We would like to thank the editor David Peres and the two anonymous reviewers for the time taken to handle, read and comment our manuscript. We provide below here our response to the editor and reviewer comments (in blue fonts), together with the actions we did on the manuscript to address these comments (in *blue italics fonts).*

Response to the Editor

Thank you for revising your manuscript. Both previous referees have provided their comments and still some criticism remains on a few points.

In addition to referee comments, I think that the manuscript could be strengthened in some aspects.

I think that the needed modifications to the manuscript can be managed by minor revisions addressing the comments by the referees and my additional ones.

So I now invite you to submit a revised version.

We thank the Editor for his comments and will reply to each comment

1. Social media info is mentioned but few details are provided about it and how it has been used for the analysis.

Indeed, in the previous version we did not include many details about this aspect. The reason is that, despite the efforts (we manually search for photos and videos - in Facebook groups and news websites - of significant flashfloods events or roadblocks during the periods of interest), we could only find one piece of information that helped us in our analyses. The only relevant piece of data that could significantly contribute to our research was a Facebook video from the afternoon of October 27, 2015. Considering this, we agree it is probably more appropriate to downplay the importance of this aspect. We thus rephrase the abstract and conclusions removing this detail, and we provided additional information in the relevant section (4.2): *"We manually searched several websites (including Facebook, YouTube and newspapers) for relevant photos and videos of floods that could help us determine the flow characteristics (debris flow or flash flood) and identify the exact timing of the event. We could only find a Facebook video showing deposits on the main road (route #90) (https://www.facebook.com/chen.sason.75/videos/1155321937831332, last accessed: Apr 25, 2022), where the road crosses the most southern stream of mapped deposits from group 1 on Oct 27, 2015. We could use this information to pinpoint a specific event and a most likely triggering cell for group 1 DFs: Oct 27, 2015 between 13:55-14:40"*

2. The definitions of "initiating storm", "triggering cell" and "antecedent precipitation" are somwhat unclear. Also, they (at least apparently) differ from other definitions in the literature. Perhaps the manuscript could benefit from a sketch explaining your definitions, as well as by some discussion on how (and why) they differ from those introduced by others.

Thanks for the suggestion. Unfortunately, there is no literature definition for such concepts, mainly due to the fact that most of the studies focus on processes such as landslides that occur on longer

temporal scales. Following the recommendation from one of the referees, we decided to slightly edit these definitions to: *triggering storm*, *triggering cell* (within a triggering storm), *antecedent rainfall* (the rainfall occurred within the triggering storms and anticipate the triggering cell). We rephrased the manuscript accordingly and we included a figure (new figure 5) to provide a visual representation of the concept.

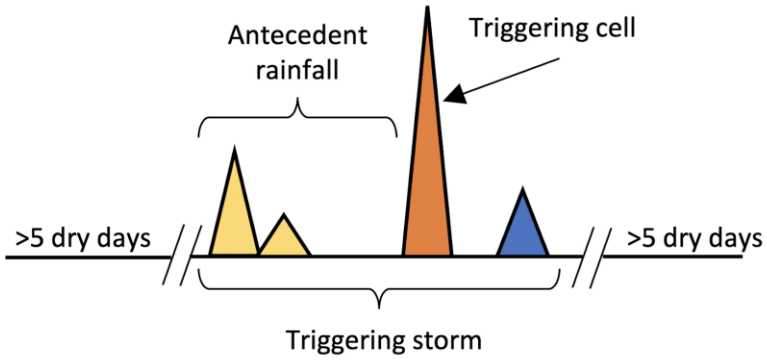

Figure 5: Schematic representation of the triggering storms showing: triggering storm, triggering cell, and antecedent rainfall. The colored triangles represent the rainfall throughout the triggering storm: antecedent rainfall (yellow), triggering cell (orange), and rainfall after the DF triggering in blue.

3. Also reading and interpretation of fig. 5 is quite difficult, so try improving the plot.

Thank you for this comment. We redrafted the figure to make it clearer and we rephrased the caption to help the reader: *"Seven sub-plots of the rainfall intensity versus time (for two-year interval) for seven DFs related to four different groups. In each sub-plot, two graphs are shown: the rain intensity in 5-min interval (in red) and the smoothed intensity observed over 30-min moving windows (in blue). Horizontal lines mark the two thresholds of 10 mm h^{-1} (red) and 5 mm h^{-1} (blue). Any storm during which the 5-min intensity exceeds the 10 mm h^{-1} thresholds or the 30-min intensity exceeds the 5 mm h^{-1} threshold, is marked by a red or blue dot, respectively. Storms exceeding both thresholds for all the DFs of a group are considered as potential initiating storms (marked by dashed light blue box), among them, the chosen storms are marked by solid light blue box".*

Response to Reviewer #1

General comment: "The second version is a much improved manuscript that better presents the originality of the work, while clarifying the points where I had concerns.

