Response to Reviewer #1

Manuscript Title: *Landslide susceptibility mapping with explainable AI techniques: Evidence from Bavaria, Germany*

Dear Editor and Reviewer,

We sincerely thank the reviewer for the very positive and constructive assessment of our manuscript. We particularly appreciate the recognition of the study's large data base and clear result presentation, as well as the helpful suggestions for further clarification.

We hope that the revisions meet the expectations.

Sincerely,

Veronika Buchauer, Marta Sapena, Christian Geiß, Patrick Aravena Pelizari, and Hannes Taubenböck

RC1: 'Comment on egusphere-2026-1647', Anonymous Referee #1

Comment 1: *Dear authors and editor,*
It has been some time since I have reviewed an article requiring so few corrections. The topic is interesting, the results are based on a very large dataset, the ROC AUC value is very high, and the results are presented clearly. I enjoyed reading it. Therefore, there are practically no critical suggestions or corrections, but please consider the following:
Provide a geological map, both a general one and, optionally, a more detailed one for a smaller region.

Response: Thank you for the suggestion. We will add the used geological map (dIGK25) in Figure 2 (b) along with the location of the landslide events in Figure 2 (a). Since the original map has 33 geological types, for simplification, we will show the 6 broader aggregated classes that were used in this study. We believe that for the readers this version is more useful, since these classes are the ones used by the model. The original geological map with all 33 classes can be viewed in the Umweltatlas.

Comment 2: *slopeunits is a GRASS GIS tool, but this is not mentioned at all.*

Response: Thank you for pointing this out. Yes, the original `r.slopeunits` tool is implemented in GRASS GIS and accessed and operated through this software. As stated in lines 195-199, we adapt the workflow and integrate `r.slopeunits` within R for parallel computing. We therefore run the `r.slopeunits` tool through an R function, which in the background calls GRASS GIS using the `rgrass` package. To clarify this, we will rephrase the sentences in lines 195-199.

Comment 3: *In Figure 3, when the DEM is zoomed in, visible squares appear across the region. Is this an artefact of a particular tool, a result of resampling to 5 × 5 m resolution, or is there another reason? Does this affect the results?*

Response: The hillshade map displays a rendering effect when zoomed in, but this does not affect the DEM data itself or its derived model input features, it only affects the display.

Comment 4: *In section 2.3.1 (Segmentation of slope units), what procedure did you follow for the karst terrains? Delineation of slope units does not really work there, but there is no information on the geological map to allow further comment on this.*

Response: We applied a uniform slope unit segmentation approach across the entire study area, without differentiating between geological settings. The delineation is based on topographic information, where flow accumulation is only used as a geometric proxy for identifying ridges and valley lines to separate the slope units.
We acknowledge that in karst terrains, surface flow accumulation does not necessarily reflect actual hydrological processes due to subsurface drainage. However, in our approach, flow accumulation is not intended

to represent real runoff pathways, but rather to support the identification of terrain features (i.e., ridges and valleys). We will add this point in the reviewed manuscript, as something that could be improved especially in local studies in karst regions.

Comment 5: *Line 238: Did you check the influence of several landslides within the same slope unit? Only the most recent was considered; however, it could also be the largest one or based on another criterion. What is the impact of this assumption?*

Response: The landslide inventory does not contain information on landslide magnitude (e.g., area or volume), which means that filtering based on the "largest" landslide is not possible. However, within the context of our modeling approach, the specific landslide selected within a slope unit is of limited importance, since the landslide metadata itself is not used during model training. Independent of which event is selected (the largest or the most recent), the resulting classification of slope units remains the same: slope units containing at least one landslide are labeled as positive (1), while slope units without landslides are labeled as negative (0). The selected event therefore mainly serves as a representative reference for the slope unit. We use the most recent event for this purpose, as some entries in the inventory date back to the 19th or early 20th century, while more recent events are assumed to better reflect current terrain conditions.
Our study focuses on landslide susceptibility mapping based on static terrain variables and aims to estimate whether terrain conditions are generally prone to landslides. For this reason, each slope unit is assigned a single binary ground-truth label indicating the presence or absence of landslides.
The alternative approach suggested by the reviewer to not filter landslides and instead use the number of events per slope unit, would fundamentally change the modeling framework. Such an approach would focus on modeling landslide frequency rather than susceptibility, for example through count-based regression or multiclass classification to predict the number of landslide events per slope unit. This would shift the objective from estimating whether a slope unit is susceptible to landslides towards predicting how often such events occur.

Comment 6: *Line 320: What is the basis for choosing these numbers of neurons?*

Response: The number of neurons per layer was finetuned together with other hyperparameters using Latin Hypercube Sampling (see lines 314–319). This approach ensures that the full range of each parameter is sampled evenly, allowing for a more systematic exploration of the hyperparameter space than simple random sampling.

Comment 7: *Line 408: Geological units also do not have a major role (geochemical rock type is more important). What is the reason for this?*

Response: As explained in lines 441–446, the SHAP analysis does not provide evidence of causal relationships, it can only explain what is learned by the model based on the sample data, and point to the features that correlate the most with our target variable. Therefore, it would be misleading to say that geological units play a minor role in landslide processes in the study area. A more appropriate interpretation is that, from the model's point of view, geological units are a less useful discriminative variable given the other input features, so that geology can not distinguish very well between high and low landslide susceptibility in this specific dataset. The geology can, of course, be an important factor in landslide processes, where a particular geological unit might increase the likelihood of a landslide. However, our current data-driven approach may not always identify these causal relationships, as it focuses more on correlations between variables. While we acknowledge that incorporating domain knowledge and causal mechanisms into the modeling would make the models more physically consistent, this was beyond the scope of the study.

Comment 8: *Figure 8: Move the right part of the figure below the left one (not side by side) and enlarge both (a) and (b), as the fonts in the figure are too small.*

Response: Thank you for the suggestion, we will rearrange the two parts of the figure below each other to make the fonts larger and increase readability.