Revision notes on Manuscript No. NHESS-2023-180

First of all, the authors thank the Editors and the reviewers for considering our manuscript and providing constructive comments to help us improve the quality of our work. We have accordingly revised the manuscript by carefully addressing or answering the comments point-by-point, summarized as follows. Following the revision, we hope we have clarified all of the points summarized by the Editor and reviewers.

Responses to the Comments Raised by Reviewer #1

1. I appreciate the authors' efforts in addressing some of my comments. They have significantly improved the structure of the paper and provided a better explanation of the novelties in their work. However, the objectives are still not clearly stated. Lines 100-111 are more an anticipation of the conclusions.

Authors' reply:

Thank you for your comments. The authors have changed the content of Line 100-111 as below and moved the content of Line 100-111 to the Conclusion part of the revised MS.

"To fill the current research and practical gaps, we aim to propose in this study a new framework based on an integrated hydrological and hydrodynamic modeling approach to more reliably estimate the rainfall thresholds of runoff-generated debris flows, i.e. providing a physically based approach to estimate trigger-cause-based rainfall thresholds. In addition, the proposed modeling framework will effectively incorporate meteorological conditions, catchment topographic properties, and grain-size distribution of debris materials, making it more suitable for application in areas with limited historical data. The rest of the paper is organized as follows: Section 2 describes the proposed framework; Section 3 introduces the case study including the flow monitoring scheme; Section 4 presents the validation results; Section 5 discusses the advantages and limitations of the proposed method, followed by brief conclusions drawn in Section 6".

The reviewer can also find the changes in Lines102-111 and Lines 790-797 in the revised MS.

2. The authors are still overstating their conclusions, which have been drawn from a generalization based on limited data (e.g. lines 505-506, 515-520, 628-630). The authors should either provide more proof or tone down their claims.

Authors' reply:

Thanks for the comments. We fully agree with the reviewer that some concluding statements are overstated. We have revised the sentences of Lines 505-506, 515-520, 628-630, as follows:

Line 505-506

The results are consistent with the actual observation, i.e. a debris flow did occur during the typhoon event, demonstrating that the proposed methodology successfully predicted the occurrence of this debris flow event.

The reviewer can also find the changes in Lines 504-506 in the revised MS.

Line 515-520

This implies that the hydrodynamic conditions necessary for triggering a debris flow are met only in a small fraction of the grid cells and the likelihood of debris flow occurrence is low. This conclusion aligns with the actual observations, i.e. no debris flow was observed during these six rainfall events. As a whole, these numerical tests demonstrate the capability of the framework including the adopted hydrodynamic thresholds in predicting six observed non-debris flow events and one actual debris flow event.

The reviewer can also find the changes in Lines 517-520 in the revised MS.

Line 628-630

From Fig. 14, it is evident that both the proposed I-D rainfall threshold and the empirical rainfall threshold can effectively distinguish one triggering and six non-triggering rainfall events, highlighting the feasibility of the proposed framework.

The reviewer can also find the changes in Lines 650-653 in the revised MS.

3. The accumulation recorded during the October debris flow (around 300 mm) and the accumulation registered during the non-triggering rainfalls shown in Table 7 (between 16 mm and 60 mm) are very different, and therefore the resulting discharges are also very different. These rainfall events have been used to decide the % of catchment area that has to exceed the hydrodynamic thresholds to trigger a debris flow. I would expect that rainfall with larger accumulation than those 6 non-triggering rainfalls would also result in larger discharges, which probably result in larger areas exceeding the hydrodynamic thresholds. How sure are you that such events would not trigger a debris flow in the catchment? Do you have any additional examples of rainfall events registered between 2013 and today that have not triggered debris flows in the studied catchment that could be used for further validation and calibration of the zone thresholds?

