NOTE : In the following documents referees comments are in normal fonts and the answer to them are in bold fonts. My response to the authors' answers are in normal fonts.

Review of Marc et al. "Towards a global database of rainfall-induced landslide inventories: first insights from past and new events" by David Milledge.

Major Comments

This is a well executed study with novel and interesting findings. I have three general comments and a large number of minor comments but neither the major nor minor comments reflect a fundamental problem in the research in my view.

I am not convinced that it is essential (or helpful) to present your inventories as the only inventories that are suitable for this type of analysis (as you seem to do on P2-3). Instead you could simply say they are one set of inventories and they demonstrate the power of this type of approach. I am not convinced of the need for landslides beneath an entire storm footprint to be mapped and am sceptical that entire storm footprints can be convincingly defined so I'm not convinced by your critique of studies that analyse far smaller study areas (other than on sample size grounds).

>> The mostly agree with the referee. We will develop the description along these lines:

1/ number is always important, in absolute term because below 50-100 landslide the reliability of any statistical treatment is uncertain but also in relative term, because a statistical study based on 500 landslides out of a storm that caused ~5000 take the risk to have biased interpretation if the (potential) specificity of the subset population are not notcied and understood (e.g., mostly large landslide ?, mostly landslide near river ? Mostly landslide in a given lithology ? Etc etc).

2/ If enough landslides are mapped within an AOI (e.g., >50-100 in total, > 75% of landslide above the resolution limit, leading to a reasonable frequency-size distribution) the inventory above the AOI is likely to be statistically usable and representative of the various processes and conditions affecting the process in the AOI. Then a partial inventory will indeed allow to study any local parameter and their variations within the AOI : e.g. landslide density, landslide size distribution, relations to slope etc.

3/ However, a comprehensive inventory may have the additionaly advantage to gather enough landslide across different areas (in terms of lithology, relief etc) potentially allowing to establish a hierarchy in controlling parameters and also allowing to study an averaged landslide response less likely to be dominated by specific site effects. comprehensive inventories are the only ones allowing to study the variations of total landsliding.

We will include a synthetic description of these point in the early part of the manuscript. These should answer the various minor comments about the importance of "comprehensive" landslide catalog, and better acknowledge the potential use of "partial" inventories.

I think this will address my concern here. I still think your points boil down to: arguments of sample size, which I agree with (point 1 & 3); and systematic sampling, i.e. all identifiable landslides larger than some minimum size, which I agree with (point 1 &2). You don't demonstrate here that multiple smaller inventories that do not cover the 'full storm footprint' are less valid than a smaller number of 'complete' inventories. It seems to me that sample size is the key here. That is fine, your sample is very large and that makes it a particularly useful contribution, I think building an argument round this rather than 'event completeness' would be more compelling.

The methodology description could be more consistent between inventories. Similar information is reported for each case but the style of the reporting differs and some key information reported in some cases is not present in others (e.g. image source, image resolution, acquisition date).

>> We will work on a higher degree of consistency.

However, we disagree on key image source/resolution/date not beeing present, as all this information is in Suppl Table 1. Is the referee suggesting that we integrate this table within the main text ? Or simply specifying some terms in the main text ?

The table is useful but it would strengthen the paper to make the reporting more consistent within the methods text. This could be done by using consistent language / terms (and covering similar material) for each inventory as you suggest.

I am not convinced that your focus on 'comprehensive' inventories is necessary nor that examination of total landslide numbers, volumes or areas are particularly meaningful in relation to rainfall triggered landslide inventories (though I think the findings on landslide density and slope are extremely interesting and thought provoking). This focus might reflect a desire for comparability to co-seismic landslides but I think the two triggers are importantly different. For example, it is extremely difficult to define the spatial and temporal limits on a single storm. In addition I find the results relating to total numbers, volumes and areas less convincing because they are predicted from a small number of point rainfall records. A clearer explanation of why 'comprehensive' inventories and total statistics are important would be a valuable addition to the paper.

This comment didn't receive a specific response but it overlaps with other discussion on how to define "rainstorm event landslide inventories" in which I think we've reached broad consensus.

Minor Comments

P2 L30: comprehensive: this term needs defining.

>> We mean that all landslides detectable above the resolution limit were mapped, and that we could observe the landslide density fading away in all direction, indicating the limits of the footprint of the high intensity part of the storm. We will specify this two criterium in the main text.

This is addressed.

P3 L7: comprehensive mapping: where do you start and finish. Your definition of a storm is very important here and I don't see it at the moment. For example shouldn't the Morakot mapping extend to the Phillipenes and China on this basis?

