

R2:

This work by Emberson et al. adds to the existing rainfall induced landslides datasets and explores different topographic parameters associated with these landslides. These parameters are used to characterizing landslide locations and differentiate scarps from the rest of the landslide body. While I think the overall analysis is well done, much work is needed in the writing to better communicate the methods and results. As it stands, these sections lack an appropriate level of organization and clarity required for publication. With that said, I do believe with some rewriting this paper will be a significant contribution to the field of landslide hazard mapping and should be published. Below I provide a few general comments and then provide some specific edits.

We greatly appreciate the effort and detailed comments from the reviewer. Their feedback has served to improve the clarity of our study and we further acknowledge the supportive comments from the reviewer about the value of our study. We broadly agree with all of the comments from the reviewer and feel that there are no major issues to revising our study to incorporate them. Below, we detail the individual responses to each of the comments.

General Comments:

The Paragraph formatting seemed a bit off. Probably just a pdf artifact, but double check to make sure the paragraphs are clearly delineated as you want them. As it stands, the apparent paragraph splits made it hard to follow in spots.

We will fix this issue when revising the paper; thank you for flagging it.

Supplemental Figure naming is inconsistent.

This will be resolved in revision.

The methods really need a lot of work. It was difficult to follow what you were talking about and that made the results also difficult to follow. I've included a few suggestions, but I believe they largely need to be rewritten to improve the flow and organization. There was a lot of jumping back and forth between ideas in the paragraphs. Another thing that would help the reader is using consistent and clearly defined terminology throughout the methods and results.

The reviewer provided a number of extremely helpful detailed comments below to help guide the revision of this section. We acknowledge that more clarity is needed and we will revise this section extensively to resolve this. See detailed comments below.

More references to your great figures need to be included in the text. There were several spots where I wanted to know which figure to look at to see what you the writers were describing but there was none.

Further figure references will be added.

There are a few instances of using round brackets for citations.

These will be resolved in revision.

Grammar could use a little work. I pointed it out in a few locations, but rereading it with only grammar in mind would help.

We will conduct a full grammatical check when revising our study.

Specific Comments:

Line 8: "4" should be superscript

We will fix this.

Line 30: This second sentence seemingly contradicts the third. Is your point that some landslides not induced by rainfall or seismic activity are more localized while those that are can have a broader footprint? Or that small localized events can occur within a broad area? Please clarify.

We recognise the potential for confusion here and will rephrase the third sentence:

"Although the impacts of individual landslides may be localized, large triggering events such as intense rainfall or seismic activity can cause large numbers of landslides across a wide region, the extent of which often mirrors the extent of the intense rainfall and seismic shaking"

Line 46: I would add Mirus et al., 2020

We will add this reference.

Line 59: Omit "for instance". It muddles the sentence.

We will remove 'for instance'.

Section 2.2: Paragraphs in this section were really distracting. If this is not a pdf error, please rework paragraphs to make the flow better.

We thank the reviewer for highlighting these issues. We will, in revision, work to fix the formatting issues and ensure the text flows appropriately.

Line 133: Is this supposed to be a new paragraph? If so, it's not a great introductory sentence since it only includes half of the new data described in the paragraph.

This was not intended to be a new paragraph, and we will amend this mistake to improve flow.

Line 134: Reference Table 1

We will reference Table 1.

Line 137: Why did you use two different approaches for making the inventories? Was it a question of data availability? Should be addressed.

We will clarify the reason with the following sentence:

"The algorithmic method was used to reduce the overall time spent mapping some of the larger new inventories."

Line 152: 'landslide' should be 'landslides'

We will change to the suggested 'landslides'

Line 152: I don't follow why debris flows were omitted. Is it just so you can estimate debris volume? That can't be done with debris flows? Or is there another reason I missed?

This is an important point; we remove debris flows to focus solely on the process of landsliding, rather than the characteristics of preferential runout paths. It also allows us to provide consistent datasets that can be used in future studies to assess landslide volumes that can then be combined into more general runout assessments. We will add the following final clause to the sentence in question:

"and permits a focus solely on the topographic characteristics of landslide source regions, rather than the characteristics of preferential runout paths."

Line 153: Please specify volume of what.

We will add the clarifying statement: "of mobilised landslide material"

Line 156: Omit comma after 'interval'

We will remove the comma.

Line 160: Should this be a new paragraph?

We will adjust the formatting here to ensure that the paragraphs are more clearly defined.

Line 163: Marc et al., [2018]

We will change this to the suggested edit.

Line 182: Consider showing a figure of this. It's hard to tell from the table.

