asian summer monsoon. Mean annual precipitation ranges from 400 to 1000 mm (Cui et al. 2005) and generally increases with elevation due to orographic effects. Majority of the rainfall occurs during the monsoon season, from May to October.

The Menqian Gully and Duo Zhao Gully constitute the primary debris flow source area in the Jiangjia watershed (Fig. 1). However, several check dams were constructed in the Duo Zhao Gully during 1979–1982, which greatly reduced debris flow activity in this sub-watershed (Zeng et al. 2009). Currently, the 13.2-km² Menqian Gully is the primary debris flow source area, with a mean slope of 32° and a maximum slope of 70°. Shallow landslides are widely distributed in this sub-watershed. Some landslides directly evolve into debris flows, while others release sediment to the channel, which is mobilized by runoff in debris flow events (Yang et al. 2022).

The trunk channel of the Jiangjia Gully originates from the confluence of the Menqian Gully and Duo Zhao Gully and terminates at the Xiaojiang River. It is over 100 m wide and 5.5 km long, with the slope decreasing from 5.1° in the transport zone to 3.7° in the deposition zone (Cui et al. 2005).

The study area is an active site for debris flows, making it an ideal location for in situ monitoring. In 1965, the Chinese Academy of Sciences installed a monitoring station, the Dongchuan Debris Flow Observation and Research Station. Staff occupies the station for 2–3 months (referred to as the debris flow observation period hereafter) during the rainy season to record the occurrence times and dynamic properties of the debris flow surges that are clearly visible in the monitoring section of the trunk channel (Fig. 1). Since the establishment of the

station, more than 500 debris flow events have been recorded, each comprising tens or even hundreds of individual surges (Guo et al. 2020).

3 Data and Methods

3.1 Debris flow and rainfall data

This study used debris flow events that occurred from 2007 to 2010 due to the high incidence of debris flow and the availability of associated rainfall data in this period. Debris flow observation periods in the four years included July 1–September 17 in 2007, July 1–August 31 in 2008, July 1–August 31 in 2009, and July 1–September 10 in 2010. A total of 29 debris flow events were observed in the monitoring section of the trunk channel, with the time of occurrence listed in Table 1. During 2007–2010, three tipping bucket rain gages were available in the Menqian Gully (Fig. 1), with elevations ranging from 2325 to 2816 m a.s.l. Rainfall depth was recorded at a 1-min interval with a resolution of 0.1 mm. Given that the sampling period of rainfall data was 5–10 min in most local studies concerning debris flow triggering (Staley et al. 2013; Bel et al. 2017; Hirschberg et al. 2021; Zhao et al. 2022; Oorthuis et al. 2023), the 1-min rainfall data were aggregated into 5-min data for comparison with existing research. Moreover, the following analysis used the average rainfall measurements from the three gages to represent rainfall conditions in the entire primary debris flow source area.

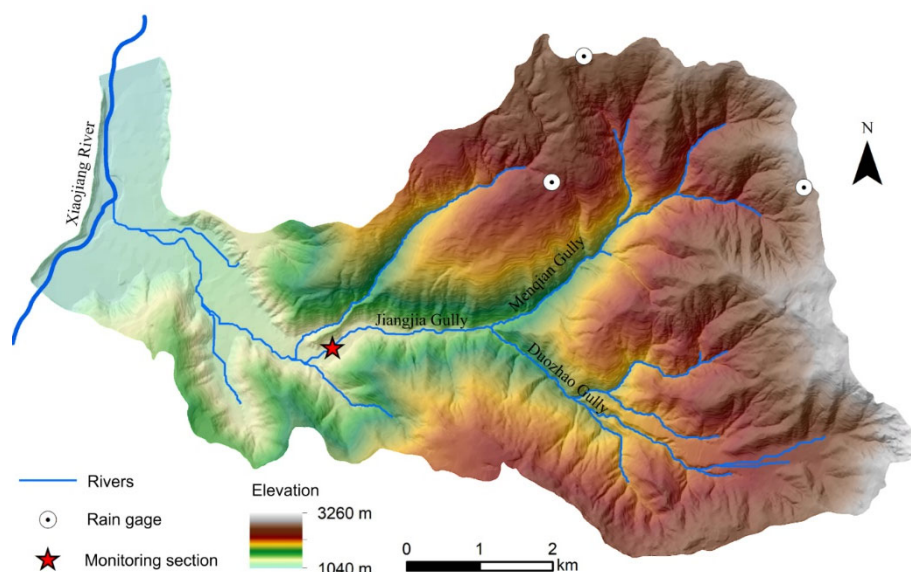


Fig. 1 Terrain of the Jiangjia Gully, in Yunnan Province, China.

Table 1 Occurrence time of debris flow events observed in the monitoring section in the Jiangjia Gully from 2007 to 2010 (mm/dd/yyyy hh:mm)

Number	Occurrence time	Number	Occurrence time	Number	Occurrence time
1	07/10/2007 04:20	11	07/05/2008 06:26	21	08/11/2008 02:33
2	07/24/2007 06:30	12	07/11/2008 06:48	22	08/17/2008 19:00
3	07/25/2007 02:36	13	07/11/2008 17:45	23	08/04/2009 05:24
4	07/25/2007 14:24	14	07/22/2008 05:00	24	07/06/2010 05:23
5	07/30/2007 05:40	15	08/01/2008 00:15	25	07/17/2010 20:39
6	08/11/2007 14:27	16	08/03/2008 04:50	26	07/22/2010 19:15
7	08/25/2007	17	08/03/2008 22:35	27	07/24/2010 19:00
8	09/14/2007 01:30	18	08/04/2008 15:37	28	08/05/2010 05:51
9	09/17/2007 15:12	19	08/05/2008 14:04	29	09/10/2010 03:26
10	07/01/2008 15:55	20	08/08/2008 03:02		

3.2 Segmentation of the rainfall time series

The first step in defining rainfall conditions that trigger debris flows was to divide the rainfall time series into individual events. However, standard criteria for this process are currently lacking in the literature (Peres et al. 2018; Jiang et al. 2021). In our study, we employed a minimum inter-event time (MIET) approach with persistently smaller rainfall intensities than a critical value, I_c (Zhou and Tang 2014). We estimated I_c using potential evapotranspiration data (Marino et al. 2020) that showed daily fluctuation between 2 to 6 mm during the study period. Accordingly, we set I_c to 0.3 mm/h. Previous studies on debris flow torrents have employed MIET values ranging from 10 min (Coe et al. 2008) to 7 h (Jiang et al. 2021). In our study, we first chose a moderate value of 3 h for evaluating rainfall threshold models and then analyzed the influence of MIET on the evaluation in the discussion section.

3.3 Threshold type and definitions

A total of 189 rainfall events were identified during the study period using MIET=3 h and $I_c=0.3$ mm/h. These rainfall events were categorized into 28 debris flow-triggering rainfall events (DFTs) and 161 non-triggering events (NDFTs). Since the two debris flow events occurred during the same rainfall event on July 25, 2007, the number of debris flow events exceeded the number of DFTs. For each rainfall event, we calculated several characteristics, including total rainfall (R_{tot} , mm), rainfall duration (D , h), mean rainfall intensity (I_{mean} , mm/h), absolute energy (E_{abs} , mm²), and the maximum rainfall intensity over different durations (I_{max_dur} , mm/h), for a total of twelve durations ranging from 10 to 120 min. I_{max_dur} was calculated with the maximum rainfall depth using

a moving time window. E_{abs} was defined as follows (Zhao et al. 2022):

$$E_{abs} = \sum_{i=1}^N R_i^2 \quad (1)$$

Here, R_i denotes the rainfall depth measured in the i^{th} sampling period, and N denotes the total number of sampling periods in a rainfall event. The E_{abs} is commonly used in analyzing time series data. Considering that the storm kinetic energy, E_s , is widely used for estimating soil loss, it was evaluated. It was computed using the following equation (Kinnell 2023):

$$E_s = \sum_{i=1}^N R_i \cdot 0.29(1 - 0.72\exp(-0.082I_i)) \quad (2)$$

Here, I_i denotes the rainfall intensity in the i^{th} sampling period. Furthermore, we evaluated the antecedent rainfall (R_a , mm) (Bruce and Clark 1966), which was determined as follows:

$$R_a = \sum_{i=1}^m k^i R_{i-24h} \quad (3)$$

Here, R_{i-24h} denotes the rainfall depth measured in the i^{th} 24 h prior to the rainfall event, m denotes the number of days considered, and k denotes the decay factor representing the outflow of the regolith (Glade et al. 2000). The suggested value for k is 0.84 (Bruce and Clark 1966). In terms of m , values employed in the literature are usually not greater than 30 (Bui et al. 2013; Garcia-Urquia 2016; Uwihiwe et al. 2020; Chinkulkijniwat et al. 2022). We first used $m=15$ in this study as suggested by Bui et al. (2013) and then analyzed the influence of m on the performance of the threshold models in the discussion section.

