Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2020-243-AC3, 2020

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

Interactive comment on "Uncertainty analysis of the estimation of stony debris flow rainfall threshold: the application to the Backward Dynamical Approach" *by* Marta Martinengo et al.

Marta Martinengo et al.

marta.martinengo@unitn.it

Received and published: 4 December 2020

We appreciate the useful comments provided by the anonymous referee #2 that will be used to improve the manuscript.

Dear Authors, Dear Editor, I have read and carefully evaluated the manuscript "Uncertainty analysis of the estimation of stony debris flow rainfall threshold: the application to the Backward Dynamical Approach" submitted for possible publication in NHESS. The manuscript applies a double Monte Carlo simulation to investigate the robustness of the recently proposed BDA model with respect to uncertainty of

Printer-friendly version

Discussion paper

input factors required to derived rainfall thresholds for stony debris flows. The idea is interesting and original and deserves attention. The English is generally clear. I identified a few shortcomings, as highlighted in my comments below. The manuscript could be considered again for publication after major revisions.

General comments

1. The topic of rainfall threshold for landslides/debris flow initiation is intensely debated and many papers are published continuously. Most of the published papers have a low content of originality. I therefore suggest to better stress the elements of novelty in the proposed research.
   As instance:
   - To my understanding, “stony debris flow” is a rather specific category. This could be briefly highlighted in the introduction and conclusion: most of the works abut rainfall thresholds mix different landslide typologies, others include DF and shallow landslides, others are addressed at DF in general (references could be easily found, e.g. with some review paper already in your reference list). You could highlight that studies explicitly addressed to stony DF are rare and thus more knowledge is needed on this field, hence suggesting the need of this test on the BDA model.
   - You could expand the state of the art review and better link it to the originality (and usefulness) of your work. You cite a few relevant papers, but many other could be cited, especially in the central part of the introduction, to better set the stage for your work. And most of all, to avoid a sentence like “As stressed in the Introduction, the rainfall intensities i(t) associated with the event are assumed to be certain. Future analysis will assess and study also the uncertainties related to this piece of data”, which seems an unnecessary justification. You could just say that uncertainty in rainfall threshold

has been already investigated for temporal resolution (Marra 2019; Gariano et al., 2020, both already in your reference list), definition of the triggering rainfall (Peres et al. 2018), rain gauge selection (Abraham et al 2020) and so on. Whereas, a study is missing for the uncertainty in the parameters used by BDA for stony debris flows thresholds: nobody did it, this is (in my opinion) your main contribution to the progress of the state of the art. Abraham, M. T., Satyam, N., Rosi, A., Pradhan, B., & Segoni, S. (2020). The Selection of Rain Gauges and Rainfall Parameters in Estimating Intensity-Duration Thresholds for Landslide Occurrence: Case Study from Wayanad (India). Water, 12(4), 1000. 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.

We agree with this comment. In the Introduction, we will highlight that we refer only to stony debris flow because of the assumptions on which the BDA is based on and we will also clarify what we mean by “stony debris flow”, as suggested by the anonymous referee #1.
The novelty of our work will be better stressed in the revised version of the manuscript to make it clearer to the reader. As correctly asserted by the referee, in the literature there are other works that investigate the uncertainty in rainfall thresholds estimate. The uncertainty analysis of these works are mainly focused on rainfall since the classical approaches, used to calibrate the threshold, estimate the rainfall condition related to an event only on the basis of the hyetograph. Instead, the BDA method computes the rainfall condition related to an occurred event on the basis of a schematic description of the phenomenon dynamics. This means that the calibration of the threshold starts from the estimate of some physical and morphological parameters and surveyed data. To assess the robustness of the BDA requires a proper method that allows to quantify the effects of the input parameters and data uncertainties on the threshold estimate.

If these effects are low, the BDA can be defined as robust. The main novelties of our work are therefore: a new method of assessing the robustness of a rainfall threshold (based on two cascade Monte Carlo simulations) and the application of this method to the BDA. It is worth noting that the developed approach can be also applied to other threshold calibrations that use physical-based parameters (e.g. the one proposed by Zhang et al., 2020).
In the Introduction, to better stress the novelties of our work, we will insert other literature references concerning the uncertainty analysis of rainfall thresholds and we will compare our work to the state of the art.

2. This paper is conceived and organized around some mathematical calculations. The risk is that the reader could perceive it as a "synthetic" experiment. I think it is important for NHESS readers to better put their minds on the specific case of study and it could be useful to add a brief description of the study case. This should include a brief description of the test area features and of the debris flows at hand. Also, some more information on the input data are needed (e.g. source of rainfall and debris flows datasets).
We agree. To make the manuscript more clear and tailored to the NHESS reader, in the revised version of the paper, we will insert a new section describing the study area and the data used in the analysis.

