

Rainfall thresholds of debris flows built with assistance of artificial intelligence in a small catchment

Siling Zhang^{a,b}, Xiaojun Guo^{a,*}, Jianyi Cheng^{a,b}, Yong Li^a

^a Key Laboratory of Mountain Hazards and Surface Process/Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610041, PR China

^b University of Chinese Academy of Sciences, Beijing 100081, PR China

ARTICLE INFO

Keywords:

Debris flow
Rainfall threshold
Rainfall parameter
Rain gauge
Uncertainty

ABSTRACT

Identification of rainfall thresholds remains a key problem in debris flow forecasting. Usually, a rainfall threshold is derived through linking rainfall and debris flow occurrence; however, the results are largely inaccurate in many cases. In this work, artificial intelligence methods were employed to select the optimal parameters for forecasting, and to evaluate the uncertainties of different parameter selections in building rainfall thresholds. The optimal procedure considers that a debris flow occurs at the peak of the rainfall event, uses average values of rainfall recorded by rain gauges installed within the source area, and incorporates more than three rainfall parameters. Selection of the rainfall parameters greatly influences the derived results, and increasing number of parameters can enhance the accuracy of predictions. The model incorporating both the antecedent rainfall and the event rainfall is proposed and thresholds are derived from the relations of antecedent rainfall and the accumulative amount and duration of event rainfall. The thresholds were found to vary with types of debris flows, suggesting effects of antecedent rainfall. Despite the “black box” nature of AI methods, the proposed approach can help select the optimal parameters and models, reduce the uncertainty in deriving rainfall thresholds, and further enhance understanding of the mechanism of debris flow formation.

1. Introduction

Debris flows are triggered primarily by heavy rainfall in catchments, and the frequency is expected to increase in response to intensification of the hydrological cycle caused by global warming (Gariano and Guzzetti, 2016). One of the primary objectives of debris flow research is to enhance capability of forecasting (Hürlimann et al., 2019), for which establishing appropriate rainfall thresholds remains an important goal. Such thresholds can be defined either by adopting physical approaches that employ a physical models (e.g., Tang et al., 2019; Berti et al., 2020; Guo et al., 2021; Hoch et al., 2021), or by employing statistical methods (e.g., Caine, 1980; Guzzetti et al., 2007, 2008).

Establishing a threshold is realized by taking rainfall data and debris flow occurrence records as “input” and producing the threshold as “output” using a generic “model” based on a single equation or a set of operations (Martinengo et al., 2021). Many critical thresholds and models have been proposed using precipitation parameters, the selection of which is dependent on the characteristics of local rainfall. Unfortunately, such thresholds are largely inaccurate and have rarely been used for warning system (Staley et al., 2013; Nikolopoulos et al., 2014;

2015). Earlier studies have discussed the uncertainties in rainfall threshold, e.g., those linked to the spatial variability and temporal resolution of data (Nikolopoulos et al., 2014; Marra et al., 2014, 2016; Gariano et al., 2020), the choice of reference rain gauges (Peres et al., 2018; Abraham et al., 2020; Crema et al., 2023), and the differences between radar rainfall and rain gauge measurements (Rossi et al., 2017; Nikolopoulos et al., 2015). Although procedures have been proposed to filter the effects of rainfall estimation uncertainty, overcoming such uncertainty remains a challenge, which contributes to the inaccuracy of forecasting (Guzzetti et al., 2007; Jakob et al., 2012).

Artificial intelligence (AI) methods have recently been used in forecasting, which can help with linking rainfall to debris flow occurrence (e.g., Mondini et al., 2023), selecting parameters (e.g., Zhao et al., 2022b), and calculating rainfall thresholds at different exceedance probabilities automatically (e.g., Melillo et al., 2018). Despite the “black box” in nature, it provides a framework for establishing better-optimized thresholds under constrained conditions. In this study, the influence of the selections in each procedure and the related uncertainties transferred to the forecasting results were evaluated, and the most important rainfall parameters were identified using AI methods. The optimal

* Corresponding author.

E-mail address: aaronguo@imde.ac.cn (X. Guo).

<https://doi.org/10.1016/j.jhydrol.2025.133440>

Received 20 September 2024; Received in revised form 29 April 2025; Accepted 1 May 2025

Available online 7 May 2025

0022-1694/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

model and specific rainfall thresholds were proposed for different types of debris flows based on actual monitoring data.

2. Case study area and data source

2.1. General introduction of Jiangjia Gully

This study focused on Jiangjia Gully (JJG) in the Xiaojiang River, a tributary of the Jinsha River in Southwest China (Fig. 1). The catchment covers an area of 48.6 km² and extends 13.9 km from the divide (elevation of 3269 m) to the outlet (1042 m). It is frequently impacted by tectonic activities (e.g., earthquakes), and characterized by deeply-cut terrain. Approximately 80 % of the exposed rocks are highly fractured and mildly metamorphosed, representing an abundant source of material for debris flows (Guo et al., 2020, 2021).

The catchment can be divided into three climatic regimes: (1) mean annual precipitation (MAP) of 850–1200 mm in the headwater area at elevation > 2200 m, (2) MAP of 700–850 mm at elevation of 1600–2200 m, and (3) MAP of 600–700 mm in the area extending from the outlet to the elevation of 1600 m. The variations in climate affect the distributions of vegetation, rock weathering, slope failure, and flow generation, thereby contributing in varying degrees to debris flow occurrence (Cui et al., 2005; Guo et al., 2020, 2021). The rainy season (May–September) accounts for > 80 % of the MAP and witnesses the frequent occurrence of debris flows.

2.2. Data source

JJG has long-term records of rainfall and debris flow. Ten rain gauges have been installed, of which four (G1 to G4) are located within the major source area (Menqian Gully; Fig. 1). These gauges measure rainfall using a 0.1-mm tipping bucket, and the data are transmitted in real time via the General Packet Radio Service.

Debris-flow parameters include the front and end times of a surge, surge numbers, flow velocity, height, and density, all are measured manually and recorded at the monitoring section (Fig. 1). The front and end time, the duration and surge number were recorded directly. The

flow velocity was determined by measuring the interval of the front passing through two sections separating 200 m in the channel, and flow height were estimated by experienced experts. Additionally, samples of the flow bodies were collected using a volume-calibrated sampling container controlled by electronic devices for density and volume sediment concentration analyses. The recorded data, especially the accurate time (1-s time scale) of debris flow detection at the monitoring section, represent an important basis for this analysis, and the errors of the source data are minimized as much as possible.

Rainfall and debris flow data recorded during 2006–2023 were used in this work. All 40 of the rainfall events known to have triggered debris flows were involved, and an additional 100 other rainfall events (daily rainfall range: 6.5–42.8 mm) that occurred during the same period but did not result in debris flow occurrence were also considered.

3. Methodology

3.1. Framework for building rainfall thresholds

The data processing includes discretizing continuous rainfall records into a series of individual rainfall events and identification of debris flow initiation time in the source regions. Identification of relevant rainfall parameters (rainfall amount, rainfall intensity, duration, antecedent rainfall, etc.) and selection of rain gauges are also important which might cause notable uncertainty to the rainfall thresholds (Hirschberg et al., 2021).

AI methods were used to analyze uncertainties in data processing and to identify the optimal combination of process and parameters, i.e., the Nonlinear Gaussian Kernel Support Vector Machine (NGK-SVM), was used to determine the optimal combination of processes, and SHapley Additive exPlanations (SHAP), was used to explain and verify the results of NGK-SVM. Two metrics, Accuracy (A_c) and the F_2 -score (F_2), were defined to evaluate the effectiveness of the parameters and selections recommended by the AI methods. The methods are explained in detail in Section 3.2. A flowchart of the overall procedure is shown in Fig. 2.

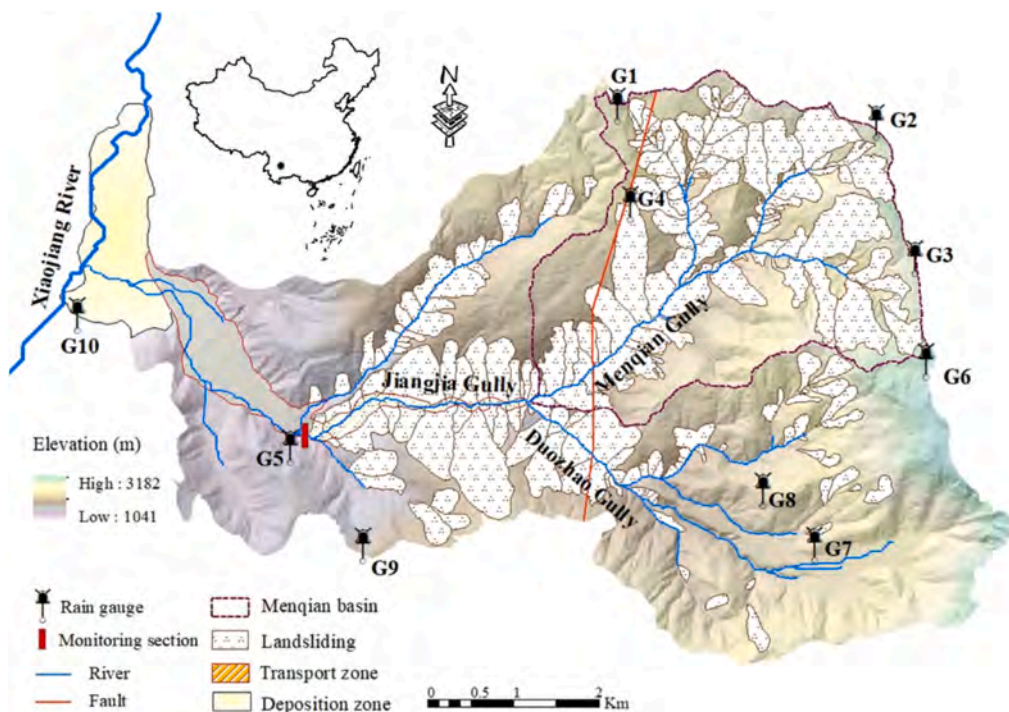


Fig. 1. Topographic map and locations of monitoring stations in Jiangjia Gully.

