Below we have responded to each comment made by the reviewer. The reviewer's comments are in bold, and our responses are in roman text. Changes we have made to the original text are in italics.

Review comment on the manuscript "Slope Unit Maker (SUMak): An efficient and parameter-free algorithm for delineating slope units to improve landslide susceptibility modeling" submitted to NHESS by Woodard et al.

GENERAL COMMENTS

Thank you for inviting me for this review. I have read the manuscript with great interest and I appreciate the effort to come up with a parameter-free tool for slope unit (SU) delineation.

In their manuscript, the authors present a newly developed tool for the automatic and parameter-free delineation of SU for landslide susceptibility mapping. In a first step, they compare SU created by their tool to SU created with the commonly used r.slopeunits algorithm by Alvioli et al., and in a second step, they compare the performance of landslide susceptibility models trained using different pixel-based and their SU-based landslide discretization methods for three case study sites to demonstrate the superiority of the SU-based approach as opposed to pixel-based methods.

I strongly believe in the advantages of SU as mapping units. I have used the r.slopeunits tool myself and found that the parametrization can be tricky and is not always transferable to other study areas. Thus, I believe that a parameter-free tool is an important step towards more objective and generalizable approaches in landslide susceptibility modelling.

As it comes to the comparison of SU vs. pixel-based approaches for landslide susceptibility modelling, it should be noted that there are already numerous publications on this topic.

Moreover, I am not totally convinced by the discussion of the SU result presented in Fig. 1 and by the attempt to demonstrate the superiority of the SU-based approach. This is partly because there is some information lacking that would enable the readers to clearly interpret the results. But also, the discussion appears a bit brief and shallow, considering the authors compared so many different approaches in three study areas. I would expect a deeper and more critical discussion not only focusing on performance metrics, but also model set-up and spatial performance for all case studies, also the Puerto Rico one. Please find some specific comments and questions below.

We thank the reviewer for their careful reading of our manuscript and are pleased that they appreciate the importance of creating a parameter-free slope unit delineation tool. We have removed our direct comparison to r.slopeunits to better focus the manuscript and to avoid any unfair comparisons between the two delineation tools. We have added more details to the manuscript to better describe our model setup and the datasets used in our analysis. Please see our responses below to your comments on section 2 for more details. A few additional adjustments to better discuss our method and results are described here.

We have revised the introduction to discuss the importance of an appropriate scaling of slope units. It reads as follows:

The scaling of slope units should not be arbitrarily set to avoid the modifiable areal unit problem (MAUP) (Openshaw and Taylor, 1983; Buzzelli, 2020; Goodchild, 2011). The MAUP occurs when the cartographic representation of data varies significantly by the scale of the mapping unit used to represent the data. MAUP is a challenging issue to overcome; however, determining a scale of the slope units so that they effectively capture the hillslope processes of interest can greatly mitigate the negative effects of the MAUP (Buzzelli, 2020). Alvioli et al. (2020) recognized this challenge, which motivated the development of their custom optimization procedure. Importantly, the optimal scale for capturing hillslope processes is spatially variant. Thus, the ideal scaling of slope units should adjust to the local topography.

We add to this point at the end of section 2.1 where we write the following:

We argue that basing the scaling of slope units used for landslide analysis on established geomorphic laws provides the best justification for their appropriate sizing and odds of mitigating the negative effects of the MAUP. Further details on how the algorithm was implemented in R are in Text S1 *and the online repository (Woodard, 2023).*

We follow up this discussion in the first paragraph of section 4. It reads as follows:

Our slope unit delineation algorithm, SUMak, has significant advantages over previous delineation methods. In contrast to other methods which use an optimization function or user-dictated setting for determining the appropriate scaling and positions of slope units, SUMak uses established geomorphic laws for determining an appropriate scale of the slope units to capture hillslope processes. *This scaling provides a non-arbitrary scaling of the slope units that are optimized to capture hillslope processes and help prevent MAUP*. Lastly, SUMak is computationally efficient compared to *some* other parameter-free algorithms. These advantages, coupled with it being open-source and easy-to-use, make it desirable for an array of geomorphic analyses.