3. Discussions are almost missing, mainly they are mixed with the conclusions. I suggest providing separate sections. Discussions should contain an interpretation of the results, while in the conclusions you should summarize the lessons learnt.
In the revised manuscript, we will divide discussion and conclusions. In the Discussion we will provide a more detailed analysis and interpretation of the results. We will also better highlight what are the characteristics of the event (e.g. the hyetograph shape) that, combined with the inputs uncertainty, mostly affect the variability in the rainfall conditions estimate (i.e. intensity I and duration D).

We will stress the impacts on the threshold estimate due to the uncertainty of the input parameters, discussing the resulting variability of the threshold parameters. Finally, we will better highlight the differences between the outcomes obtained performing the standard calibration (i.e. the reference values) and the mean values of the Monte Carlo simulations outputs, both in term of rainfall conditions and threshold.
In the Conclusions, we will better highlight the resulting robustness of the BDA method. We will also stress that this robustness could be dependent on the events characteristics used in the estimate. Indeed, applying the BDA method to a different study area, if many of the considered events are characterised by the features that tend to increase the outputs variability (we will describe them in the Discussion), the robustness of the estimate can be undermined. In this case, utmost care in estimating the input parameters is recommended.
We will also stress the possibility to apply the developed method for the uncertainty analysis both to different study area (using the BDA) and other thresholds whose calibration is based on physical parameters, as mentioned in a previous answer.
Finally, we will emphasize the importance of a validation analysis, not dealt with in this manuscript, to assess the forecast capability of a threshold, once proved to be robust.

In particular, some points in my opinion are not clear enough: how does this study help us in predicting stony debris flows?
Perhaps we were not clear enough in explaining the purpose of this work and actually, some sentences, regarding the forecast skills, were perhaps a bit misleading. The uncertainty analysis aims to quantify the effects of the input uncertainty on the output in order to check the robustness of a method. Instead, to asses the forecast capability of a threshold is the purpose of a validation analysis. This kind of analysis requires a completely different approach than

the robustness analysis. This work aims to study the robustness of the BDA approach to calibrate the threshold and, therefore, it does not allow to say anything about the forecast capability of the threshold. We will better clarify these aspects in the Introduction of the revised manuscript.

Does it prove that BDA is robust, or does it prove that the utmost care should be put in calibrating/measuring the input parameters required?
The results prove that the approach used to estimate the threshold is robust. Indeed, even if some events are characterised by high variability in the rainfall condition (namely in the duration D and/or intensity I) due to the uncertainty of the inputs, the resulting threshold has low variability. However, as described in a previous answers, it is advisable to put utmost care in the estimate of the input parameters in some cases, depending on the considered events characteristics. This concept will be inserted and explained in the Conclusions of th revised manuscript.

How does the uncertainty is reflected in the forecasting effectiveness of the resulting threshold? The latter point, in particular, is very important and some tests about that should be shown in the revised version of the manuscript.
We agree with the referee that the validation of a calibrated threshold is crucial to prove its forecasting effectiveness and to make this tool operational. We are currently working on the validation, developing a suitable method that will be the object of a forthcoming manuscript. Nevertheless, as said before, validation is not the aim of this paper.
We preferred to perform a robustness analysis of the calibration method before performing a validation analysis. In our opinion, it would have been useless to validate a threshold obtained from a method later proved not robust.
We will better clarify the aim of this work in the Introduction of the revised manuscript and we will stress the importance of the validation in the Conclusions

highlighting that this analysis will be the object of a future work.

Specific comments

- L17. About review casualties, I suggest to add also the review "Dowling, C. A., & Santi, P. M. (2014). Debris flows and their toll on human life: a global analysis of debris-flow fatalities from 1950 to 2011. Natural hazards, 71(1), 203-227."
We thank the referee for the suggestion. We will insert this reference.
- L23 I suggest "Usually, rainfall thresholds for debris flows initiation are power laws that link the rainfall duration to the rainfall cumulated or Intensity". This is because when dealing with other landslide typologies, other parameters are often used (e.g. antecedent rainfall indexes).
We will fix the sentence during the revision.
- L25 I would add also the work of Caine (1980), who started this methodology of analysis. Also, I would add some review. And I would substitute the work by Guzzetti by his work published the year after, reviewing ID thresholds (thus, more strictly related to your research).
We agree. Other references concerning the rainfall thresholds will be inserted in the revised manuscript and the work by Guzzetti will be substituted by his following work.
- L59 The Backward Dynamical Approach (BDA)...
We will follow the suggestion.
- L106: Since the study area is located in the Alps, this equation seems very low. Has this threshold been validated before (e.g. in Rosatti et al., 2019)? Could you report the validation result? Before going on with the reading, the readers should know how reliable this threshold is.

