L29 When I have read the abstract, I thought that treating shallow landslides as erosional processes could have been a bit of a stretch because the two mechanisms are not the same. This sentence is particularly relevant and I would stress the same concept once more in the conclusion.

Thank you, we will separate the two concepts clearly, in the abstract and conclusions.

L47 This sentence is not entirely correct, as you are also missing:

Amato, G., Eisank, C., Castro-Camilo, D., & Lombardo, L. (2019). Accounting for covariate distributions in slope-unit-based landslide susceptibility models. A case study in the alpine environment. Engineering geology, 260, 105237. (This one is particularly relevant because it was done in the Alps and proves LASSO to be performing much better than the common Stepwise procedure found in landslide studies).

Lombardo, L., & Tanyas, H. (2021). From scenario-based seismic hazard to scenario-based landslide hazard: fast-forwarding to the future via statistical simulations. Stochastic Environmental Research and Risk Assessment, 1-14.
(This is of minor relevance to the paper although they still use LASSO).

Tanyaş, H., Kirschbaum, D., & Lombardo, L. (2021). Capturing the footprints of ground motion in the spatial distribution of rainfall-induced landslides. Bulletin of Engineering Geology and the Environment, 80(6), 4323-4345.
(This is of minor relevance to the paper although they still use LASSO).

Thank you for pointing this out. We will consider and include the missing citations.

L82 I know this is not the aim of the paper, but I think it would be very interesting to report the results of the U-Net.
Even just by mentioning the accuracy of this deep learning tool would be a nice addition.

We will add the accuracy scores of U-net for comparison. They are presented in Samarin et al. 2020, but you are right; it will be helpful to present them here as well.

L117 This operation is absolutely correct and yet almost nobody implements it within the landslide community. It ensures that any residual spatial dependence in the data is dissected and therefore removed.
In the landslide literature the only researchers that do it systematically gravitate around Alexander Brenning. I believe, this is very well explained in:

Goetz, J. N., Brenning, A., Petschko, H., & Leopold, P. (2015). Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers & geosciences, 81, 1-11.

Steger, S., Brenning, A., Bell, R., Petschko, H., & Glade, T. (2016). Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps. Geomorphology, 262, 8-23.

and I think they deserve to be mentioned here.

Thank you for pointing this out. We will include these authors at this point.

L132 Interesting.. I did not know about it.

L148 So, you opted for a balanced presence-absence data. This has some implications in the resulting probability values and their overall distributions which have been nicely explained in:

H. Petschko, A. Brenning, R. Bell, J. Goetz, T. Glade Assessing the quality of landslide susceptibility maps–case study Lower Austria. Nat. Hazards Earth Syst. Sci., 14 (1) (2014), pp. 95-118.

Frattini, P., Crosta, G., & Carrara, A. (2010). Techniques for evaluating the performance of landslide susceptibility models. Engineering geology, 111(1-4), 62-72.

I suggest to mention them in the text.

Thank you for these references. We will include these publications in our text.

Fig 4 & 5 A very elegant way to represent the LASSO results!

Thank you.

Fig 6 & 7 As you may have understood by now, I really liked your work. But, if I may suggest something, I would remove Figures 6 and 7 and substitute them with a single figure where the boxplots plotted for Chrauchtal and Val Piora share the same plotting space. You could plot one set in one color, say red and the other in blue and help the reader comparing the effect of each covariate in your model as the test site changes.

As the two figures are right now, it is almost impossible to make a clear comparison and yet it would be an interesting step from a geomorphological perspective.

I do not know how feasible it is but you could even think of plotting the regression coefficients estimated for all test sites, pretty much as you have done for the ROC curves below.

I know it will be messy so, it is just a consideration that I am doing with you. Maybe you could split the figure into subpanels, in one (row?) you can plot the morphometric covariates, in another you can plot the climatic ones, then lithology and then aspect. This could ensure that you have some horizontal space to add the boxplots from the other sites.

Maybe it will still be too messy and difficult to read, so I will leave it up to you.

Thank you for this idea. We will try it out and see if this improves these figures and increases readability.

L264 That's also very interesting, I would have considered running a all-but-Slope model to see if in a multivariate scheme the slope retains its explanatory power or if it is captured by other morphometric properties. In the end, they all derive from the DEM and part of the spatial information conveyed by the slope may be captured by other covariates.

This is an interesting idea, and we will look into this possibility! Thanks for the suggestion.