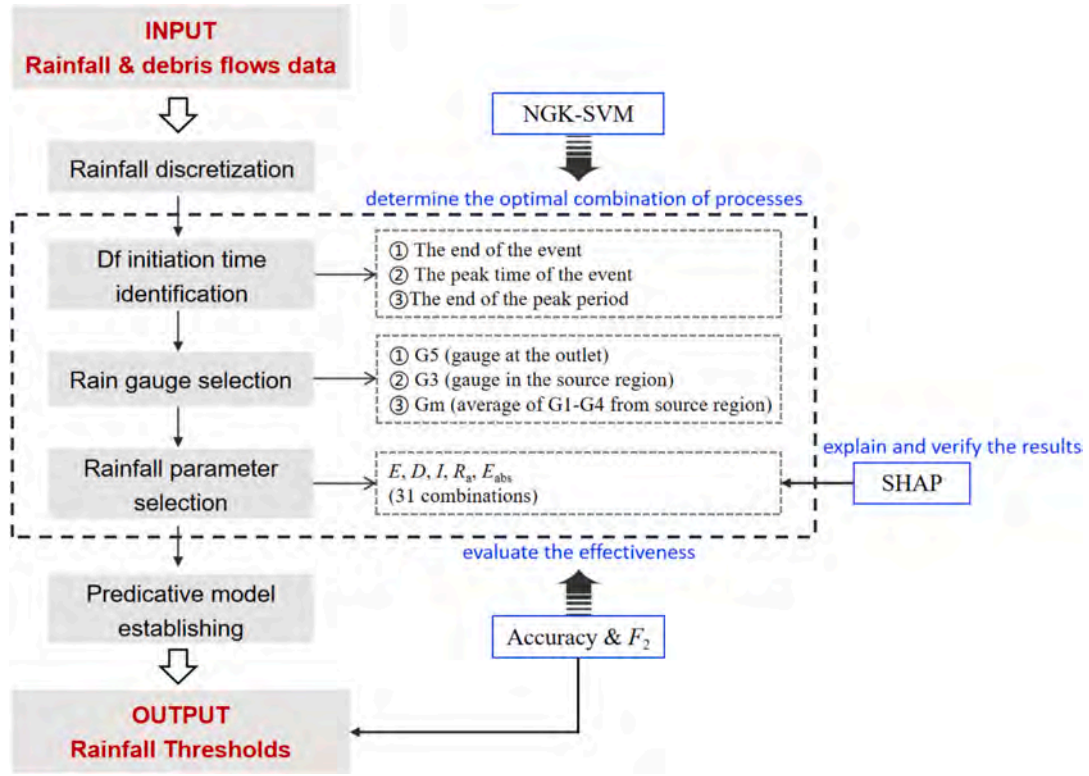


Fig. 2. Flowchart of the process of rainfall threshold establishment. The content within the blue solid boxes refer to the methods of building rainfall thresholds (explained in Section 3.2). The content within the grey dashed boxes represent parameters/selections in the process of building rainfall thresholds (explained in Section 4).

3.2. Methods for building rainfall thresholds

(1) Nonlinear Gaussian Kernel Support Vector Machine for selecting parameters.

NGK-SVM is a type of machine learning model that can be used to evaluate the performance of different selections with respect to debris flow initiation time, rain gauges, and the rainfall parameters. It can also help to analyze the impact on the results of those selections at each step

of the process.

A traditional linear Support Vector Machine (SVM) can derive a threshold with linear form; however, it does not distinguish between events occurring or not (Fig. 3a); and a nonlinear SVM can do this by providing an ideal threshold. But this might be overfitted and cause errors (Fig. 3b). In comparison, NGK-SVM can handle more complex nonlinear relationships (Vapnik, 1998; Bishop and Nasrabadi, 2006). The input dataset is transformed into a spatial dataset using the Gaussian

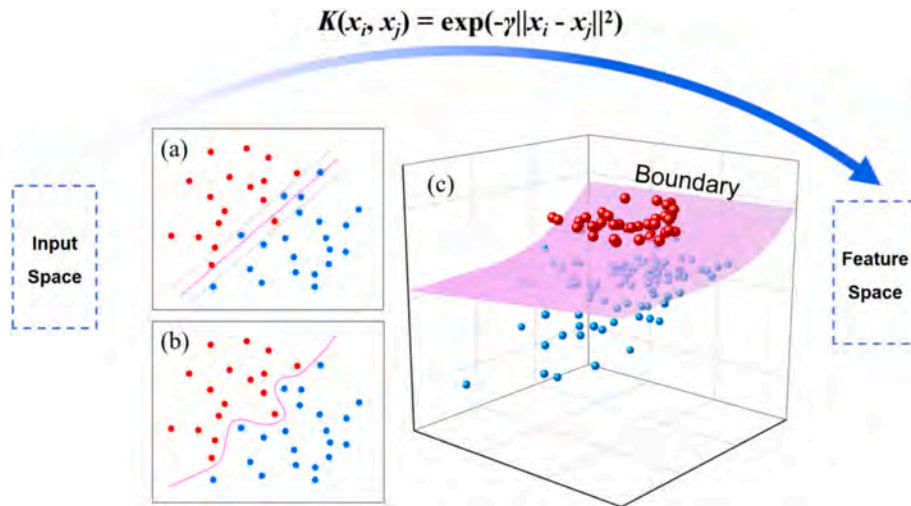


Fig. 3. Modeling principle of NGK-SVM. (a) threshold with linear form, (b) an ideal threshold using nonlinear SVM, and (c) a curved surface boundary obtained by NGK-SVM. Input Space: the original space where data are represented. For example, for a dataset with features such as rainfall and time, the input space is a two-dimensional plane where each point represents a combination of rainfall and time. Feature space: the space where data are represented after being transformed or mapped from the input space. In the context of NGK-SVM, this transformation is achieved using a Gaussian kernel function. Red and blue dots represent debris flow and non-debris flow events, respectively.

kernel function, and thereby a curved surface boundary is obtained as a threshold to distinguish occurrence and non-occurrence events (Fig. 3c). However, as a “black box”, the boundary is insufficient to elucidate the physical importance between the parameters, and can sometimes be excessively intricate as a threshold. Therefore, NGK-SVM is employed only to help determine the optimal procedures and parameters, rather than to establish a threshold.

(2) SHapley Additive exPlanations for evaluating rainfall parameters.

The SHAP method provides a unified framework that explains the output of any machine learning model by attributing the prediction to the individual feature in a way that is both consistent and interpretable (Lundberg and Lee, 2017). It assigns a value to each feature that represents its contribution to the model's output for a given prediction. It was used primarily to assess the importance of rainfall parameters in this study, and the value can be calculated as follows:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \quad (1)$$

where ϕ_i is the SHAP value for rainfall parameter i in NGK-SVM model. N is the set of all parameters. $|N|$ is the number of parameters in set N . S is a subset of parameters without containing parameter i . $|S|$ is the number of parameters in subset S . $f(S)$ is the model's output under the condition of subset S and $f(S \cup \{i\})$ is model's output with parameter i added to the subset S .

For example, For instance, consider a scenario where we have three rainfall parameters $N = \{D, E, I\}$, and the aim to calculate the SHAP value for parameter $i = D$. The possible subsets S of parameters excluding parameters D are: $S_1 = \emptyset$, $S_2 = \{E\}$, $S_3 = \{I\}$, $S_4 = \{E, I\}$. Then ϕ_i for every subset is calculated as:

$$\phi_{D, S_1} = \frac{0!(3 - 0 - 1)!}{3!} [f(S_1 \cup \{D\}) - f(S_1)] \quad (2)$$

$$\phi_{D, S_2} = \frac{1!(3 - 1 - 1)!}{3!} [f(S_2 \cup \{D\}) - f(S_2)]$$

$$\phi_{D, S_3} = \frac{1!(3 - 1 - 1)!}{3!} [f(S_3 \cup \{D\}) - f(S_3)]$$

$$\phi_{D, S_4} = \frac{2!(3 - 2 - 1)!}{3!} [f(S_4 \cup \{D\}) - f(S_4)]$$

The SHAP value of parameter D is therefore obtained by accumulating through Eq. (3):

$$\phi_D = \phi_{D, S_1} + \phi_{D, S_2} + \phi_{D, S_3} + \phi_{D, S_4} \quad (3)$$

(3) Indicators for evaluating the thresholds

The metrics A_c and F_β were used as the major indicators (Sokolova and Lapalme, 2009) for assessing the quality of the rainfall thresholds: A_c for the performance of classification models, and F_β for the performance with precision (P) and recall (R).

Specifically, A_c measures the proportion of correctly classified instances out of the total number in the dataset:

$$A_c = (TP + TN) / (TP + TN + FP + FN) \quad (4)$$

where TP (True Positives) and TN (True Negatives) represent the number of positive and negative instances correctly predicted, respectively, and FP (False Positives) and FN (False Negatives) represent the number of negative and positive instances incorrectly predicted as positive and negative, respectively.