To better describe the details on the model set-up and spatial performance we augmented section 2.2 in several locations. The first paragraph now reads as follows:

Several papers have evaluated the relative effectiveness of slope units over grid mapping units in statistical landslide susceptibility models (Jacobs et al., 2020; Steger et al., 2017; Zêzere et al., 2017; Van Den Eeckhaut et al., 2009; Martinello et al., 2022). However, none of these studies has thoroughly evaluated the effectiveness of slope units for better *visualizing* the *imprecise* susceptibility model outputs caused by inconsistent input data or their advantages in displaying *near real-time or forecasted landslide occurrence* maps. To demonstrate these benefits, we use the Middle Umpqua and Calapooia 10-digit hydrologic unit code (HUC) watersheds (U.S. Geological Survey, 2004) in the State of Oregon (U.S.) and the island of Puerto Rico which have areas of 257 km^2, 743 km^2, and 8,870 km^2, respectively. *Each area's landslide catalog includes* *an assortment of landslide types (slumps, debris flows, rockfalls, deep-seated landslides, and others) which are not differentiated in this study*. The landslide data from the Oregon were collected over decades using a combination of 1-m DEM data and its derivatives, geologic maps, orthophotos, aerial photography, and field reconnaissance and consists of both point and polygon data (Burns and Madin, 2009). *The Oregon landslide catalogs contains no temporal constraints on landslide occurrence. The Umpqua dataset contains 941 points and 3213 polygons, while the Calapooia dataset contains 33 points and 456 polygons*. In this dataset, polygons cover the extent of the landslide affected area while points are placed at the centroid of the landslide affected areas. All data were reviewed for accuracy after their initial mapping. The areas of the individual landslides mapped using polygons are highly variable, spanning 30-4.4×10^6 m^2 and 1500 - $1.88\text{x}10^7$ m^2 in Umpqua and Calapooia, respectively. This data variability can lead to problems when using grid mapping units because the landslide data is standardized to a consistent format for the creation of the landslide susceptibility models. The Puerto Rico landslide dataset consists of *71,431* point locations of the centers of landslide headscarps that occurred during Hurricane Maria on September 20-21, 2017 (Hughes et al., 2019). *Headscarps were manually identified using high-resolution (15-50 cm), post-event imagery and quality checked by three experienced supervisors. Importantly, the output of the landslide models for Puerto Rico are not a susceptibility map, rather a landslide occurrence map. That is, the models output the probability of a landslide occurring during Hurricane Maria. This type of output is similar to the landslide models developed for near real-time or forecasted assessment of event-specific landslides (Nowicki Jessee et al., 2018; Nowicki et al., 2014; Tanyas et al., 2019; Kirschbaum and Stanley, 2018). Our example from Hurricane Maria is intended to show how event-specific model outputs might differ between slope unit and pixel-based assessments.* Thus, the Oregon watersheds and Puerto Rico datasets are used to demonstrate the benefits of slope units when using inconsistent and event-based input data, respectively.

The middle of the second paragraph was augmented as follows:

Creating multiple points within the polygons allows us to capture some of the variability in the large landslides' measured attributes *without eliminating the influence of landslides originally mapped as points. Using all the raster cells within the polygons would essentially oversaturate the model with data from the landslide polygons and omit any influence of the landslides originally mapped as points.* Finally, we sample a point within each polygon at the median elevation value using the 10m DEM. In the case of multiple points per polygon, we select the point with the highest slope. This dataset is used to verify that the chosen statistics in the slope unit-based approach did not bias the results *and to make the standardization more compatible with the Oregon point data.*

The final paragraph of section 2.2 reads, in part, as follows:

Following *common practice (e.g., Molinaro et al., 2005), we use 70% of the data to perform* a 10-fold cross-validation procedure with ten iterations to optimize the models parameters and obtain representative distributions of the ROC-AUC and Brier score metrics, *while reserving 30% of the data as a final test set. Model development and post-processing is conducted within R (R Core Team, 2016).*

We also added a figure to better illustrate the slope unit binary classification of landslide existence (see our response to Line 200 below).