>> The pragmatic answer is that the landslide response in Philippines and China was negligible compared to the one in Taiwan (We doubled check that quickly by looking at Landsat images where no hyperpicnal flows or alluviations in stream exiting hilly areas are visible, contrarily to Taiwan where these processes are very clear). Same is true if we look at landsliding in the rest of Japan hit by Typhoon Talas progressing Northward after hitting the Kii peninsula. This may be in part due to topographic difference, but I think this is also due to the fact enormous amount of rainfall was poured on this topography, probably largely because of orographic effects (cf Chien and Kuoa 2009, Taniguchi et al., 2009). Thus preceding rainfall on less high topography (e.g. in the Philippines) probably received much less rainfall, and the following rainfall over China (or Japan for Talas) was also likely less simply because little or no recharge ove the ocean was possible and a significant fraction of the typhoon moisutre was used up.

A theoretical answer is more difficult to find and would require a proper definition of a storm event. A tentative meteorological definition could consider a mass of moving moisture with a single source of moisture. A typhoon or (afternoon) convective cell could be such an object, that can then travel and pour its accumulated moisture as rainfall (and or snowfall) over an area limited in space an time. However, for the landslide community, only the relatively high intensity part of this rainfall matters, (as the part being below the landslide threshold can be

neglected) and the spatial and temporal limits of a storm event could be further limited (as in the case of Morakot and Talas, where the orographic rainfall effect limit greatly the part of the typhoon relevant for landsliding).

Chien, F.-C. and Kuo, H.-C.: On the extreme rainfall of Typhoon Morakot (2009), J. Geophys. Res., 116(D5), D05104, doi:10.1029/2010JD015092, 2011.

I think summarising this discussion within the text would be a good way forward so that it is clear what your working definition of a storm (and a storm footprint) is and what complexities might be neglected in this.

L30: this gets at a difficult issue, what do you include as a landslide? I think you need a clear definition that can be applied across all inventories and I don't see one at present. Divergence from the definition in different inventories will introduce bias to your results.

>> We tried to avoid mapping (or remove in the already mapped inventories (J11, TW9) deposition and erosion in the fluvial system, broadly defined as the areas with permanent flows, visible in the high resolution image. This meant that debris flow on hillslope would be mapped but not its prolongation within the fluvial system. We considered bank collapses as a disturbance that would be localized, usually not symmetric and not necessary linked to a landslide/debris flow on the hillsllopes. Clearly we may miss some bank collapse, and where to put the limit between a debris flow on a hillslope and its continuation on the fluvial channel is difficult and somewhat subjective. However, if amalgamation is avoided, the width estimate (and thus scar area and volume) will be relatively insensitive to these issues, that mainly affect the total runout and aspect ratio.

This explanation is useful and goes some way to (and even beyond) addressing my comment. I think it would be useful to include in the text something along the lines: "Here we define a landslide as…" it might include something that captures your distinction above that fluvial transport and bank collapse are not included in this definition though the precise boundary between these and a landslide may be fuzzy. It would also be useful to highlight that these are rapid landslides – you don't include landslides that do not remove the covering vegetation I guess so you won't see earth flow activity (and probably wouldn't want to include it anyway since you would then need to talk about velocity changes rather than presence or absence).

L20-21: maximal forcing: this doesn't seem to be consistent with your argument for the importance of complete landslide footprints. You are comparing the forcing at a single location within the footprint to the properties of the entire footprint.

>> Well the maximum forcing is taken as a "storm magnitude", and it is compared to total landsliding and peak landslide density close of this maxima l forcing, so we think the approach is reasonable. We think that the issue of the referee is that in the introduction we push for "comprehensive landslide inventory" and later do not make fully use of it.

This is for 2 reasons: 1/ We have access to extensive rainfall data constraining the spatial pattern of rainfall for only a few cases, and the analysis of the spatial pattern is beyond the scope of this study and left for a future study (Marc et al., in Prep).

2/ Generally a comparison of storm magnitude with total landsliding requires an accurate order of magnitude of the total landsliding. We think that the fact that storm magnitude correlates well with total landslidng suggest some internal correlation between the peak total rainfall, and the mean rainfall and its variability within the storm footprint. Still we acknowledge that such correlation may not hold for all type of rainfall events: In our database, small events are likely brief convective thunderstorm (C99,C15), while large ones are typhoons (M02,TW8,TW9, J11), which are very large singular system that are loaded during their displacement over ocean, and unloaded on landfall, even more importantly when hitting high relief. They fit well in a tentative definition of a storm event based on a single process / source of moisture accumulation and subsequent downpour on a given spatial/temporal zone (cf comments P3 L7).