F_β evaluates the performance of classification models with consideration of both precision (P) and recall (R), which represent false alarms and missed detections, respectively:

$$F_\beta = (1 + \beta^2) \cdot P \cdot R / (\beta^2 \cdot P + R) \quad (5)$$

$$P = TP / (TP + FP) \quad (6)$$

$$R = TP / (TP + FN) \quad (7)$$

where β is a weight parameter.

The values conventionally used for β are 0.5, 1, and 2 (Witten and Frank, 2002). A bigger β value gives a higher weight of R . Because the impact of a missed detection is much greater than that of a false alarm in debris flow prediction, β was set to 2 in this study. The formula for calculation of F_2 can be simplified to the following:

$$F_2 = 5TP / (5TP + 4FN + FP) \quad (8)$$

4. Procedure for deriving rainfall thresholds

The key steps in deriving rainfall thresholds include identification of rainfall events, determination of debris flow initiation times, selection of representative rain gauges and rainfall parameters, and deriving a threshold based on the parameters.

4.1. Identification of rainfall events

In most cases, rainfall is a complex signal exhibiting fluctuations and irregular behavior on multiple temporal scales. Only after identification of a rainfall event can the required rainfall parameters (e.g., amount, duration, and intensity) be derived. A common method is to define a minimum inter-event duration ($D_{\min}$) and maximum inter-event amount ($P_{\max}$) that discretize continuous rainfall records into a series of individual rainfall events with rainfall amount less than $P_{\max}$ in $D_{\min}$ (Guo et al., 2016; Jiang et al., 2021; Hirschberg et al., 2021).

The definitions of $D_{\min}$ and $P_{\max}$ vary between catchments and depend on local hydrogeological conditions. For small catchments, $D_{\min}$ ranges from 10 min to 6 h and $P_{\max}$ ranges in 0.1–1.0 mm (e.g., Bel et al., 2017; Berti et al., 2020; Hirschberg et al., 2021). For the JJG, antecedent rainfall events have been distinguished previously as occurring within 3 h (or more) before rainfall of < 0.1 mm or within 6 h of rainfall of < 0.5 mm (i.e., $D_{\min} = 3$ h, $P_{\max} = 0.1$ mm; or $D_{\min} = 6$ h, $P_{\max} = 0.5$ mm), identified by local meteorological conditions (Guo et al., 2013, 2020; Zhuang et al., 2015). Following this criteria, 140 rainfall events were determined in the period 2006–2023, of which 40 triggered debris flows.

4.2. Identification of debris flow initiation time

Although a debris flow can be identified precisely at a point, the exact time of its initiation in the source area remains unknown. The initiation time is set as a peak before the debris flow occurrence time, with consideration of the debris flow travel time from the source area to the monitoring section (e.g., Berti et al., 1999; Arattano and Moia, 1999; Tecca and Genevois, 2009; Okano et al., 2012).

Debris flows are generally triggered by rainfall that occurred shortly before their appearance. This study considered the highest rainfall intensity within approximately 30 min are plausible to trigger the debris flow based on the velocity and travel length of debris flows (Guo et al., 2020). The potential triggering rainfall for debris flows at each gauge can be estimated. The following three scenarios for the triggering rainfall period were established (Fig. 4):

Scenario 1: from rainfall beginning to the end of the event (rainfall amount: E_1 , duration: D_1);

Scenario 2: from rainfall beginning to the peak time of the event (rainfall amount: E_2 , duration: D_2);

Scenario 3: from rainfall beginning to the end of the peak period (rainfall amount: E_3 , duration: D_3).

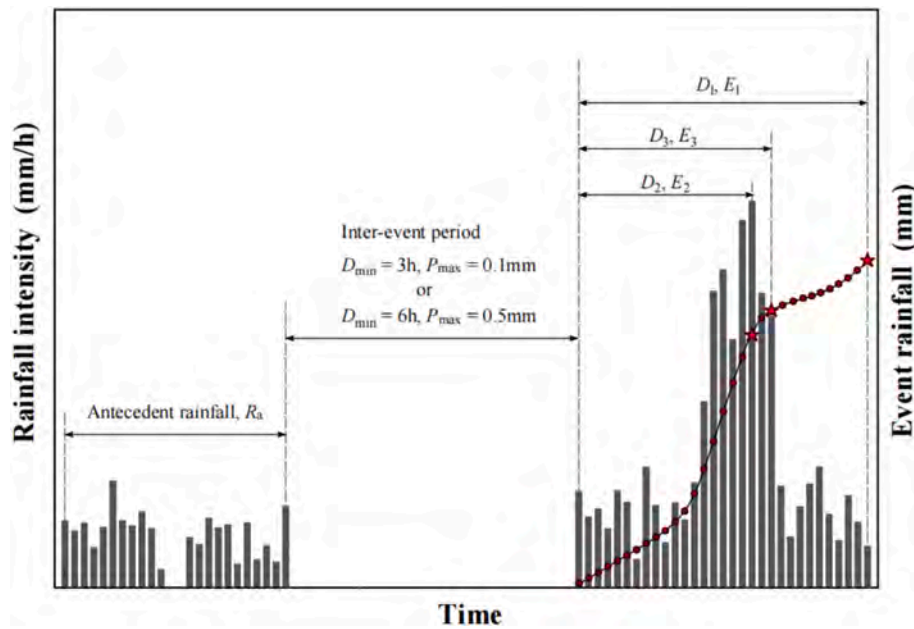


Fig. 4. Definition of the relevant rainfall and duration for each of the three scenarios.

4.3. Selection of representative rain gauges

In many cases, rain gauge selection is obligated because installing rain gauges in mountainous areas presents substantial challenges, and therefore the reference rain gauge is often the only gauge available. However, many studies have indicated that relying solely on a single rain gauge may not be the optimal approach, particularly in the context of localized convective storms. Rain gauges close to the initiation may record rainfall that differ significantly from those at the catchment outlet due to the spatiotemporal variability of rainfall (e.g., Smith et al., 2007; Staley et al., 2017; Guo et al., 2021; Crema et al., 2023).

Previous studies suggested that the distance between a debris flow source area and a rain gauge station should typically be limited to within 3–6 km within a mountainous region. For example, it is proposed that the distribution of triggering rainfall varies considerably over distance of < 5 km in alpine regions (Smith et al., 2007; Panziera et al., 2011). On average, it is observed that the depth of the triggering rainfall at a distance of 5 km is approximately 70 % of that estimated at the initiation point, whereas the estimation accuracy reduces to 40 % at a distance of approximate 6–10 km (Marra et al., 2016; Nikolopoulos et al., 2015).

The records in JJG indicate that rainfall is highly spatially heterogeneous, e.g., the gauges in Menqian Gully only 2-km apart present spatial variation in some events. Uncertainties arise notably when the distance between failure locations and rain gauge sites is > 3 km (Guo et al., 2021). The following three options were evaluated to select the optimal rain gauge:

- 1) G5, the lowest rain gauge (represents the station at the outlet);
- 2) G3, one of the rain gauges in the high-elevation headwater region;
- 3) Gm, the mean values of G1 to G4, representing the average rainfall conditions in the source regions.

4.4. Selection of rainfall parameters

Selection of appropriate rainfall parameters is important both for building rainfall thresholds and for improving prediction accuracy. The rainfall parameters that are generally considered include rainfall intensity (I , mm/h), rainfall duration (D , h), and event rainfall (E , mm), based on which I - D and E - D are the parameter pairings used most commonly for threshold building (e.g., Caine 1980; Aleotti 2004; Guzzetti et al., 2007, 2008; Guo et al., 2020). This study also considered the

additional parameters of antecedent rainfall (R_a , mm) and absolute energy (E_{abs} , mm²) when selecting the optimal triggering factor.

Antecedent rainfall (R_a) is generally calculated as follows:

$$R_a = \sum_{i=1}^n R_i(K)^i \quad (9)$$

where R_i is the rainfall amount recorded during the preceding n days, i indicates the number of days before the triggering rainfall event of the debris flow ($1 \leq i \leq n$), and K is an attenuation coefficient of the i -th day that represents the role of evaporation; for JJG, the value of K is taken as 0.8 (Cui et al., 2007).

Generally, n is set to 30 days (e.g., Cui et al., 2007; Tien Bui et al., 2013; Saadatkhah et al., 2015); however, its value was set to 3, 5, 7, 10, 15, 21, and 30 days for this evaluation. The increasing rate of the cumulative antecedent rainfall was used to assess the impact of additional rainfall on soil moisture as the value of n increased. The increasing rates

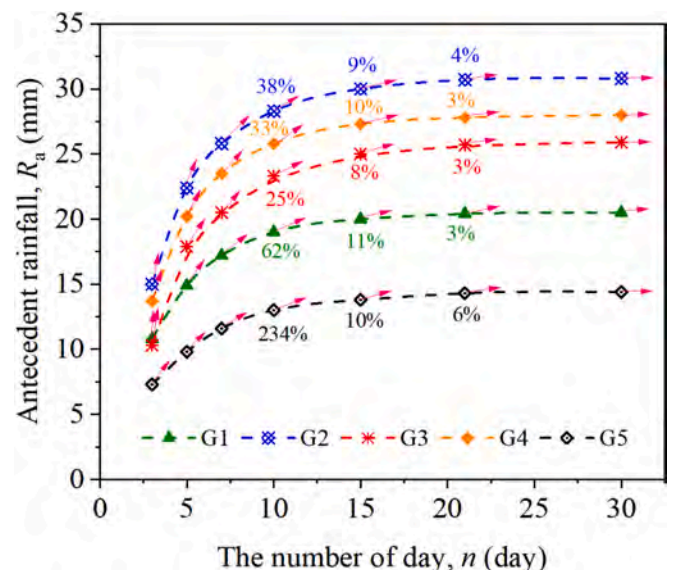


Fig. 5. Growth rate of antecedent rainfall.

of cumulative antecedent rainfall were $< 10\%$ for all the gauges of G1 to G5 after $n = 15$ (Fig. 5). Therefore, $n = 15$ was used in subsequent analysis.

