REFEREE #3

ABSTRACT

Page 1 Line 1 Landslide and debris flows are just a part of the geo-hydrological phenomena that can be impacted be fires. Please mention also the other phenomena or clarify that these are only some of them.

> We have addressed this concern by changing the following text on line 1 of the original manuscript:
>
> > "Wildfires change the hydrologic and geomorphic response of watersheds, which has been associated with cascading hazards that include shallow landslides and debris flows."
>
> To:
>
> > "Wildfires change the hydrologic and geomorphic response of watersheds, which has been associated *with cascades of additional hazards and management challenges. Among these continuing post-wildfire events are shallow landslides and debris flows."*
>
> We further return to this statement by the following reference to the Introduction (line 23):
>
> > "Here, we focus on a particular sequence of cascading natural hazards known as the post-wildfire landslide. In these events, wildfires are followed by intense precipitation leading to mass movements *such as shallow landslides, or debris flows. The impact of wildfires, which themselves occur more frequently and severely as a consequence of higher temperatures and increasingly widespread drought, can lead to a variety of geo-hydrological hazards including increased snowmelt, water contamination, increased erosion rates, and decreased infiltration (AghaKouchack et al., 2020).* Post-wildfire landslides in particular occur when wildfires are followed by intense precipitation, leading to mass movements such as a sediment-laden floods, shallow landslides, or debris flows."
> >
> > AghaKouchak, A., Chiang, F., Huning, L. S., Love, C. A., Mallakpour, I., Mazdiyasni, O., ... & Sadegh, M. (2020). Climate extremes and compound hazards in a warming world. *Annual Review of Earth and Planetary Sciences*, *48*, 519-548.

Page 1 Line 4 Please specify why GLC should facilitate regional inter-comparison?

> We appreciate this comment and include the following explanation:
>
> > "Landslide events are selected from the NASA Global Landslide Catalog (GLC). *Since this catalog contains events from multiple regions worldwide, it allows a greater degree of inter-regional comparison than many highly localized catalogs."*

Page 1 Line 8 The authors here speculate on the seasonality of "mass movement-triggering storms" but actually this should be read as the seasonality of mass movements triggered by rainfall. In addition, please specify what "other rainfall-triggered mass movements" means; here is too generic.

> We thank the reviewer for this comment and clarify as follows:

> "An analysis of the seasonality of mass movements at burned and unburned locations shows that mass movement-triggering storms in burned locations tend to exhibit different seasonality from rainfall-triggered mass movements *in areas undisturbed by recent fire."*

Page 1 Line 12 "... characteristics of rainfall-triggered mass movements ...".in general, or only fire-related?

We clarify this text as follows:

> "Overall, this manuscript offers an exploration of regional differences in the characteristics of rainfall-triggered mass movements *at burned and unburned sites* over a broad spatial scale and encompassing a variety of climates and geographies."

Page 1 Line 15 Please specify what the author intend for "path"? Propagation, runout path?

We thank the reviewer for this comment, and will replace this sentence with:

> "M*ass movements are destructive when they occur near vulnerable areas, causing damage to or failure of buildings, utility lines, and roadways (Highland and Bobrowski, 2008).* Landslide mitigation costs in the United States (US) are approximately 2 billion USD annually, with worldwide costs much higher (Schuster and Highland, 2001). *There can also be indirect impacts, such as aggradation of the streambed, or the formation of landslide dams (Glade and Crozier, 2005)."*

Glade, T., & Crozier, M. J. (2005). The nature of landslide hazard impact. *Landslide hazard and risk*, 43-74.

Page 1 Line 24 "sediment-laden floods" or "sediment-laden flows", since the authors use the term mass- movements, I assume they refer to the second.

We appreciate this point since floods would not be included in the inventory, and edit the text to:

> "Here, we focus on a particular sequence of cascading natural hazards known as the post-wildfire landslide. In these events, wildfires are followed by intense precipitation leading to mass movements *such as shallow landslides, or debris flows*."

Page 1 Line 29 Since this is a generic/general statement should be referred not necessarily to US recent works. In addition, I suggest to avoid defining "meteorology" or "length of time since the most recent fire" as factors, since those may be just indirect ways to refer to proper landslide conditioning factors. AND Page 2 Line 36 This sentence is really cryptic. What is the "relative magnitude of triggering precipitation events", why this should be a proxy for the susceptibility? Is this a result of the study or is an initial assumption?

We have reconfigured these last two paragraphs of section 1 to address the two concerns above:

> "There are numerous local studies demonstrating a relationship between wildfire occurrence or severity and the amount of precipitation that triggers a mass movement (Cannon et al., 2008; Gartner, 2005; Reneau et al., 2007; Riley et al., 2013). The impact of wildfire on landslide hazards can also vary on the basis of local factors such as vegetation, and soil type (Cannon et al., 2010; Staley et al., 2018). In general, the lack of complete landslide inventories including a wide variety of

climates and ecoregions presents an obstacle to an evaluation of the role of fire in rainfall-triggered landslides, and regional studies are exceptional (Klose, 2015b).

This study seeks to test the hypothesis that wildfire consistently increases mass movement susceptibility across six global regions. In this large sample of mass movement events (n=5313), higher expected frequency of precipitation is indicative of greater mass movement susceptibility. Though we cannot draw conclusions about the susceptibility leading to any particular event, higher median expected frequency suggests that the threshold for triggering a mass movement was lowered, e.g. susceptibility was greater. A second purpose of this study is to explore the possibility that the relationship between wildfire history and the expected frequency of landslide-triggering precipitation varies by regions."

Page 4 Line 90 Here use the term "widely recognized relationship" in place of "statistically significant positive relationship", since you are mostly referring to literature and not to specific statistical tests results.

We appreciate this suggestion and have made the following change to the text:

"The *widely recognized* relationship between mass movements and burn severity suggests that mass movement susceptibility increases after wildfires in the Western US,"

Page 4 Line 100 "is a universal phenomenon" seems a bit ambitious here.

We modify this sentence to:

"A global study by Riley et al. (2013) comparing post-wildfire a non-fire-related debris flows found that the volumes of the post-wildfire debris flows tended to be smaller. This finding suggests that the increase in debris flow hazard and frequency after wildfires *occurs in a variety of environments*."

Page 5 Line 124 "precision" or "accuracy", here and at line 132 you are using the wo worlds but they do not refer to the same problem. In addition, completeness may not be necessarily a problem when using inferential statistics.