I have listed below a few small ideas/comments for improvement. Note that I refer to the line numbers in the track change version."

We thank the reviewer for his positive feedback. Please find our detailed reply below.

Response to minor comments:

Comment #1: Line 30: order of references is not correct.

Thank you for this comment. Actually, NHESS guidelines for authors (https://www.natural-hazards-and-earth-system-sciences.net/submission.html#references) give freedom for what concerns the order of the references: "In terms of in-text citations, the order can be based on relevance, as well as chronological or alphabetical listing, depending on the author's preference". The only constraints are on the references list and on the way single and multi-author contributions are cited.

Comment #2: Line 47: "Albeit rare". This is rather subjective I think. Any reference to support this?

Thank you for pointing this out. We revised the sentence removing "Albeit rare,": *"DF activity in this region can be particularly dangerous to human settlements, main roads and nature reserve infrastructures."*

Comment #3: Line 64, reference is made to tropical areas (Brand, 1992). It am not sure to understand the reasoning behind the highlight on this type of climate in particular. In addition, would this text be kept in the final version, there is certainly more recent work on rainfall triggering conditions in these regions. I think for example in regions such as Puerto Rico, Rio De Janeiro, Central Africa and the Philippines. But again, if extra info is provided on rainfall conditions in the tropics, we shall expect that it is used to support the discussion.

We thank the reviewer for this comment. Indeed, we are not interested to highlight tropical areas. We rewrote the sentence to better explain this point: *"It was suggested that in areas where slopes are covered by permeable soils or sediments, antecedent rainfall is less important in reducing the potential of failure (Rahardjo et al., 2001; Brand, 1992; Corominas and Moya, 1999)."*

Comment #4: Line 71: "to our best knowledge". I would suggest the authors to explore (if not done yet) the research carried out by Sepúlveda and co-authors (see references below) to make

sure that such a statement is as robust as possible. Overall, the works carried out in Latin America by these researchers is worth to be mentioned in the state of the art and possibly in the discussion.

We thank the reviewer for suggesting these references. They are definitely relevant to our introduction, although they do not address the specific topic of the sentence. We rephrased the relevant part in the Introduction to account for them: *"In the arid northern Chile, Sepulveda et al. (2006) identified possible relations between large-scale climatic drivers, such as El Nino episodes, and debris flow triggering. Similarly, Sepulveda et al. (2014) closely examined a debris flow triggering by an extreme rainstorm in the same region."*

Comment #5: Figure 1. The background map displays location names that cannot be (easily) read. I know that this is not trivial for the purpose of the paper, but that would be logical to not show names that cannot be read.

We changed the background map to improve the readability of the figure.

Comment #6: Line 167. "… DF deposits THAT occurred…"

We updated the revised manuscript.

Comment #7: Figure 4. The font size of the elevation quotes in the legend "DSM differences" may be too small.

Thank you, we increased the font size.

Comment #8: Line 224. I know that in this region, precipitation is mostly related to rainfall. However, to stick to the terminology used in the title I would pay attention that the word "precipitation" is not used too often. In this sentence specifically, I would write: "… of the rainfall that led to DF initiation".

Thank you, we updated the revised manuscript accordingly.

Comment #9: Lines 225-226. Definition of the "initiation storm" based on the 5 days criteria can be backed-up by the literature, or is this a custom-based definition? Maybe a word on this should be added.

The definition is custom for our study case. We look for a separation that is sufficient to completely dry the soils of the area. Please note that given the climatology of the region (only few storms occur every year), there is basically no sensitivity to this choice for separations greater than a few days. We now specify this: *"To this end, storms are defined as wet periods separated by at least 5 days of dry weather (i.e., 120 hours with less than 0.1 mm h-1 in the radar data; note that given the climatology of the area with very few storms per year, there is little sensitivity to this choice)."*

Comment #10: Lines 239-241. How can it be said that rain gauge information is less reliable (underestimate) than radar data? Is this because the rain gauges are not directly in the DF? I think that this paragraph could be presented in a clearer way to say that radar data are used to assess the rainfall triggering conditions and that rain gauge data are used to constrain/validate radar data.