The cases of B08 and B11 are more complex as they may results from interactions between multiple oceanically sourced moisture and specific meteorologic conditions on land. As mentionned in the main text, other rainfall period such as a monsoon could also rather be characterized as the sum of repeated convection events then transported across India, not allowing to differentiate individual meteorological event.

We will try to add some elements of this discussion into section 2.3

This would be a useful addition to the paper and addresses my comment.

P9 L5: how, and where, did you measure landslide width?

--> The width was initially measured by GIS on a limited number (~50) of randomly selected slides in Colombia and Japan. To make this point more robust we proceded as follow: Text added in the revised manuscript:

"We measured systematically the width of 419 randomly selected landslides across all range of polygon area and aspect ratio, in the following inventories : J11, TW8, B11, and C15. The width was measured on the upper part of the landslide only, the likely scar, and ~4 width measurements made in arcGIS were averaged. When compared to the equivalent width obtained through our runout correction, 72% of the polygons are within 30% of the measured width and 96% within a factor of 2 (Fig Suppl X). We do not observed a trend in bias with area nor aspect ratio, except perhaps for the automatically mapped landslide in B11, where high aspect ratio correlates with underestimated width"

I had misunderstood your area correction method. You don't actually calculate scar area based on measured width and assumed aspect ratio. You calculate scar area based on mapped area and perimeter by estimating width assuming landslides are elliptical then using an assumed scar aspect ratio to translate to area. The width test figure that you include is useful and gives some sense of error in width estimates. Perhaps you could also clarify your method in the manuscript.

P10 L24: to what extent is the landslide distribution area constrained by your study area (i.e. the extent of available images). Taking this to an extreme did Typhoon Morakot trigger landslides in China or the Phillipenes and should these also be included? This again reflects something that I think you need to discuss somewhere, the differences between rain storms and earthquakes as triggers: where are they similar enough to borrow frameworks from one another and where do they differ?

>> Well it is clear that, contrarily to earthquake that have a well delimited source (across the fault), the rainfall forcing is moving together with the storm and can travel other significant areas.

Agreed, and discussed elsewhere in this review-response document.

P11 L24: peculiar distributions: are these distributions peculiar if you are seeking power laws but not if other alternatives are considered? Have you tried a log-normal distribution? Negative curvature of the tail in log-log space sometimes indicates better fits for log-normal distributions?

>> The question is difficult to solve and not so important for our studies: Fitting log-normal distribution by MLE we obtain better agreement for some distributions and worst for others (comparing the Kolmogorov-Smirnov Test statistic and Anderson Darling test statistic obtained for log-normal or IGD fit obtained by MLE). In other work the Inverse Gamma distributions has been found to provide the best fit to 3 large landslide catalogue (Malamud 2004). Further some work on the theoretical emergence of landslide size distribution also predict power-law decaying tail (Stark and Guzzetti 2009), with a tail related to the mechanical properties of the medium, implying the debate may not only be a question of goodness of fit, especially given that some datasets may be affected by artifacts. Although we can mention these facts, solving such a debate is clearly out of the scope of our work. We will of course double check wether or not LogNormal fit parameters (Mu and Sigma) are correlated to rainfall

parameters. Figure for the discussion: Comparison of lognormal distribution (Solid) and Inverse gamma distribution (dashed) for the best fit of landslide scar size distribution. Fit are obtained by MLE.

I agree that this is difficult to solve and not central to this study. I think noting that the distributions are not always a particularly good fit to IGD but in some cases may be better fit by LogNormal would probably be enough. I think theoretical work explaining power law tails is a distraction since that work explicitly set out to explain power law tails.

L29: aspect ratio below 2: why below 2? What are the specifics of the equation? I had understood it to be A=w^2, which would give an aspect ratio of 1.

--> This is now updated based on Domej 2017 (Cf comments above). This is a reduction to an aspect ratio of 1.5.

This issue is resolved.

L11: focussing on scar areas seems sensible but this particular approach seems strange and the choice of modal topographic slope somewhat arbitrary, could you provide a more robust explanation for this choice? Alternatively couldn't you have used your previously defined scar area (w^2) to identify scars as the highest w^2 area of each polygon? This would be consistent with your previous definition and would avoid introducing an arbitrary slope threshold which could bias the results.