The absolute energy (E_{abs}) was also introduced to reflect the intensity of changes in rainfall magnitude, which is defined as:

$$E_{abs} = \sum_{k=1}^m R_k^2 \quad (10)$$

where R_k is the rainfall amount (mm), and m the number of sampling periods in a rainfall event. For a given duration and total rainfall amount, rainfall events characterized by more abrupt variations exhibit higher E_{abs} values and pose a greater potential for triggering debris flows (Zhao et al., 2022b).

4.5. Methods for establishing rainfall thresholds

The objective of building a rainfall threshold is likely to define the lower boundary for debris flow occurrence based on the selected rainfall parameters. Numerous methods have been adopted for drawing rainfall thresholds, e.g., regression and frequentist methods, return time calculations, and Bayesian analysis (Brunetti et al., 2010; Peruccacci et al., 2012; Chen and Huang, 2010; Berti et al., 2012). These methods are well-suited for delineating the critical boundary between occurrence and non-occurrence of debris flows, particularly when the relationships among parameters are ambiguous. AI methods, with their advantages of offering objective data analysis, automation, and adaptability to evolving data, provide more flexible and adjustable solutions that can be optimized and improved as new data become available (Mondini et al., 2023). However, in this context, given the “black box” nature of AI methods, we derived a threshold model incorporating both the power relationship of I - D and the linear relationship of E - R_a (Jan and Lee, 2004; Jan and Chen, 2005; Caine et al., 1980; Brunetti et al., 2010; Peruccacci et al., 2012), based on the optimal parameters selected by NGK-SVM, to develop a model with physical significance.

5. Results

Specific identification of the parameters in each step (e.g., initiation time, rainfall parameters, and rain gauges) results in various combinations of the parameters available for building a rainfall threshold. In this case, 279 combinations were produced and evaluated, and the optimal combination for establishing rainfall thresholds was determined (Fig. 6). Moreover, the influence of each key step was examined, especially the importance of the selection of rainfall parameters to the performance of the thresholds. The rainfall thresholds were ultimately built based on the

optimal process and optimal combination of parameters.

5.1. Optimal process for establishing rainfall thresholds

The threshold boundary was obtained by the NGK-SVM and the predictive effect was evaluated by A_c and F_2 . The A_c of 92 % of the combinations exceeded 0.50, with a maximum of 0.84 and a mean of 0.66, suggesting that most of the results were reasonable. However, F_2 showed a range of variation and therefore it was used as the major indicator in the following evaluation.

The predictive performances are shown in Fig. 7. Each combination provided a certain threshold, and its efficiency was evaluated using the F_2 metric. The results were classified into six levels: (I) excellent: $F_2 \geq 0.80$, (II) good: $0.80 \geq F_2 > 0.75$, (III) satisfactory: $0.75 \geq F_2 > 0.70$, (IV) mediocre: $0.70 \geq F_2 > 0.60$, (V) poor: $0.60 \geq F_2 > 0.50$, and (VI) ineffective: $F_2 < 0.50$. The numbers in the boxes represent the quantity of F_2 at a certain level, and the histograms represent the proportion exceeding the satisfactory level ($F_2 > 0.70$).

The performance of these processes and the selection of the parameters are investigated on the basis of the criteria of exceeding the satisfactory ($F_2 \geq 0.70$) and the highest (I, $F_2 \geq 0.80$) levels.

In identifying debris flow initiation time, at the level of $F_2 \geq 0.70$, the proportion for Scenarios 1–3 is 30 %, 36 %, and 34 %, respectively, i.e., the result for Scenario 2 represents a slightly higher proportion. At level I, the performance of Scenario 2 is similar to Scenario 3, each accounting for 38 %, and the performance of both is notably superior to that of Scenario 1 (23 %). This result suggests that all three scenarios can provide predictive performance that is reasonably comparable; however, Scenario 2 offers the best effectiveness.

In selecting rain gauges, at the level of $F_2 \geq 0.70$, the prediction efficiency of Gm is slightly higher than that of G3 (42 % versus 40 %), and the value of each is much better than that of G5 (18 %). Notably, the performance of Gm is markedly superior to that of G3 at level I. This suggests that the selection of Gm and G3 would be effective, but that Gm would provide the optimal predictive effect.

The F_2 values of different combinations of parameters (an individual parameter and combinations of two and multiple (≥ 3) parameters) are used to evaluate the predictive performance. Combination of multiple (≥ 3) parameters yields much better performance, and individual parameter provides the worst results. At the level of $F_2 \geq 0.70$, combination of multiple parameters presents an evident advantage (64 % at level III, 70 % at level II) in comparison with the other two options. Specifically, the proportion increases as F_2 rises, reaching up to 96 % at level I. This suggests that employing a combination of multiple parameters markedly enhances predictive effectiveness.

Overall, the results of the evaluation are consistent at the levels of

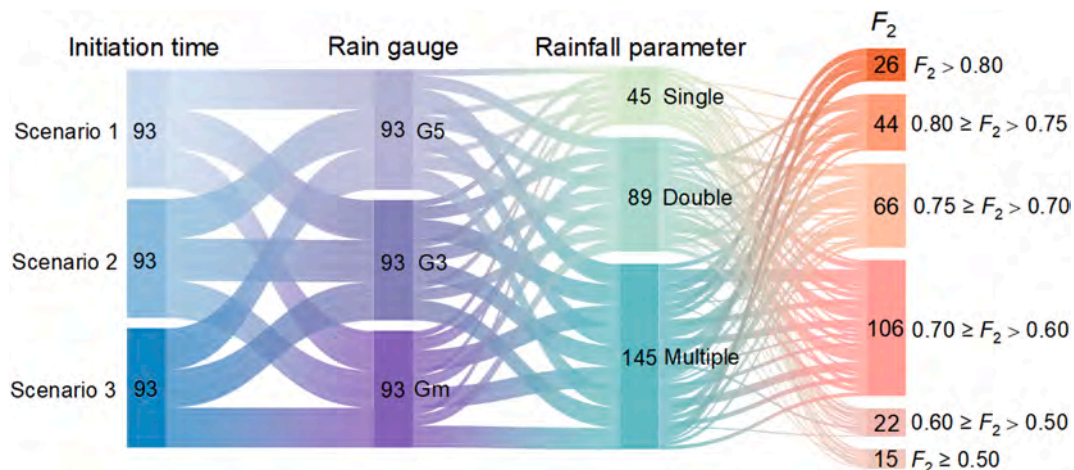


Fig. 6. Procedure and selections involved in deriving rainfall thresholds.

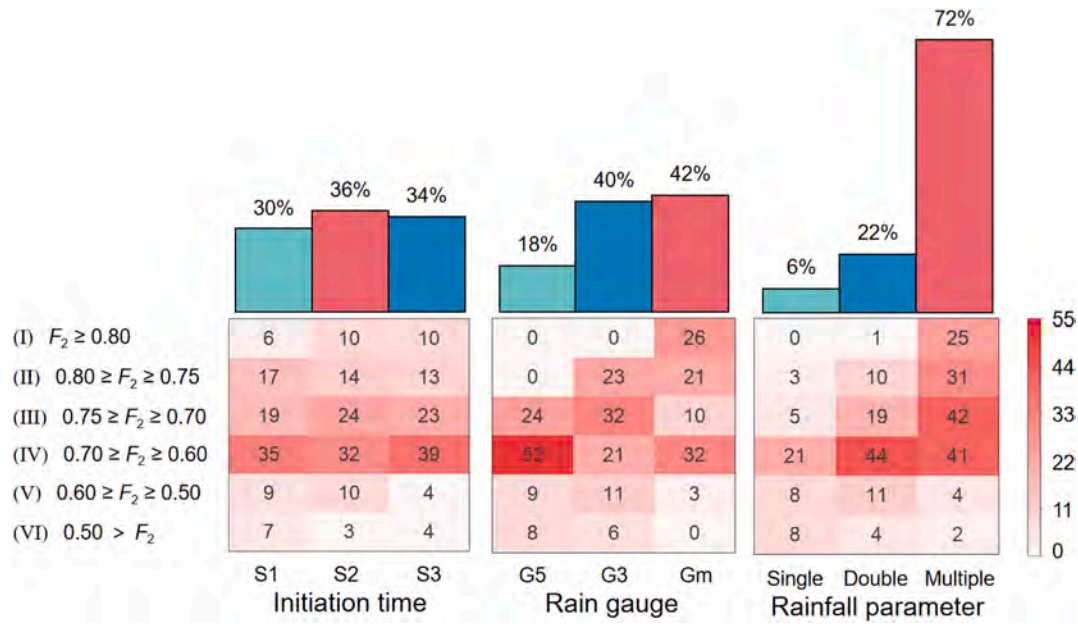


Fig. 7. Number and proportion of different selections at each step across different levels.

both $F_2 \geq 0.70$ and $F_2 \geq 0.80$, indicating satisfactory and excellent expectations, respectively. Ultimately, it is suggested that the optimal selection for establishing a rainfall threshold is to take the rainfall peak time as the debris flow initiation time, consider the average of rain gauge data from the source region as the data source, and combine multiple rainfall parameters.