Also, some maps are in my opinion not ideally composed.

We have removed figure 1 from the manuscript to better focus our analysis. We have remade figure 2 (now figure 1) per the reviewer's recommendations below and moved the slope units of the entire watersheds to larger figures in the supplemental information document (Figure S4). We also adjusted the color of the slope units in Figure 3 to improve their visualization.

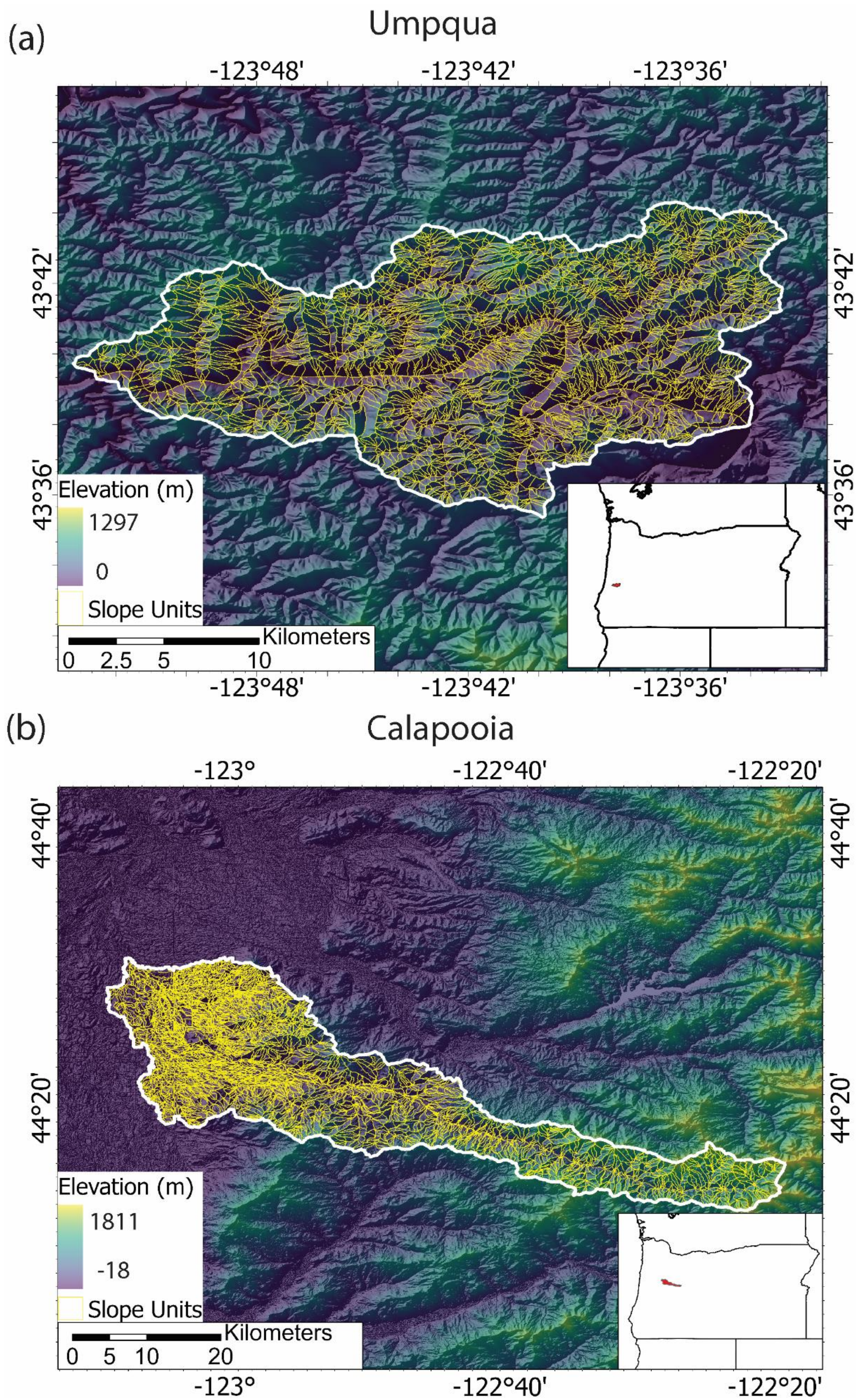

(a)
Umpqua
-123°48'
-123°42'
-123°36'
43°42'
43°36'
Elevation (m)
1297
0
Slope Units
Kilometers
0
2.5
5
10
(b)
Calapooia
-123°
-122°40'
-122°20'
44°40'
44°20'
Elevation (m)
1811
-18
Slope Units
Kilometers
0
5
10
20

Figure S4. SUMak delineated slope units over the (a) Umpqua and (b) Calapooia watersheds, Oregon.