We thank the reviewer for this comment and adjust the text as follows:

For this study, we chose to use the NASA Global Landslide Catalog (GLC, Kirschbaum et al., 2010). As with the few other regional and global databases available, the broad domain of the GLC comes coupled with issues of *location error and spatial bias. For each landslide location, the GLC contains an estimate of the area in which the landslide occurred, labeled the "location accuracy". For consistency, we refer to this parameter using the same name.*

…

Despite limitations, the GLC was chosen for this study primarily because as of this writing it offers the largest spatial range of any catalog.

Page 5 Line 145 Here and in the rest of the text, the readers have the impression that the authors just considered part (mostly US) of the literature. I'm not a specific expert of fire related mass movements, but just a quick search on the main scientific literature search engines revealed also specific studies in other part of the world. Since the authors use this lack of local studies as one if not the main justification for the work, this problem is relevant. Hence, please account also the other studies in different countries and modify the text accordingly.

> Finally, in contrast to post-wildfire mass movement studies focused on a specific regions like the western US (Cannon and DeGraff, 2009), southern California (Gartner et al., 2014)*, Western Canada (Jordan, 2014), Korea (Jong-Ook et al., 2018)* or southeast Australia (Nyman et al., 2011), this study combines the GLC with globally-observed fire and precipitation data to offer unique insights into the role of fire on mass movement susceptibility in diverse regions across the globe.

Jordan Peter (2016) Post-wildfire debris flows in southern British Columbia, Canada. *International Journal of Wildland Fire* **25**, 322-336.

Lee, J.-O., Lee, D.-K., & Song, Y.-I. (2019). Analysis of the potential landslide hazard after wildfire considering compound disaster effect. *Journal of the Korea Society of Environmental Restoration Technology*, *22*(1), 33–45. https://doi.org/10.13087/KOSERT.2019.22.1.33

Page 5 Line 151 What is the rationale behind the choice of "the seven-day running total precipitation depth percentile for the 30 days surrounding the day of the year"? Which percentile do the authors refer to? Which day of the year is used? Need to clarify why these should be used as a "proxy for mass movement susceptibility". The rest of the paragraph till the beginning of section 2.1 give a series of details that are just confusing the reader. This is really cryptic, for this reason I suggest the authors to simplify this part, identifying the methodology with "understandable" steps and demanding the specificities and explanations of methods to the dedicated sections.

> First, the seven-day running total precipitation depth percentile for the 30 days surrounding the day of the year and across the total 38-year record (see Sect. 2.4) was used as a proxy for mass movement susceptibility. *We assume here that greater susceptibility results in a lower precipitation threshold to trigger a landslide. An observation, therefore, of lower precipitation percentile values triggering storms across a sample of sites suggests that susceptibility is generally higher in that group. This principle is illustrated in the susceptibility-based rainfall threshold model developed by Monsieurs et al. (2019), in which the predicted threshold of antecedent rainfall resulting in a landslide is adjusted according to susceptibility factors.*

Monsieurs, E., Dewitte, O., & Demoulin, A. (2019). A susceptibility-based rainfall threshold approach for landslide occurrence. *Natural Hazards and Earth System Sciences*, *19*(4), 775-789.

Page 6 Line 162 It will be more correct defining this a "sample" and not a "large sample".

> This change has been made

Page 6 Line 174 What does "recorded locations" refer to? Fire location?

> We have clarified this text as follows:
>
> > *Only records with location uncertainty of 10 km or less* were included,

Page 6 Line 180 Please explain how the "hierarchical clustering algorithm" based solely on latitude and longitude is able to highlight/account for climate differences. This will be an important information to complete the description of the procedure.

> To compare the differences in mass movement triggers in different climates, we divided the mass movements into regions (see Fig. 1 panels (a)and (b)). Regions were determined using the AGglomerative NESting (AGNES) hierarchical clustering algorithm (Kaufman and Rousseeuw, 2009) considering the latitude and longitude of the mass movements, and clusters were subsequently combined, split, or eliminated on the basis of equalizing sample sizes as described below. *Though the regions are still large enough to encompass considerable variability in climate, the spatial clustering helps to ensure that the variability across regions – particularly in latitude - is larger than the variability within.*

Page 9 Line 205 Event the opposite case is possible. Please comment in the text.

> We thank the reviewer for this observation and correct the text as follows:
>
> > *Due to uncertainty in the locations of many of the mass movement locations, both false positive and false negative errors in burn history classification are possible. Some mass movements classified as burned may have occurred near a recent fire but not within the fire perimeter, or conversely some mass movements classified as unburned may in fact have been located inside a fire perimeter but near the edge. However, by classifying mass movements as burned if any part of the potential location was burned limits the potential for false negative errors while increasing the possibility of false positive errors.* For this reason we refer to mass movements as 'burned' instead of post-wildfire in this analysis. *Also important to note is that false positive burned classification* is a function of both the burned fraction and the conditional probability of mass movement occurrence given that a fire has occurred.

Page 9 Line 218 What does "significantly significant differences" mean?

> We appreciate this observation and have corrected the text as below:
>
> > The number of days with *statistically* significant differences in precipitation percentile in the 14 days prior to the mass movement and 7 days are computed in each group.

Page 10 Line 231 It is not clear how a "7-day running average of antecedent precipitation" is able to highlight "storms of different lengths and intensities"? Please specify.

> "Mass movements can be triggered by intense and short storms, by long storms of lower intensity, or somewhere in-between. A seven-day running sum of antecedent precipitation was computed to enable direct comparison of the mass movements triggered by storms a range of durations."

Page 10 Line 233 Please specify what type of "7-day antecedent rainfall indices".

> We clarify the text below and provide additional support for this choice:
>
> > While including an estimate of the soil moisture was outside the scope of this study, 7-day antecedent rainfall indices *consisting of a weighted average of precipitation over the 7-day time period* have been used by other modelling studies as a surrogate for soil moisture in a combined indicator of landslide susceptibility (James and Roulet, 2009; Kirschbaum and Stanley, 2018). *Furthermore, 7-day sums of precipitation have been found to perform better than other durations in threshold models of landslide occurrence (Krcak et al., 2017; Garcia-Urquia et al., 2015).*

Garcia-Urquia, E., & Axelsson, K. (2015). Rainfall thresholds for the occurrence of urban landslides in Tegucigalpa, Honduras: an application of the critical rainfall intensity. *Geografiska Annaler: Series A, Physical Geography*, *97*(1), 61-83.
Krkač, M., Špoljarić, D., Bernat, S., & Arbanas, S. M. (2017). Method for prediction of landslide movements based on random forests. *Landslides*, *14*(3), 947-960.