5.2. Impact of the rainfall threshold procedure on the results

It is important to evaluate the impact of the procedures, including the identification of rainfall period, adoption of rain gauges, and selection of rainfall parameters, on the derived rainfall thresholds. The mean values of F_2 for the selections in each step were calculated (Table 1), which varies with scenarios of initiation-time determination and makes little differences (only of 0.01 in F_2 value). F_2 was improved by 0.04 and 0.05, respectively by selecting rain gauges of G5, G3, and Gm. And F_2 increased by 0.06 when rainfall parameters were selected from single to combination of two and multiple parameters. This indicates that changing the selections at different steps influences the results to unequal extents.

The increase proportion of F_2 resulting from individual step selections was used to assess the impacts (Table 2). Altering the debris flow initiation time had a minimal effect on F_2 , contributing only 1.4 % and 1.5 %. In contrast, the selection of rain gauges significantly influenced the results, accounting for an increase of 6.3 % and 7.2 % in F_2 . Notably, the selection of rainfall parameters exerted the greatest influence, contributing 9.8 % and 9.0 % to the overall increase in F_2 . This indicates that the rainfall models and parameters remain the most critical factor affecting prediction reliability. Furthermore, the selection of rain gauges also warrants significant consideration. Data from the lowest station introduces a much lower F_2 value compared to using the average data of rainfall gauges within source regions, which significantly enhances forecasting accuracy (0.65 versus 0.74). This indicate that a rain gauge

Table 1
Mean value of F_2 for different selections within the key steps.

Debris flow initiation time	Rain gauge	Rainfall parameter
0.68 (Scenario1)	0.65 (G5)	0.61 (Single)
0.69 (Scenario 2)	0.69 (G3)	0.67 (Double)
0.70 (Scenario 3)	0.74 (Gm)	0.73 (Multiple)

Table 2
The increase proportion of F_2 .

Debris flow initiation time	Rain gauge	Rainfall parameter
1.5 % (Scenario 1 → 2)	6.2 % (G5 → G3)	9.8 % (Single → Double)
1.4 % (Scenario 2 → 3)	7.2 % (G3 → Gm)	9.0 % (Double → Multiple)

located in the lower part of the catchment may not accurately represent the actual rainfall conditions that trigger debris flows due to the spatial heterogeneity of rainfall.

5.3. Influence of rainfall parameters on threshold predictions

In the following, the order of the importance of the selected parameters, and the influence on the thresholds are examined using SHAP values by taking A_c and F_2 as the major evaluation indices.

The SHAP values of the selected parameters were calculated (Fig. 8a) by taking the rainfall peak time (Scenario 2) as the debris flow initiation time and the average of the data from gauges (Gm) as the data source.

For each parameter, a positive (negative) SHAP value reflects positive (negative) impact. A positive impact means that a debris flow occurs and a negative impact means that a debris flow does not occur. The values of R_a , E , I , and E_{abs} are all positively correlated with the SHAP value, indicating that higher values of R_a , E , I , and E_{abs} are associated with greater likelihood of debris flow occurrence. In contrast, D is negatively correlated with the SHAP value, suggesting that larger values of D are associated with reduced likelihood of debris flow occurrence.

To express the relative importance of the parameters, the absolute average SHAP values for each parameter were calculated using all 140 samples, indicating the order of importance: $R_a > D > E > I > E_{abs}$ (Fig. 8b), which is broadly consistent with the order based on A_c and F_2 . The SHAP value evaluates the importance of a parameter by increasing or decreasing the specified parameter, and A_c and F_2 represent the predictive effective realized when using this specific parameter. Therefore, the derived order of parameter importance is considered reasonable.

It is evident that R_a and D rank as the most important parameters, and previous studies have similarly highlighted the relevance of R_a and D in building rainfall thresholds (Guo et al., 2013). The role of D is particularly distinctive because it exhibits negative correlation with

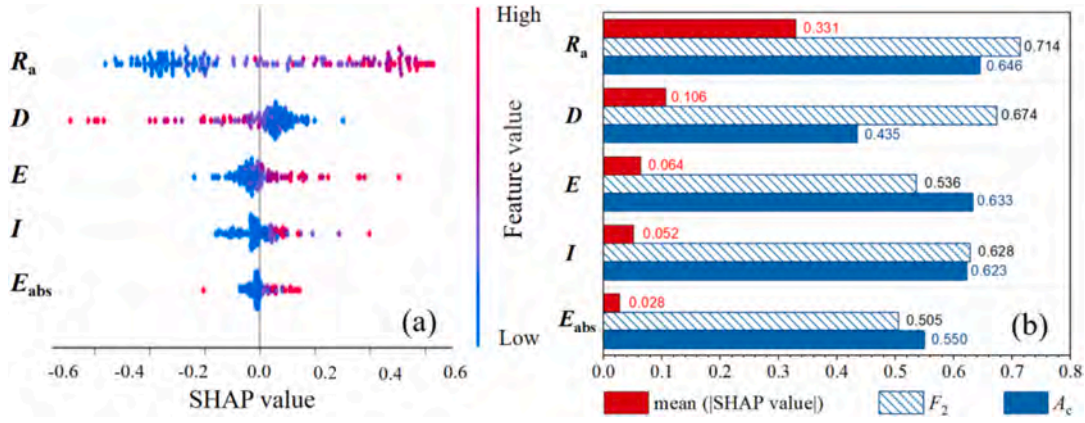


Fig. 8. SHAP values of different rainfall parameters: (a) SHAP values and (b) absolute averages of the SHAP values, A_c and F_2 for each rainfall parameter.

SHAP on the one hand, and demonstrates exceptionally low A_c (0.44) on the other hand. Specifically, the low A_c value attributes to the high false alarm rate ($P = 0.32$), and suggests that D cannot act as the forecasting parameter independently. Conversely, the high SHAP value indicates that its importance is reflected through combination with other parameters. The most commonly used combinations are E - D and/or I - D , in which E and I are usually regarded as directly responsible for triggering debris flows. When using E - D to build the thresholds, both A_c and F_2 are improved to 0.54 and 0.69 (Table 3), representing an improvement of 25.1 % and 2.8 %, respectively. The combination of D , E , and R_a further improves A_c and F_2 by 40.3 % and 5.7 % compared with the combination of D and E , reaching values of 0.76 and 0.75, respectively. The remarkable improvement of A_c emphasizes the important role of D in the combinations of parameters through reducing false alarms. Additionally, the combinations of D , E , I , and D , R_a , I are comparatively less effective than the combination of D , E , and R_a ($A_c = 0.60$, $F_2 = 0.69$, and $A_c = 0.73$, $F_2 = 0.73$, respectively).

However, if the remaining parameters (I and E_{abs}) are included in the combination, further improvements in A_c and F_2 are minimal. For example, when adding I to the D - E - R_a combination, A_c increased by only 0.9 % and F_2 increased by 1.7 %. When adding E_{abs} to the D - E - R_a - I combination, there was an increase of only 1.5 % for A_c and 1.0 % for F_2 . Despite the improvement, incorporating additional parameters inevitably increased the complexity of building rainfall thresholds. From this perspective, the D - E - R_a combination is regarded as the optimal combination for rainfall thresholds.

5.4. Optimal rainfall thresholds

(1) Rainfall thresholds.

The optimal rainfall thresholds for the JJG catchment were then derived based on the importance of the parameters. R_a yields the best prediction performance, and can acts as a threshold by its own. The result is shown in Fig. 9a:

$$R_a = C \quad (11)$$

Table 3

Performance of various combinations of parameters.

Rainfall parameter	A_c	F_2
D	0.44	0.67
D - I	0.54	0.66
D - E	0.54	0.69
D - E - I	0.60	0.69
D - I - R_a	0.73	0.73
D - E - R_a	0.76	0.75
D - E - R_a - I	0.77	0.77
D - E - R_a - I - E_{abs}	0.78	0.78

where C is a constant.

Then, D is combined with R_a to build a threshold owing to its important role in a combination of parameters. An exponential relation of D and R_a , proposed as Eq. (12), was used to build a threshold following previous research (Guo et al., 2013), and the result is shown in Fig. 9b:

$$D = \alpha \cdot \exp(\beta \cdot R_a) \quad (12)$$

where α and β are parameters.

Then E is added to the combination and with R_a and D . The relation between R_a and E is generally presented as a linear function in the expression of the rainfall threshold for debris flow occurrence (Jan and Lee, 2004), indicating that antecedent rainfall and event rainfall both make substantial contributions in parallel to debris flow occurrences. Therefore, the relation is preliminarily proposed as:

$$R_e = R_a + E \quad (13)$$

where R_e is the effective rainfall (mm).

The commonly used E - D relation assumes that the threshold curve follows a power law (Brunetti et al., 2010; Peruccacci et al., 2012):

$$E = aD^b \quad (14)$$

The relations of D , R_a , and R_e are then expressed as in Eq. (15) to build a rainfall threshold, and the result is shown in Fig. 9c:

$$R_e = aD^b + R_a \quad (15)$$

where a and b are parameters.

The given model was fitted to the debris flow data, followed by iterative adjustments of the boundaries (line and/or surface). The model parameters were then determined upon achieving the maximum F_2 -score.

The R_a threshold yields a high false alarm rate ($P = 0.39$) at $R_a = 6.8$. When taking $R_a = 11.5$, which is associated with the maximum F_2 (0.72), A_c is 0.54 and three debris flow events were missed detections (Fig. 9a).

