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 #2

No.	Comment	Response
1	The term "process-based method" should be changed in the more correct "physically-based model" to define deterministic methods of assessment of slope stability. This correction has to be inserted throughout all the paper.	Agreed, changes made throughout paper.
2	Several examples of both data-driven techniques and physically-based models, and related references, could be add in the Introduction section.	We provided literature review in a separate data-driven and physically-based model review paragraphs in the Introduction of the paper. If the referee knows of relevant papers to cite in our paper that we missed and shares them with us, we would consider including them in the Introduction of our paper and discuss their key aspects.
3	I disagree with the choice of the Authors of considering the entire landslides bodies, both triggering and accumulation zones, as predictor variable of the data-driven method. Landslides runout and accumulation zone are related to other predisposing factors than the ones influencing the the landslides triggering. Instead, I know that the approach of using the entire landslide body in a data-driven approach is very common in the literature. Thus, I suggest to add the reasons why the Authors have chosen this approach and to discuss about the potential limits of this choice.	As the referee notes, considering the entire mapped landslide is a common approach in data-driven hazard identification. We too agree with the limitation of this approach. Thus, we developed two other methods that used landslide source areas and a single landslide type to study how the first could improve a physically-based model and to allow comparisons with these previous studies. Often, only the entire landslide or portions of the landslide are mapped as part of an inventory, and many inventories lack information on types of landslides. Thus, we wanted to explore and demonstrate the differences in the site characteristics associated with these various types of datasets. This types analysis can provide insights into the value of more specific inventories, depending on the goals of the hazard identification study. Some studies are content with identifying landslide prone areas regardless of the type of landslide or landslide feature. However, our analysis demonstrates the variability in results, depending on the landslide dataset used, given

		the same site attributes. The limitations with these datasets and resulting hazard maps, relate to the objectives of the study and intended use. For example, using all landslide types may highlight general areas where landslide activity is possible, but it does a poor job at identifying where landslide may initiate. More explanation of these important choices and limitations will be added to the results and conclusion in the manuscript.
4	It is necessary to describe the main features and the main outputs of the Landlab model considered for the implementation of the physically-based approach. In particular, how the rainfall features are inserted and considered by this model?	In this paper we directly use predicted landslide probability from a physically-based shallow landslide model reported in Strauch et al. (2018). The landslide model developed in Landlab has been detailed extensively in Strauch et al. (2018). However, we agree that additional detail on the main features and outputs could be added to the text in Sect. 2.2 Model Integration. Precipitation is considered in the landslide model through its use as input to a macro-scale hydrology model, Variable Infiltration Capacity model. This model produces a spatially distributed recharge field which is used to form subsurface flow in Landlab. A probability distribution of recharge is used to determine soil relative wetness within Monte Carlo simulations of factor-of-safety.
5	A more detailed description of the bedrock geological features and on the main properties of the soil type are required in the presentation of the study area.	The bedrock geology in North Cascades National Park is dominated by gneiss and granite, with lower grade metamorphic rocks schist and phyllite on the western edge of the park, and Mesozoic sedimentary rocks on the eastern flank (Tabor and Haugerud, 1999). Placement of granite at depth along faults led to hydrothermal alteration of some overlying rocks, and the clustering of large landslides. Soils in the park are generally coarse-grained and relatively young due to active slope processes, but soil age, thickness and distribution are highly variable. Soils formed in glacial deposits from the last ice age are also widespread, and many soils are classified based on the amount of volcanic ash they contain. This detail will be added to the study area description.

6	Considering in the same inventory rockfalls/topple and debris flows/avalanches is not really correct. These phenomena are characterized by different kinematic behaviors their predisposing factors can be different. Even if the combined probability model between data-driven and physically-based approaches have been obtained only taking into account for the source areas of debris flows/avalanches, I advise to add an explanation of why you consider different typologies of landslides in the same inventory of your study area.	Please see response to comment #3.
7	For a further validation of the data-driven model, it could be useful calculating a statistical index such as the Area Under ROC Curve or the values of False Positives/True Positives. This would strengthen the reliability of the proposed model.	We included the physically-based model and the integrated model in Fig. 10b. Our intent with the ROC curves was to seeing if the empirical information could improve the physically-based model results by providing some unknown information missing from the infinite slope model. An ROC curve from the data-driven model would show a well performing model by definition because it is derived from the observations used to develop the susceptibility index or probability. The AUC for the empirical model alone will be added to the text with explanation on its comparison to the other two models.
8	It could be useful presenting also the results of the application of the physically-based probabilistic model implemented in the study area and its validation.	This information is provided in an earlier study by Strauch et al. (2018) and is not repeated here for the sake of brevity.
9	Why did the Authors choose those ranges of probability to consider a slope as relatively stable (< 0.1) or highly unstable (> 0.9). Several Authors identified other ranges for the classification of the probability distribution. Please, discuss about this aspect.	The terms relative stable and highly unstable were terms chosen by the authors to identify where the cumulative distribution curve generally shifts direction. In between these probabilities, a small portion of the landscape is modeled to have a widely range of failure potential. We removed the labels from the figures and instead, added the corresponding return periods of 10 years and 1.1 years to provide a sense of the hazard distribution, similar to the plotting used in Strauch et al.

		(2018). Additionally, we modified the text to further clarify interpretation of the figure.
10	It is necessary adding a section where the Authors will discuss about the main advantages and the limitations of their proposed approach, in particular compared with the typical methodologies used for the assessment of landslides susceptibility or hazard.	Many articles have described the advantages and disadvantages to data-driven and physically-based models (e.g., Ercanoglu and Sonmez, 2019; Reichenback, et al., 2018; Hungr, 2018; Claque and Stead, 2012; Aleotti and Chowdhury, 1999). Our approach attempts to benefit from the strengths of both traditional modeling methods. While empirical models validate well with given mapped landslides, they lack a mechanistic explanation for the susceptibility level. Parsimonious physical models predict failure based on forces within the soil, but they may miss properties demonstrated by failure or lack of failure on the landscape. Our approach is limited to areas where landslides have been mapped. Additional text and references will be added to explain the main advantages and limitations of our integrated approach.