Response letter for
“Projecting changes in rainfall-induced landslide susceptibility across inhabited areas of China under climate change”
(egusphere-2025-3834)

Response to Editor Comments

Editor: I have carefully read the revised manuscript and the authors' responses to the previous round of review, as well as the comments raised by the two additional reviewers. Despite the revisions made, I find that the study still presents a major methodological concern that must be addressed before the manuscript can be considered for publication, in addition to the minor points raised by the reviewers.

Thank you very much for your thoughtful and constructive comments on our manuscript. We have carefully considered all of your suggestions and revised the manuscript accordingly. We are truly grateful for the time and expertise you devoted to reviewing our work. Your valuable feedback has greatly helped us improve the manuscript.

Sincerely,
Jinqi Wang, on behalf of all co-authors.

1. Editor: My main concern relates to the statistical representativeness of the landslide inventory used as input for the susceptibility model. Estimating landslide susceptibility requires that the inventory be statistically representative of the actual spatial distribution of landslides across the study area. This does not imply that the inventory must be complete, but it does require that the documented failures adequately represent the distribution of landslides across the different terrain units or zones within the area. This condition is necessary for the model to be trained in a way that allows it to forecast landslide occurrence also in areas where no observations are available.

Examining the spatial distribution of mapped landslides, it is evident that large portions of the territory lack any records. A brief inspection in Google Earth is sufficient to verify that several of these areas — such as Tibet — are in fact densely affected by landslides, yet none are mapped there. The same observation applies to other areas with no mapped landslides, excluding the plains. The most plausible explanation is that mapping efforts were concentrated in inhabited areas, where landslide events attract attention and documentation, while uninhabited regions were systematically overlooked. As a result, the inventory reflects a mapping bias rather than the true landslide distribution.

Given this, I see two viable paths forward. The first is to restrict the analysis to inhabited areas only, where the mapping approach can be considered reasonably coherent and homogeneous, and where the assumption of spatial representativeness is more defensible. The second is to conduct targeted new mapping to document landslide occurrences in all areas currently lacking data, so that the inventory achieves adequate spatial coverage.

This issue also has direct implications for the validation procedure. The authors have sampled "no-landslide" points from areas with no mapped landslides, which is only a valid approach if the inventory can be assumed to be spatially complete — that is, if the absence of mapped landslides genuinely reflects the absence of landslides on the ground. This assumption does not hold here. As a consequence, the validation dataset is compromised, and the reported model performance metrics cannot be considered reliable. In short, both the susceptibility model and its performance estimators lack a sound empirical basis given the data they are built upon.

Response:

We sincerely thank the editor for this important and constructive comment. We have therefore reframed the study as an assessment of rainfall-induced landslide susceptibility across inhabited areas of China under climate change, rather than a national-scale geomorphological landslide susceptibility assessment for all terrain units. We fully agree that the statistical representativeness of the landslide inventory is fundamental for landslide susceptibility modelling, especially when the model is expected to be extrapolated to areas without documented landslides. We also agree that the absence of mapped landslides in some remote mountainous areas cannot be interpreted as the true absence of landslides on the ground.

Following the editor's suggestion, we have chosen the first proposed path. Accordingly, the title, abstract, methodological framework, results, figures, and discussion have all been revised to clarify that the study no longer aims to predict landslide susceptibility across the entire national territory, but focuses on the inhabited-area modelling domain where the inventory is more coherent and spatially defensible.

The landslide inventory used in this study is an operational post-event geohazard record compiled by the Ministry of Natural Resources of China from 2008 to 2023. These records were obtained after landslide-related geohazard occurrences through field investigation, verification, and reporting by professional geological survey teams organized by local governments, supported where applicable by remote-sensing interpretation and UAV surveys. Therefore, the inventory is not a complete geomorphological landslide inventory for all terrain units in China. Instead, it mainly documents landslide-related events that affected, or had the potential to affect, people, infrastructure, settlements, and assets. For this reason, its spatial representativeness is more defensible in areas with human presence than in remote uninhabited regions.