The R_a - D threshold improves the forecasting effectiveness, e.g., 22.3 % in terms of A_c and 5 % in terms of F_2 when compared with the effectiveness realized using the threshold based only on R_a (Fig. 9b); and the D - E - R_a threshold further improves the effectiveness with $A_c = 0.81$ and $F_2 = 0.82$. In comparison with the R_a - D model, A_c and F_2 are improved by 21.5 % and 9.2 %, respectively (Fig. 9c). This threshold is preferable not only for its higher effectiveness, but also because of its physical relevance in considering both the event rainfall and the antecedent rainfall, and its combination with the commonly used E - D and E - R_a relations.

Comparing with the E - D model, which results in $A_c = 0.36$ and $F_2 = 0.62$ (Fig. 10a), the E - D - R_a model improves results remarkably, with A_c and F_2 increasing by 121.7 % and 32.2 %, respectively.

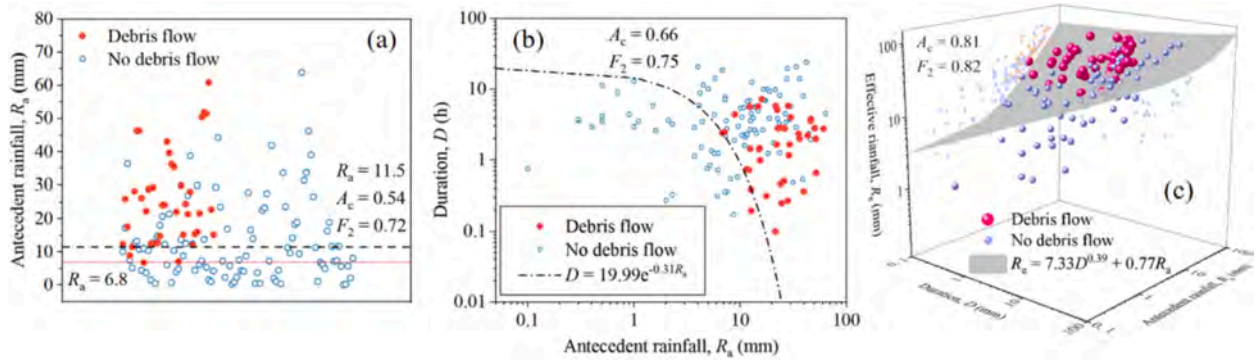


Fig. 9. Rainfall thresholds of the three models: (a) R_a , (b) R_a - D , and (c) E - D - R_a .

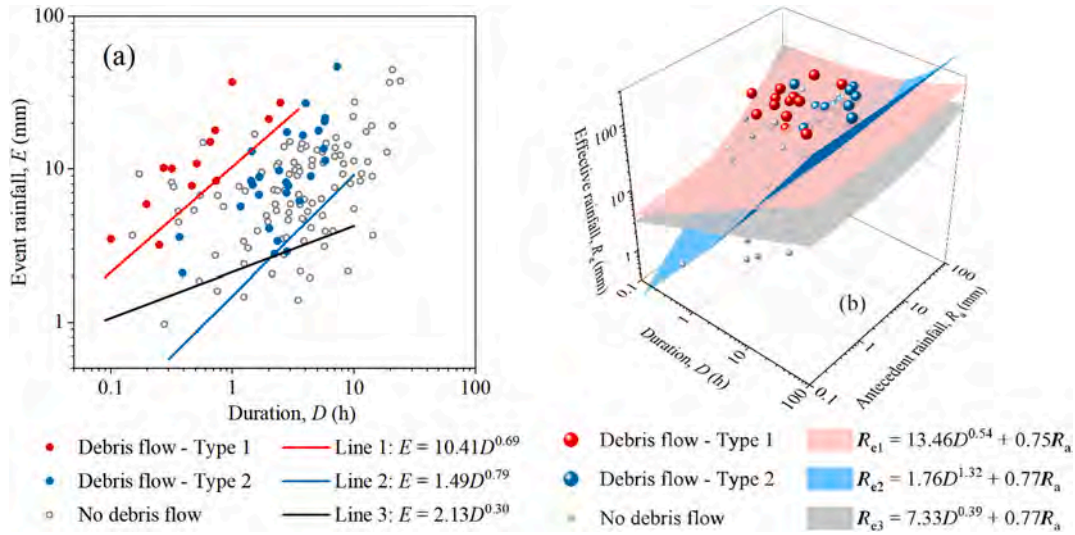


Fig. 10. Thresholds for different types of debris flows: (a) E - D . Lines 1, 2, and 3 represent the E - D thresholds for debris flows of type 1, 2, and all events, and (b) D - E - R_a . R_{e1} , R_{e2} , and R_{e3} represent the D - E - R_a thresholds for debris flows of type 1, 2, and all events.

(2) Rainfall thresholds for debris flows with different formation type.

Influence of forming types of debris flow has been usually ignored in previous studies. In some cases, debris flows act as normal hydrological processes supplied by instantaneous shallow soil failures (Type 1), whereas other cases involve large landslides and blockage-breaking phenomena (Type 2). These can be distinguished by the lag time between the rainfall peak and the flow occurrence, and the ratio of the peak discharge of the debris flow to the water flow (Guo et al., 2021).

The thresholds using the E - D and E - D - R_a models for the two types are shown in Fig. 10. Type 1 requires a much higher threshold than Type 2. The results are evaluated using A_c and F_2 , as listed in Table 4. Slight improvement is achieved for Type 1 when accounting for antecedent rainfall. However, the results for Type 2 are improved markedly, i.e., A_c and F_2 are improved from 0.29 to 0.59 and from 0.42 to 0.66, indicating that antecedent rainfall has much greater impact on Type 2, and thus the accurate predication of Type 2 is much more difficult.

Table 4

Test results of rainfall thresholds for different types of debris flows and models.

Formation type	Parameter	Accuracy	F_2
1	E - D	0.92	0.84
	E - D - R_a	0.93	0.85
2	E - D	0.29	0.42
	E - D - R_a	0.59	0.66

6. Discussion

Recent studies have indicated that AI can play an important role in the assessment and prediction of geohazards (Guzzetti et al., 2008; Mondini et al., 2023; Zhao et al., 2022a). This study further substantiates AI's potential to assist developing rainfall thresholds for debris flows in small catchments, including the precise identification of debris flow initiation times, the strategic selection of representative rain gauges and optimal rainfall parameters, and the quantitative evaluation of these selections.

Statistical methods integrated with expert knowledge of debris flow formation processes can effectively identify key parameters through meticulous manual selection and construct a reasonable predictive model (Chien-Yuan et al., 2005; Guo et al., 2020). However, these approaches often incorporate subjective interpretations and assumptions, which may introduce bias and uncertainty (Guo et al., 2020). In contrast, AI offers significant advantages in processing large and complex datasets objectively, uncovering hidden patterns and relationships that might not be readily apparent, thereby reducing the uncertainty associated with debris flow threshold construction. For instance, although antecedent rainfall was recognized in debris flow occurrence in this specific catchment, its role in threshold has proven challenging (Guo et al., 2013). This study leverages AI to assess the significance of various rainfall parameters (including antecedent rainfall), and develops a comprehensive threshold incorporating multiple parameters, demonstrating superior accuracy compared to those derived from traditional

empirical methods (“ E - D - R_a ” vs “ I - D ”, and 0.76 vs 0.54).

Nevertheless, AI exhibits certain limitations. A primary concern is its black-box nature, where models operate in ways that are not easily interpreted. Additionally, while one of AI’s strengths lies in processing large volumes of data and complex relationships, insufficient or poor-quality data can compromise predictive accuracy due to the reliance on data integrity. Another consideration is that, despite the inherent advantages of any AI method, the selection should be based on the specific conditions of the basin and available data, as well as practical forecasting requirements. While numerous AI approaches are currently available, not all are suitable for this specific scenario. For example, deep learning methods such as convolution neural networks (CNN) and recurrent neural networks (RNN) have been effectively applied in the data analysis of geohazards (e.g., Ghorbanzadeh et al, 2019; Ma and Mei, 2021). CNN is more appropriate for processing continuous spatial data but is less suitable for rainfall data in this catchment due to the relatively sparse distribution of rain gauge networks. RNN, which excels at capturing spatiotemporal dependencies in rainfall data, requires larger datasets, increasing the risk of overfitting during model training when using the limited dataset available. Therefore, these methods may be more appropriate for large-scale, long-term spatiotemporal data analysis. Given the limited data availability typical of small catchments, this study selected a relatively simpler model NGK-SVM, which features a more straightforward structure, lower data requirements, and enhanced interpretability, making it particularly suitable for this context. It is utilized to identify the most critical rainfall parameters and the most appropriate forecasting model. Subsequently, SHAP is employed to validate the results of NGK-SVM, allowing for further interpretation in conjunction with our understanding of debris flow physical processes.

It is attempted to utilize the NGK-SVM method to derive a series of data points that collectively delineate the boundary between debris flow occurrence and non-occurrence. Ultimately, the data points fit a polynomial relationship representing the threshold value. Although the A_c and F_2 yielded results similar to those proposed by Eq. (15), it is evident that the polynomial relationship lacks a physical explanation. From this perspective, when constructing an appropriate model, we should not over rely on AI methods to generate arbitrary models, but instead integrate the relationships between parameters to propose a model with physical significance. In this work, the AI methods offered an objective and effective approach for parameters selection and uncertainty assessment, while the thresholds were still determined based on the physically meaningful interpretations relationships. Overall, human expertise remains indispensable for comprehending the physical mechanisms driving debris flows, thereby enabling more judicious selection and application of AI technologies. From this perspective, AI can assist and complement, but not replace traditional methods in constructing debris flow rainfall thresholds.