While we acknowledge the reviewer's point here, we prefer to avoid including any figures describing the rainfall – landslide density relationship since it would potentially confuse the more focused points we are making regarding topographic relationships. The rainfall-landslide density relationship deserves significant time and research focus and our belief is that it would be preferable to do so in a more expansive way, with another research paper focused on this in the future.

Line 191: What characteristics? How does this prevent inconsistencies? I found it hard to see the context of this paragraph. More detail I think may help.

We will rephrase this sentence to the following to avoid confusion:

"We have analyzed the topographic characteristics of landslide locations for the event inventories, using global satellite datasets to ensure consistency across each site."

Table 3: "Analysis datasets. Explanation of each of the variables in found is the accompanying text."

We are not clear what the reviewer means by flagging this text.

Line 196: Direct comparison of what?

We will rephrase this sentence to the following:

“this means we do not have to resample either dataset when conducting a raster-based analysis at this scale”

Line 205: Insert “(TPI)”

We will insert ‘(TPI)’ as suggested.

Line 211: An equation describing this parameter (and TPI) or a more precise definition would be helpful. Why do you call CTI the 'wetness index' later on? I would delete that unless you mention it here.

We will add definitions for both of these parameters. We will continue to include the clarification that CTI is sometimes referred to as wetness index, since that provides some context for readers who may have seen one or other of the terms.

Line 215: You never describe TRI. Please do so.

Thank you for flagging this mistake! We will add the following description:

“We also calculate Topographic Ruggedness Index (TRI), a measure of the local surface roughness. It is defined as the root mean squared difference in elevation between a central pixel and each of its eight neighbouring pixels.”

Line 218: The difference between Figure S1 and S2 is not clear. Please provide more descriptive captions.

We will add the following sentences to the caption for S1 and S2 to provide clarification:

“The difference between figure S1 and S2 is that figure S1 shows the data solely for the pixels containing landslides, whereas S2 shows the data for all pixels in the landscape. This allows comparison of the differences between areas containing landslides and the more general characteristics of the landscape.”

Line 223: How? Are you referring to the relative ratio analysis or the LASSO? Better flow and organization would help readers follow this great work. Consider outlining this portion of the methods in an introductory paragraph. Something needs to be done to help keep the reader oriented in this section.

We acknowledge that the lack of sign posts to the relevant sections may confuse the reader. We will add in references to the relevant sections (bivariate analysis in section 3.1 for the analysis of single parameters, and multivariate analysis in section 3.2 for the LASSO work) as well as clarify that this is the bivariate analysis (see comment for line 285, below).

Line 226: Following up on the previous comment, the remainder of the methods is very difficult to follow. I believe this is largely due to poor organization and unclear language.

We acknowledge that the flow and clarity likely needed improvement. In revision, and thanks to the comments of the reviewer, we intend to make significant changes to resolve these issues. These are detailed in the responses below.

For example, what is meant by topography in this context? I think I figured it out by reading Marc et al., 2018 (which I shouldn't have to do to understand this paper) but it's not clear where the bounds are of the topography you're considering. The whole island, within 10 m of the landslide?

This is a really important point that we did not adequately explain. We will add the following text to clarify:

“The topography values are calculated for all pixels within the area in which landslides were mapped. Since we lack data on the mapped areas for all of the inventories, we assume that the convex hull (minimum bounding polygon) for the landslide polygons represents the mapping area”

Why do you start this paragraph talking about the methods for figure 3, then go into the GLM, then later back to talk about the methods for figure 3 in more detail. Maybe I'm completely missing the mark here. If so, all the more reason to rework this section.

The reviewer raises a good point here. The flow of this section was not working well, so in revision we will move the entire GLM / LASSO discussion to after the bivariate methodology discussion, and add references to where the results are found.

Line 232: Cite figure.

We will add a reference to Figure 3.

Line 235: Please add specific description of your choice of priors.

Thank you so much for this notification. We fully agree with the reviewer comment that the choice of the priors need to be specified in Bayesian analysis. However, in the final version of the manuscript we did not use Bayesian GLM but binomial GLM. This is to say that thanks to your comment we realized that we forgot revising the text. We intended to use Bayesian GLM in the very initial draft of the manuscript but then we switched back to its binomial counterpart. We will revise the text and remove the line mentioning Bayesian GLM, which was obviously used by mistake. We apologize if we have caused confusion during the reading/reviewing process.