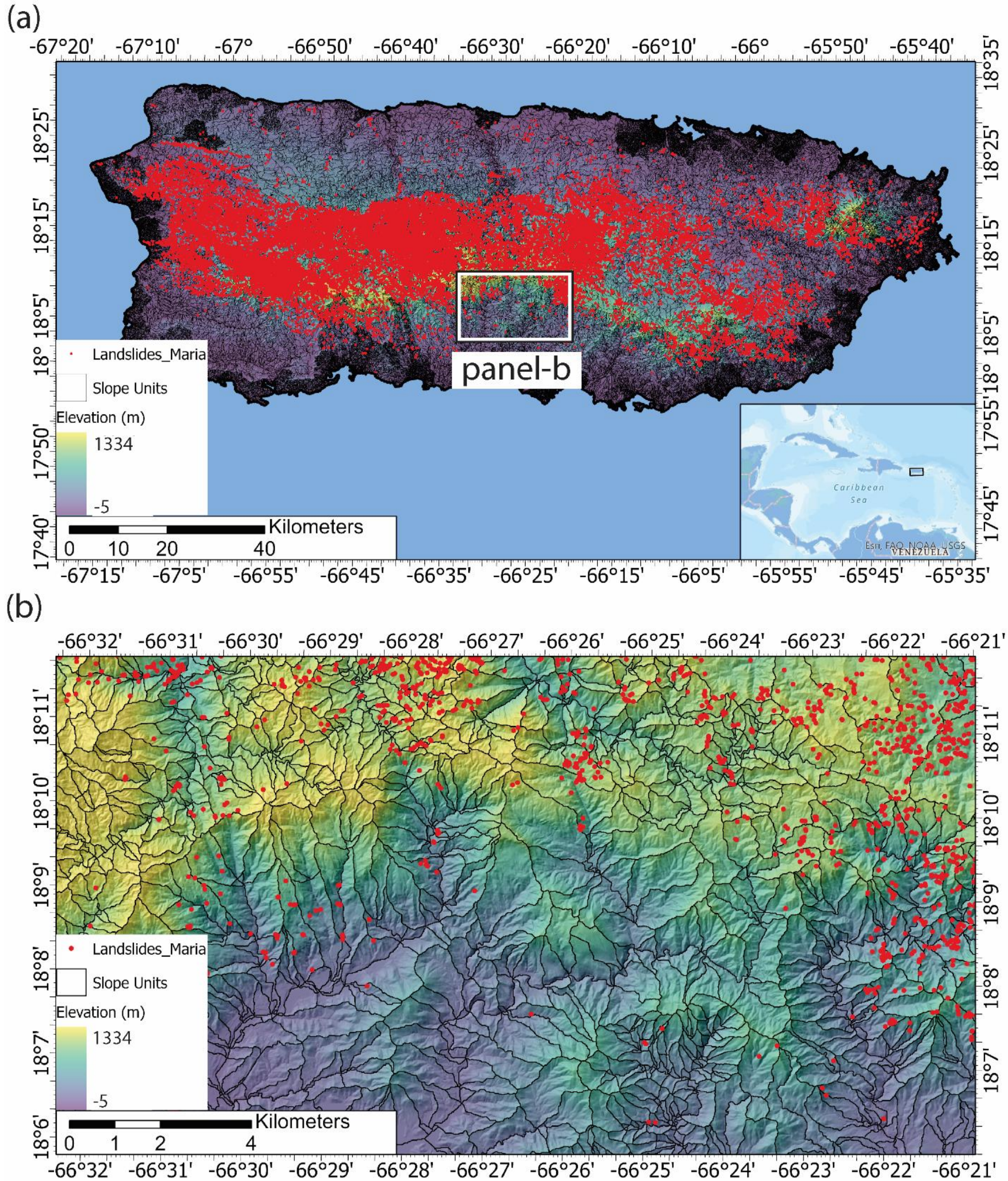


Figure 2: Island of Puerto Rico. (a) Slope unit delineation and mapped landslide points from Hurricane Maria. (b) Zoomed--in portion of the island.

Apart from that, the English language is flawless, the paper is well structured, and the references are complete. Multiple references in the text should be sorted alphabetically though.

We appreciate the positive feedback. We have double checked the sorting of references.

I think that after the additional information on the methodology has been provided and the discussion improved, the paper could be published.

SPECIFIC COMMENTS

Abstract

I would suggest to mention the software the proposed algorithm runs on in the abstract (and also in the main text in section 2.1). This is relevant information for the readers.

We rewrote part of the abstract to specify the software used by the algorithm. It now reads as follows:

We introduce a computationally efficient algorithm for the parameter-free delineation of slope units *that leverages tools from within TauDEM and GRASS, using an R interface.*

In section 2.1 we specify how the SUMak algorithm leverages different software packages.

1 Introduction

Lines 120-127: Could you mention here which methods were used for landslide susceptibility modelling?

We inserted the following sentence into this section.

Landslide models were developed using logistic regression and XGBoost machine learning algorithms.

2 Methods

Lines 130-143: Since the new method is presented as "easy-to-use" I would expect a little more information and instructions on where and how to run it for readers who are not so familiar with GRASS or R for geospatial analyses, or at least a reference to the repository where more detailed instructions can be found.