It is also crucial to recognize that while integrating multiple parameters enhances predictive accuracy, it also increases model complexity, which may impede practical adoption in operational forecasting. In practice, an effective strategy for forecasting involves balancing predictive accuracy with computational feasibility. From the perspective of machine learning models, identifying the minimal set of parameters necessary is fundamentally a dimensional reduction process aimed at simplifying the model while preserving its predictive power. Besides NGK-SVM, other techniques such as Principal Component Analysis (PCA) can reduce data dimensions by analyzing the variance of the original data, thereby mitigating misjudgments caused by outliers or extreme values (Jolliffe and Cadima, 2016). However, PCA is better suited for datasets with linear relationships between parameters, whereas rainfall parameters exhibit non-linear relationships (e.g., E - D and I - D follow power functions). In contrast, NGK-SVM is more appropriate for this application. The optimized model selected by NGK-SVM is E - D - R_a , which introduces only one additional dimension compared to traditional two-parameter models (E - D , I - D), yet significantly improves

forecasting accuracy, thus achieving a balance between effectiveness and usability. It is also worth noting that the forecast also represents a trade-off among evaluation metrics. Theoretically, an increase in the false positive rate will inevitably result in a decrease in the false negative rate. Therefore, the selection of models and threshold identification should be guided by actual requirements.

The conclusion of this study that R_a is the most important rainfall parameter in the JJG catchment differs from the findings of many previous studies, most of which considered E to be the principal parameter and adopted E - D (or I - D) as the most frequently used parameter combination (Zhuang et al., 2015; Guo et al., 2020; Zhang et al., 2020). However, the results of both A_c and F_2 of the models that consider R_a showed marked improvement in comparison with the results of the E - D model, which neglects antecedent rainfall. Additionally, the influence of antecedent rainfall varies substantially with types of debris flows. The diversity of debris flow formation types and the complexity of the associated mechanisms further complicate the establishment of rainfall thresholds.

7. Conclusions

This study derived the optimal rainfall threshold for the debris flow using AI methods to identify the most important influencing parameters. Based on 140 rainfall events in the JJG catchment during 2006–2023, of which 40 triggered debris flows, the metrics of A_c and F_2 were employed to evaluate the influence of debris flow initiation time, rain gauge selection, and rainfall parameters in each step of the procedure.

It is suggested that the optimal selection for establishing the rainfall threshold is to take the rainfall peak time as the debris flow initiation time, consider the average of rain gauge data from the source region as the data source, and combine the parameters of event rainfall (E), rainfall duration (D), and antecedent rainfall (R_a). The optimal rainfall threshold, derived using the E - D - R_a model, can be expressed as $E = 7.33D^{0.39} - 0.23R_a$.

Rainfall parameter selection has the greatest influence on the rainfall threshold, and the results in terms of A_c improved from 0.54 for the model using the signal parameter of R_a , to 0.66 for the model using the combination of R_a and the D , to 0.81 for the model using the combination of R_a , D and E . Incorporation of further parameters would improve forecasting precision but inevitably increase the complexity of building rainfall thresholds.

Compared with the traditional E - D threshold, the AI-derived E - D - R_a threshold presents advantages not only from the perspective of precision evidenced by marked improvement in terms of both A_c and F_2 , but also from the perspective of its physical meaning, which considers the effect of both antecedent rainfall and event rainfall.

Due to the “black box” nature of AI methods, they are difficult to effectively interpret the relationship between rainfall parameters and the physical significance of the threshold models. However, they can help with the selection of key parameters and evaluation of the influence of such selections on the results, thereby assisting in building more reasonable rainfall thresholds.

CRedit authorship contribution statement

Siling Zhang: Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Xiaojun Guo:** Writing – review & editing, Validation, Funding acquisition, Conceptualization. **Jianyi Cheng:** Software. **Yong Li:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was supported by the NSFC, China (42322703, U21A2008 and 41925030), the Second Scientific Expedition on the Qinghai-Tibet Plateau (2019QZKK0903-02), Science and Technology Project, Sichuan, China (2022JDJQ0008), Western Light of Young Scholars, CAS, China and the Youth Innovation Promotion Association, CAS, China (Grant No. 2022379).

We thank the Dongchuan Debris Flow Observation and Research Station (DDFORS) of the Chinese Academy of Sciences for providing the field observation data. We thank James Buxton MSc, from Liwen Bianji (Edanz) (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript.

The code used in this study can be found through the following link: <https://github.com/An-jing-99/HYDROL63040R1>

Data availability

Data will be made available on request.