The section will read as follows:

“We also apply a feature selection algorithm to identify the significant and irrelevant variables to feed the GLM. For this purpose, we use Least Absolute Shrinkage and Selection Operation (LASSO) technique (Tibshirani, 1996). This method is particularly suggested for landslide susceptibility assessment to reduce the large number of highly correlated predictors without losing parameter interpretability (e.g., Camilo, et al., 2017). GLM fitting with a LASSO implementation is carried out by using the R (Team R, 2013) “glmnet” library, which was made available by Friedman et al. (2009). We apply this method and couple it with the 10-fold cross-validation to remove non-informative covariates and to assess the modeling performance based on the area Under the Receiver Operating Characteristic curve (AUC) calculated for each landslide inventory (Hosmer and Lemeshow, 2000). From each model we built, we store the information related to the regression coefficients. Before fitting the regression model, we apply a mean zero and unit variance normalization to all variables (e.g., Lombardo et al., 2018), which are expressed in different ranges and scales. This normalization allows us to better examine the modelling results in terms of contribution of each variable. In this scheme, larger absolute values of the regression coefficients refer to relatively large contribution of variables.”

Line 251: comma after secondly.

We will correct accordingly.

Line 256: "than" should be "as"

We will adjust as suggested.

Line 273: Median of which dataset? the topography, landslide? In the figure you say "landscape" is that just topography, or a combination of topography and landslide. Are the two mutually exclusive?

We will clarify this with the following edits to the statement:

"we calculate the median slope value for all pixels within the mapped area for each inventory "

Line 285: When you describe the bivariate analysis in the methods I would call it by that to better orient the reader.

We will describe the bivariate analysis appropriate above (see response to comment about line 223).

Line 285: In this paragraph, and really throughout the paper, more references to the figures your discussing need to be included.

We thank the reviewer for pointing this out. We will add references to the figures more effectively throughout this section.

Line 288: What figure(s) are you see this in?

We will clarify that this is seen in Figures 4, 5, and 7.

Line 292: I would explicitly say that this is where the y-axis = x-axis = 1. I would also add lines to the plots showing this.

We thank the reviewer for pointing this out. We will add a sentence to clarify that this is case (see below) and also add the lines to the plots.

"This can be observed in Figure 4-8, where the probability ratio of 1 for almost all inventories occurs at approximately the median value of the parameter for the entire landscape."

Line 297-298: I think this needs to be explained better using the figure as a reference. What are you looking at to pull this out?

We will add a sentence to clarify that this can be observed in Figure 5.

Line 312: Should this be a new paragraph?

Thank you for highlighting this, it should not be a new paragraph. We will correct this formatting issue.

Line 325: I think drawing lines between the points of a given location would help the reader see this better. It would also make the plots generally more interpretable. I suggest doing it to all these plots.

We have tested including a line to connect the points in the figures but we feel that it is not appropriate both for clarity (since the large number of lines become somewhat hard to distinguish) and because it may suggest a more clear relationship particularly at higher values when one does not exist.

Line 338: Please reference the appropriate figure.

We will remove the discussion of the long-wavelength TPI values since the data themselves show no discernible pattern and does not add to the discussion within the paper. This section will be removed.

Line 444: That the caption for Figure S5 says 'difference' is a bit misleading. Maybe substitute 'comparison'?

We will change this to the suggested 'comparison'

Line 491: Explain why the evolution of the regolith matters.

We will clarify in revision that the hillslope regolith is an important control on the susceptibility to landslides.

The new text will read: "The Evolution of the hillslopes' regolith state, which acts as an important control on landslide susceptibility, under climatic forcing is predicted by geomorphological models of hillslope stability coupled with stochastic rainfall forcing (Dietrich et al., 1995, Iida 1999, 2004)."

Line 501: I suggest omitting "Anyway"

We will omit this.

Figure 3: Is landscape slope the same as topography slope?

We will adjust Figure 3 in line with comments from reviewer 1, above, and this will explain that it is the slope of the entire topography. We apologise for the confusion.

This is just an example of how the other figures are made, correct? If so, why do you label the axes differently in the other figures?

Both reviewers highlighted this oversight and we will fix it in revision.

Figure 5: Here and in other figures does "vs Inventory Median" mean the same thing as "/Inventory Median"? If it does, please change it to be consistent. If it doesn't, this needs to be clarified in the text.

As mentioned by reviewer 1, there is a some inconsistency in the figure naming. This will be fixed in revision.

Figure 6: In this figure you don't use the acronym, in others you do. Be consistent.

This will be fixed in revision, thank you for flagging.

Figure 7: Make sure you define what TRI is in the text.

We will add this.

Citation: <https://doi.org/10.5194/nhess-2021-250-RC2>