However, for DFTs, using the entire rainfall time series from beginning to end may lead to the overestimation of the rainfall threshold (Abancó et al. 2016; Bel et al. 2017). Therefore, we only considered

the rainfall time series recorded before occurrence time of the debris flow event (i.e. time listed in Table 1) for calculating the rainfall characteristics.

When using a single rainfall property, X , to define debris flow-triggering conditions (hereafter referred to as the univariate threshold model), the threshold type was as follows:

$$X = c \quad (4)$$

Here, c denotes a constant.

When multiple rainfall properties were considered in the rainfall threshold (referred to as the multivariate threshold model), we first utilized a linear combination of these properties:

$$\sum_{i=1}^n a_i X_i = c \quad (5)$$

Here, a_i denotes the linear coefficient of the i^{th} variable, X_i , and n denotes the number of variables. Additionally, we incorporated the power function of each property by multiplying them to account for the power-law relationship between different properties (e.g., I_{mean} and D):

$$\prod_{i=1}^n X_i^{a_i} = c \quad (6)$$

The coefficients in the multivariate threshold models were determined using the logistic regression method, an effective tool for classification problems. In this method, the probability of debris flow occurrence (p ; where $p=0$ for NDFTs and $p=1$ for DFTs) was expressed as a Sigmoid function of the linear combination of the explanatory variables:

$$\ln\left(\frac{p}{1-p}\right) = a_0 + \sum_{i=1}^n a_i Y_i \quad (7)$$

Here, $Y_i=X_i$ for Eq. (5) and $Y_i=\ln(X_i)$ for Eq. (6). Linear independence between variables is required in the logistic regression model. Thus, we calculated the Pearson correlation coefficient (CC) for each pair of rainfall properties, and only properties with CC between -0.5 and 0.5 were used to establish multivariate threshold models.

3.4 Skill scores for performance evaluation

We employed the receiver operating characteristic (ROC) analysis to evaluate the performance of different threshold models (Staley et al. 2013; Bel et al. 2017; Oorthuis et al. 2023). First, we assigned different thresholds (i.e., different values of c in Eqs. (4), (5),

and (6)) to the model to be evaluated. For each threshold, the rainfall events were classified into four groups: true positives (TP; DFTs with conditions above the threshold), true negatives (TN; NDFTs with conditions below the threshold), false positives (FP; NDFTs with conditions above the threshold), and false negatives (FN; DFTs with conditions below the threshold). Second, we calculated two skill scores, probability of detection (POD) and probability of false detection (POFD), for each threshold:

$$\text{POD} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

$$\text{POFD} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (9)$$

Third, we constructed an ROC curve by plotting the POD against the POFD. We used the area under the ROC curve (AUC) to evaluate the performance of different threshold models, where a value of 0.5 indicates no improvement over random guessing and a value of 1.0 indicates perfect discrimination. Finally, we determined the optimal threshold by maximizing the true skill statistic (TSS; Allouche et al. 2006), also known as the Hanssen–Kuipers discriminant (Hanssen and Kuipers 1965), which is expressed as follows:

$$\text{TSS} = \text{POD} - \text{POFD} \quad (10)$$

The TSS of the optimal threshold was also used as an evaluation metric of the different threshold models.

4 Results

4.1 Performance of univariate threshold models

Fig. 2 depicts the AUC and maximum TSS of the univariate threshold models and Table 2 lists the optimal thresholds. Models using E_{abs} , E_s , or $I_{\text{max_dur}}$ as an indicator performed better than the other models, with AUC and maximum TSS values greater than 0.90 and 0.70, respectively. Among them, the $I_{\text{max_30min}}$ and $I_{\text{max_40min}}$ models had the maximum TSS (0.815) while the $I_{\text{max_40min}}$ model also had the maximum AUC (0.938). Therefore, the $I_{\text{max_40min}}$ model exhibited the best performance. Additionally, the E_{abs} model performed better than the E_s model. R_{tot} and I_{mean} also exhibited strong ability in distinguishing triggering from non-triggering conditions of debris flow, with AUC and maximum TSS greater than 0.87 and 0.68, respectively. Furthermore, the R_{tot} model performed

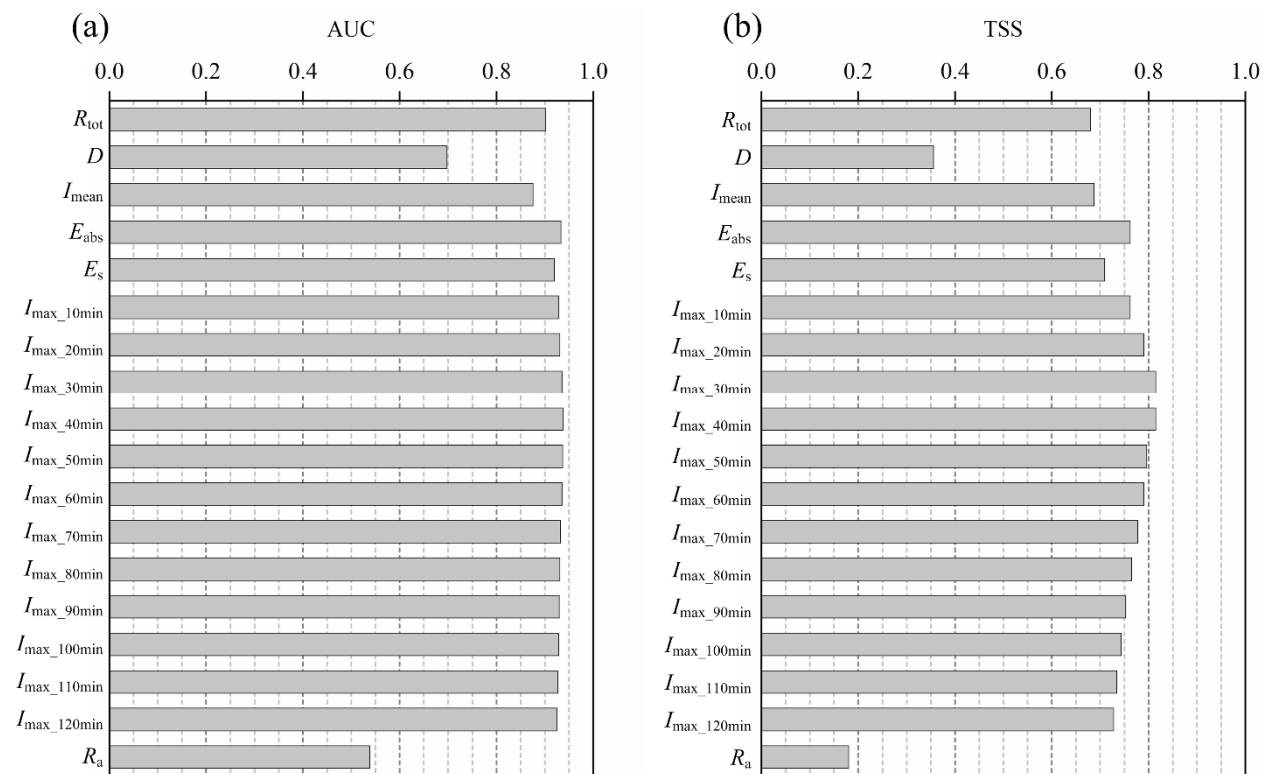


Fig. 2 Performance metrics, including (a) the area under the ROC curve (AUC) and (b) maximum true skill statistic (TSS), of different univariate threshold models.

better than the I_{mean} model in terms of AUC, whereas the opposite was true regarding TSS. However, using D or R_a as an indicator resulted in considerably poorer performance, particularly in terms of FP. Specifically, the AUC of the R_a model was 0.538, indicating that it was only slightly better than random guessing.

4.2 Performance of multivariate threshold models

Fig. 3 depicts the correlation matrix between different rainfall characteristics. R_{tot} , I_{mean} , E_{abs} , E_s , and all I_{max_dur} variables were found to be linearly dependent, whereas D and R_a were generally independent of these variables, despite the linear dependence between D and R_{tot} . Therefore, D , R_a , and both were combined separately with R_{tot} , I_{mean} , E_{abs} , E_s , and each I_{max_dur} variable to establish multivariate threshold models using Eqs. (5) and (6), resulting in a total of 96 multivariate models that were established and evaluated.