We amended this paragraph to include a reference to the online repository that details how to run the software within R. The last sentence of this paragraph now reads as follows:

Further details on how the algorithm was implemented in R are in Text S1 *and the online repository (Woodard, 2023).*

Lines 144-147: Which parameters were used for the r.slopeunits algorithm?

Per the recommendation of the other reviewer, we have omitted this direct comparison with r.slopeunits.

Lines 163-169: What types of landslides were included in the invetories?

We included the following sentence to address this point:

Each area's landslide catalog includes an assortment of landslide types (slumps, debris flows, rockfalls, deep-seated landslides, and others) which are not differentiated in this study.

Lines 167-168, lines 179-180: It is a bit unclear to me. What I understand is that the landslide inventories were mixed, with some landslides represented as points, and others as polygons, and the points were mapped at the centroids of the landslides. How many points and polygons, respectively, did each of the landslide inventories contain? How did you deal with landslides that were originally mapped as (centroid?) points for the different sampling strategies that put the points at the scarp or randomly within a landslide body?

We added the number of landslide points and polygons to this paragraph. It now reads, in part, as follows:

The Umpqua dataset contains 941 points and 3213 polygons, while the Calapooia dataset contains 33 points and 456 polygons.

Further down the paragraph, it now reads,

The Puerto Rico landslide dataset consists of *71,431* point locations of the centers of landslide headscarps that occurred during Hurricane Maria on September 20-21, 2017 (Hughes et al., 2019).