Page 10 Line 234 "more equal comparison of mass movement triggers which fall within throughout this spectrum of storm intensity"? This is really cryptic. Please specify. In addition, "is less sensitive to small errors in precipitation" is really tautologic, since this is an aggregated measures; better removing it. It is unclear why being "less sensitive to mass movement date accuracy" is an advantage for this type of analysis. Maybe this hides important causal relations between rainfall and post-fire landslide occurrences, in line with what you have mentioned before about the importance of runoff related phenomena compared to the infiltration related once.

We have removed this sentence in favor of the explanation above.

Page 10 Line 239 The sentence is not clear! Please rephrase!

We clarify below:

> "Upon computing the CHIRPS precipitation measurements for each event, we noted that some of the ostensibly rainfall-triggered mass movements had no antecedent precipitation in the 7-day window. We removed such mass movements from the analysis."

Page 11 Line 256 How do you exactly normalized these value (i.e. which kind of calculation did you do)? Please also explain exactly why the normalization you are performing should "facilitate the comparison of mass movement-triggering events across a variety of seasons and climates". Is this based on previous study? Please also justify why "this statistic controls for geographic and seasonal differences across mass movement events". Which is the rationale behind that?

In an initial analysis of the precipitation data, we were unable to distinguish between normal seasonal increases in precipitation and the mass movement-triggering precipitation. In order to isolate that triggering storm, it was important to normalize for both location and time of year. We accomplished this by computing a 30-day rolling percentile of the 7-day running precipitation values based on 38 years of historical precipitation climatology from 1981–2019 for each location. Percentiles have been used to compare landslide-triggering precipitation across larger, e.g. country-sized regions (Kirschbaum et al., 2020 and Araujo et al., 2022) in order to control for differences in climate or precipitation data source. For this study, the percentile produced a uniform distribution of precipitation ranging from 0 to 1, controlling for geographic and seasonal differences.

Araújo, J.R., Ramos, A.M., Soares, P.M.M. *et al.* Impact of extreme rainfall events on landslide activity in Portugal under climate change scenarios. *Landslides* **19**, 2279–2293 (2022). https://doi.org/10.1007/s10346-022-01895-7

Kirschbaum, D., Kapnick, S. B., Stanley, T., & Pascale, S. (2020). Changes in extreme precipitation and landslides over High Mountain Asia. *Geophysical Research Letters*, 47, e2019GL085347. https://doi.org/10.1029/2019GL085347

Page 11 Line 265 Again, here and hereafter, it's unclear why the percentile should serve as a proxy for relative mass movement susceptibility?

Please see the additional explanation above

Page 11 Line 270 Mann–Whitney test does compare ranks and not median directly, or at least it may compare medians under certain circumstances and distribution assumptions (i.e. the two samples should have the same shape). Please check.

Please see the correction below:

"The null hypothesis of the Mann–Whitney test was that the distribution of precipitation percentile of the burned sites is generally greater than or equal to the distribution of precipitation percentiles of the unburned sites. (Helsel et al., 2020)"

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020, Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chap. A3, 458 p., https://doi.org/10.3133/tm4a3. [Supersedes USGS Techniques of Water-Resources Investigations, book 4, chap. A3, version 1.1.]

Page 11 Line 272 The authors do not provide any evidence on the fact that data are uniformly distributed, and is unclear how Sect 2.4 should guarantee this. In addition, Mann–Whitney test does require any uniform distribution. Instead it requires that the two analysed samples should be measured on a ordinal or continuous scale and should not normally distributed.

See the clarification below:

Percentiles are by definition uniformly distributed rather than normally distributed, making the Mann–Whitney test, which does not require normal distribution, the most appropriate hypothesis test for these data.

Page 11 Line 280 It is unclear what " ...with the actual sample number adjusted by region so that all sites were selected evenly" mean.

We clarify the text below:

(n= the smallest number above 100 that ensured each site was included in the same number of samples).

Page 11 Line 283 What do you intend whit "lead time"? Is this the time of occurrence of the mass movement?

See the clarification below:

These samples are representative of precipitation for a particular *number of days before* the mass movement and serve as a control dataset with which to compare the pre-landslide precipitation.

Page 11 Line 283 If I well understood you "control dataset" include those rainfall cumulative values in the selected accumulation period (i.e. the 7 day period) in a period of 15 days before and after the date of landslide. This assume a pretty constant rainfall characteristics over time and in particular over the same 30 days period across all the years. Given the general rainfall variability I'm not sure this 30-days period length is large enough to considered all the possible climatic variability in the selected area, and I believe it should be better to consider a seasonal period to estimate the "control" references. Alternatively, the authors could show the influence of the selection of the length of the "control" period on their results. This comment is someway related to the previous comment "Page 10 Line 231", please consider them related (the larger is this accumulation period, the larger should be the control period).

We appreciate this concern. The choice of 30 days was motivated by

a) preliminary analysis in which it was difficult to distinguish landslide-triggering precipitation from ordinary seasonal changes. For example, precipitation typically increases substantially over time during the fall months in California, and so a seasonal percentile would result in a typical but misleading amount of increase in precipitation percentile that
b) Climatology data are frequently monthly

Page 11 Line 295 What do the authors intend for seasonality? Please give a definition or refer to a reference one.

2.7. *Mass movement* seasonality experiment

The probability of landslide occurrence in a given temporospatial domain varies throughout the year (Stanley et al., 2020); we refer to this annual pattern for a given domain as mass movement seasonality. We hypothesize that wildfire alters mass movement seasonality. To test this hypothesis, we estimated precipitation frequency at the mass movement sites over time by computing the fraction of sites in the burned and unburned groups that had precipitation on any given day.

Stanley, T. A., Kirschbaum, D. B., Sobieszczyk, S., Jasinski, M. F., Borak, J. S., & Slaughter, S. L. (2020). Building a landslide hazard indicator with machine learning and land surface models. *Environmental Modelling & Software*, *129*, 104692.

