Response to Reviewer 1 (Odin Marc).
Many thanks Dr Marc for your constructive and useful review. Your two major comments are both useful we respond to each in bold below. Many of your line by line comments are straightforward so we include responses only where there is more to say.

MAJOR COMMENTS
My first comment is about the normalization of several of the hazard metrics : I am convinced that a substantial part of the difference between the hazard curves could be removed by plotting the hazard against a landscape metric : For example for slope, each landscape as likely a modal slope, that may be interpreted as the result of geolechanical difference (for steady state landscape at least). Thus curves may be plotted against S-mode(S) , somewhat normalizing for difference between two landscape. I can understand the author may still want to express their rules in terms of absolute values of slope or other variables, but I suspect this normalization would clarify and strengthen the result and their analysis (as this did in other studies). I make suggestion for the other variables in my inline comments.
This is an excellent suggestion and something that we will certainly explore. As you suggest, we have been keen to express our rules in terms of absolute values to keep them simple and to enable estimates to be made without additional equipment or information. However, your suggested normalization does promise improved explanatory power our only concern is whether reporting this analysis will dilute the focus of the paper.

My second concern is that their maybe some over-interpretation of the data scatter towards the extremity of the hazard curves. And the author do not provide clear metrics or indication of the validity of individual datapoint. This is not an easy task but the work of Rault et al., which I co-authored, recently proposed a method to do exactly that. I would suggest the author to apply these criterium and check. In this work we consider the probabilty p of the whole topography, and the one resulting from the landslides affected area only p_L. To assess whether p_L is significantly different from p we compute the confidence interval Ip associated to the random drawing of n (n the number of landslides) pixels out of the landscape distribution. If p_L belongs to [p-Ip : p+Ip] then we cannot exclude that the difference between p and p_L just comes from random fluctuations and it is likely not significant. Given landslides remain rare in the whole topography, the drawing can be assumed independent, and similar to a Bernoulli sampling. Provided the central limit theorem is respected (i.e. n>30, np>5 and n(1-p)>5) the 90% confidence interval can be estimated as: Ip = p – 1.96 (p(1-p)/n)^0.5 ; p + 1.96 (p(1-p)/n)^0.5.
Some additional details can be found in the supplementary methods of Rault et al., 2018. Basically n is large (n>1000-10,000) so the authors should obtain very narrow Ip until they reach p<0.001 – 0.0001 but I expect these low probability to be reached in the tail of the distribution (Fig 3,4, 6) and the cut off will vary for the different landscape with higher or lower p or n. The authors could compute Ip as well as the convergence criterium and show the points which may be insignificant in shaded / transparent ?
Thankyou for pointing us to this approach. We had struggled to find a way to account for sample sizes in our analysis and this approach looks perfectly suited to the problem!

LINE BY LINE COMMENTS
L155: True they pertain to initiation, but vast majority of studies highlighting their role or quantifying statistical relations between these predictors and landslide use total area and therefore are combining both initiation and runout.
We agree. However, the mechanistic justification for the factors is almost always initiation based. Our point here is that when these variables are used for landslide hazard prediction they are used because to represent controls on landslide initiation. We propose a modification to the text along the following lines: "The potential predictors described above are primarily chosen in hazard models for their perceived link to the probability of coseismic landslide initiation."

L180: I have the impression it should be the minimum skyline angle, not intersecting topography, …. Indeed a maximum reach angle. Cf comment on Fig 1
You could phrase this either as: *'The maximum angle from horizontal to the skyline'* or *'the minimum angle from horizontal that does not intersect the skyline'*. We have chosen the former because it is shorter and because we are concerned that the latter is more open to misinterpretation. The misinterpretation would be that people may not think of cones at increasing angles and so may misunderstand or ignore the second clause.

L458: Shalrun-EQ= Probability of mobilization convolved with connection probability. Average in the above area. So hazard area is basically the number of pixel where debris flow can occur and reach the interest

cell (say Nhaz)... in the contributing area, times pixel area, and divided by the contour length, i.e. the square root of contributing area. Although I am confused because in Fig 2 : Hazard Area seem to be Nhaz times Pixel resolution (or Nhaz.a/sqrt(a)). But then smallest vales should be 0 and 30 (as it does in Fig 2). But in Fig 6 it goes from 0.1 to 1e3… So there seems to be a problem between the 2 definitions. Please check.
Your interpretation of SHALRUN-EQ is correct, the reason for hazard areas less than one pixel is our use of multiple flow path routing, which we had not properly explained. We will explain this more clearly.

L512-516 : Ok, simplicity is important and it is difficult to integrate other effect mentioned here. But what about checking the actual evolution of probability with slope, for both initiation and stop? A reasonable estimate of scar area can be obtained by selecting the highest elevation pixel in your landslide, and selecting as many as needed to reach a scar area with an aspect-ratio of 1.5 (Domej et al., 2017) and a mean width representing of your polygon (see Marc et al., 2018 for how to do that). Doing so you could check if a plateau develop in your probability ratio after 39Æ or so... Interestingly you could reverse the idea and take the lowest N pixel (N ~ Width / 30 for 30m resolution DEM) of your landslides to obtain a probability of stopping.
We have actually done this analysis already and the results are interesting. They suggest that landslides initiate on a fairly narrow range of slopes but stop on a much broader range (consistent with our observations here) modal values are somewhat similar to the optimized values used here. However, these (unreported) results indicate the slopes on which landslides initiate and stop rather than the probability of initiation and stopping given slope so we need to be careful in connecting the two sets of results. A careful examination of this connection is outside the scope of the paper since we are focused here on coming up with simple rules.