Essentially, the problem stems from two factors: (1) convective precipitation has sharp spatial variability and (2) rain gauges measure precipitation in the wrong location (not above the triggering location but usually few km away). Since DF triggering needs high-intensity rainfall, the two aspects above concur in creating the systematic rain gauge underestimation. We rephrased this part to better explain this aspect: *"In fact, rain gauges tend to systematically underestimate the DF triggering rainfall because they sample precipitation fields away from the triggering locations and because these fields are characterised by sharp spatial variability and necessarily have high intensities over the triggering locations (Marra et al., 2016; Destro et al., 2017; Marra et al., 2014; Nikolopoulos et al., 2014, 2015)."* In addition, we added a sentence to provide the information suggested by the referee: *"Weather radar data are used to assess the rainfall conditions above the DFs, and rain gauge data are used to constrain and validate the radar data on a regional scale."*

Comment #11: Lines 275. Replace "eleven" by "11".

We updated the revised manuscript.

Comment #12: Line 362. "… most likely triggering stormS identified in this study ARE marked…"

We updated the revised manuscript.

Response to Reviewer #2

General comment: I appreciate the efforts of the authors to improve the manuscript following my prior suggestions. However, the manuscript requires further work to improve the structure and scientific English. Although much work has been put into the new version of the manuscript, it can sometimes be difficult to read, informal and vague in parts (e.g. lines 151, 164-166, 302-310, 319-328, 339-340). Especially sections 4.2 and 4.3, which were extensively modified in the new version of the manuscript, require major revisions to improve the readability.

Additionally, in the new manuscript version, the authors use many abbreviations throughout the paper which were not used in the original version of the manuscript. Some are introduced once and then not employed again in the entire manuscript (DoD, DEM, DSM). Others are introduced and then applied inconsistently (DF). Mathematical signs are used embedded within the text to, for example, refer to approximately, lower/greater than, etc. I suggest avoiding abbreviations when terms are not used recurrently throughout the paper. I also suggest avoiding mathematical expressions in the middle of sentences.

We thank the reviewer for the thorough review. We will response to three points raised in this general comment. We went over the text and in the revised version we suggest some changes in order to clarify and improve the text, especially where the reviewer points on the specific line numbers.

We removed the abbreviations for DEM and DSM and only left DoD which is used extensively in the text.

Similarly, we removed the in-text mathematical expressions and replaced them with appropriate words.

Response to minor comments:

Comment #1: Line 87 and lines 399-400: This should be discussed in the context of water content/saturation of the soil using the well-established concepts described in the principle of effective stresses (Terzaghi 1943).

The current paper focuses on the rain conditions leading to the triggering of DFs in arid regions and not the specific failure mechanism which is the goal of our followed-up future study. Therefore, we choose to add a short sentence about sediment strength reduction, including two fundamental references: *"We speculate that antecedent rainfall or the flow of water in the steep channels caused by antecedent rainfall could reduce the sediment strength by wetting. The effects of pore pressure increase and the reduction of apparent cohesive strength most likely cause this strength reduction (Takahashi, 2014; Terzaghi, 1943). The strength reduction may be followed by enhanced incision or direct failure of the sediments within the channel."*

Comment #2: Line 168 and 171: what is N?

Here like in other situations, N is used with the meaning of "number". We rephrased with words.

Comment #3: Line 201: You use the term "triggering storm". In my opinion, the term triggering storm is a much more intuitive concept than the terms introduced in the definitions in the 1st paragraph of section 4.

Thanks for the suggestion. Unfortunately, there is no literature definition for such concepts, mainly due to the fact that most of the studies focus on processes such as landslides that occur on longer temporal scales. Following your recommendation and the editor's, we decided to slightly edit these definitions to: *triggering storm*, *triggering cell* (within a triggering storm), *antecedent rainfall* (the rainfall occurred within the triggering storms and the triggering cell). We rephrased the manuscript accordingly and we included a figure (new figure 5) to provide a visual representation of the concept.