Authors' reply:

Due to the failure to obtain high-quality observed hydrological data in the small and unstable channels, the monitoring station has not been maintained and has since been out of work after September 2013. So, we, unfortunately, do not have additional data for both occurred and non-occurred events between 2013 and today to further validate the zone thresholds in the study area. Due to the lack of data, the authors proposed rainfall thresholds for the study area with different assumed zone thresholds (see Fig. 15 in revised MS). Sensitivity analysis on the rainfall thresholds regarding the zone thresholds was also conducted. The authors acknowledge such a major limitation in this study, as the application and testing of the framework were limited to a small number of events in a small catchment. This highlights the need for broader testing in similar catchments to enhance the robustness of the framework. This limitation was explicitly discussed in the Discussion Section of the original manuscript. This is actually the challenge we would like to address in this study, i.e. by proposing a framework based on physical principles, which is more suitable for areas with limited historic debris flow data. The main objective of this manuscript is to explore the feasibility of a coupled hydrological and hydrodynamic modeling framework for estimating rainfall thresholds for debris flow occurrence. Due to the limited availability of the observation data, we have used one debris flow event and six non-debris flow events to validate the proposed framework, but not to determine the values of zone thresholds. Since it is a physically based approach, the users may propose their own rainfall thresholds within the framework if they can establish the relationship between rainfall and runoff and given corresponding values for zone thresholds. Actually, the selection of zone thresholds should vary depending on the definition of debris flow occurrence. For example, choosing a 10% zone threshold means that debris flow occurrence is considered positive when 10% of the computed domain is identified as triggering cells. Choosing a 15% threshold means that debris flow occurrence is considered positive when 15% of the computed domain is identified as triggering cells. Such a threshold may be more reliably defined when data is available for different applications, and the selection of zone thresholds is linked to varying degrees of conservatism or adventurousness in rainfall thresholds. A smaller zone threshold corresponds to a lower rainfall threshold, reflecting a more conservative approach, while a larger value indicates a higher rainfall threshold, suggesting a more adventurous approach. Different zone thresholds correspond to different levels of warning, with smaller values corresponding to lower warning levels and larger values to higher warning levels. So, we recommend that the selection of zone thresholds for a specific catchment should be conducted based on risk management decisions and formally established by governmental authorities in advance. This approach allows for the implementation of a variety of risk management strategies that can be customized to meet the desired level of caution or preparedness.

4. The same triggering rainfall, and non-triggering rainfall events that have been used to select the zone thresholds have been used to check the performance of the proposed thresholds. This is an important limitation. The recorded discharges of such events are significantly different, and of course it is possible to separate very well between the events and no-events (Fig 13 and Fig 14). Do you have proof that all rainfall events resulting in larger discharges than the one that triggered the debris flow in 2013 would also result in debris flow events? Similarly, what about rainfall events that result in slightly lower discharges than the rainfall that triggered the debris flow, are you sure

any of them could trigger a debris flow?

Authors' reply:

In this manuscript, one triggering debris flow event and six non-triggering debris flow events were utilized to validate the effectiveness of the proposed framework, rather than to determine the zone thresholds, which should be user-defined with more observations available and for different applications. The rainfall thresholds presented in Figures 13 and 14 were established based on various design hyetographs generated from Intensity-Duration-Frequency (IDF) analysis, rather than from observed rainfall events. This implies that the rainfall events used to assess the performance of the proposed thresholds differ from those used to develop them. It's noteworthy that the discharge calculated from some design hyetographs is notably consistent (see Table 1), in contrast to the variability observed in recorded discharge data.

Table1 the peak discharge under different IDF analysis

IDF	1h-100y	3h-20y	1h-10y	3h-5y	3h-3y	1h-5y
Peak discharge(m^3/s)	0.38	0.37	0.16	0.17	0.12	0.11

To further demonstrate the feasibility of the threshold framework, we have also conducted several extra scenario simulations. Specifically, we modified the input rainfall events, varying the cumulative amount from 10% to 110% of the original rainfall that triggered the 2013 debris flow. The rainfall distribution is presented in Fig. 1. The cumulative input rainfall ranges from 38 mm to 418 mm, with the original value that triggered the 2013 debris flow being 380 mm. The calculated discharge for each scenario is shown in Fig. 2. Additionally, the calculated peak discharge and the corresponding proportion of triggering cells based on the Thresholds G are detailed in Table 2.