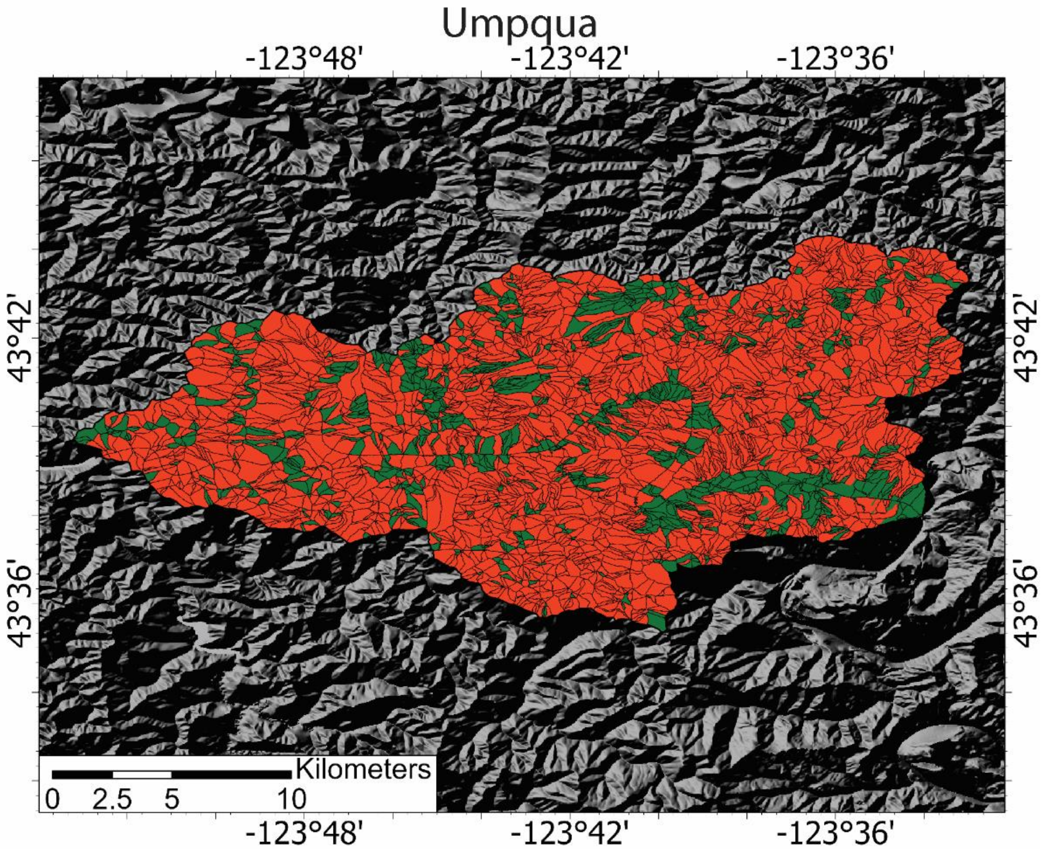
The median point location conversion method allows the landslide data to be more compatible with the centroid mapping method used in the Oregon datasets. However, as only a minority of landslides in the different datasets were mapped as points (29% and 7% for Umpqua and Calapooia, respectively), we did not pursue an additional standardization method using the centroids of all the landslide polygons.

Lines 200-201 and lines 202-209: It would be very helpful for the interpretation of the modelling results if you could provide some statistics. How many samples did each dataset contain? How many SU were delineated in each study area? What was the original positive to negative ratio, especially for the SU?

We have added the number of landslides of each datatype to section 2.2. Please see our response to your previous comment for details. We have included the number of slope units to Table 1. Finally, we added additional figures to the supplemental (Figures S2 and S3) that shows which slope units do or do not contain a landslide, per your recommendation in your final comment below. This illustrates the positive to negative ratio and facilitates interpretation of the model results.

Table 1. SUMak performance metrics.

Location	**Area (km^2)**	**Coastline**	**DEM Resolution (m)**	**Compute Time (minutes)**	**Slope Unit Count**	**Time per area (seconds/km^2)**	**Time per Slope unit (seconds)**
Umpqua	257	No	10	3.11	3841	0.7	0.05
Calapooia	743	No	10	9.97	6990	0.8	0.09
Puerto Rico	8870	Yes	10	383.28	140367	2.6	0.16