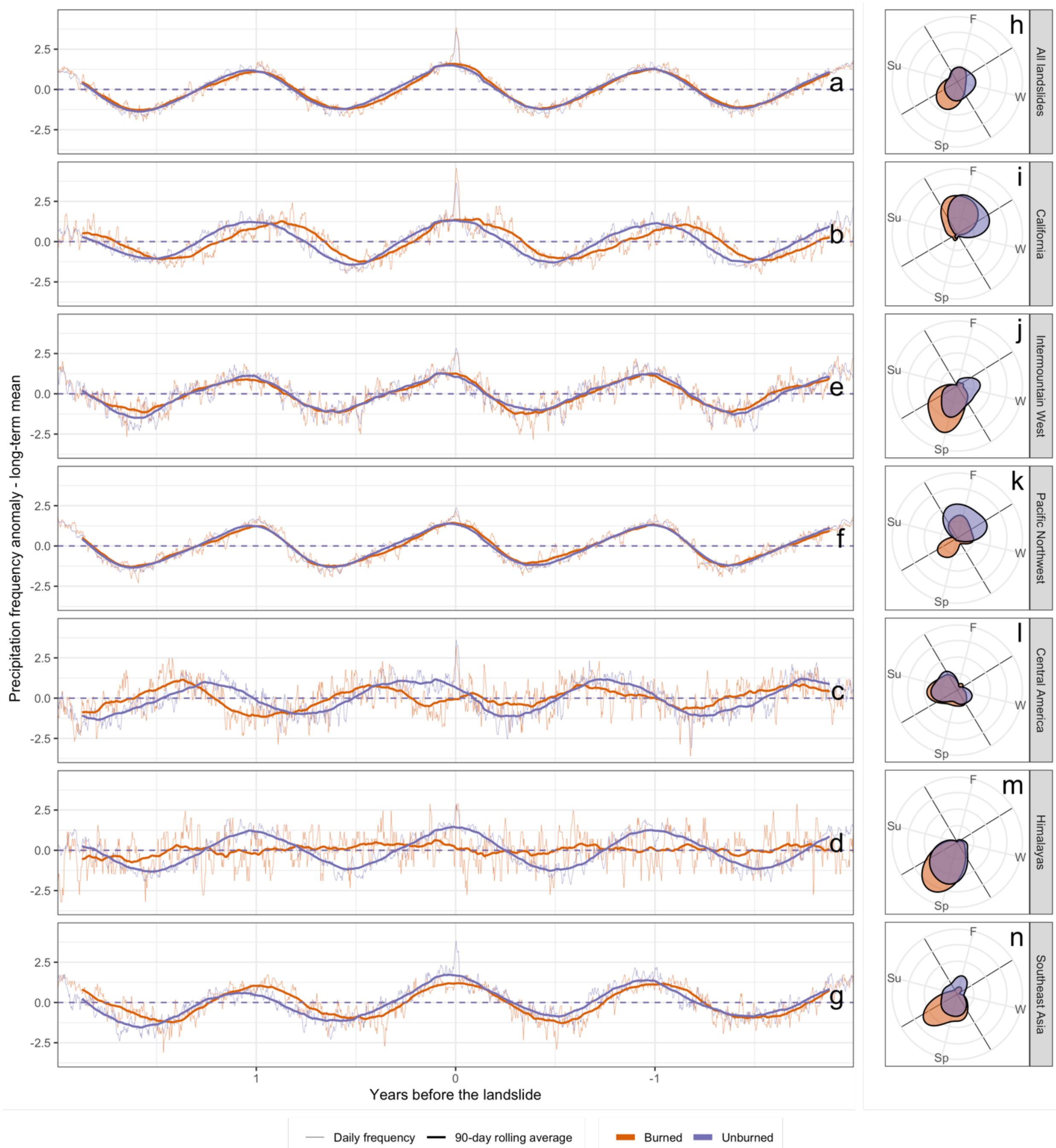
Page 12 Line 301 "frequency" of what?

We found that in most regions there was a long-term difference in the mean annual *precipitation* frequency

Page 12 Line 301 "These persistent differences between burned and unburned sites were removed by subtracting the mean precipitation frequency for both the burned and unburned groups": is the intent of the authors to perform a variable scaling? Why don't they do a full variable standardization dividing by the standard deviation? Please explain better how this should be useful in the analysis?

We thank the reviewer for this suggestion, and have revised the analysis to standardize the data:

> These persistent differences between burned and unburned sites were removed by standardizing the mean precipitation frequency for both the burned and unburned groups, that is to say subtracting the mean and frequency and dividing by the standard deviation.



Page 13 Line 319 From here to the end of the paragraph: so basically, you got confirmation on the effect or fires only in the area in US with previous studies, while your main hypothesis is to test this possible effect worldwide. Please comment

See the following changes:

> *At first glance,* this difference supports the overarching hypothesis that wildfire increase mass movement susceptibility, since mass movements in the period after a fire can be triggered by less precipitation than might be required in the absence of fire. *However,* an examination of each region separately reveals that the difference in precipitation percentile between burned and unburned sites is present in some

