

Review: Referee 1

General Comments

1. The abstract is too short. I strongly recommend expanding it, providing more details on what has been done and the results obtained.

Answer: The abstract was expanded (#words in revised abstract: 200) to provide details on what has been done and to better describe our findings.

2. In my opinion, the introduction has two main shortcomings and should be reorganized:
a) A literature review should be added to describe the existing state of the art and to account for previous works in the field of rainfall thresholds. You can start with some recent reviews (Segoni et al., 2018a; Guzzetti et al., 2008; Guzzetti et al., 2007) and some recent relevant works (Cannon et al., 2011; Jakob et al., 2012; Lagomarsino et al., 2015; Peruccacci et al., 2017; Rossi et al., 2017). Also, NHESS recently published a special issue on a related subject and it could be another starting point. b) The part concerning the description of the case study should be placed into another section devoted to the study area description, merged with the 1.1 section and re-numbered as 2.

Answer: Answer: Introduction was modified in accordance with the suggestions made by Reviewer 1. A literature review was provided to describe the existing state of the art and previous works in the field of rainfall thresholds. All the suggested reviews (Segoni et al., 2018a; Guzzetti et al., 2008; Guzzetti et al., 2007) and recent relevant works (Cannon et al., 2011; Jakob et al., 2012; Lagomarsino et al., 2015; Peruccacci et al., 2017; Rossi et al., 2017) were included in the revised introduction. Please refer to the revised Introduction in Pages 2 and 3 (P2L29–P3L20). The part concerning description of study area is included in section 2 as requested.

3. (a) What happens when you have two or more landslides during the same storm and they are located in two different pixels? Are they characterized by different I/D values?

Answer: When one storm causes multiple landslides which is often the case, they are represented by same I/D values. The text has been modified to reflect this statement in P8L15–18.

(b) You should better explain if you use the mean rainfall intensity or the peak intensity. In other words: do you consider the total duration of the whole rainfall event or do you identify a shorter section of the event in which you have the higher intensity?

Answer: We used the shorter section of the event with higher intensity. This has been clarified in P8L15–18 in the revised text.

(c) The comparison with literature threshold is questionable. First, I find little meaning in a graphical comparison among thresholds defined for completely different settings (e.g. the Alps), different processes (e.g. post-fire debris flows), and different background (e.g. Caine used only debris flows with a relevant impact, that's why his global threshold is much higher than yours). Second, it is not clear how you assess which threshold is the most similar to yours. Similar intercept? Similar gradient? To sum up, I suggest either deleting this part or normalizing the threshold equations by the mean annual precipitation

(this should allow a straightforward comparison).

Answer: As per the suggestion, this part was omitted from the revised text.

4. The discussion of the results needs to be strengthened. I can suggest some inputs, feel free to explore other directions: a. After your validation, I understand that you have 6 true positives and 7 false positives. How many false negatives? If there is none, that's something that deserves to be stressed. And these results should be further discussed. E.g. the balance between false positives and false negatives is usually a trade-off that could be modified adjusting the calibration of the curves. b. It would be very interesting for the community working on rainfall thresholds, to highlight what difference in forecasting accuracy is obtained (i) when the only Faifa curve is used as a generic threshold for the whole area; (ii) when your more refined pixel-based approach is used. c. The results shown are maybe too weak for the implementation of a EWS (more false positives than true positives). This is something that should be clearly accounted for. I suggest stating that this work represents just a starting point towards the implementation of a prototype EWS.

Answer: The discussion was added to the revised validation section to explain the large number of false positives compared to the false negatives. Please refer to P19L18–P20L9 in the revised text.

5. Just an advice: in the title and in the text you use the term “intensity-duration curve”. This is correct, however in the international community the term “threshold” is more used than “curve”. Your choice is correct but it could penalize your work in on-line search engines and databases indexing.

Answer: As per the advice, the “curve” in the title was changed to “threshold”. Also, the word “threshold” has been used at multiple places throughout the manuscript.

Specific Comments

P1L11(and elsewhere in the manuscript): If I have understood correctly, your methodology cannot be used for a EWS, because it doesn't use precipitation forecasts. Instead, there are two possibilities: either you state that your methodology can be used for nowcasting of landslide hazard in near-real time, or you state that your methodology represents a prototype version of a EWS that could be implemented in the (near?) future. The back-analyses you performed tested the potentiality of the prototype for future EWS applications. Please, rephrase the text where appropriate if my comment is correct, otherwise please describe better the possibility of providing forecasts for a EWS.

Answer: We have clarified in the revised abstract and throughout the manuscript that our work represents significant steps towards the development of an early warning and nowcasting systems for landslides (EWNSL) over Faifa (P1L9–11, P20L9 P22L1).

P1L19 The second type results. . .

Answer: A word “type” was added after the word “second” in the revised text (P1L25).

P2L9: Please, remove blank line

Answer: The blank line was removed in the revised text (P2L15).

P2L10-20: You use rainfall data from two different satellite missions (TRMM and GPM), which use different sensors. In general, such circumstances should be analyzed carefully because different sensors may produce slightly different measurements. Can you be sure that the precipitation values coming from these two different datasets are consistent.

Answer: This has now been explained in P2L23–25.

