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

Reviewer #1 Comments

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. – round 2

GENERAL COMMENTS
Thank you for inviting me to reevaluate this manuscript in the second round of reviews. I checked the replies of the authors and the modified manuscript carefully. I want to thank the authors for providing more explanations and improving the manuscript and figures according to my suggestions. I agree with most of the changes, with a few follow-up questions below.

We thank the reviewer for their initial review and their careful follow up.

My main point of critique is still the discussion, which is only based on comparing statistical performance measures and the granularity of the resulting maps (in my opinion a no-brainer when comparing pixel-based to SU-based models). It has not been improved much in the revised version. There is no connection between the statistics and the geological phenomenon. In which areas did the false positives and false negatives occur and what might be geological explanations for it? Which role does the original landslide distribution that is varying significantly between the three study areas, play? These and many other questions could be addressed, in order to generate new understanding of spatial patterns of landslides. Anyway, as it is, the discussion does serve the purpose of demonstrating the superiority of the SU-based models, so maybe I am asking too much.

The reviewer poses many interesting questions that could provide many insights into the landslide controls at our three study sites. However, as the reviewer notes, these questions are ancillary to the main objectives of our manuscript, which are to introduce and demonstrate the utility of SUMak for delineating slope units. Consequently, we do not address these questions at this time.

Please check for some grammar artifacts that were now added to the text.

We have reread through the manuscript several times to correct the grammatical errors in the revised manuscript. Please see these changes in the tracked-changes document for edits.

SPECIFIC COMMENTS
Original comment:
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?

Follow-up question: I still do not understand what happened to the landslides that were only in the database as centroid points. Were they omitted in the different pixel-based sampling approaches? Or were they omitted in all approaches? I just wanted to make sure all models were based on the same distribution of landslide cases.

At the beginning of the second paragraph of section 2.2 we state the following:

"We evaluate four different methods of standardizing landslide polygons to points for grid-based susceptibility maps in the Oregon watersheds. Each method converts the polygons to points *which* are combined with the landslides originally mapped as points."

There was a typo at the beginning of this paragraph which may have led to some confusion. However, to reiterate, the landslides initially mapped as points were retained for all the different sampling approaches. Landslide polygons converted to points were thereafter consistent with landslides mapped as points, and merged with the original, points-based dataset.

Original comment:
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?

Follow-up question: Could you please also add the percentages of the SU classified as positive and negative, respectively and the number of training samples each model was based on? I understand that they all had a 50:50 distribution, but did all the models also have the same number of training samples?

We added the following to the middle of the third paragraph of section 2.2:

In the Umpqua, Calapooia, and Puerto Rico study sites, 68%, 28%, and 4% of the slope units contained landslides, respectively.

We also added a table to the supplemental materials that provides the other requested statistics.

Table. S1 Number of landslide samples

Location	**Sampling Method**				
	10m	10m_med	10_m_multi	30m	SU
Umpqua	3499	3499	3707	3090	1237
Calapooia	485	485	4824	484	1983
Puerto Ricc	NA	NA	NA	71431	6263

Reference to this table is found at the end of the second paragraph of section 2.2 and reads as follows:

Table S1 shows the number of points for each study site and sampling method, respectively.

Fig 5 and 6: What are the red and blue dots?

We inserted the following into the caption of figure 5:

the red and blue dots show the data outlying the whiskers;

Reviewer #2 Comments

I believe the manuscript is substantially improved with respect to the initial submission, the authors followed both reviewers' suggestions, and removing the detailed comparison with one existing model is probably a good thing. I still believe that the slope unit maps produced using the proposed software and the other one - in one of the removed figures - are not equivalent at all. This is crucial if one is to discuss the processing time of two software programs: if the output is different, or substantially different, the comparison is not justified. If the differences have an impact on landslide susceptibility is a different question, of course. I could further discuss about the issues of optimization, optimization algorithms, and the advantages of parametrization, but this is not the venue for that: I believe the manuscript deserves publication at this stage. One only suggestion, about one of the paragraphs in the supplementary material, also quoted in the authors' reply. The paragraph discussing the role of intermediate watersheds and the possibility of providing custom watersheds belongs to the main text, in my opinion. I am referring to the paragraph starting with "Creating intermediate watersheds allows ..." and ending with "each intermediate watershed in parallel."

We thank the reviewer for their additional comments on our revised manuscript. The comparison to between r.slopeunits and SUMak was removed from the manuscript due to the reviewers' initial suggestions. There is no direct comparison between our method and any other slope unit software in the manuscript.

We moved the description of intermediate watersheds to section 2.1 in the main text. The new paragraph reads as follows:

If the domain of interest has significant variation in topography, TauDEM may choose a threshold that doesn't adequately characterize every area within the domain. Thus, SUMak provides different options for subdividing the domain in preparation for the application of the slope unit optimization procedure described above. We refer to these preliminary subdivisions as intermediate watersheds. Intermediate watersheds must be small enough to limit the variation in topography but large enough to avoid significantly reducing computational efficiency. While experimenting with different watershed dimensions on the topographically diverse regions of Sicily, Puerto Rico, and the Umpqua and Calapooia watersheds, we found an accumulation threshold of ~100 km^2 to adequately strike this balance. This threshold can be adjusted to meet the user's needs, or SUMak has an option to input predetermined intermediate watersheds. After appropriate intermediate watersheds are created, the algorithm runs the rest of the processing steps individually for each intermediate watershed in parallel as detailed in Text S1 and the online repository (Woodard, 2023).