L536 : Did you check the curve appearance when using gradients, that is tan(Theta) (with theta the slope in degree) ? Because the tan(theta) does appear mostly exponential over a large range of Theta. Thus a linear function of tan(theta) (the relevant parameter for landslide stability) may appear exponential when plotted against theta.
Yes, the kink in some curves at low slopes suggested that tan(theta) might be a better predictor and it is attractive for its connection to landslide mechanics. However, we found that for most inventories the slope of the tan(theta) relationship was too gentle to provide a good fit to the data at high slopes. Given the additional complexity of the tangent function it is not well suited to a simple rule so we chose not to report it here. However, the misfit between tan(theta) and landslide probability is clearly interesting and merits further examination in future.

L538-542: Northridge and Haiti are shifted compared to other. They both become > average probability around 20Æ , vs 30 for others. This roughly correspond to modal slopes of these areas. It would be interesting to re-plot all curves not against slope, but slope – Sm the modal slopes. This collapsed curves on a similar analysis for rainfall (cf Marc et al 2018) Similarly is there a large variety of drainage area distribution ? Haiti and Northridge are very peculiar again compared to the other cases. Some normalization by the mode of the landscape drainage area may be important.
This is a good suggestion in terms of improving the explanation of the dataset as a whole but is problematic in terms of developing a simple rule because it would require knowledge of the slope distribution for a location. We will certainly try the analysis and look for opportunities to report the findings here but we are concerned about retaining our focus on the simple rules. In the case of upslope contributing area we already seek to explain the variability in terms of hillslope length. The best solution may be to point out modal slopes in the same way then include the normalized analysis for both slope and UCA in supplementary information. If not this may need to wait for a paper more focused on the use of these metrics within a more conventional hazard mapping exercise.

The definition you take for hazard area gives 0 hazard area for the reference in all cases and then a decrease. It does not seem that shift in the horizontal direction would do any good, and the vertical shift seems due to the proportion of zero hazard area in the landscape, so maybe computing a landscape PDF ignoring the zero would be insightful?
We suspect that normalisation may not be particularly informative in this case, we could introduce a normalisation in each rule i.e. normalized initiation and stopping angles but we are then very far from a simple rule.

L694: I am a bit surprise by the term of channel inside this rule. I guess it derives from the fact that the hazard consider upslope contributing areas defined from flow algorithm. But the hazard area at many intermediate locations on hillslopes may be a channel for your analysis but not for the resident and deciders of the area. Because a channel is defined on finer scale than the DEM. You already say that this metrics is anyway difficult to estimate and handle for application, but this terminology would also complexify the problem for deciders or policy makers.
We will think about whether we can come up with a better term, we are talking about areas of convergence so channel seemed to work but perhaps we have to stay general. The key consideration here is that the expression must be precise but also make sense to people on the ground.

L699 : This is fortunate indeed, almost surprising.
Agreed we expected more sensitivity to the parameters here.

L711: Interesting. Do you think this could be somewhat validated by making skyline and hazard graph for landslide above and below a certain threshold (say 5e3 m2 or even better above a certain width...) ?
The idea of examining whether different metrics preferentially predict larger or smaller landslides is interesting and your suggested approach would be a good way to do that. However, this analysis is probably outside the scope of the current paper given its focus on simple rules. It is difficult to see how our answer to this question could be worked into an improved simple rule.

L822-23: I do not understand what you mean by “we expect the length scales over which this occurs to be long (order kilometres) relative to the other factors examined here” Do you mean that main lithological units are usually big (regional scales) and thus significant part of a landscape will have homogeneous lithology, whereas topographic attribute change at the scales of 10s of meter ? Then it is the length scale for the variability of lithology that you want to mention. Anyway please clarify.
Your interpretation is correct but we will try to clarify our point.

On a side comment, normalizing each landscape slope by their modal slope would be somehow a step toward normalizing difference in landscape that can be due to major lithological or geomechanical attributes (Korup 2008).
Agreed but as discussed above including this in a simple rule would be problematic.

L830: And they likely do, given that large landslide (likely to travel further away as you recall in the introduction) are usually reported closer of the fault or at larger shaking values (Khazai and Sitar (2004), for the Chi-Chi earthquake (1999), Massey et al. (2018) for Kaikoura or Valagussa et al 2019 for systematic evaluation of PGA and landslide size distribution. So future exploration of the behavior of your hazard curve split for specific lithology of different area class should be done.
Agreed, both landslide size and lithology are interesting topics for future work but are outside the scope of this paper.

L834 : I would say we can reasonably expect strong differences : given that hazard increase strongly with local slope for EQ (Fig 4) but not for the rainfall induced landslides : as shown by the anaysis similar to your Fig 3 in Marc et al., 2018. Further, the longer runout (due to lower stopping angles) and stronger dependence on contributing areas are additional changes.
We agree that strong differences are possible but we think it is fair to say that the strength of these differences is not yet clear.