Fig. 4 presents the performance metrics of the

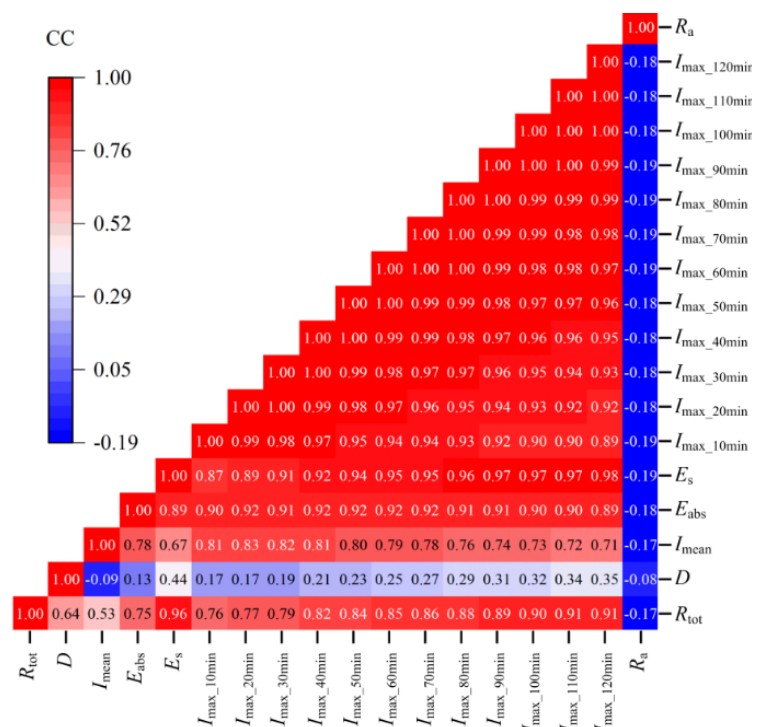


Fig. 3 Pearson correlation coefficient (CC) between different rainfall characteristics.

multivariate threshold models. Models using Eq. (6) generally outperformed those using Eq. (5); therefore,

Table 2 Optimal threshold and its corresponding skill scores for each univariate threshold model

Optimal threshold	TP	TN	FP	FN	POD	POFD	TSS	AUC*
$R_{tot}=11.4$	22	144	17	6	0.786	0.106	0.680	0.901
$D=2.83$	19	109	52	9	0.679	0.323	0.356	0.698
$I_{mean}=2.36$	25	128	33	3	0.893	0.205	0.688	0.876
$E_{abs}=5.03$	25	140	21	3	0.893	0.130	0.762	0.934
$E_s=1.85$	23	143	18	5	0.821	0.112	0.710	0.920
$I_{max_10min}=12.00$	25	140	21	3	0.893	0.130	0.762	0.929
$I_{max_20min}=7.70$	27	133	28	1	0.964	0.174	0.790	0.931
$I_{max_30min}=6.73$	27	137	24	1	0.964	0.149	0.815	0.936
$I_{max_40min}=5.65$	27	137	24	1	0.964	0.149	0.815	0.938
$I_{max_50min}=4.68$	27	134	27	1	0.964	0.168	0.797	0.937
$I_{max_60min}=4.00$	27	133	28	1	0.964	0.174	0.790	0.936
$I_{max_70min}=3.42$	27	131	30	1	0.964	0.186	0.778	0.933
$I_{max_80min}=3.05$	27	129	32	1	0.964	0.199	0.766	0.931
$I_{max_90min}=2.73$	27	127	34	1	0.964	0.211	0.753	0.930
$I_{max_100min}=3.44$	25	137	24	3	0.893	0.149	0.744	0.929
$I_{max_110min}=2.23$	27	124	37	1	0.964	0.230	0.734	0.927
$I_{max_120min}=2.05$	27	123	38	1	0.964	0.236	0.728	0.926
$R_a=19.09$	24	52	109	4	0.857	0.677	0.180	0.538

Note: *AUC is the skill score of the corresponding threshold model; true positives (TP)/false negatives (FN) mean DFTs with conditions above/below the threshold; true negatives (TN)/false positives (FP) mean NDFTs with conditions below/above the threshold; probability of detection (POD) and probability of false detection (POFD).

they were further compared with univariate threshold models. For models dominated by R_{tot} or I_{mean} , after adding both D and R_a , they significantly outperformed the univariate models. They also performed better when only D or R_a was included, where including D outperformed including R_a for I_{mean} while including D showed similar performance with including R_a for R_{tot} . Compared to models that only used R_{tot} or I_{mean} , including both D and R_a resulted in an increase in AUC and TSS exceeding 0.05 and 0.10, respectively. For models dominated by E_{abs} or E_s , after adding R_a , they performed much better than the univariate models, with the increase in AUC and TSS greater than 0.02 and 0.05, respectively. They also performed better than models that considered D and had slightly poorer performance than models that considered both D and R_a . Finally, the model dominated by I_{max_dur} and included R_a performed approximately with the one that included both D and R_a . Compared to the univariate I_{max_dur} models, the inclusion of R_a averagely increased AUC by 0.024 and TSS by 0.067.

In the multivariate models using Eq. (6), the $E_{abs}-R_a$, $I_{max_40min}-D-R_a$, and $E_{abs}-D-R_a$ models had the maximum AUC (0.961) while the $I_{max_20min}-R_a$ and $I_{max_40min}-R_a$ models had the maximum TSS (0.865).

However, all $I_{max_dur}-R_a$ and $I_{max_dur}-D-R_a$ models had good performances, with $AUC > 0.95$ and $TSS > 0.795$. The E_s-D-R_a model also had approximate performance ($AUC=0.960$, $TSS=0.812$). These models outperformed the widely used $I_{mean}-D$ or $R_{tot}-D$ models ($AUC=0.926$ and $TSS=0.739$) and their R_a -included type (i.e., $I_{mean}-D-R_a$ or $R_{tot}-D-R_a$ models; $AUC=0.953$ and $TSS=0.790$). Considering that including D in the $I_{max_dur}-R_a$, $E_{abs}-R_a$, and E_s-R_a models brought about no or marginal improvement in models' performances, we recommend the $I_{max_dur}-R_a$, $E_{abs}-R_a$, and E_s-R_a models for debris flow early warnings in the study area. Among the $I_{max_dur}-R_a$ models, the $I_{max_50min}-R_a$ and $I_{max_60min}-R_a$ models had the best AUC (0.960) while the $I_{max_20min}-R_a$ and $I_{max_30min}-R_a$ models had the best TSS (0.865). The $I_{max_40min}-R_a$ model had the second best AUC (0.959) and TSS (0.852). Therefore, it was selected as a typical example of the $I_{max_dur}-R_a$ model. The optimal thresholds of the $I_{max_40min}-R_a$, $E_{abs}-R_a$, E_s-R_a models and the traditional $I_{mean}-D$ model are shown in Fig. 5.

These results reveal the importance of R_a in determining the triggering conditions of debris flow in the Jiangjia watershed. A detailed investigation of the classification results of the optimal $I_{max_40min}-R_a$, $E_{abs}-R_a$, and E_s-R_a thresholds are shown in Table 3. Generally, including R_a in the threshold could decrease FP or increase TP while avoiding a substantial increase in FP. Compared to the best I_{max_40min} threshold that resulted in 24 FP, the $I_{max_40min}-R_a$ threshold reduced FP to 18; correspondingly, the POFD decreased from 0.149 to 0.112. The best E_{abs} threshold yielded 25 TP and 21 FP. In comparison, the $E_{abs}-R_a$ threshold increased by 2 TP, resulting in an increase in the POD by 0.071, while the increase of 3 FP led to a small increase in the POFD (0.019). Compared to the best E_s threshold, the E_s-R_a threshold increased by 2 TP and decreased by 4 FP; therefore, the TSS was increased from 0.710 to 0.806.

5 Discussion

5.1 The best duration for calculating the maximum intensity

Compared to R_{tot} , I_{mean} , E_{abs} , and E_s , I_{max_dur} demonstrated stronger discriminatory power between triggering and non-triggering rainfalls in the study area. This finding is consistent with previous research (Bel et al. 2017; Hirschberg et al. 2021; Tsunetaka et al.

