R1:

This is an interesting and highly timely manuscript that will make a strong contribution to our understanding of rainfall-triggered landslide occurrence. It builds logically on prior work by the authors and others on large-scale analysis of the growing body of landslide inventories, and also adds to those inventories by providing some additional datasets to the community. I strongly recommend publication after fairly minor revisions. Most of my comments are aimed at clarifying aspects of what the authors have done. The writing is mostly clear, but in places there are some minor inconsistencies or elements of the explanation that feel less complete or a little rushed. I've made some comments and suggestions on the annotated PDF; I won't repeat all of those here, but instead I've included the more substantive comments below, tied to line number.

We thank the reviewer for their supportive comments and appreciate their effort to review our study. We feel that the comments made are both fair and constructive, and we will in revision work to incorporate their input. Below, we provide detailed responses to the individual comments.

Line 80: I agree with all of the statements in this introduction, but at the same time I feel like part of the argument is missing. The previous paragraphs make the case for why inventories are important and why landslide location data are useful, but there isn't really a problem statement - what is the gap that the authors are trying to fill? The abstract makes the point that prior work (like that of Milledge et al. NHESS, and also work by Jeandet and Gallen and colleagues) has looked at the characteristic locations of coseismic landslides to try to derive some underlying patterns or rules that can be useful for understanding exposure. It's logical to extend that work to look at rainfall-triggered landslides, for all of the reasons listed above, but that point isn't made explicitly here. I think the paper would be strengthened by making that case in the intro, to help convince the reader of why this is needed. This need not be a lengthy addition, but a few extra sentences could make a big difference.

This is a great point. We will include an additional sentence near the close of the introduction:

"Finally, recent studies have sought to derive underlying simple topographic rules to understand hazard associated with earthquake-triggered landslides [e.g. Milledge et al. 2019], and it is important that we extend this kind of analysis to rainfall triggered events to provide comparative data.

83-84: 'We suggest that...' - I agree, but for the reasons outlined above I think this also undersells what you are doing. If there really are patterns or rules that can be gleaned from this kind of analysis, then that is incredibly valuable for hazard assessment in areas where more detailed models or investigation have not or cannot take place. Similarly, if there are no such patterns, that's also valuable. Milledge et al. 2019 NHESS made this point explicitly and I think you could do the same here.

Again, we appreciate the reviewer pushing us to amplify the value of our study. We will add an additional sentence to the closing paragraph of the introduction:

"Moreover, with a set of simplified rules for landslide hazard, researchers can support hazard assessment in areas where more detailed models may be unavailable."

119: This point about image resolution is a tricky one, because it depends upon the definition of completeness - even 1 m imagery will not catch every event, especially in areas of either sparse

vegetation or very dense vegetation (where for example small events are still hidden under the canopy, or contained within channels). I mostly agree with this statement, but it is an oversimplification to say that higher resolution = 'better' in all cases. It's also true that high resolution imagery suffers more from issues with rectification (and sometimes thus georeferencing), which can have an impact on inventory creation and analysis - again, see Williams et al. 2018 NHESS for an example of this.

We will add a caveat statement to the end of this paragraph to clarify that higher resolution imagery does not always lead to better quality inventories:

"It is important to note that although high resolution imagery can provide more accurate mapping in some cases, it can also be more challenging to ortho-rectify, which can limit the quality of landslide inventories generated [Williams et al. 2018]"

138: 'We therefore consider...' - I'm sure you're right, but this is a pretty broad-brush statement. Quality in terms of landslide location, or information on size and geometry, or both? Does it matter for 'quality' that two of your new inventories were produced with Sentinel-2 and the others with Planet imagery?

This is a fair comment, and we will remove the statement in revision.

140: Does the 'Date' column in Table 1 represent the dates of the triggering rainfall? What were the time windows over which images were collected (given that those are almost certainly longer)? This has implications for the relationship between rainfall data (which themselves are aggregated or simplified compared to what actually hit the ground) and landslide occurrence – put simply, for many inventories we cannot be certain of when landslides occurred unless the images very closely bracket the dates of the storm. This is relevant because of the discussion of the potential differences between extreme and persistent rainfall in triggering landslides in different parts of the landscape (lines 505-510).

The date in Table 1 does indeed refer to the date of triggering rainfall. For the newer inventories, Planet data allows us to closely bracket the rainfall (within a few days) and many of these inventories are the result of rainfall that exceeds annual averages by a significant margin, suggesting that the bulk of landslides are generated during short intense bursts of rain that we consider. However, the reviewer is correct that this issue may be important, and so we will include the following statement to highlight this potential issue:

"It is important to note that the date of the triggering rainfall is not identical to the dates the imagery used to map the landslides was obtained. Although we have selected events where the triggering rainfall significantly exceeds historical peak rainfall (and therefore is likely to be the dominant trigger for landslides) some events may have occurred as a result of lesser rainfall before or after. While the new inventories generated for this study that utilise Planet imagery that closely brackets the rainfall events (within 1 week either side), the older inventories may be more subject to this challenge."