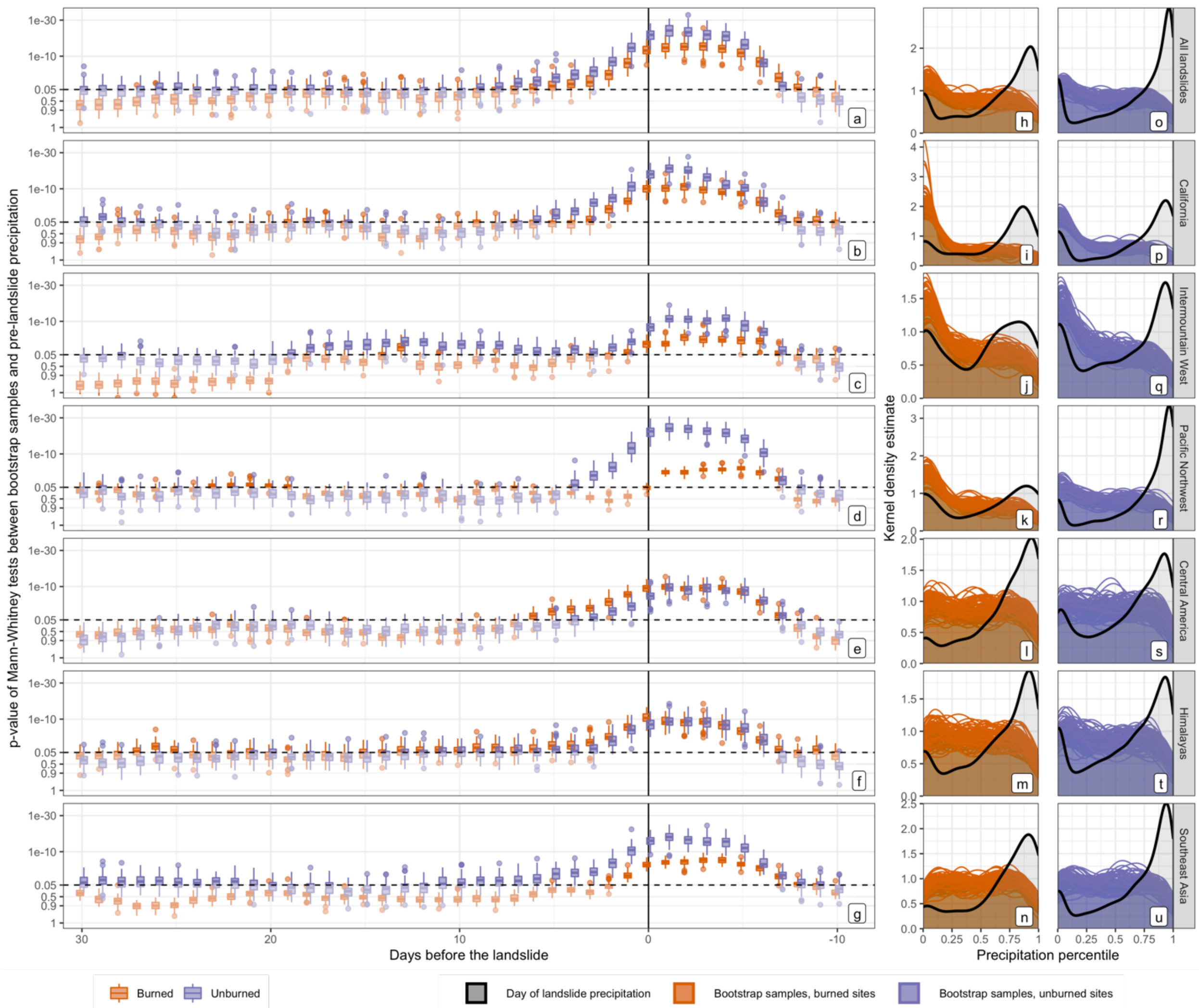
regions but not in others (see Fig. 3). The California area (Fig. 3 panel (b)) has a particularly strong signal, whereas tropical regions do not show any significant *decrease* between precipitation at burned and unburned sites. *In total, these initial results suggest that post-wildfire landslides are isolated to areas, such as California, where such cascading hazards have been repeatedly observed.*

Page 13 Line 326 Figure 4 and figure 5 are only showing the p-value. Without boxplots it is impossible to check whether a p-value greater than 0.05 correspond to a precipitation percentile in burned areas lower than the unburned ones. This is a relevant information for the analyses. I suggest to realize plots or multiple plots similar to that in fig 3 or modifying the plot in a way to show such information. Please homogenize the p-value threshold descriptions in the figure captions.

We are unable to display a box plot or spread for the p-value, since for these plots represent a direct comparison of the full burned and unburned groups. We could consider a bootstrap analysis such as in Figure 6, if the reviewer feels this would add to the analysis.

Page 17 Line 343 I assume the figure 6 refers to the output of Mann–Whitney test even if in they axis is specified ”Wilcox” test, which actually should read a “Wilcoxon”. Maybe this is just a refuse since the Mann-Whitney test is also called Wilcoxon rank-sum test, but there is a completely different Wilcoxon test called “Wilcoxon signed-rank test” which test something different. Since some of the interpretation of p-value done by the authors make me think on the use of this last, I ask them to check. Indeed, it seems (i.e. al least observing plots styles) that these test have been done in R which uses the same function (i.e. wilcox.test() fuction) to perform different types of Wilcoxon tests. Please highlight the meaning of the different colour tones in the caption/tex.t

The statistics were computed with R’s wilcox.text with paired=False, making it a Mann-Whitney test. Please see the modified label in the plot below:



Page 17 Line 349 “with lower values”: similarly, to comment Page 13 Line 326, p-value may just highlight some difference but not their positive or negative sing. This information in figure 6 can be only appreciated in panels (h)-(u) but only for the case of the day-of-landslide precipitation.

We have modified the text as follows

> In panels (a)-(g), box plots of p–values represent the degree to which the mass movement-triggering precipitation differed from climatological precipitation with lower p-values indicating a more significant difference between the two precipitation distributions. The Mann-Whitney tests were directional, so only differences where the precipitation is greater than would be expected result in low p-values.

Page 17 Line 357 “with implications for potentially region-specific physical processes associated with mass movement triggers”: the analysis of timing of rainfall/storms in different regions cannot say anything in this regard, these may only say something on the rainfall characteristics leading to landslides which presumably depends from regional climatic difference.

We have modified the text as follows:

> Different storm timing is apparent among the regions, and between the burned and unburned sites of the same region.

Page 17 Line 358 “in the Himalayas and Southeast Asia (Fig. 6 panels (f) and (g)) precipitation rises at a similar rate for each group”: It is really difficult to appreciate this from the figure! Indeed, there seems to be substantial differences between burned and unburned areas. In addition, from here to the end of the paragraph all the speculations/comments of the rainfall intensity do not find support from the analysis of the figure, or at least the authors do not provide all the information to appreciate this (only panels (h)-(u) may provide such information). For instance, why “In the Pacific Northwest and California (Fig. 6 panels (d) and (b)), the burned sites exhibit shorter but more intense storms than the unburned sites in the week preceding the mass movement”?