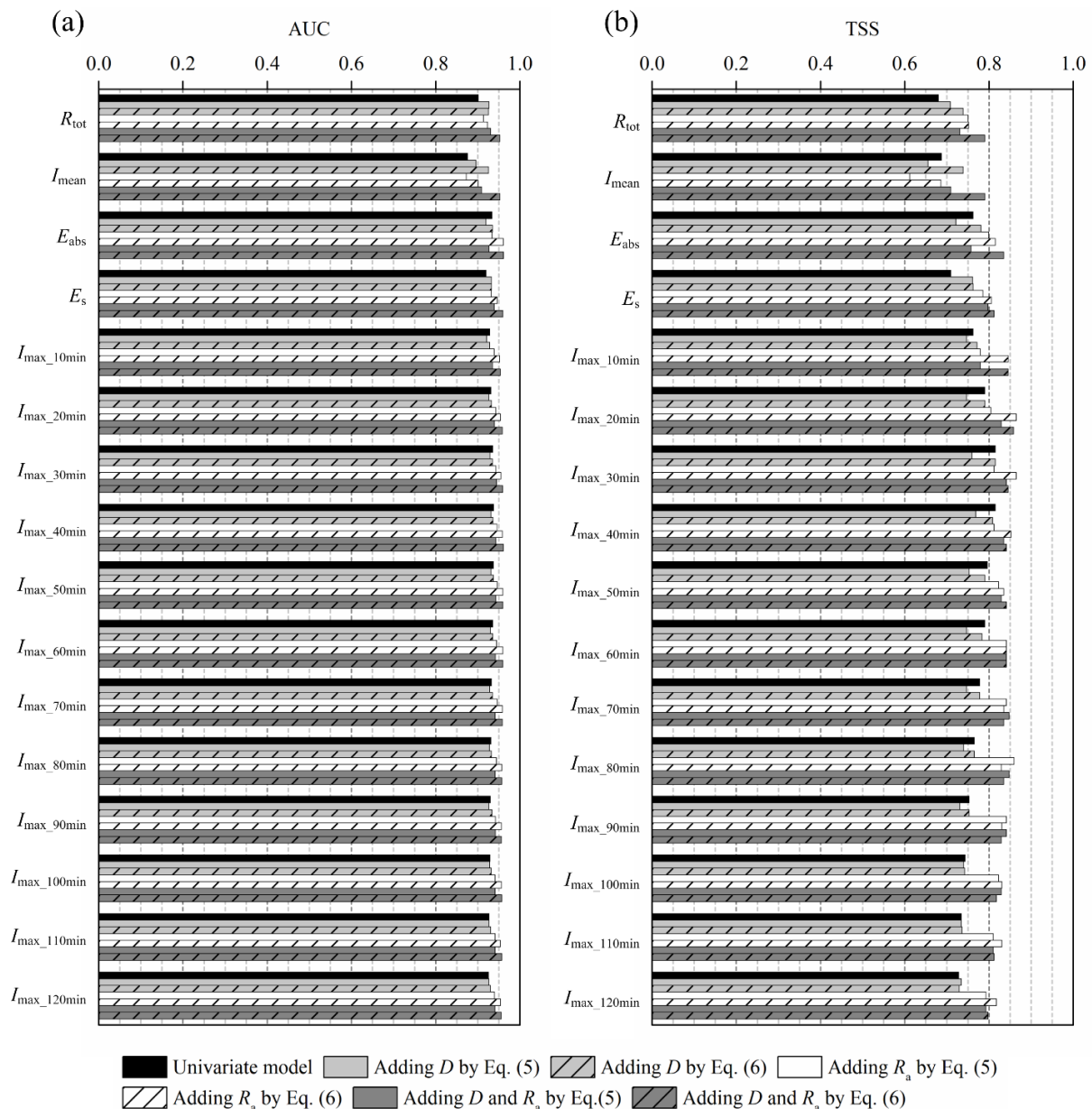


Fig. 4 Performance metrics, including (a) AUC and (b) maximum TSS, of the multivariate threshold models established by separately adding D , R_a , or both D and R_a to the models dominated by R_{tot} , I_{mean} , E_{abs} , E_s , or I_{max_dur} .

2021). However, in the present study, the I_{max_dur} model performed best at a 40-min duration, which is longer than the best durations reported in other studies, ranging from 10 to 30 min (Staley et al. 2017; Hirschberg et al. 2021; Tsunetaka et al. 2021; Oorthuis et al. 2023). This inconsistency is likely due to differences in the drainage area of the investigated sites (Table 4). Durations of 10–20 min were reported in the Ichino-sawa torrent and the Rebaixader catchment, which are smaller than 1 km², and in a regional study of post-fire debris flow torrents ranging from 0.02 to 8 km². In the 4.83-km² Illgraben

catchment, the best duration was 30 min.

Runoff is crucial in triggering debris flow in these watersheds (Cannon et al. 2008; Berger et al. 2011; Imaizumi et al. 2019; Guo et al. 2020; Pastorello et al. 2020). Therefore, debris flow occurrence is likely attributed to a critical discharge at the debris flow initiation site in each watershed. Since peak discharge occurs when all segments of the drainage area contribute to the runoff of the site, the best duration may be related to the time of concentration (T_c) at the debris flow initiation site, T_{c_ini} . Considering that T_c at the outlet of the watershed, T_{c_out} , is the upper limit for

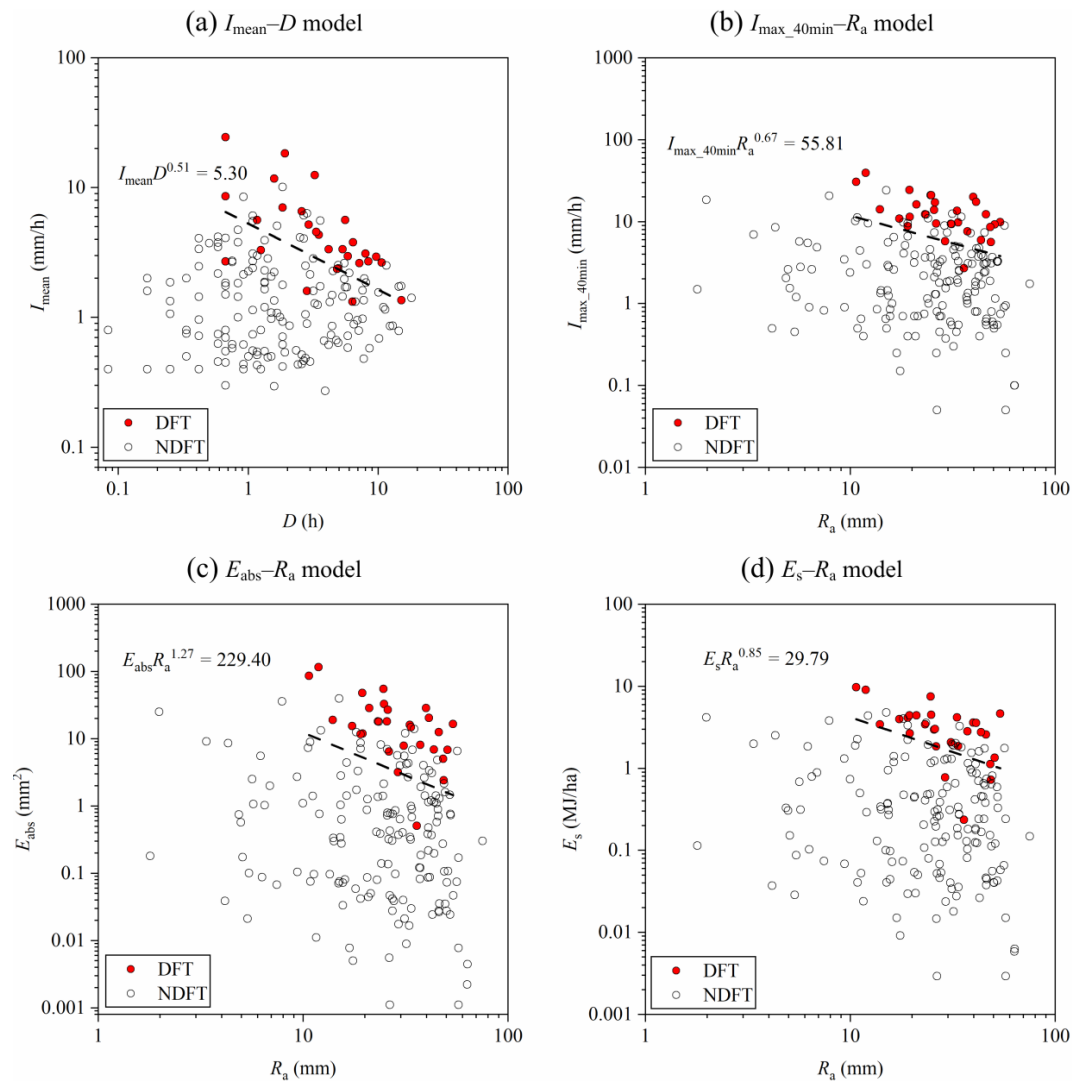


Fig. 5 Optimal thresholds of the (a) $I_{\text{mean}}-D$, (b) $I_{\text{max_40min}}-R_a$, (c) $E_{\text{abs}}-R_a$, and (d) E_s-R_a models.

Table 3 Comparison of the classification results associated with the optimal thresholds of the $I_{\text{max_40min}}-R_a$, $E_{\text{abs}}-R_a$, E_s-R_a models, and the corresponding univariate models

Optimal threshold	TP	TN	FP	FN	POD	POFD	TSS	AUC*
$I_{\text{max_40min}}=5.65$	27	137	24	1	0.964	0.149	0.815	0.938
$I_{\text{max_40min}}R_a^{0.67}=55.81^{**}$	27	143	18	1	0.964	0.112	0.852	0.959
$E_{\text{abs}}=5.03$	25	140	21	3	0.893	0.130	0.762	0.934
$E_{\text{abs}}R_a^{1.27}=229.40$	27	137	24	1	0.964	0.149	0.815	0.961
$E_s=1.85$	23	143	18	5	0.821	0.112	0.710	0.920
$E_sR_a^{0.85}=29.79$	25	147	14	3	0.893	0.087	0.806	0.947

Note: *AUC is the skill score of the corresponding threshold model.

****:** For multivariate threshold models that included two explanatory variables, the original form of the threshold, $X_1^{a_1}X_2^{a_2} = c$, was expressed with a more concise form, $X_1X_2^a = c'$, where $a = a_2/a_1$ and $c' = c^{1/a_1}$.