235: Does this mean that you have rasterised all of your inventories? Are you assuming some kind of majority rule to go from landslide polygons (in your inventories) to pixels (for comparison to the continuous raster variables)? And is the resolution of those rasters 1 arcsec? Are you therefore censoring the smaller events? I think this needs to be clarified. Further down in the ms it emerges that in fact the inventories have been rasterised and with a 'presence' rather than 'majority' rule, but to address the questions that I had at this point in the ms, I think that information needs to come sooner.

We will move the paragraph explaining the GLM methodology to later in the text to ensure that the description of landslide pixelation occurs prior to where it is explained in the GLM context.

253-254: 'We utilize the method...' - I think it would be useful to give a very brief description of how this is done, so that the reader does not need to go to the Marc et al. (2018) paper to understand this.

We will add the following explanatory text to describe the method:

"We utilize the method of Marc et al. [2018] to extract the scar areas, which uses the perimeter and area (A) of landslide polygons to calculate aspect ratio of an equivalent ellipse, K, (Marc and Hovius 2015) and the associated width (W) , according to the formula $W \cong \sqrt{(4A/\pi K)}$ (Marc et al., 2018). The scar area is defined as $\sim 1.5W^2$ based on a global database of scar aspect ratio (Domej et al 2017)."

Related to this: I think you use 'scar' and 'headscarp' interchangeably throughout the ms, and I would suggest focusing on one or the other. To me, the headscarp is a particular part of the scar area - they are not synonymous

We will replace all uses of 'headscarp' with 'scar'; thank you for flagging this.

406: 'While we have attempted…' - this does raise the question, though, of how we 'ground truth' landslide inventories in a meaningful and rational way. I think that is an issue that's beyond the scope of this ms, but at the same time there is an implicit assumption in this kind of phrasing that inventories can be 'corrected' by hand and presumably brought closer to a true representation of landslide occurrence. Whether that's the case or not and what we mean by 'better' are both open to discussion. I'm not suggesting that you address these points in any detail, but some recognition that 'correction' is challenging might be useful here.

The reviewer raises a good point here. We agree that hand-correcting is not necessarily the gold standard, with subjectivity an important problem. We will rephrase the highlighted sentence as follows:

"Although we have used hand-corrections to reduce the impact of polygon amalgamation from algorithmic mapping methods, some inconsistencies may still exist."

450-453: 'Characterizing these parameters…' - this is a little confusing, because you HAVE used a globally available rainfall product as well as a globally available forest cover and loss product. So I don't see how the final sentence is correct. Or am I misunderstanding what you mean?

We acknowledge the potential for confusion here and will rephrase this sentence as follows:

"Although global data for rainfall, soil type and geological parameters exists, the resolution of these datasets is too low to allow for consistent comparison of landslide and non-landslide areas at the scale of the analysis described here (~100 m)."

506-510: 'Nevertheless, we also suggest…' - I agree that this would be a really instructive comparison, and it's great to point that out here. But this also raises the issue that I flagged above, of the difference between (1) the time window over which high-intensity rainfall occurred in these areas and (2) the time window over which landslides were mapped. If the latter is much greater than the former, then I think you need to be a little careful about the inferences you draw from the landslide patterns. I don't think this invalidates what you have done, but I think it's important enough to warrant a mention.

We thank the reviewer for raising this important point. We will add the following clarifying sentence to the end of the paragraph to ensure this point is communicated effectively:

"In any future comparative study of low-intensity and high-intensity rainfall events, it will be necessary to carefully select landslide inventories where the imagery used to generate them closely brackets the start and end of the rainfall events, to ensure only landslides triggered by an individual event are analysed."

Figures

Despite the fact that Figs 3-8 all show essentially the same thing, you use variable labels for both x- and y-axes. I suggest choosing a consistent naming convention and sticking to it, to make things easier for the reader. Otherwise, it's necessary to repeatedly work out what is being shown.

The reviewer raises a good point here, and we will revise the figures to ensure consistency.

Also, the color scheme used in these figures for the different inventories is really hard to distinguish. Maybe this doesn't matter, if the point is to show that all datasets show similar performance, but picking out an individual dataset and distinguishing it from the others (which is required to evaluate some of the statements in the text) is really tough. Perhaps use a progressive colour scheme instead (e.g., blue to white to red)?

Although we appreciate the point the reviewer is raising, we have tried multiple different colour schemes and still prefer the current system we are using. Progressive colour schemes with so many different inventories make it extremely challenging to pick out which dataset is which.

Fig 3: the x axis labels are not very clear or informative on panels a and b. I think these could usefully be made more precise.

We will revise these labels to ensure clarity in revision.

Figs 4-5: it's not clear why the x axes are reported to different levels of precision here (integers on the one hand, real numbers on the other).

We will revise these figures to ensure that there is a consistent choice for data – we will use real numbers for all data.

Fig 11: this figure isn't really explained in the text, and the axis labels here are identical to those for Figs 6 and 9. The reader could therefore be forgiven for not recognising what you are showing, and not really understanding how this figure is distinct from those earlier figures. I think it would be good to provide a more clear explanation of what this figure shows and how it is distinct from the other figs. The caption is factual but confusingly worded, and as a result I don't think this is very effective.

Citation: https://doi.org/10.5194/nhess-2021-250-RC1