P2L19-22: I suggest cutting this part (and a similar part which is found later in the manuscript). I think that the value of a research paper is to propose alternative approaches that have not been experienced before, therefore in my opinion the fact that you don't propose a "classic" ID threshold is not a drawback of your manuscript. On the contrary, it potentially makes your work more original and interesting. Concerning antecedent rainfall, I share your opinion of poor constraints with debris flow triggering. I suggest moving this sentence in the state of the art review. You can write that despite many recent works on rainfall thresholds took advantage of the use of antecedent rainfall or other hydrologic constraints (e.g. Posner et al., 2015; Bogaard et al., 2018; Segoni et al., 2018b), there is a general agreement that in case of debris flows on granular terrain (with relatively high hydraulic conductivity), the triggering time is well correlated with peak intensity and duration of the triggering rainfall (Caine, 1980; Guzzetti et al., 2008; Chen et al., 2015).

Answer: Lines (P2L19-20 and P20L4-5) were deleted and the part concerning the antecedent rainfall was moved to discussion section. Refer to P21 L33–34 in revised text.

Page 3, Figure 2: I suggest deleting this figure. A figure based on such a large area is not needed for your work. It would be better to add a zoom on the study area as a third panel of figure 1.

Answer: Figure 2 was deleted, and a zoom was provided as Fig. 1b in revised text.

Page 6, Figure 4, step II, Block 2, 2nd box: please, change "paxe" with "pixels".

Answer: "paxe" was replaced by "pixels". Please refer to Figure 3 in the revised text.

Page 7, lines 17-18: "the threshold. . . the ID curve". This sentence is not clear and I think it is redundant with the previous one. I suggest deleting.

Answer: The referenced sentence was omitted from the revised text (P8L24).

Page 8, Table 1. Please remove the blank row.

Answer: Blank row was omitted from the revised Table 1.

Page 8, Table 1. Something in this table is not clear to me. When you have a pair of Google Earth images bracketing two or more storms, how can you assess which storm triggered the landslides? In the text you explain that the landslides are given to the larger storm (lines 8-9). However, from the table it seems that landslides are shared among all the storms (see e.g. entries 2-3 and 4-5-6).

Answer: We have clarified the situation in P9L13–15.

P9L5: How did you extract his curve? Is it just a subjective manual sketch to low-bound the experimental data or did you use some more robust approach? In case you use

subjective visual approach, you should comment in the discussion or conclusion that this is the weakness of the methodology and that this should be another improvement to carry out during next phase of the research, when more data will be available.

Answer: We used the manual approach as mentioned in P10L7-8. We indicated in the revised introduction section (P3L13–20), methods (P18L21–22) and in the revised conclusion that this is a weakness of the methodology that should be addressed during the next phase of the research.

P9L5: Please, write the equation of the curve.

Answer: We did. Refer to P10L5 in the revised text.

P10Fig5: Please add in the legend a key explaining the meaning of the dots. Add in the figure the threshold equation. Remove these information from the caption.

Answer: We did. Please refer to the revised Figure 4.

P10L8: There is no need to recap what you obtained in the previous chapter. If you want to introduce the section, you could briefly state what is the objective of the next steps.

Answer: The redundant introductory statement was omitted from the revised text. Please refer to section 3.2 in the revised text.

P13L9: Instead of colors, please use class names that are related with the physical meaning of the classes.

Answer: We did. The physical meaning of the classes was used instead of the colors in the revised text (P14L11–12 and P15L4–6).

P14L4-6: In my opinion this part is interesting and I would like to have more details, to evaluate quantitatively the outcomes. Can you provide quantitative statistics? E.g. the percentage of pixels $>3\sigma$, $>2\sigma$, and $>1\sigma$ that you found outside landslides and inside landslides.

Answer: We did. The requested statistics was added in the revised text. Please refer to P15L8–10).

P14L5: If you find a single pixel with these characteristics, is it enough to define a landslide? Or do you need to find a cluster of pixels? Please, clarify.

Answer: We clarified in the revised text that two conditions are needed: cluster of pixels with $>1SD$ on the difference images. Please refer to P15L4–7 in the revised text.

P14L6: These landslides were not “predicted”. I would rephrase with “identified with the proposed method”.

Answer: This has been rephrased in the revised text (P15L10).

P15L7: Maybe the correct reference is to Fig. 3a.

Answer: With merging of Figs. 2 and 3, this has now become Fig. 2a. Please refer to Table 2 in revised text.

P16L3: Please delete “/line”.

Answer: We did. Please refer to P18L2 and P18L4 in the revised text.

Section 2.5: A sounder mathematical approach would help understanding the methodology. How were the A and B curves defined? I guess you translated upward the general Faifa curve to fit experimental data. Is that correct? Please, provide explanations

Answer: We further explained in the revised text how each pixel threshold was identified and how it could be improved in the future. Please refer to P18 (L10–25).

Figure 9: Please, modify the figure to make it as much self-explaining as possible. E.g., write the curve equations, provide a key in the legend to explain the meaning of the colors of the points.

Answer: We did. Please refer to Figure 8 in the revised text.

Section 2.6. Is it possible to show a figure depicting the pixels with correctly identified landslides (true positives) and the pixel forecasted as unstable (further subdivided into true positives and false positives)?

Answer: As per the suggestion, Figure 9 has been added.

P19L3: Maybe “limited size” is more appropriate. By the way: does a landslide smaller than 25m represent a significant hazard in your study area?

Answer: In the revised text, we changed the word “limited distribution” to “limited size” as suggested (P21L12–13). Only if houses and infrastructure were in the way, could a small landslide (<25 m) represent a significant hazard.

P19L10-23: Usually a research describes what has been done and what will be done in the future should be summarized in a few lines. I suggest deleting the part dealing with future work. At least, reduce it consistently. In addition, most of these sentences could be moved in the state of the art description in the introduction (see some of my previous comments).

Answer: The referred section has been broken into discussion and conclusions in the revised text.