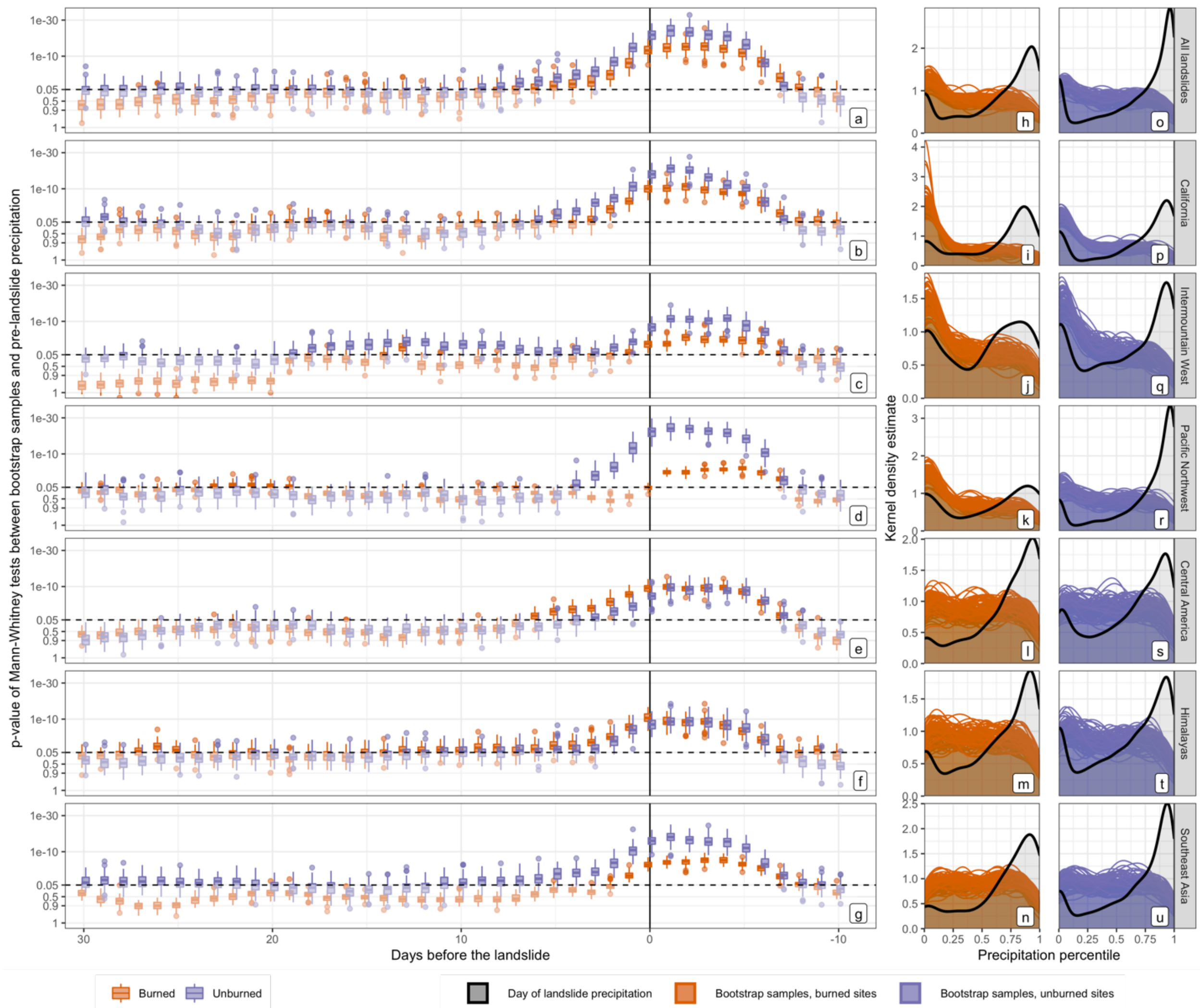
We have revised this paragraph as follows:

> Firstly, in the California and Southeast Asia (Fig. 6 panels (b) and (g)), we see a similar pattern where relative precipitation at unburned sites is consistently higher than at burned sites. Nonetheless, for both burned and unburned sites, the rise in precipitation takes place over a similar amount of time (approximately 5 days). Curiously, unlike in California, the bootstrap analysis reveals a long-term difference between burned sites and unburned sites in the Mann–Whitney p–value for Southeast Asia despite location-specific normalization, suggesting that the mass movements at unburned locations might be primarily triggered in years that are wetter than usual on a monthly or seasonal scale. In the Pacific Northwest (Fig. 6 panel (d)) the precipitation at the burned sites does not become significantly larger than usual until the day of the mass movement. The Mann–Whitney p–values for the burned group remain well above 0.05 just days before the mass movement as the p–value for the unburned group begins to fall. Under the assumption that shorter, storms are associated with runoff-driven mass movements while longer storms that allow more time for the soil column to saturate are associated with infiltration-driven mass movements, this difference in storm timing could reflect that in the Pacific Northwest the burned mass movement locations are largely runoff-driven while rainfall-triggered mass movements at unburned locations are infiltration-driven (Cannon and Gartner, 2005).

Page 19 Line 371 This does not seems the case for all the panels, but is almost impossible to appreciate correctly. Please add 0.9 p-value line or at least their ticks in Figure.

This observation was intended to apply only to the Intermountain West region. We have added the 0.9 tick mark and clarified the text:

> In this region, twenty to thirty days before the mass movement p–values for burned sites are consistently above 0.9, suggesting a high likelihood (> 90 %) that there was less precipitation than usual during that time. During the same period, the p-values at unburned sites remain close to 0.05.



Page 19 Line 378 Fig.5 should probably read as Fig. 7.

We have fixed this error

Page 19 Line 381 4854%? Please check.

We thank the reviewer for noticing this error, and correct it below:

> The mass movements are distributed such that 53 % occurred within a one year after the fire.

Page 19 Line 384 Maybe hereafter, it will be worthwhile to mention the rainfall season. This is the really relevant information to speculate on landslide occurrence (e.g. in the Himalaya the monsoon period is generally in June and bring almost the entire yearly rainfall leading to landslides, but this is not the same for other regions characterized by different seasonal rainfall regimes and distributions). How the analysis on seasonality account this?

For example, California has a long fire season and a shorter mass movement season. When fires occur at the end of winter, immediately after mass movement season, there is typically a longer delay before the mass movement than when fires occur immediately before mass movement season. *In general, the mass movements occur during the period of most rainfall, such as the winter in California and the summer in the Himalayas (see Figure 1 for regional precipitation climatology).*

The seasonal pattern of post-wildfire mass movements is to some degree determined by an interaction between fire seasonality and precipitation seasonality.

Page 19 Line 388 Please see previous comments on this type of plots (Figure 4 and 5), and on their possible interpretation regarding rainfall.

See comment above.

Page 19 Line 393 Why? p-value > 0.05 only for "less then 1 year"

We believe this statement to be correct, since it is a p-value of < 0.05 that indicates statistical significance in this case.

Page 19 Line 395 Why? The p-values are always > 0.05 for all the lines. Here the authors seem using a p-value interpretation opposite to what in Fig 4 and 5 which should be the correct one. Is something missing here for the interpretation?

See previous comment

Page 19 Line 415 This is certainly visible for California, but not for Himalayas!

We agree with the reviewer and have modified the text as follows:

In California (Fig. 9 panel (b)) burned mass movements are clearly shifted to earlier in the time of year with more frequent precipitation.

Page 19 Line 427 In the view of the comments above, the discussion may be revised carefully and have significant changes. In the section some points are really too much speculative with no support from the data and results described in the manuscript: please revise these parts and indicate appropriate references to support them.