Umpqua
-123°48'
-123°42'
-123°36'
43°42'
43°36'
Kilometers
0
2.5
5
10

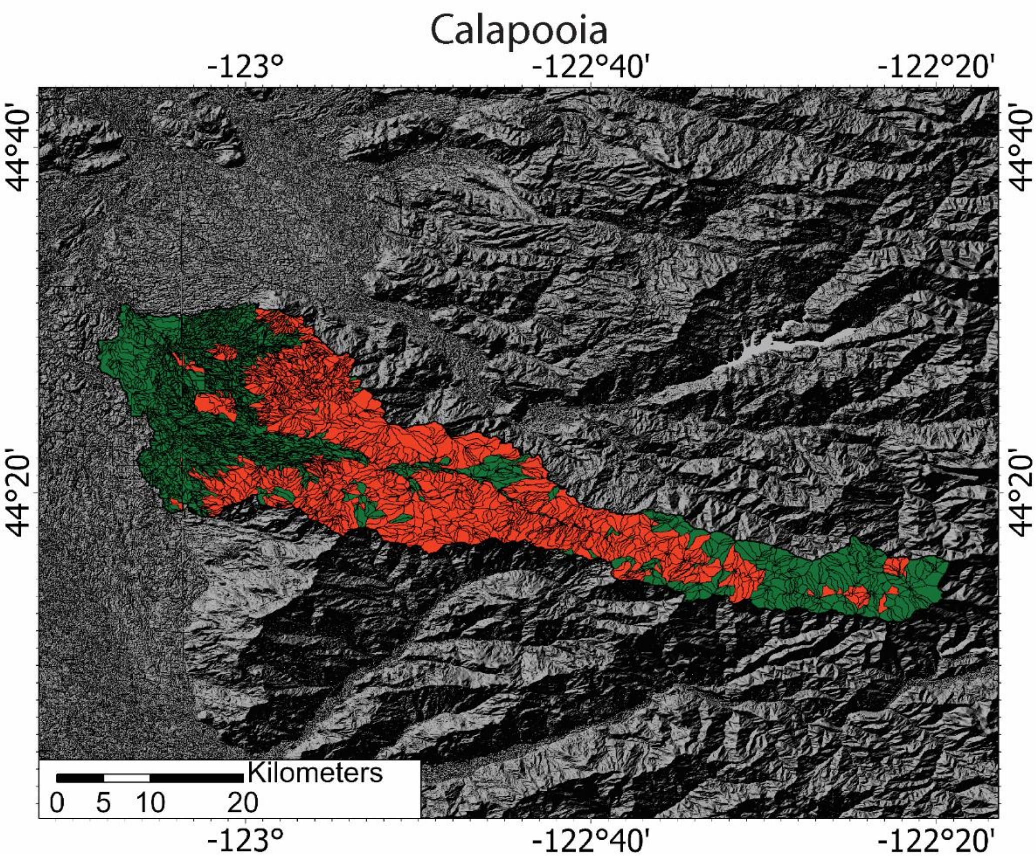

Calapooia
-123°
-122°40'
-122°20'
44°40'
44°20'
Kilometers
0
5
10
20

Figure S2: Maps illustrating the existence (red) or non-existence (green) of a landslide within each slope unit over the Umpqua and Calapooia watersheds, Oregon.