Table 4 Best duration for calculating the maximum intensity in different research

Study site	Best duration (min)	Drainage area (km ²)	Source
Ichino-sawa torrent	10	0.22	Tsunetaka et al. 2021
Post-fire debris flow torrents	15	0.02–8	Staley et al. 2017
Rebaixader catchment	15–20	0.53	Oorthuis et al. 2023
Illgraben catchment	30	4.83	Hirschberg et al. 2021
Menqian Gully	40	13.20	This study

T_{c_ini} , the best duration is expected to be $<T_{c_out}$. To test this assumption, we used five empirical formulas (Tropeano et al. 1996) suitable for steep mountain catchments to estimate T_{c_out} for the Menqian Gully. T_{c_out} was not estimated in other watersheds due to a lack of morphometric parameters. The results are listed in Table 5. The estimated T_{c_out} ranged from 40 to 109 min in the Menqian Gully; the best duration was the lower limit of this range. It indicates that the estimated T_{c_out} has a potential to be used for calculating the maximum intensity when the initiation location is unknown.

Table 5 Time of concentration at the outlet (T_{c_out} , h) of the Menqian Gully estimated with different empirical formulas

Formula (Tropeano et al. 1996)	T_{c_out}
$T_{c_out} = (4A^{0.5} + 1.5L)(H_m - H_o)^{-0.5} / 0.8$	1.04 (62 min)
$T_{c_out} = 0.396Ls^{-0.5}(AL^{-2}s^{0.5}s_v^{-0.5})^{0.72}$	1.58 (95 min)
$T_{c_out} = 0.055Ls^{-0.5}$	0.66 (40 min)
$T_{c_out} = 6L^{2/3}(H_{max} - H_o)^{-1/3}$	1.82 (109 min)
$T_{c_out} = 0.127A^{0.5}s^{-0.5}$	0.87 (52 min)

Note: $A=13.20$ km² is the basin area, $L=6.42$ km is the headwater basin length, $H_{max}=3002$ m is the basin's maximum elevation, $H_m=2366$ m is the average basin elevation, $H_o=1526$ m is the basin outlet elevation, $s=0.283$ is the average channel gradient, and $s_v=0.624$ is the average slope gradient.

5.2 Duration for calculating antecedent rainfall

The results of this study indicated that threshold models performed better when R_a was included, although the univariate R_a model had poor performance. This suggests that high antecedent soil moisture is not required for the initiation of debris flow in the Jiangjia Gully. However, high antecedent moisture levels decrease the triggering rainfall conditions, including R_{tot} , I_{mean} , E_{abs} ,

E_s , and I_{max_dur} . This is consistent with findings in Chalk Cliffs (Coe et al. 2008) and the Rebaixader catchment (Abancó et al. 2016; Oorthuis et al. 2023). Antecedent moisture facilitates debris flow in the study area in two ways. First, the increase in soil moisture decreases the shear strength of the soil (Hu et al. 2011), making the soil mantle more prone to landslides (Hawke and McConchie 2011) and the bed sediment more prone to erosion (McCoy et al. 2012). Second, higher antecedent moisture means a smaller soil water deficit and lower infiltration rate in the watershed, both leading to higher runoff ratios (Penna et al. 2011; Schoener and Stone 2019).

R_a has been widely used to determine the rainfall threshold for landslides, with durations within 30 days commonly considered (Bui et al. 2013; Garcia-Urquia 2016; Uwihirwe et al. 2020; Chinkulkijniwat et al. 2022). The duration used in this study was 15 days. To investigate the influence of duration on the performance of the $I_{max_40min}-R_a$, $E_{abs}-R_a$, and E_s-R_a models, R_a was calculated with different durations ranging from 1 to 30 days, and the corresponding AUC and TSS of the three models were computed (Fig. 6). For the $I_{max_40min}-R_a$ and $E_{abs}-R_a$ models, the AUC and TSS increased with duration when the duration was <7 days, while they remained approximately stable when the duration was ≥ 7 days, reaching maximum values at 15 and 10 days, respectively. For the E_s-R_a model, the AUC and TSS stabilized when the duration was greater than 20 and 15 days, respectively. This is because later rainfall had a greater weight in Eq. (3). Specifically, when R_a was calculated with a duration of 30 days (denoted as R_{a_30d}), rainfall that occurred within 15 days prior to the rainfall events averagely contributed 29.02 mm to R_{a_30d} (31.19 mm) and accounted for 93%. Accordingly, the Pearson correlation coefficient between R_a calculated with durations ≥ 15 days and

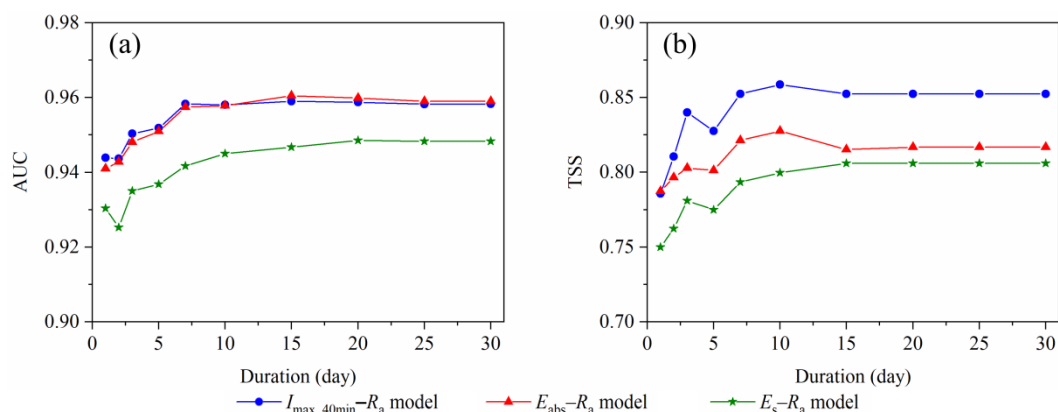


Fig. 6 Performance metrics, including (a) AUC and (b) maximum TSS, of the $I_{max_40min}-R_a$, $E_{abs}-R_a$, and E_s-R_a models when different durations were used for calculating R_a .

R_{a_30d} was high (>0.99) (Fig. 7). Therefore, durations ≥ 15 days are recommended for computing R_a in regions with similar environments to the study area.

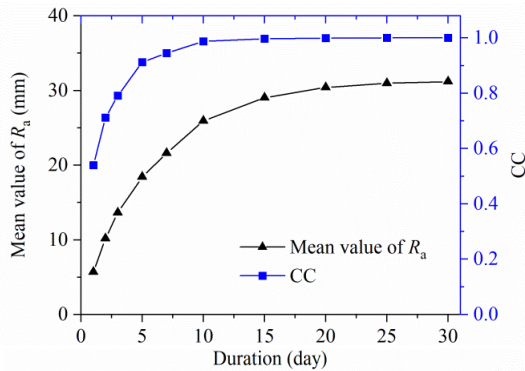


Fig. 7 Mean value of R_a for all the 189 rainfall events when different durations were considered and Pearson correlation coefficient (CC) between R_a calculated with different durations and the 30-day R_a .

5.3 Influence of the rainfall temporal resolution and MIET

The temporal resolution of rainfall used in this study was 5 min. However, ground-based rainfall measurements used to determine regional rainfall thresholds usually have temporal resolutions of ≥ 1 h (Brunetti et al. 2018; Jiang et al. 2021; Rana et al. 2022). Since the sampling interval affects the values of rainfall characteristics and might further influence the relative performance of different threshold models, the performance of the $I_{\text{mean}}-D$, $I_{\text{max_dur}}-R_a$, $E_{\text{abs}}-R_a$, and E_s-R_a models were reevaluated at the 1-h interval. The results are listed in Table 6. In this case, 1 h was used to calculate the $I_{\text{max_dur}}$. The AUC and TSS of the

$I_{\text{max_1h}}-R_a$ model were 0.959 and 0.848, respectively, which were similar to the scores of the $I_{\text{max_40min}}-R_a$ model at the 5-min interval. The $E_{\text{abs}}-R_a$ model had slightly better performance at the 1-h interval, with an increase of 0.016 in the TSS. In contrast, the TSS of the E_s-R_a model decreased by 0.023. Nonetheless, these three models still outperformed the $I_{\text{mean}}-D$ model.

Table 6 Performance of the threshold models using 1-h interval rainfall data

Threshold model	AUC	Optimal threshold		
		TSS	c'	a
$I_{\text{mean}}D^a=c'$	0.929	0.713	4.39	0.46
$I_{\text{max_1h}}R_a^a=c'$	0.959	0.848	31.81	0.62
$E_{\text{abs}}R_a^a=c'$	0.959	0.831	2886.60	1.22
$E_sR_a^a=c'$	0.945	0.783	23.54	0.83

Note: For multivariate threshold models that included two explanatory variables, the original form of the threshold, $X_1^{a_1}X_2^{a_2} = c$, was expressed with a more concise form, $X_1X_2^a = c'$, where $a = a_2/a_1$ and $c' = c^{1/a_1}$.