To address this issue, we defined an inhabited-area modelling domain using WorldPop population data. Specifically, human presence was defined using the WorldPop R2025A population dataset. The population data were converted into a binary inhabited-area mask, in which cells with human presence were assigned a value of 1 and cells without human presence were assigned a value of 0. This mask was then used to constrain the modelling, validation, prediction, and non-landslide sample generation domains. In addition, a 3 km buffer around each recorded landslide-related point was used as an exclusion zone to avoid selecting pseudo-absence samples too close to documented occurrences, where the absence of mapped landslides would be uncertain.

The test samples are drawn from the same inhabited-area modelling domain, rather than from areas without mapped landslides across the whole country. Therefore, the reported performance metrics are interpreted as an internal evaluation of model performance within the inhabited-area domain, rather than as evidence that the model can reliably predict landslide susceptibility in remote uninhabited areas where the inventory is incomplete or spatially biased.

Overall, the revised study no longer assumes that the absence of mapped landslides in remote areas represents true landslide absence. Instead, it explicitly acknowledges the spatial bias of the operational inventory and limits the modelling framework to inhabited areas, where the mapping and reporting process is more consistent with the purpose and coverage of the source data. We believe that this revision directly addresses the editor's concern regarding the statistical representativeness of the landslide inventory and the reliability of the model validation.

2. As a minor but important point: throughout the manuscript, the word "disaster" is used as a synonym for "landslide." These are not interchangeable terms, and this usage is scientifically incorrect. A landslide is a geomorphological process; a disaster refers to the societal consequences of a hazardous event. Please revise all instances accordingly, ensuring consistent use of precise scientific terminology throughout.

Response:

We sincerely thank the editor for this important clarification. We fully agree that “landslide” and “disaster” are scientifically distinct terms and should not be used interchangeably. A landslide refers to a geomorphological process, whereas a disaster refers to the societal consequences of a hazardous event. Accordingly, we carefully revised the terminology throughout the manuscript. All inappropriate uses of “disaster” as a synonym for “landslide” have been replaced with more precise expressions, such as “landslide”, “landslide-related records”, “landslide-related geohazards”, “documented landslide occurrences”, or “slope failures”, according to context. The term “disaster” is now used only when referring to societal impacts, such as casualties, losses, disaster prevention, or risk management.

We believe these revisions have improved the scientific accuracy and consistency of the terminology throughout the manuscript.

Response letter for
“Projecting changes in rainfall-induced landslide susceptibility across inhabited areas of China under climate change”
(2025-3834)

Response to Reviewer4 Comments

Reviewer #4: The article brings interesting novelty into the topic of this Special Issue and the manuscript quality increased over previous revisions. I have a few requests:
Thank you for your thoughtful and constructive comments on our manuscript. We have carefully considered your feedback and made the necessary revisions while ensuring that the overall content and structure of the manuscript remain intact. It has been an honour to benefit from your expertise throughout this process, and we sincerely appreciate your valuable insights, which have significantly contributed to improving the quality of our work.

Sincerely,
Jinqi Wang, on behalf of all co-authors.

1. Elaborate mode sampling (row 104).
Response:

Thank you for this helpful comment. We have clarified the meaning and implementation of mode sampling in Section 2.2. Specifically, for categorical variables such as lithology and land cover, mode sampling was used to preserve their discrete class information during resampling. Taking land cover as an example, the original land-cover data had a spatial resolution of 30 m. For each 1 km grid cell, all 30 m land-cover pixels within the cell were counted by category, and the category with the highest frequency was assigned as the representative land-cover type of that 1 km cell. This majority-rule approach avoids generating artificial class values and ensures that the resampled categorical variables retain their physical meaning.

2. in section 3.2., you state non landslide points are twice the landslide points. Also, that the landslide points were buffered to a 3km radius. From a perspective that the same pixel (not point, but pixel) amount “should” be equal for both landslide and non-landslide samples, how do you explain your choice of sampling? E.g. cca 100k landslide points when buffered turn into maybe 30 times bigger number of pixels? and you sampled 100k × 2 = 200k non-landslide pixels. Please elaborate and improve the text accordingly.