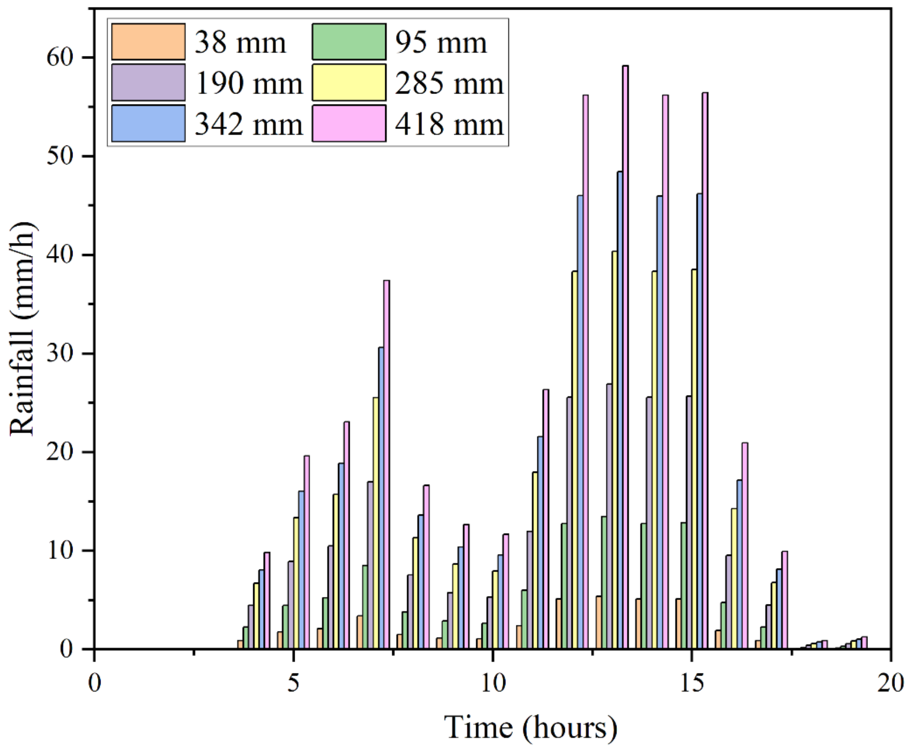

Fig.1 Rainfall distribution of the scenario simulations.

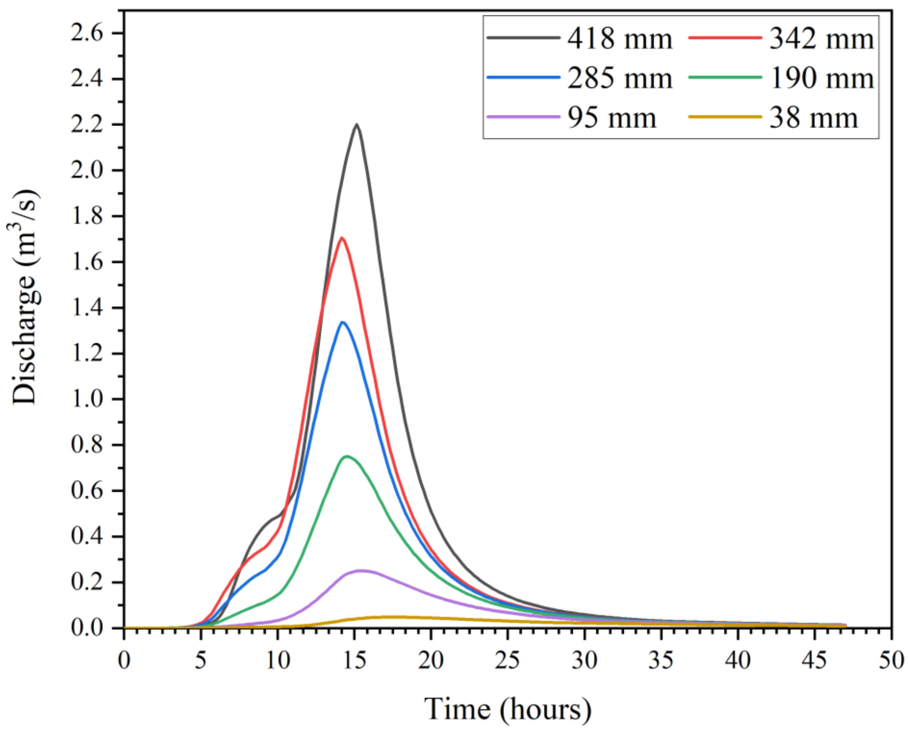

Fig.2 Calculated runoff of the scenario simulations.