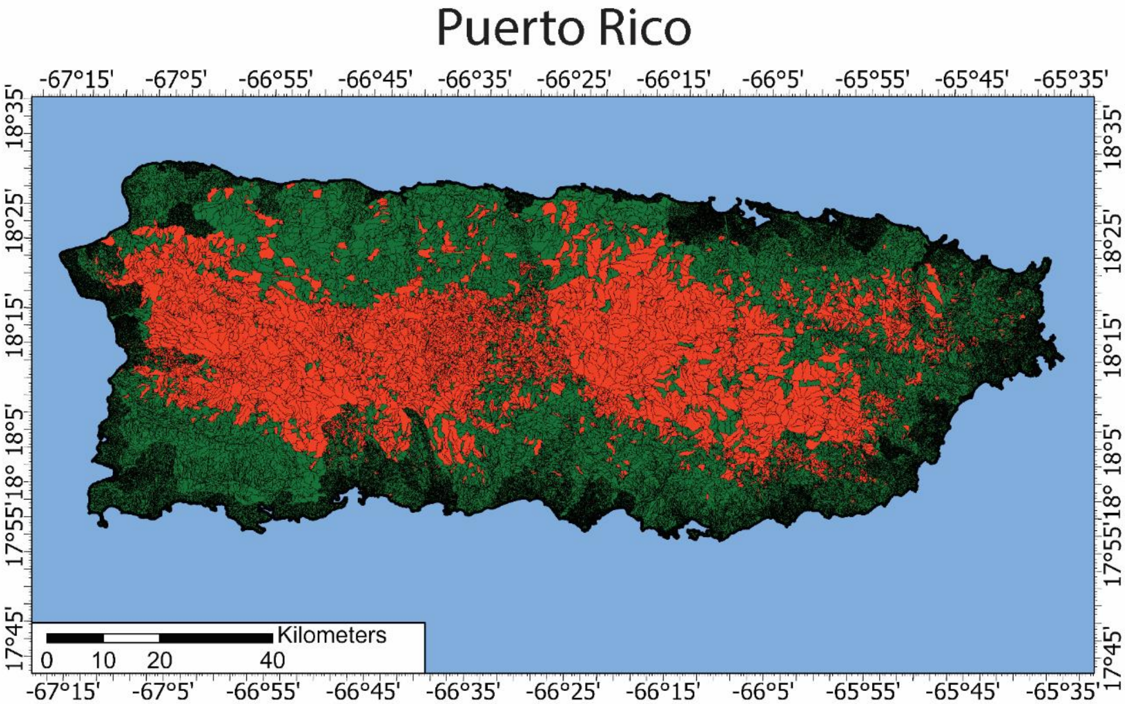


Figure S3: Maps illustrating the existence (red) or non-existence (green) of a landslide within each slope unit over Puerto Rico.

Line 210-229: Which software was used for the susceptibility modelling? Did you conduct any data preparation, such as scaling? Why didn't you use lithology as an input parameter?

We inserted the following at the end of section 2.2:

Model development and post-processing was conducted within R (R Core Team, 2016).

We did not perform any data preparation beyond what we describe in the text.

We did not use lithology as an input parameter because we did not have a consistent lithological map of all the areas that has sufficient resolution to be useful.

Section 2.2: How were the final landslide susceptibility maps generated? Were the trained models applied to all pixels in the study area in the pixel-based approaches? And for the SU-based approach, did you apply the trained model on a pixel-basis or SU-basis?

We applied the trained models to entire study areas to create the final maps. Models were applied to all the pixels or slope units depending on the mapping unit used to train the model. To clarify this point we added the following to the end of the 4th paragraph of section 2.2.

After the models are trained we generated maps by applying the trained models to the entire study areas.

3 Results

Fig. 1: The different scales of the two excerpts are confusing. What are the colors in map c? In case they are SU, its unrecognizable. A plain hillshade or DEM could work better.

We have removed this figure per the recommendation of the other reviewer.

Lines 259-260 and Fig. 1: To me the SUMak SU look much more heterogenous than the r.slopeunits ones. Some SU are larger, and then there are some areas containing many small ones. Could you explain this in more detail? Is the result really so similar to the r.slopeunits one? Here it would also help to know which parameters were used for the latter, see my previous comment.

We have removed this figure and text per the recommendation of the other reviewer.

Fig. 2 a, b, Fig. 3 a: at these scales it is impossible to recognize the SU. I would suggest to enlarge the maps or omit them. Then again, for being able to interpret the performance of the landslide susceptibility maps, it would be helpful to see maps with the distribution of positive and negative SU.

We have removed fig. 2a,b, instead putting a larger figure of the slope units of the entire watersheds in the supplemental (Figure S4). We changed the colors of the slope units in Figure 3. While the slope units are still difficult to see in 3a, it provides a reference map for figure 3b and provides a figure illustrating the distribution of landslide points associated with Hurricane Maria. We provide zoomed in portions of the slope unit maps to make the slope unit more recognizable (2c,d and 3b). We have also created an additional figure that shows a binary categorization of the slope units for comparison to the final landslide maps (Figures S2 and S3). See these new figures above.