The selection of MIET impacts the separation of rainfall events and may influence the relative performance of different rainfall threshold models. To investigate the effects of MIET on the performances of rainfall thresholds in the study area, different MIET values ranging from 1 to 12 h were used to divide the 5-min interval rainfall data. The performance of the $I_{\text{mean}}-D$, $I_{\text{max_40min}}-R_a$, $E_{\text{abs}}-R_a$, and E_s-R_a models using different MIET values is illustrated in Fig. 8. Generally, the performance metrics of these models decreased with MIET when MIET was ≤ 7 h, while they had relatively small fluctuations when MIET was >7 h. The $I_{\text{max_40min}}-R_a$ and $E_{\text{abs}}-R_a$ models consistently performed better than the $I_{\text{mean}}-D$ model, while the E_s-R_a model had a moderate performance.

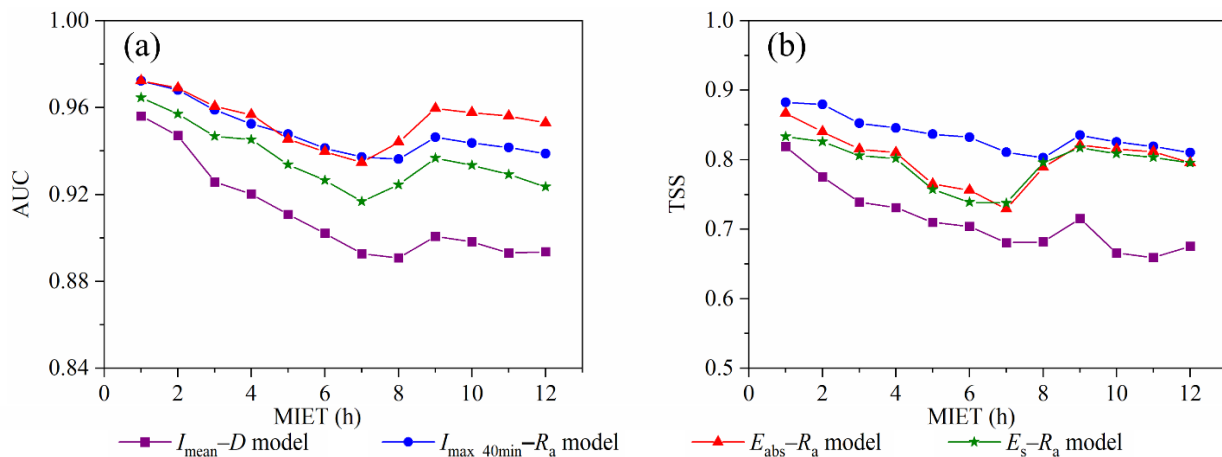


Fig. 8 Performance metrics, including (a) AUC and (b) maximum TSS, of the $I_{\text{mean}}-D$, $I_{\text{max_40min}}-R_a$, $E_{\text{abs}}-R_a$, and E_s-R_a models when different MIET values were used to separate the rainfall time series.

5.4 Comparison with rainfall threshold models in other monitored watersheds

Systematic evaluation of the discriminatory power of various rainfall properties has been conducted in some monitored watersheds and the best threshold model has been proposed, as listed in Table 7. These studies demonstrated the strong ability of $I_{\max_dur}$ in indicating debris flow occurrence. In the Goulinping catchment, E_{abs} was a good indicator for debris flow triggering (Zhao et al. 2022). Results in the Jiangjia Gully were consistent with existing research. In the Réal Torrent and the Goulinping catchment, the combination of R_{tot} with $I_{\max_dur}$ or E_{abs} could improve model performance (Bel et al. 2017; Zhao et al. 2022). In the Jiangjia Gully, compared with the univariate models, adding R_{tot} to the models dominated by $I_{\max_dur}$ or E_{abs} had no improvement for the strong correlation between these variables. In addition, as a proxy of sediment recharge, the number of days elapsed since the end of winter (n_d) also had relevance to occurrence of debris flow in the Réal Torrent. In the other watersheds, only variables related with rainfall were studied and n_d was not considered.

The role of R_a or soil water content (SWC) was investigated in these watersheds except the Goulinping catchment. SWC was not important in the post-fire debris flow torrents, which may be attributed to the relatively porous hillslope soil and bed material, where it was difficult to sustain high water contents (Kean et al. 2011). In the Illgraben catchment, although R_a did not appear to be a significant precondition for debris flow triggering, it affected the magnitude of debris flow (Hirschberg et al. 2021). In the Réal Torrent, the Rebaixader catchment and our study, R_a or SWC played an important role in debris flow occurrence (Bel et al. 2017; Oorthuis et al. 2023).

5.5 Limitations of the study

In this study, rainfall properties that were included in multivariable threshold models were selected based on the correlation matrix between various rainfall

properties (Fig. 3). Since the correlation matrix is influenced by rainfall patterns, the correlation matrix in other regions will be different from the one in the Jiangjia Gully for different rainfall patterns. Therefore, some rainfall threshold models that were not recommended in this study may have good performance in other regions. For instance, the correlation coefficient between R_{tot} and $I_{\max_60min}$ was 0.85 in the study area, indicating a positive relationship. Compared to the univariate $I_{\max_60min}$ model that had AUC=0.936 and TSS=0.790, the inclusion of R_{tot} had no improvement in this study, with AUC=0.931 and TSS=0.753 when Eq. (5) was used and AUC=0.936 and TSS=0.790 when Eq. (6) was used. However, the $I_{\max_1h}-R_{tot}$ model performed well in the Wenchuan earthquake-affected region (Jiang et al. 2021). Therefore, results in this study still need to be tested in other watersheds. This is the major limitation of the present research.

Another limitation is associated with uncertainty in the rainfall data used in this study. Debris flows observed in the monitoring section of the Jiangjia Gully in the study period were primarily triggered in the Menqian Gully for presence of check dams in the Duo Zhao Gully. Therefore, rainfall data monitored in the Menqian Gully had stronger ability in distinguishing triggering from non-triggering conditions of debris flow in the Jiangjia Gully (Yang et al. 2023). Thus, they were used in this study. However, debris flows might also be triggered in the Duo Zhao Gully in the study period, and some of them might be observed in the monitoring section. In this case, using rainfall data monitored in the Menqian Gully may lead to weaker representativeness of the triggering rainfall. However, these cases were rare.

Moreover, this study only evaluated the discriminatory power of different rainfall threshold models. Stability of the coefficients in the threshold model is also important for an operational warning system. Therefore, a more systematical study that also considers the stability of the coefficient needs be performed in further research.

Table 7 Best rainfall threshold models in different monitored watersheds

Study site	Best threshold model	Source
Post-fire debris flow torrents	$I_{\max_dur}=c$	Staley et al. 2013
Réal Torrent	$a_1 I_{\max_5min} + a_2 R_a + a_3 R_{tot} + a_4 n_d = c$	Bel et al. 2017
Illgraben catchment	Random forest model using I_{mean} , D , $I_{\max_30min}$	Hirschberg et al. 2021
Goulinping catchment	$E_{abs}=c_1$ and $R_{tot}=c_2$	Zhao et al. 2022
Rebaixader catchment	$I_{\max_dur}=aSWC+b$	Oorthuis et al. 2023
Jiangjia Gully	$I_{\max_40min} R_a^a = c_1$ or $E_{abs} R_a^b = c_2$	This study

Note: SWC is soil water content, and n_d is the number of days elapsed since the end of winter.

6 Conclusions

We evaluated the performance of different univariate and multivariate rainfall threshold models for identifying triggering conditions of debris flows that occurred during 2007–2010 in the Jiangjia Gully, Yunnan Province, China, where the univariate models used single rainfall properties as indicators and the multivariate models used at least two rainfall properties as indicators.

Among the univariate models, the $I_{\max_dur}$ and E_{abs} models performed the best, followed by the E_s , R_{tot} , and I_{mean} models. The D and R_a models had poor performances for substantial FP. The results reemphasized the important role of the maximum intensity over short durations in debris flow occurrence. They also demonstrated the strong ability of E_{abs} in identifying debris flow-triggering conditions. In the $I_{\max_dur}$ models, which used different durations for calculating the maximum intensity, the best performance was obtained at a 40-min duration, which is longer than durations reported in other watersheds for a larger drainage area of the study watershed.

Although using D or R_a alone had minimal influence in distinguishing triggering from non-triggering rainfalls, adding these indicators to the models dominated by E_{abs} , E_s , R_{tot} , or I_{mean} generally improved the discriminatory power of these models, specifically when D was combined with I_{mean} or when R_a was combined with E_{abs} or E_s . Including R_a in the $I_{\max_dur}$ model performed better than the univariate $I_{\max_dur}$ model. The selection of MIET influenced the performance of $I_{mean}-D$, $I_{\max_dur}-R_a$, $E_{abs}-R_a$, and E_s-R_a models. Nonetheless, the $I_{\max_dur}-R_a$ and $E_{abs}-R_a$ models always had better performance than the traditional $I_{mean}-D$ model, while the performance of the E_s-R_a model was moderate.

