A new approach to mapping landslide hazards: a probabilistic integration of empirical and processed-based models in the North Cascades of Washington, U.S.A.
MS No.: nhess-2019-104

Referee #1 comment response by authors

No.	Comment	Response
1	Why an independent training and testing dataset was not used for this approach. This is typical practice and it would be more robust if there was a separate validation dataset.	We did not perform a formal "validation" study. Instead, the improvements gained by the proposed methods in predicting landslide probability was obtained by a comparative ROC analysis. The focus of the study was to determine if an empirical-based model of landslide hazard could be used to improve an existing physically-based model for shallow landslide probability. A major reason for not separating the data to training and validation was that the performance of the statistical model improves with the size of the observational data used to train the model. The idea is to capture more spatial variability and geologic controls on observed landslides by using all the data we obtained for this region. Validation using ROC rather than training and testing datasets was used to assess this as has been used in similar studies (e.g., Kirschbaum et. al 2012). Future research could carry out validation approaches such as the training and testing approach suggested by the referee. This will be made more clear in the manuscript and suggested for future studies.
2	Provide a brief discussion on the accuracy and comprehensiveness of landslide inventories with respect to representative landslides over this region	The empirical data on landslides was obtained from a series of reports published by the National Park Service (Riedel and Probala, 2005). This comprehensive inventory across the 684,000 acre national park was conducted at 1:24,000 scale and based on 10 m DEMs, a series of large scale stereo air photographs taken since the 1960s, field verification, and from Lidar in a few basins. Where areas were mapped by traditional methods, and Lidar later became available, the original approach captured most (~75%) of the landslides. Dense vegetation cover and a lack of access also limited identification of some existing landslides. Larger, more recent debris avalanches that left large deposits on the valley floor were more

		easily recognized and mapped. Presence of observable landslide features were part of the mapping criteria, rather than mapping topographic imprint of landslides (e.g., Strauch et al., 2018). Ancient landslides that occurred before the last glacial period 16,000 years ago were generally not mapped because their deposits were buried or reworked by subsequent continental glaciation. This text will be incorporated into the revised paper.
3	Discuss relevance of the methodologies for other regions.	See response to comment #5.
4	From the maps in 8e and f it is clear that accurately characterizing the entire landslide using the current methodologies is challenging. Can the authors comment a bit more on how this may be improved with differentiating source area from possibly considering a runout model to develop probabilistic estimates of runout?	Maps in 8e and 8f show the probability of initiation of slope failures, only applicable for slopes steeper than half of the internal friction angle of soils, which is the failure criterion for saturated soils. These maps reflect probability of landslide initiation, which was the focus of the integrated model as described in Section 2.2. The model test using the ROC analysis in Fig. 10 is also conducted only for source areas. Thus, the current integrated methodology is not developed for characterizing the probability of an entire landslide. As the occurrence of runout is conditioned on the failure of source areas, these two models can be developed separately and linked for applications. Differentiating the source area from the transport and deposition portions of landslides, in an inventory could improve the characterization of site features and conditions conducive to failure initiation, transport and deposition. In our paper, only source areas are identified by the integrated model; transport and depositional zones are not addressed as the physical model doesn't apply. We agree with the reviewer that developing a probabilistic runout component would improve the prediction of hazards from the entire landslide disturbances. There could be several different ways of developing a runout model. We could map runout zones of landslides from our inventory and train a rule-based runout model, or develop a purely statistical model based on the occurrence of runout in relation to geologic and topographic attributes. If well-developed and tested, combining a runout model

		with the initiation methodology we proposed in this paper should improve prediction of hazards from entire landslides. Additional text will be included in the manuscript to suggest these model advancements.
5	It would be helpful to have a bit more discussion on the applicability of these methods to other regions, including the size of the region over which this methodology could be applied and other considerations.	The applicability of this approach to characterize shallow landslides hazard is limited by the quality of the site-specific data on soils and vegetation, extent of hydrologic modeling, as well as the comprehensiveness of the landslide inventory. Accurate data for environmental variables such as rock, soils, and vegetation would be as important as comprehensive landslide data as our method relates landslide risk to the environmental variables. The spatial scale of data is another issue we have not studied yet with this method. Larger data sets would profoundly improve predictions, however they could also increase uncertainty of predictions. We argue that this method should be used along with other mapping methods and its performance should be compared against other methods using ROC analysis or other tests. It could potentially be applied over large areas, even continental scales, if these data are available, complete, and validated. The design of the methodology described and demonstrated in this paper allows broad application and is not limited to use at the specific location within Washington, U.S.A. Advancements in surface terrain delineation and in distributed hydrologic modeling specifically contribute to the broad applicability of this approach. We will add additional text to capture these applications.
6	Specific comments on manuscript	Suggested edits and clarifications called out within the draft manuscript are helpful for improving the writing and clarity of the findings. For example, additional explanation on the association with developed landscape is provided in Section 3.1 as well as additional thoughts on the importance of mapping accuracy. Suggested figure improvements were also appreciated and will be updated in the final manuscript.