Please see the revisions to the Discussion below:

The results of this study suggest that while post-wildfire mass movements are associated with shifts in the magnitude, timing, and seasonality of storms relative to other mass movements, these effects are not consistent across regions. Globally, there are clear differences in the percentiles of mass movement-triggering storms (see Fig. 3), with mass movements in burned areas often triggered by comparatively smaller storms. *At first glance*, this supports the hypothesis that fires increase mass movement susceptibility, since a smaller precipitation trigger is sufficient to cause a mass movement. However, this trend is largely driven by the California region and to a lesser extent the Intermountain West and Pacific Northwest of North America. In Central America/Caribbean, Southeast Asia, and the Himalayan regions there is an increase in *rainfall relative to climatology* leading up to the mass movement, but there is no significant difference between relative precipitation depths based on fire history. The original percentile analysis includes only wet days; the bootstrap analysis takes into account both wet and dry days.

Differences in the mass movement-triggering storms relative to their precipitation climatology shown by the bootstrap analysis (Fig. 6) *confirm the results of the original percentile analysis, with two notable exceptions. In Southeast Asia, the bootstrap analysis indicates that the burned sites had smaller precipitation triggers relative to climatology despite no significant difference in the first analysis between the wet-day precipitation percentiles. This discrepancy suggests* *that there may be some kind of precipitation frequency bias between burned and unburned sites in this region.* In addition, burned locations the Pacific Northwest appear to be associated with rainfall that began closer to when the landslide occurred (Fig. 6 panel (d)). *This raises the possibility that mass movements at burned sites in this region are caused more often than in unburned locations by runoff instead of infiltration. More information is needed, for example about the antecedent soil moisture at these locations.* This result is consistent with previous research suggesting that post-wildfire debris flows are predominantly triggered by runoff-driven erosion as a result of shorter and more intense storms in the Western US (76 % Cannon and Gartner, 2005); *however in that case we would have expected to see a similar pattern in the California and Intermountain West regions.*

Many of the landslides at burned locations in the Intermountain West (Fig. 6 panel (c)) appear to follow a pattern of short-duration storms that occur after a dry spell stretching from twenty to thirty days before the mass movement and possibly beyond. Fig. 9 panel (c) shows low frequency precipitation at burned locations followed by a sharp spike which also supports this pattern. One possible explanation is that the combination of dry and recently burned soil promotes mass movements in this region. *An example of drought conditions contributing to a landslide is described by Handwerger et al. (2019).* These differences could also be due in part to the different regional climates, with the California and Pacific Northwest regions having more clearly defined longer-duration rainy seasons, relative to the more variable and sporadic precipitation seasonality of the Intermountain West.

Different combinations of fire season, mass movement season, and any overlap between the two may be an important driving factor in the degree to which fires increase mass movement susceptibility. For example, in places where the wet season begins towards the end or immediately after fire season, such as the Intermountain West, California, and the Himalayas, the landscape has no time to recover from the fire before mass movement season begins and therefore burned locations may be much more susceptible (see Fig. 7 panels (b), (c), and (f)). On the other hand, in regions like the Pacific Northwest, Central America, and Southeast Asia (Fig. 7 panels (d), (e), and (g)), where mass movement season is not as well defined, it is more likely that the landscape could at least partially recover before a triggering storm occurs. Some of the regions that did not display a significant difference in percentile nonetheless showed a shift in the timing of burned mass movements relative to their respective annual pattern of precipitation (see Fig. 9 panels (h)–(n)). The various types of shifts in landslide seasonality are likely reflective of the different effect of fires. A shift of the mass movement season to slightly earlier in the year, such as was noticeable in California and the Himalayas (see Fig. 9 panels (i) and (m)) supports the hypothesis that wildfire increases mass movement susceptibility because it suggests that fewer or smaller precipitation events earlier in the season are sufficient to trigger a mass movement. The Intermountain West (Fig. 9 panel (j)) also has a pronounced seasonal shift, but in this case the shift is much larger: burned mass movements appear to occur an entire season later than unburned mass movement, falling in the driest part of the year instead of the wettest. This corresponds to the evidence from the bootstrap analysis suggesting that dried out soil or slow vegetation regrowth may be an important part of the post-wildfire mass movement mechanism in this region. Vegetation regrowth as a main control of mass movement susceptibility is supported by a study of mass movement occurrence in the San Gabriel mountains of the US by Rengers et al. (2020), in which the authors found that hillslopes with slower vegetation regrowth were more likely to have mass movements. A similar trend to the Intermountain West in terms of seasonal shift is visible for some, but not all, of the mass movements in the Pacific Northwest (Fig. 9 panel (k)), suggesting perhaps that some of mass movements in that region would have been better categorized as part of the Intermountain West region. In Southeast Asia (Fig. 9 panel (n)) there also appears to be a seasonal shift similar to that of the Intermountain West, but it is not matched by a shift relative to the annual precipitation frequency pattern (Fig. 9 panel (g)). This suggests that the seasonality "shift" in Southeast Asia is *due to spatial bias in fire occurrence. Further study of variation in climate across this region is needed.* Finally, Central America (Fig. 9 panel (l) has very similar precipitation in burned and unburned locations. Since there is little difference between the precipitation frequency or magnitude (see Figs. 3 panel (e), 9 panel (e)) in this area, it is possible that there are many misidentified false positive post-wildfire mass movements in Central America, perhaps due to the generally low location accuracy in that region. It is also possible wildfire does not have as much of an effect on mass movement susceptibility in that region.

The timing of mass movements relative to wildfire may also influence the magnitude of triggering storms. While in some regions, such as California and the Pacific Northwest, timing does not have a major impact on precipitation percentile differences, the Intermountain West of the US displays two distinct behaviors depending on the timing of mass movements relative to wildfire. In the year immediately after a fire, the precipitation percentile is lower than for mass movements at unburned locations in the seven-to-three days before the mass movement, before rising to match precipitation percentile at unburned locations (see Figure 8 panel (c)). This pattern matches the result from Figure 9 panel (c) in which post-wildfire mass movements in this region appear to manifest as a large storm preceded by a period of infrequent precipitation. In contrast, timing appears to make little difference to the precipitation percentile in other regions.