Although high antecedent soil moisture is not required for the initiation of debris flow in the Jiangjia Gully, our evaluation showed that including R_a in

debris flow rainfall threshold models improved models' performances because high antecedent moisture levels decrease the triggering rainfall conditions. This highlighted the importance of systematically investigating the role of R_a in establishing rainfall thresholds for debris flow occurrence.

Acknowledgments

The authors sincerely appreciate the valuable comments from the anonymous reviewers. Dongchuan Debris Flow Observation and Research Station, Chinese Academy of Sciences is acknowledged for supplying the data. This research was supported by the National Key R&D Program of China (No. 2023YFC3007205), the National Natural Science Foundation of China (Nos. 42271013, 42077440) and Project of the Department of Science and Technology of Sichuan Province (No. 2023ZHCG0012).

Author Contribution

YANG Hongjuan: Formal analysis, Writing-original draft. ZHANG Shaojie: Conceptualization, Supervision, Funding acquisition. HU Kaiheng and WEI Fangqiang: Writing-review & editing. LIU Yanhui: Writing-review & editing, Funding acquisition.

Ethics Declaration

Availability of Data/Materials: The datasets generated during this study are available from the corresponding author on reasonable request and within the framework of cooperation agreements and scientific research projects.

Conflict of Interest: HU Kaiheng is the scientific editor of Journal of Mountain Science. He was not involved in the journal's review of, or decisions related to, this manuscript. The authors declare no conflicts of interest.

References

- Abancó C, Hürlimann M, Moya J, et al. (2016) Critical rainfall conditions for the initiation of torrential flows. Results from the Rebaixader catchment (Central Pyrenees). *J Hydrol* 541: 218–229.
<https://doi.org/10.1016/j.jhydrol.2016.01.019>
- Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* 43(6): 1223–1232.
<https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- An H, Ouyang C, Wang F, et al. (2022) Comprehensive analysis and numerical simulation of a large debris flow in the Meilong catchment, China. *Eng Geol* 298: 106546.
<https://doi.org/10.1016/j.enggeo.2022.106546>
- Baum RL, Godt JW (2010) Early warning of rainfall-induced shallow landslides and debris flows in the USA. *Landslides* 7: 259–272.
<https://doi.org/10.1007/s10346-009-0177-0>
- Bel C, Liébault F, Navratil O, et al. (2017) Rainfall control of

- debris-flow triggering in the Réal Torrent, Southern French Prealps. *Geomorphology* 291: 17-32.
<https://doi.org/10.1016/j.geomorph.2016.04.004>
- Berenguer M, Sempere-Torres D, Hürlimann M (2015) Debris-flow forecasting at regional scale by combining susceptibility mapping and radar rainfall. *Nat Hazards Earth Syst Sci* 15(3): 587-602.
<https://doi.org/10.5194/nhess-15-587-2015>
- Berger C, McArdell BW, Schlunegger F (2011) Sediment transfer patterns at the Illgraben catchment, Switzerland: implications for the time scales of debris flow activities. *Geomorphology* 125(3): 421-432.
<https://doi.org/10.1016/j.geomorph.2010.10.019>
- Berti M, Bernard M, Gregoretti C, et al. (2020) Physical interpretation of rainfall thresholds for runoff-generated debris flows. *J Geophys Res-Earth* 125(6), e2019JF005513.
<https://doi.org/10.1029/2019JF005513>
- Bruce JP, Clark RH (1966) *Introduction to hydrometeorology*. Pergamon, Oxford.
- Brunetti MT, Melillo M, Peruccacci S, et al. (2018) How far are we from the use of satellite rainfall products in landslide forecasting? *Remote Sens Environ* 210: 65-75.
<https://doi.org/10.1016/j.rse.2018.03.016>
- Bui DT, Pradhan B, Lofman O, et al. (2013) Regional prediction of landslide hazard using probability analysis of intense rainfall in the Hoa Binh province, Vietnam. *Nat Hazards* 66: 707-730.
<https://doi.org/10.1007/s11069-012-0510-0>
- Cannon SH, Gartner JE, Wilson RC, et al. (2008) Storm rainfall conditions for floods and debris flows from recently burned areas in southwestern Colorado and southern California. *Geomorphology* 96(3-4): 250-269.
<https://doi.org/10.1016/j.geomorph.2007.03.019>
- Chang M, Dou X, Hales TC, et al. (2021) Patterns of rainfall-threshold for debris-flow occurrence in the Wenchuan seismic region, Southwest China. *B Eng Geol Environ* 80: 2117-2130.
<https://doi.org/10.1007/s10064-020-02080-7>
- Chinkulkijniwat A, Salee R, Horpibulsuk S, et al. (2022) Landslide rainfall threshold for landslide warning in Northern Thailand. *Geomat. Nat Haz Risk* 13(1): 2425-2441.
<https://doi.org/10.1080/19475705.2022.2120833>
- Coe JA, Kinner DA, Godt JW (2008) Initiation conditions for debris flows generated by runoff at Chalk Cliffs, central Colorado. *Geomorphology* 96(3-4): 270-297.
<https://doi.org/10.1016/j.geomorph.2007.03.017>
- Cui P, Chen X, Wang Y, et al. (2005) Jiangjia Ravine debris flows in the south-western China. In: Jakob M, Hungr O (Eds), *Debris-flow hazards and related phenomena*. Springer, Berlin, pp 565-594.
https://doi.org/10.1007/3-540-27129-5_22
- Cui P, Zhou GGD, Zhu XH, et al. (2013) Scale amplification of natural debris flows caused by cascading landslide dam failures. *Geomorphology* 182: 173-189.
<https://doi.org/10.1016/j.geomorph.2012.11.009>
- Devoli G, Tiranti D, Cremonini R, et al. (2018) Comparison of landslide forecasting services in Piedmont (Italy) and Norway, illustrated by events in late spring 2013. *Nat Hazards Earth Syst Sci* 18(5): 1351-1372.
<https://doi.org/10.5194/nhess-18-1351-2018>
- Dowling CA, Santi PM (2014) Debris flows and their toll on human life: a global analysis of debris-flow fatalities from 1950 to 2011. *Nat Hazards* 71: 203-227.
<https://doi.org/10.1007/s11069-013-0907-4>
- García-Urquía E (2016) Establishing rainfall frequency contour lines as thresholds for rainfall-induced landslides in Tegucigalpa, Honduras, 1980–2005. *Nat Hazards* 82: 2107-2132.
<https://doi.org/10.1007/s11069-016-2297-x>
- Glade T, Crozier M, Smith P (2000) Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical “Antecedent Daily Rainfall Model”. *Pure Appl Geophys* 157: 1059-1079.
<https://doi.org/10.1007/s000240050017>
- Guo X, Li Y, Cui P, et al. (2020) Intermittent viscous debris flow formation in Jiangjia Gully from the perspectives of hydrological processes and material supply. *J Hydrol* 589: 125184.
<https://doi.org/10.1016/j.jhydrol.2020.125184>
- Guzzetti F, Peruccacci S, Rossi M, et al. (2008) The rainfall intensity–duration control of shallow landslides and debris flows: an update. *Landslides* 5: 3-17.
<https://doi.org/10.1007/s10346-007-0112-1>
- Hanssen AW, Kuipers WJA (1965) On the relationship between the frequency of rain and various meteorological parameters. *Meded Verh* 81: 2-15
- Hawke R, McConchie J (2011) In situ measurement of soil moisture and pore-water pressures in an ‘incipient’ landslide: Lake Tutira, New Zealand. *J Environ Manag* 92(2): 266-274.
<https://doi.org/10.1016/j.jenvman.2009.05.035>
- Hirschberg J, Badoux A, McArdell BW, et al. (2021) Evaluating methods for debris-flow prediction based on rainfall in an Alpine catchment. *Nat Hazards Earth Syst Sci* 21(9): 2773-2789.
<https://doi.org/10.5194/nhess-21-2773-2021>
- Hu M, Wang R, Shen J (2011) Rainfall, landslide and debris flow intergrowth relationship in Jiangjia Ravine. *J Mt Sci* 8: 603-610.
<https://doi.org/10.1007/s11629-011-2131-6>
- Hürlimann M, Coviello V, Bel C, et al. (2019) Debris-flow monitoring and warning: review and examples. *Earth-Sci Rev* 199: 102981.
<https://doi.org/10.1016/j.earscirev.2019.102981>
- Imaizumi F, Masui T, Yokota Y, et al. (2019) Initiation and runout characteristics of debris flow surges in Ohya landslide scar, Japan. *Geomorphology* 339: 58-69.
<http://doi.org/10.1016/j.geomorph.2019.04.026>
- Iverson RM (1997) The physics of debris flows. *Rev Geophys* 35(3): 245-296.
<https://doi.org/10.1029/97RG00426>
- Iverson RM, Reid ME, Lahusen RG (1997) Debris-flow mobilization from landslides. *Annu Rev Earth Planet Sci* 25: 85-136.
<https://doi.org/10.1146/annurev.earth.25.1.85>
- Jiang Z, Fan X, Subramanian SS, et al. (2021). Probabilistic rainfall thresholds for debris flows occurred after the Wenchuan earthquake using a Bayesian technique. *Eng Geol* 280: 105965.
<https://doi.org/10.1016/j.enggeo.2020.105965>
- Kean JW, Staley DM, Cannon SH (2011) In situ measurements of post-fire debris flows in southern California: comparisons of the timing and magnitude of 24 debris-flow events with rainfall and soil moisture conditions. *J Geophys Res* 116(F4): F04019.
<https://doi.org/10.1029/2011JF002005>
- Kean JW, McCoy SW, Tucker GE, et al. (2013) Runoff-generated debris flows: observations and modeling of surge initiation, magnitude, and frequency. *J Geophys Res* 118: 2190-2207.
<https://doi.org/10.1002/jgrf.20148>
- Kean JW, Staley DM, Lancaster JT, et al. (2019) Inundation, flow dynamics, and damage in the 9 January 2018 Montecito debris-flow event, California, USA: opportunities and challenges for post-wildfire risk assessment. *Geosphere* 15(4): 1140-1163.
<https://doi.org/10.1130/GES02048.1>
- Kinnell PIA (2023) An I3o focused approach to estimating event erosivity in Australia. *Catena* 226: 107052.
<https://doi.org/10.1016/j.catena.2023.107052>
- Leonarduzzi E, Molnar P, McArdell BW (2017) Predictive performance of rainfall thresholds for shallow landslides in Switzerland from gridded daily data. *Water Resour Res* 53(8): 6612-6625.
<https://doi.org/10.1002/2017WR021044>
- Li Y, Meng X, Guo P, et al. (2021) Constructing rainfall thresholds for debris flow initiation based on critical discharge and S-hydrograph. *Eng Geol* 280: 105962.
<https://doi.org/10.1016/j.enggeo.2020.105962>