Table 2 Calculated peak discharge and the corresponding proportion of triggering cells based on Threshold G

Cumulative amount of input rainfall (mm)	418	342	285	190	95	38
Calculated peak discharge (m^3/s)	2.2	1.7	1.3	0.75	0.25	0.048
Proportion of triggering cells (%)	48	44	39	33	17	3

The calculated peak discharge from the rainfall that triggered the 2013 debris flow is 2.0 m^3/s. As shown in Table 2, rainfall events resulting in discharges greater than the one in 2013 could also result in debris flow events, with the calculated proportion of triggering cells for a 418 mm rainfall event being 48%. Even slightly lower discharges than the 2013 event may trigger a debris flow, as indicated by the 342 mm rainfall event, which has a calculated triggering cell proportion of 44%. This demonstrates that the calculated proportion of triggering cells is consistent with the change in the input rainfall.

The reviewer can also find the changes in Lines 526-549 in the revised MS.

5. In line 680, it is stated that the reliability of statistical I-D thresholds depends on the data being used and that long-term observations are essential to derive reliable thresholds. I think your approach also depends on the quality of the data being used to obtain the thresholds. To be able to select adequate and objective zone thresholds, you also need to have a long time series of observations of debris flow events and no-events.

Authors' reply:

We fully agree with the reviewer that a long time series of observations is essential to validate the true values of the zone threshold of a specific catchment. However, as mentioned in Comment 3, the selection of zone thresholds depends on the definition of debris flow occurrence. These specific values are not predetermined for individual catchments. Therefore, to calculate the rainfall thresholds based on the proposed framework, governmental decision-makers or other users must assign the values for zone thresholds according to applications. Consequently, the zone threshold should be established by users to reflect catchment settings and applications. Whilst it is important to show the feasibility of the approach, it is not entirely essential to be validated by multiple debris flows or non-debris flow events in the current case study.

The reviewer can also find the changes in Lines 748-752 in the revised MS.

6. Line 66-67: In the literature you can find a number of examples (e.g. Oorthuis et al., 2023, Hirschberg et al., 2021, Bel et al., 2017, Abancó et al., 2016), of reliable statistical rainfall thresholds defined for specific catchments where monitoring data exists (and thus a complete register of debris flow occurrences).

Authors' reply:

The authors fully agree that reliable statistical rainfall thresholds for a specific catchment could be derived when an adequate dataset of debris flow events is available. The authors have revised the sentences as below:

"So, the reliable statical rainfall thresholds for a specific catchment may be derived when sufficient high-quality data of debris flow events is available (Oorthuis et al., 2023, Hirschberg et al., 2021, Bel et al., 2017, Abancó et al., 2016). However, in the areas with limited data availability, it will be technically challenging to reliably define the statistical I-D thresholds of debris flows. It may be more useful to propose a physical-based approach to estimate the rainfall thresholds in such data-scarce area".

The reviewer can also find the changes in Lines 66-71 in the revised MS.

7. Line 475, and section 4.3: What is the return period of the October 2013 rainfall event? How long did the event last?

Authors' reply:

The total rainfall amount of October 2013 rainfall event is 380 mm with the rainfall duration being 16 hours. The rainfall conditions of the Typhoon Fitow (October 2013) event are also included in Fig 15 of revised MS, which has a return period of about 100 years.

The reviewer can also find the changes in Lines 471-472 in the revised MS.

8. Line 505-506: It proves that you can distinguish this debris-flow event.

Authors' reply:

The authors have revised the sentence as below:

The results are consistent with the actual observation, i.e. a debris flow did occur during the typhoon event, demonstrating that the proposed methodology can be used to predict the occurrence of such an event.

The reviewer can also find the changes in Lines 504-506 in the revised MS.

9. Line 629: The analysed 6 non-triggering rains and 1 triggering rain.

Authors' reply:

The authors have revised the sentence as below:

From Fig. 14, it is evident that both the proposed I-D rainfall threshold and the empirical rainfall threshold can effectively distinguish the one triggering and six non-triggering rainfall events, highlighting the feasibility of the proposed framework.

The reviewer can also find the changes in Lines 650-653 in the revised MS.

10. Line 680: Clearly is a very strong word.

Authors' reply:

The authors have revised the sentence as below:

This indicates that a statistical approach may be difficult to be reliably applied in ungauged catchments where high-quality historical data is missing.

The reviewer can also find the changes in Lines 706-708 in the revised MS.