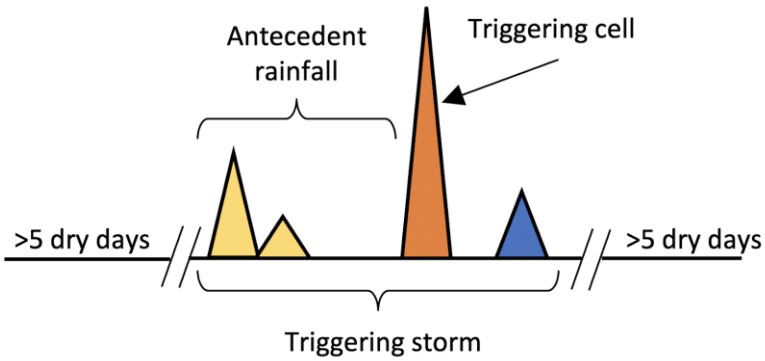

Figure 5: Schematic representation of the triggering storms showing: triggering storm, triggering cell, and antecedent rainfall. The colored triangles represent the rainfall throughout the triggering storm: antecedent rainfall (yellow), triggering cell (orange), and rainfall after the DF triggering in blue.

Comment #4: Section 4.2: How many rainfall events have you considered for your analysis at the end?

As mentioned in the text: "Combining the above information, we ended up identifying the four triggering cells that, based on the information at our hand, most likely triggered the DFs in the four groups (Table 1, right column)." we considered four triggering cells that were developed during three different triggering storms.

Comment #5: Lines 203-211 state that the debris flows have been divided into groups according to their triggering time and the distribution of the events over the area. It would be nice to know how many debris flows are included in each group and the deduced approximate triggering times. A table including such information would be handy.

Following this suggestion, we added these details to the revised Table 1.

Comment #6: Line 242-243: Is it possible that some of the debris flows had been triggered by the other rainfall events?

In general, yes, as we don't have direct observations we cannot exclude it. However, we did our best to consider all the available knowledge and reduce its likelihood. We added a statement about it in the revised text: *"In total, we identify 11 potential triggering storms that satisfy the above intensity conditions over the DFs during the periods of interest (Table 1). Nevertheless, only seven of these 11 storms satisfied the conditions over all the mapped DFs of a group of interest (marked by bold text in Table 1 and blue rectangles in Fig. 6). Therefore, only seven storms can be considered as potential triggering storms. Note, that one of them (Oct 26-29, 2015) is a potential triggering storm for the DFs of two groups. We should remember that the process of identifying the triggering cell could be subject to errors so that a different storm or cell, or more than a single convective cell may have triggered the mapped DFs in a group. The procedure we adopt, however, closely follows our knowledge of DF triggering in general and of local climatology of storms in particular, so this possibility is expected to be unlikely."*

Comment #7: Line 266-270: Isn't analysing the information in social networks a method to identify debris flows?

Yes, we agree. When such information is vastly available social networks alone can be sufficient to identify and label natural hazards. In our case we could only use it to confirm the type of flow for a given event and to pinpoint the time of occurrence.

Comment #8: Line 285: Is that telling us some information about one specific storm that triggered one group of debris flows or all of them? This paragraph requires some clarification.

Indeed, we have done this check for all four groups, but we present the results in Fig. 8 and the detailed explanation in the beginning of section 4.3, only for group 3. The text was revised as follows to clarify this paragraph: *"To further explore this point, we plotted for each group the spatial distribution of antecedent rainfall over the study area and compared it with the spatial distribution of the total rainfall yield of the triggering cells, and with their maximum intensities over 5- and 30-min intervals (an example of such a plot is given in Fig. 9). Figures 9b through 9d present the triggering cell characteristics of group 3 (in the center of the map), and suggest that two additional convective cells (warmer colors) north and south to the main cell passed through the studied area at the same time."*

Comment #9: Line 306: Is there a specific reason why a different definition of antecedent rainfall is used here? The authors state that the same triggering storms can be picked up in most cases using this new definition. What happens with the debris flows for which the triggering storm changes?

Thank you. Yes, we defined the antecedent rainfall in a slightly different way: while for the other storms we could identify the storm period after identifying the storm (with 5 days separation), here we run an automatic search over the data and had to define automatically the antecedent rainfall over a moving window. However, we use a 5-day window (120 hours) that, given the climatology of the area, makes the two definitions nearly identical, with the only rare exceptions of storms lasting more than 5 days, for which this definition may lead to lower amounts of antecedent precipitation. We clarify this in the text: *"Here, the antecedent rainfall is calculated automatically over the entire time series as all the rain accumulated on the deposit pixel over the 120 hours preceding the specific measured intensity. This automatic definition provides estimates of the antecedent rainfall that are equivalent to the ones used earlier, except for the rare storms lasting longer than 5 days prior to the peak intensity; in these cases, this automatic definition might be underestimated."*

It is not fully clear what the referee asks with the second question, we hope the clarification above solves the issue.