- Liu Z (2023) Evaluation of rainfall thresholds triggering debris flows in western China with gauged- and satellite-based precipitation measurement. *J Hydrol* 620: 129500. <https://doi.org/10.1016/j.jhydrol.2023.129500>
- Marino P, Peres DJ, Cancelliere A, et al. (2020) Soil moisture information can improve shallow landslide forecasting using the hydrometeorological threshold approach. *Landslides* 17: 2041-2054. <https://doi.org/10.1007/s10346-020-01420-8>
- Marra F, Nikolopoulos EI, Creutin JD, et al. (2016) Space-time organization of debris flows-triggering rainfall and its effect on the identification of the rainfall threshold relationship. *J Hydrol* 541: 246-255. <https://doi.org/10.1016/j.jhydrol.2015.10.010>
- Martinengo M, Zugliani D, Rosatti G (2023) Validation and potential forecast use of a debris-flow rainfall threshold calibrated with the Backward Dynamical Approach. *Geomorphology* 421: 108519. <https://doi.org/10.1016/j.geomorph.2022.108519>
- McCoy SW, Kean JW, Coe JA, et al. (2012) Sediment entrainment by debris flows: in situ measurements from the headwaters of a steep catchment. *J Geophys Res* 117(F3): F03016. <https://doi.org/10.1029/2011JF002278>
- Nieto N, Chamorro A, Echaveguren T, et al. (2021) Development of fragility curves for road embankments exposed to perpendicular debris flows. *Geomat Nat Haz Risk* 12(1): 1560-1583. <https://doi.org/10.1080/19475705.2021.1935330>
- Nikolopoulos EI, Destro E, Maggioni V, et al. (2017) Satellite rainfall estimates for debris flow prediction: an evaluation based on rainfall accumulation-duration thresholds. *J Hydrometeorol* 18(8): 2207-2214. <https://doi.org/10.1175/JHM-D-17-0052.1>
- Oorthuis R, Hürlimann M, Vaunat J, et al. (2023) Monitoring the role of soil hydrologic conditions and rainfall for the triggering of torrential flows in the Rebaixader catchment (Central Pyrenees, Spain). *Landslides* 20: 249-269. <https://doi.org/10.1007/s10346-022-01975-8>
- Osanai N, Shimizu T, Kuramoto K, et al. (2010) Japanese early-warning for debris flows and slope failures using rainfall indices with Radial Basis Function Network. *Landslides* 7: 325-338. <https://doi.org/10.1007/s10346-010-0229-5>
- Pastorello R, D'Agostino V, Hürlimann M (2020) Debris flow triggering characterization through a comparative analysis among different mountain catchments. *Catena* 186: 104348. <https://doi.org/10.1016/j.catena.2019.104348>
- Penna D, Tromp-van Meerveld HJ, Gobbi A, et al. (2011) The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. *Hydrol Earth Syst Sci* 15(3): 689-702. <https://doi.org/10.5194/hess-15-689-2011>
- Peres DJ, Cancelliere A, Greco R, et al. (2018) Influence of uncertain identification of triggering rainfall on the assessment of landslide early warning thresholds. *Nat Hazards Earth Syst Sci* 18(2): 633-646. <https://doi.org/10.5194/nhess-18-633-2018>
- Rana H, Babu GLS (2022) Regional back analysis of landslide events using TRIGRS model and rainfall threshold: an approach to estimate landslide hazard for Kodagu, India. *B Eng Geol Environ* 81: 160. <https://doi.org/10.1007/s10064-022-02660-9>
- Schoener G, Stone MC (2019) Impact of antecedent soil moisture on runoff from a semiarid catchment. *J Hydrol* 569: 627-636. <https://doi.org/10.1016/j.jhydrol.2018.12.025>
- Siman-Tov S, Marra F (2023) Antecedent rainfall as a critical factor for the triggering of debris flows in arid regions. *Nat Hazards Earth Syst Sci* 23(3): 1079-1093. <https://doi.org/10.5194/nhess-23-1079-2023>
- Staley DM, Kean JW, Cannon SH, et al. (2013) Objective definition of rainfall intensity-duration thresholds for the initiation of post-fire debris flows in southern California. *Landslides* 10: 547-562. <https://doi.org/10.1007/s10346-012-0341-9>
- Staley DM, Negri JA, Kean JW, et al. (2017) Prediction of spatially explicit rainfall intensity-duration thresholds for post-fire debris-flow generation in the western United States. *Geomorphology* 278: 149-162. <https://doi.org/10.1016/j.geomorph.2016.10.019>
- Thomas MA, Lindsay DN, Cavagnaro DB, et al. (2023) The rainfall intensity-duration control of debris flows after wildfire. *Geophys Res Lett* 50(10): e2023GL103645. <https://doi.org/10.1029/2023GL103645>
- Tropeano D, Casagrande A, Luino F, et al. (1996) Processi di mud-debris flow in Val Cenischia (Alpi Graie): osservazioni nel bacino del T. Marderello. *Quaderni Di Studi E Di Documentazione* 20: 5-31 (In Italian).
- Tsunetaka H, Hotta N, Imaizumi F, et al. (2021) Variation in rainfall patterns triggering debris flow in the initiation zone of the Ichino-sawa torrent, Ohya landslide, Japan. *Geomorphology* 375: 107529. <https://doi.org/10.1016/j.geomorph.2020.107529>
- Uwihirwe J, Hrachowitz M, Bogaard TA (2020) Landslide precipitation thresholds in Rwanda. *Landslides* 17: 2469-2481. <https://doi.org/10.1007/s10346-020-01457-9>
- Yang H, Hu K, Zhang S, et al. (2023) Feasibility of satellite-based rainfall and soil moisture data in determining the triggering conditions of debris flow: the Jiangjia Gully (China) case study. *Eng Geol* 315: 107041. <https://doi.org/10.1016/j.enggeo.2023.107041>
- Yang H, Zhang S, Hu K, et al. (2022) Field observation of debris-flow activities in the initiation area of the Jiangjia Gully, Yunnan Province, China. *J Mt Sci* 19: 1602-1617. <https://doi.org/10.1007/s11629-021-7292-3>
- Zeng QL, Yue ZQ, Yang ZF, et al. (2009) A case study of long-term field performance of check-dams in mitigation of soil erosion in Jiangjia stream, China. *Environ Geol* 58: 897-911. <https://doi.org/10.1007/s00254-008-1570-z>
- Zhao Y, Meng X, Qi T, et al. (2022) Extracting more features from rainfall data to analyze the conditions triggering debris flows. *Landslides* 19: 2091-2099. <https://doi.org/10.1007/s10346-022-01893-9>
- Zhou W, Tang C (2014) Rainfall thresholds for debris flow initiation in the Wenchuan earthquake-stricken area, southwestern China. *Landslides* 11: 877-887. <https://doi.org/10.1007/s10346-013-0421-5>