In Rosatti et al., 2019, two thresholds have been calibrated for the study area following two different approaches to estimate the rainfall condition of the analysed debris flow events: the Critical Duration Method (CMD) (e.g. Restrepo-Posada and Eagleson, 1982; Bonta and Rao, 1988) and the Backward Dynamical Approach (BDA). The obtained CDM-based threshold is the following:

$$I = 4.91 D^{-0.7} \tag{1}$$

and most of the events have durations in the interval $[1\ h, 10\ h]$. *This threshold is consistent with the ones obtained by Marra et al., 2014 and by Idanza et al., 2016 for a comparable study area. The BDA-based threshold:*

$$I = 6.2 D^{-0.67} \tag{2}$$

is higher than the CDM-based one at least by 25% *(the variation changes based on duration) and most of the durations related to the events belong to the interval* $[0.1\ h, 1\ h]$. *These durations are coherent with the time scale of the observed debris flow.*
As mentioned in a previous answer, we are currently working on the validation of the threshold.

How I said, this value seems very low to me and maybe another information should be provided about how the rainfall threshold is operated. E.g. within a long rainfall event a shorter but more intense burst of rain could easily reach the hourly peak intensity of 6.2 or higher. In that case is the threshold exceeded or not?
According to the BDA method, the intensity and duration associated to an occurred debris flow event are computed starting from the surveyed deposited volume. Then, the rainfall volume per unit area E, *Eq.* (4) *of the manuscript, has to be identified in the event hyetograph. If the example described by the*

referee refers to an occurred event, the $I - D$ couple has to be computed starting from the deposited volume and the other input parameters. If the example refers to the forecasting phase, the referee has highlighted one of the most problematic aspect of the threshold use in the forecast: how to define the exceeding conditions of the threshold if the rainfall strictly pertaining to a debris flow (not all rainfall) and its duration is not known a priori? As mentioned above, we are currently developing a suitable method to validate and use the threshold in the forecast.

- L112:"some": which one? The ones in tb. 1?
  Exactly. We will make the sentence more clear.

- L157-160. This is not clear to me. If I understood correctly, in section 3.1 you set N=100, obtaining 100 random "points" in the I/D plane. Now, you randomly pick one of the I/D points, and you do it 5000 times. I guess most of the points are sampled many times. Because 5000 Âż 100. Moreover I do not understand how you can generate a threshold for each one of the selected points (5000 thresholds), since you should use many points to define a threshold (in short: many ID couples are needed to define a single a-b couple). I think I misunderstood something in this part, therefore I suggest to rephrase or to explain better.
  Perhaps we were not clear enough in explaining this part of the method. Performing the first Monte Carlo (MC) simulation (explained in Sect. 3.1), we have obtained $N = 100$ (I, D) couples for each occurred debris flow event. In the case study, we have considered 84 events. This means that the total amount of (I, D) couples obtained from the first MC are $100 \cdot 84 = 8400$. As regards the second MC (described in Sect. 3.2), an input sample is generated randomly selecting one of the possible 100 (I, D) couple for each event. Hence, one sample consists of 84 (I, D) couples. Following this procedure, among the possible combinations, 5000 samples, each comprising 84 (I, D) couples, have been generated. We will rephrase this part of the method to make it clearer.

- L175. For what concerns = regarding?
  We will follow the suggestion.

- L209: please check: the flow of the text is broken by the table and the image. It is hard to follow (similar issues elsewhere)
  We agree. We will optimize the position of the figures and tables.

- L238-239: please check the text: possible issues.
  We will fix these sentences inserting a better description of figure 8.

Graphical improvements

- Figure 2: I strongly suggest adding another panel to this figure, where the DTM of the study area could be shown together with other relevant data (e.g. debris flows locations).
  We will follow the suggestion.

- Please check the text immediately before and after images and tables. Sometimes sentences are split.
  We agree. We will optimize the position of the figures and tables so as not to split sentences.

Bibliography

Bonta, J.V., Rao, A.R., 1988. Factors affecting the identification of independent storm events. J. Hydrol.98, 275–293.
Iadanza, C., Trigila, A., Napolitano, F., 2016. Identification and characterization of rain-fall events responsible for triggering of debris flows and shallow landslides. J. Hydrol.541, 230–245.
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.
Restrepo-Posada, P.J., Eagleson, P.S., 1982. Identification of independent rainstorms. J. Hydrol.55, 303–319.
Zhang, S. J., Xu, C. X., Wei, F. Q., Hu, K. H., Xu, H., Zhao, L. Q., and Zhang, G. P.: A physics-based model to derive rainfall intensity-duration threshold for debris flow. Geomorphology, 351, 106930, 2020.