References

- Abraham, M.T., Satyam, N., Kushal, S., Rosi, A., Pradhan, B., Segoni, S., 2020. Rainfall threshold estimation and landslide forecasting for Kalimpong, India Using SIGMA Model. *water* 12 (4), 1195. <https://doi.org/10.3390/w12041195>.
- Aleotti, P., 2004. A warning system for rainfall-induced shallow failures. *Eng Geol* 73, 247–265. <https://doi.org/10.1016/j.enggeo.2004.01.007>.
- Arattano, M., Moia, F., 1999. Monitoring the propagation of a debris flow along a torrent. *Hydrol. Sci. J.* 44 (5), 811–823. <https://doi.org/10.1080/02626669909492275>.
- Bel, C., Liébault, F., Navratil, O., Eckert, N., Bellot, H., Fontaine, F., Laigle, D., 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>.
- Berti, M., Genevois, R., Simoni, A., Tecca, P.R., 1999. Field observations of a debris flow event in the Dolomites. *Geomorphology* 29 (3–4), 265–274. [https://doi.org/10.1016/S0169-555X\(99\)00018-5](https://doi.org/10.1016/S0169-555X(99)00018-5).
- Berti, M., Martina, M.L.V., Franceschini, S., Pignone, S., Simoni, A., Pizzoli, M., 2012. Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. *J. Geophys. Res. Earth* 117 (F4). <https://doi.org/10.1029/2012JF002367>.
- Berti, M., Bernard, M., Gregoretti, C., Simoni, A., 2020. Physical interpretation of rainfall thresholds for runoff-generated debris flows. *J. Geophys. Res. Earth* 125 (6). <https://doi.org/10.1029/2019JF005513>.
- Bishop, C. M., Nasrabadi, N. M., 2006. Pattern recognition and machine learning (Vol. 4, No. 4, p. 738). New York: Springer.
- Brunetti, M.T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., Guzzetti, F., 2010. Rainfall thresholds for the possible occurrence of landslides in Italy. *Nat. Hazards Earth Syst. Sci.* 10 (3), 447–458. <https://doi.org/10.5194/nhess-10-447-2010>.
- Caine, N., 1980. The rainfall intensity-duration control of shallow landslides and debris flows. *Geogr. Ann. Ser. B* 62 (1–2), 23–27. <https://doi.org/10.1080/04353676.1980.11879996>.
- Chen, S.C., Huang, B.T., 2010. Non-structural mitigation programs for sediment-related disasters after the Chichi earthquake in Taiwan. *J. Mt. Sci.* 7, 291–300. <https://doi.org/10.1007/s11629-010-0221-3>.
- Chien-Yuan, C., Tien-Chien, C., Fan-Chieh, Y., Wen-Hui, Y., Chun-Chieh, T., 2005. Rainfall duration and debris-flow initiated studies for real-time monitoring. *Environ. Geol.* 47, 715–724. <https://doi.org/10.1007/s00254-004-1203-0>.
- Crema, S., Coviello, V., Cesca, M., Dainese, R., Marchi, L., Pasuto, A., & Cavalli, M., 2023. When instrument location makes the difference on rainfall thresholds definition: Lessons learned at Cancia, Dolomites. In Proceedings of the 8th International Conference on Debris Flow Hazard Mitigation (DFHM8), E3S Web of Conferences (Vol. 415, p. 03008). EDP Sciences.
- Cui, P., Chen, X.P., Wang, Y.Y., Hu, K.H., Li, Y., 2005. Jiangjia Ravine debris flows in the southwestern China. In: Jakob, M., Hungr, O. (Eds.), *Debris-Flow Hazards and Related Phenomena*. Springer-Verlag, pp. 565–594.
- Cui, P., Zhu, Y.Y., Chen, J., Han, Y.S., Liu, H.J., 2007. Relationships between antecedent rainfall and debris flows in Jiangjia Ravine, China. In: Chen, C.L., Major, J.J. (Eds.), *Debris-Flow Hazard Mitigation-Mechanics, Prediction, and Assessment*. Millpress, Rotterdam, pp. 1–10.
- Gariano, S.L., Guzzetti, F., 2016. Landslides in a changing climate. *Earth Sci. Rev.* 162, 227–252. <https://doi.org/10.1016/j.earscirev.2016.08.011>.
- Gariano, S.L., Melillo, M., Peruccacci, S., Brunetti, M.T., 2020. How much does the rainfall temporal resolution affect rainfall thresholds for landslide triggering? *Nat. Hazards* 100 (2), 655–670. <https://doi.org/10.1007/s11069-019-03830-x>.
- Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S. R., Tiede, D., Aryal, J., 2019. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sensing*, 11(2), 196. <https://doi.org/10.3390/rs11020196>.
- Guo, X.J., Cui, P., Li, Y., 2013. Debris flow warning threshold based on antecedent rainfall: a case study in Jiangjia Ravine, Yunnan, China. *J. Mountain Sci.* 10 (2), 305–314. <https://doi.org/10.1007/s11629-013-2521-z>.
- Guo, X., Cui, P., Li, Y., Ma, L., Ge, Y., Mahoney, W.B., 2016. Intensity-duration threshold of rainfall-triggered debris flows in the Wenchuan earthquake affected area, China. *Geomorphology* 253, 208–216. <https://doi.org/10.1016/j.geomorph.2015.10.009>.
- Guo, X.J., Li, Y., Cui, P., Yan, H., Zhuang, J.Q., 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>.
- Guo, X.J., Chen, X.C., Song, G.H., Zhuang, J. Q., Fan, J.L., 2021. Debris flows in the Lushan earthquake area: formation characteristics, rainfall conditions, and evolutionary tendency. *Natural Hazards*, 2021, 106(3), 2663–2687. <https://doi.org/10.1007/s11069-021-04559-2>.
- Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P., 2007. Rainfall thresholds for the initiation of landslides in central and southern Europe. *Meteorol. Atmos. Phys.* 98, 239–267. <https://doi.org/10.1007/s00703-007-0262-7>.
- Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P., 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>.
- Hoch, O.J., McGuire, L.A., Youberg, A.M., Rengers, F.K., 2021. Hydrogeomorphic recovery and temporal changes in rainfall thresholds for debris flows following wildfire. *Journal of Geophysical Research-Earth Surface* 126 (12). <https://doi.org/10.1029/2021JF006374>.
- Hirschberg, J., Badoux, A., McArdell, B.W., Leonarduzzi, E., Molnar, P., 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>.
- Hürlimann, M., Coviello, V., Bel, C., Guo, X., Berti, M., Graf, C., Hübl, J., Miyata, S., Smith, J.B., Yin, H.Y., 2019. Debris-flow monitoring and warning: Review and examples. *Earth Sci. Rev.* 199, 102981. <https://doi.org/10.1016/j.earscirev.2019.102981>.
- Jan, C.D., Lee, M.H., 2004. A debris flow rainfall-based warning model. *J. Chin Soil Water Conserv* 35 (3), 275–285 in Chinese.
- Jan, C.D., Chen, C.L., 2005. Debris flows caused by Typhoon Herb in Taiwan. In: Jakob, M., Hungr, O. (Eds.), *Debris Flow Hazards and Related Phenomena*. Springer, Berlin Heidelberg, pp. 363–385.
- Jakob, M., Owen, T., Simpson, T., 2012. A regional real-time debris-flow warning system for the District of North Vancouver, Canada. *Landslides* 9, 165–178. <https://doi.org/10.1007/s10346-011-0282-8>.
- Jiang, Z., Fan, X., Siva Subramanian, S., Yang, F., Tang, R., Xu, Q., Huang, R., 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>.
- Jolliffe, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. *Phil. Trans. R. Soc. A* 374, 20150202. <https://doi.org/10.1098/rsta.2015.0202>.
- Lundberg, Scott M. and Su-In Lee., 2017. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 4768–4777.
- Ma, Z., & Mei, G., 2021. Deep learning for geological hazards analysis: Data, models, applications, and opportunities. *Earth-Science Reviews*, 223, 103858. <https://doi.org/10.1016/j.earscirev.2021.103858>.
- Marra, F., Nikolopoulos, E.I., Creutin, J.D., Borga, M., 2014. Radar rainfall estimation for the identification of debris-flow occurrence thresholds. *J. Hydrol.* 519, 1607–1619. <https://doi.org/10.1016/j.jhydrol.2014.09.039>.
- Marra, F., Nikolopoulos, E.I., Creutin, J.D., Borga, M., 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., 2021. Uncertainty analysis of a rainfall threshold estimate for stony debris flow based on the backward dynamical approach. *Nat. Hazards Earth Syst. Sci.* 21 (6), 1769–1784. <https://doi.org/10.5194/nhess-21-1769-2021>.
- Melillo, M., Maria Teresa Brunetti, Silvia Peruccacci, Stefano Luigi Gariano, Anna Roccati, Fausto Guzzetti., 2018. A tool for the automatic calculation of rainfall thresholds for landslide occurrence. *Environmental Modelling & Software*, Volume 105, 2018, Pages 230–243, ISSN 1364-8152. <https://doi.org/10.1016/j.envsoft.2018.03.024>.
- Mondini, A.C., Guzzetti, F., Melillo, M., 2023. Deep learning forecast of rainfall-induced shallow landslides. *Nat. Commun.* 14, 2466. <https://doi.org/10.1038/s41467-023-38135-y>.
- Nikolopoulos, E.I., Crema, S., Marchi, L., Marra, F., Guzzetti, F., Borga, M., 2014. Impact of uncertainty in rainfall estimation on the identification of rainfall thresholds for debris flow occurrence. *Geomorphology* 221, 286–297. <https://doi.org/10.1016/j.geomorph.2014.06.015>.
- Nikolopoulos, E.I., Borga, M., Creutin, J.D., Marra, F., 2015. Estimation of debris flow triggering rainfall: Influence of rain gauge density and interpolation methods. *Geomorphology* 243, 40–50. <https://doi.org/10.1016/j.geomorph.2015.04.028>.
- Okano, K., Suwa, H., Kanno, T., 2012. Characterization of debris flows by rainstorm condition at a torrent on the Mount Yakedake volcano, Japan. *geomorphology* 136 (1), 88–94. <https://doi.org/10.1016/j.geomorph.2011.04.006>.
- Panziera, L., Germann, U., Gabella, M., Mandapaka, P.V., 2011. NORA–Nowcasting of Orographic Rainfall by means of Analogues. *Q. J. R. Meteorol. Soc.* 137 (661), 2106–2123. <https://doi.org/10.1002/qj.878>.

- Peres, D.J., Cancelliere, A., Greco, R., Bogaard, T.A., 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>.
- Peruccacci, S., Brunetti, M.T., Luciani, S., Vennari, C., Guzzetti, F., 2012. Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in central Italy. *Geomorphology* 139, 79–90. <https://doi.org/10.1016/j.geomorph.2011.10.005>.
- Rossi, M., Luciani, S., Valigi, D., Kirschbaum, D., Brunetti, M.T., Peruccacci, S., Guzzetti, F., 2017. Statistical approaches for the definition of landslide rainfall thresholds and their uncertainty using rain gauge and satellite data. *Geomorphology* 285, 16–27. <https://doi.org/10.1016/j.geomorph.2017.02.001>.
- Saadatkah, N., Kassim, A., Lee, L.M., 2015. Hulu Kelang, Malaysia regional mapping of rainfall-induced landslides using TRIGRS model. *Arab. J. Geosci.* 8, 3183–3194. <https://doi.org/10.1007/s12517-014-1410-2>.
- Smith, J.A., Baek, M.L., Meierdiercks, K.L., Miller, A.J., Krajewski, W.F., 2007. Radar rainfall estimation for flash flood forecasting in small urban watersheds. *Adv. Water Resour.* 30 (10), 2087–2097. <https://doi.org/10.1016/j.advwatres.2006.09.007>.
- Sokolova, M., Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* 45 (4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>.
- Staley, D.M., Kean, J.W., Cannon, S.H., Schmidt, K.M., Laber, J.L., 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, D.M., Negri, J.A., Kean, J.W., Laber, J.L., Tillery, A.C., Youberg, A.M., 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>.
- Tang, C., Tanyas, H., van Westen, C.J., Tang, C., Fan, X., Jetten, V.G., 2019. Analysing post-earthquake mass movement volume dynamics with multi-source DEMs. *Eng. Geol.* 248, 89–101. <https://doi.org/10.1016/j.enggeo.2018.11.010>.
- Tecca, P.R., Genevois, R., 2009. Field observations of the June 30, 2001 debris flow at Acquabona (Dolomites, Italy). *Landslides* 6, 39–45. <https://doi.org/10.1007/s10346-009-0145-8>.
- Tien Bui, D., Pradhan, B., Lofman, O., Revhaug, I., Dick, Ø.B., 2013. Regional prediction of landslide hazard using probability analysis of intense rainfall in the Hoa Binh province. *Vietnam. natural Hazards* 66, 707–730. <https://doi.org/10.1007/s11069-012-0510-0>.
- Vapnik, V., 1998. *Statistical learning theory*. John Wiley & Sons Google Schola 2, 831–842.
- Witten, I.H., Frank, E., 2002. Data mining: practical machine learning tools and techniques with Java implementations. *ACM SIGMOD Rec.* 31 (1), 76–77. <https://doi.org/10.1145/507338.507355>.
- Zhang, S.J., Xu, C.X., Wei, F.Q., Hu, K.H., Xu, H., Zhao, L.Q., Zhang, G.P., 2020. A physics-based model to derive rainfall intensity-duration threshold for debris flow. *Geomorphology* 351, 106930. <https://doi.org/10.1016/j.geomorph.2019.106930>.
- Zhao, Y., Meng, X., Qi, T., Li, Y., Chen, G., Yue, D., Qing, F., 2022a. AI-based rainfall prediction model for debris flows. *Eng. Geol.* 296, 106456. <https://doi.org/10.1016/j.enggeo.2021.106456>.
- Zhao, Y., Meng, X., Qi, T., et al., 2022b. 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>.
- Zhuang, J., Cui, P., Wang, G., Chen, X., Iqbal, J., Guo, X., 2015. Rainfall thresholds for the occurrence of debris flows in the Jiangjia Gully, Yunnan Province, China. *Eng. Geol.* 195, 335–346. <https://doi.org/10.1016/j.enggeo.2015.06.006>.