Handwerger, A.L., Huang, MH., Fielding, E.J. *et al.* A shift from drought to extreme rainfall drives a stable landslide to catastrophic failure. *Sci Rep* **9**, 1569 (2019). https://doi.org/10.1038/s41598-018-38300-0

REFEREE #2

The study areas are limited only in the USA, Central America, the Himalayas, and Southeast Asia. As a result, discussion and conclusions should focus on the differences of rainfall intensity and duration that exist in each area and distinguish the respective thresholds that initiate landsliding in each region.

As the reviewer points out, the database contains a variety of types of landslides from different geological settings. We are unable to draw any region-specific conclusions about intensity-duration thresholds because the precipitation conditions leading to landslides are highly variable on the basis of soil, geologic, and anthropogenic factors. We focus instead on relative precipitation amounts, e.g. the degree of abnormality of the triggering storm events. Furthermore, the precipitation dataset used is only available at daily resolution, which would be insufficient to determine triggering intensity. We have included additional references to studies which use precipitation percentile rather than the more standard intensity-duration threshold in order to encompass regional or global data:

> Percentiles have been used to compare landslide-triggering precipitation across larger, e.g. country-sized regions (Kirschbaum et al. 2020, Araujo et al.2022) in order to control for differences in climate or precipitation data source.

The geological setting which is one of the most important predisposing landslide factors is entirely absent from the paper. For example, geological conditions and landslide mechanism is entirely different in the Himalayas and North America. Authors should comment on that. In addition, they should categorize all landslides in surficial and bedrock type landslides since rainfall intensity and duration as a trigger mechanism is completely different between these types. Furthermore, authors should explain if the data sample of n=5313 includes both landslides and rockfalls.

The sample of 5313 does include both landslides and rockfalls, although rockfalls are less than 5%. Many of the landslides in the GLC are labeled as a non-specific 'landslide' type, which could refer to rock falls as well as well as any other type of mass movement. We have clarified this in the text:

> Finally, though the GLC does contain some information about the mass movement mechanisms that would allow mass movements to be classified, for example, as debris flows or shallow landslides, the majority of the events in the GLC are labeled as a non-specific 'landslide' type, which could refer to any type of mass movement.

It is not clear if earthquake or snowmelt triggered landslides have been eliminated from the data sample of n=5313. Which was the original sample number, how many earthquakes or snowmelt triggered landslides were eliminated and which are the respective maps, with the sampling areas before and after the landslide elimination?

The sample size, 5313, represents the sample size after elimination, and is the sample shown in all plots. We have clarified this as follows:

> The GLC contains a large sample of mass movements from across the globe (n = 11,377, 5313 of which meet study requirements – see Section 2.1)

The data set used in the analysis has no uniformity. It should be on the same temporal range, following the range of the NASA Global Landslide Catalog 1988–2015, or 2007-2015

This is an error in table 1 – the version of the NASA GLC used here spans from 1988-2018. We have also included the following text to specify the study time period:

> We first describe the mass movement data (Sect. 2.1), the study regions (Sect. 2.2) and fire data (Sect. 2.3). Mass movements were included only if precipitation data and at least 3 years antecedent fire data were available. The mass movements occurred between 2007 and 2019, with corresponding precipitation and fire data extending as far as 2004-2019 so as to capture antecedent conditions.

Since the authors have used MODIS Terra and Aqua satellites 2000-2020, which is the percentage of potential post-wildfire landslide events among the period 2000-2015 that coincides with the NASA Global Landslide Catalog?

> Table 1 indicates that the MODIS data are available from 2000-2020, but only data 2004-2019 were used. All the potential post-wildfire landslide events were taken from the NASA Global Landslide Catalog. See clarification above.

According to the authors landslides were classified as burned if any part of the area where the mass movement occurred was burned at some point within the three years prior to the event to capture both waves of increased susceptibility without over-identifying mass movements areas where fires occur every few years. Authors should explain why they used there years prior to the event and not more or less than three years

> We direct the reviewer to the cited De Graff and Gartner studies which justify the 3-year period:
>
> > Previous studies suggest that the post-wildfire increase in mass movement susceptibility peaks within the first six months, but that a second period of increased susceptibility can appear at 3 years or even longer as a result of root decay (DeGraff et al., 2015; Gartner et al., 2014). Landslides were classified as burned if any part of the area where the mass movement occurred was burned at some point within the three years prior to the event to capture both waves of increased susceptibility without over-identifying mass movements areas where fires occur every few years.

According to the authors 489 mass movements (9:2%) were categorized as potential post-wildfire events. Which of these mass movements belong to landslides or rockfalls?

> A small percentage of the mass movements were specifically labeled as rock falls (3.3% of burned sites; 4.4% of unburned). However, it is possible that the generic 'landslide' category contains additional rockfalls.

Precipitation analysis (intensity or duration) is different in different types of mass movements. Authors should clarify what types of mass movements they are focusing on.

> We add this paragraph at the end of Section 2.1 to clarify this:
>
> > The GLC contains a variety of types of rainfall-triggered mass movements with different physical mechanisms, including debris flows, shallow landslides, and rock falls. The majority of included mass movements (65%), however, are categorized simply as 'landslide', which according to the dataset authors can mean any type of mass movement. Since most of the mass movements are of an unknown type, we did not exclude data on the basis of category. Of the specific types of mass movements, most are labeled mudslides (25%), with the next largest category being rockfalls at 4%. This uncertainty as to landslide mechanism is currently a necessary trade-off for large spatial scales. This limitation highlights the need for large-scale catalogs for specific types of mass movements, such as debris flows or shallow landslides.

Finally, it is not clear if rainstorms or rainfall intensity have triggered landslides only in the burned areas, since rainfall intensity is a common triggering mechanism in both burned and burned areas. Authors should explain this more thoroughly.

> The GLC claims to consist solely of rainfall-triggered landslides, regardless of fire history. We have clarified this as follows:

This study seeks to test the hypothesis that wildfire increases mass movement susceptibility across six global regions by detecting and characterizing differences in mass movement-triggering precipitation at both burned and unburned sites. The relative magnitude of triggering precipitation events at large sample of *rainfall-triggered* mass movement sites (n = 5313) is used here as a proxy for mass movement susceptibility.

The GLC provides a large collection of *rainfall-triggered landslides* taking place in a variety of climates such that, in combination with spatially continuous observations of fire (500m Moderate Resolution Imaging Spectroradiometer [MODIS] Burned Area by Giglio et al., 2018) and precipitation (5.5km Climate Hazards group InfraRed Precipitation with Station data [CHIRPS] by Funk et al., 2015) data, it is well suited for comparing the diverse precursors under which post-wildfire mass movements occur.