--> We note that for C99 there is not much difference between initiation point and the steeper part of the landslides (i.e, after the mode). Additionally we simply do not interpret what happens below the topogaphic mode but can certainly show it. Reducing all landslides polygon to their scar would require some work but is possible and we will try to examine whether or not it creates any difference in the results.

This would be a good addition.

P13 L14: Could you use line thickness to indicate the slope beyond which small numbers of cells in the value range preclude interpretation of the line? It would be useful for the reader to know where that point is for each dataset. Also could you colour the lines in Fig 5 by storm duration? This might make it easier to pick out the behaviour you are identifying in the text and to make a connection between 5A and 5B.

>> Actually we plan to use the notion of prediction interval of the landscape slope distribution to assess if the landslide-affected slope distribution is staistically different from the landscape one. With this method we will indicate which part of the distribution can be robustly interpreted and which one are less robust. We will change the colour-code to reflect the duration, this is a good suggestion.

This sounds good.

P17L9: Total storm rainfall: These results are extremely interesting. They suggest that absolute rainfall properties are good predictors for landslide properties. In the rainfall threshold literature there has been debate over whether absolute rainfall properties are driving failure or whether it is the degree of deviation from normal conditions (e.g. expressed as percentiles). It might be useful if you could reflect on this in relation to your findings. Would a plot of rainfall percentiles for these storms look very similar to the plot of absolutes that we see here?

>> By deviation of from normal conditions, do the referee means the comparison between the sotm rainfall and for example mean annual rainfall or mean seasonnal rainfall ? Or more something like the estimated return period of such a storm?

I was thinking largely of percentile based approaches, which use somewhat more information than a straight deviation from the mean but are generally less complete than a return period.

It may be difficult to estimate one or the other for a number of events but we can try (Mean annual or seasonnal rainfall may be tractable). It is clear that although total rainfall may be a good predictor of the relative amount of landsliding between different storm (as shown in Fig 6 and 7) the control on landsliding must be more compliacted as in the surrounding area similar rainfall occur without triggering landslides (Taiwan, Japan), or in the same season similar total rainfall did not trigger landslide (in Colorado), so either antecedent rainfall or some constraints on intensity will be needed to generalize /strengthen the results we found. In any case we will add such caveat somewhere in the discussion.

Adding something like this would be useful and would address my comment. The other analysis that you suggest would be extremely interesting but may not be possible here and that would be fine by me.

P18L17: we have no clear physical explanation: isn't this something that either extreme rainfall community or the hurricane community have thought about? It would be useful to point readers to key reference from that literature here even if you don't strongly back one particular explanation.

>> We will look at this literature to try to suggest interesting reads. Some work indeed study the correlation between the total rainfall on land of hurricane and tropical storm with their diameter and travel velocity (Jiang, et al., 2008). We will try to relate to such work. Jiang, H., Halverson, J. B., Simpson, J. and Zipser, E. J.: Hurricane "Rainfall Potential" Derived from Satellite Observations Aids Overland Rainfall Prediction, J. Appl. Meteor. Climatol., 47(4), 944–959, doi:10.1175/2007JAMC1619.1, 2008.

This sounds good.

From R2 comments – where the authors seek reviewers' opinions:

The title largely makes sense, although I find the last section stating "first insights from past and new events" confusing since all datasets are within the past twenty years and the youngest event occurred in 2015, and I'm not sure what that phrase adds to the description of the research.

2/ Some edits of the title : We could indeed drop the second statement. An interesting alternative, slightly more descriptive could be "Towards a global database of rainfall-induced landslide inventories: first insights on landscape scale landsliding caused by rainfall event" This option includes somewhat the notion of global magnitude of landsliding (e.g., Fig 6,7) and spatial distribution witin the landscape (e.g., Fig 5). I would welcome comments of Referees and AE on such a title, and if they oppose it, we could simply stop with "Towards a global database of rainfall-induced landslide inventories", although it does not leave a hint than the paper do not only report on collating data but also analyze and interpret them.

I would prefer something that indicates your key findings, which to me are: the breakdown in local slope dependence and that landslide density depends on storm total rainfall. Perhaps:

"Initial insights from a global database of rainfall-induced landslide inventories: the weak influence of slope and strong influence of total storm rainfall"

These are fairly strong claims but I think they reflect your findings. Alternatively I am fine with:

"Towards a global database of rainfall-induced landslide inventories: insights on landscape scale landsliding caused by rainfall events"

I removed first because it is a little ambiguous whether you mean initial insights from these events (which would be fine) or the first ever insights on landscape scale landsliding (which would not).