Response:

Thank you for pointing out this important issue. We agree that the previous description could have caused ambiguity regarding the role of the 3 km buffer. In our sampling design, the buffer was not used to convert landslide points into landslide pixels. Each landslide record was treated as one positive point-based sample for model training. The 3 km buffer was used only as an exclusion zone to prevent non-landslide samples from being selected too close to known landslide locations, where the absence of landslides may be uncertain. Therefore, the 1:2 ratio refers to the number of point-based training samples, rather than to the total number of pixels within landslide buffers. We have revised Section 3.2 to clarify this point. We also added references to previous studies showing that the selection of non-landslide samples can affect the performance and reliability of landslide susceptibility models.

3. how are landslide susceptibility classes defined? e.g. Figure 5, should be mentioned

Response:

Thank you for this helpful suggestion. We have clarified how landslide susceptibility classes were defined in Section 3.7. The classification was based on the mean susceptibility map of the historical baseline period. Specifically, annual landslide susceptibility maps from 1950 to 2014 were first averaged pixel by pixel to generate the baseline mean susceptibility map. The Jenks natural breaks method was then applied to this baseline map to classify landslide susceptibility into five levels: very low, low, median, high, and very high. To ensure comparability across different future periods and climate scenarios, the classification thresholds derived from the baseline period were consistently applied to all subsequent susceptibility assessments.

4. Considering improving uncertainty for the susceptibility.

Response:

Thank you for this valuable suggestion. We agree that uncertainty should be more explicitly discussed in future landslide susceptibility projections. In the revised manuscript, we clarified that uncertainty in future projections was partly quantified using the spread among selected GCMs. Specifically, we used a multimodel ensemble and reported the 95% confidence interval of the ensemble mean. In addition, we expanded the Discussion to acknowledge other sources of uncertainty, including the landslide inventory, sampling strategy, RF model training process, and susceptibility classification thresholds. We also clarified that the projected susceptibility changes should be interpreted as scenario-based tendencies rather than deterministic predictions.

5. I suggest adding more references to previous work done considering landslide susceptibility and sampling strategies as key words.

Response:

Thank you for this helpful suggestion. We have added additional references related to landslide susceptibility modelling and sampling strategies. In Section 3.2, Zhu et al. (2019) and Guo et al. (2024) were added to support the discussion that the selection of pseudo-absence, or non-landslide, samples is a key methodological issue in landslide susceptibility modelling and may affect model performance and prediction reliability. Accordingly, we clarified our sampling strategy by restricting non-landslide samples to the inhabited-area modelling domain and excluding areas within 3 km of documented landslide-related records.

6. Can you elaborate spatial validation, e.g. K-fold, from the perspective of your work and the topic of this special issue? Are there future plans including this perspective? Can or should it be implemented to your research? Was it considered at some point and, if yes, why was it rejected?

Response:

Thank you for this constructive comment. We agree that spatial validation is important for evaluating the spatial transferability and robustness of landslide susceptibility models, especially when spatial autocorrelation may exist among samples. In the present study, model performance was evaluated using a random training-testing split, which provides a general assessment of predictive performance. The main objective of this study, however, was to assess large-scale and long-term changes in landslide susceptibility within the inhabited-area modelling domain under future climate scenarios. Therefore, a dedicated spatial transferability design, such as spatial-block or region-based K-fold validation, was not implemented in the current analysis. We have now clarified this issue in the Discussion. We state that spatially structured validation methods, such as spatial block or region-based cross-validation, could be further explored in future work to provide additional insights into model transferability and the spatial robustness of landslide susceptibility projections. This revision also better connects the manuscript with the topic of the Special Issue, particularly regarding the influence of inventory quality and spatial sampling on susceptibility-map reliability.

We sincerely appreciate your valuable comments once again. If there are any further issues with the revised manuscript or if we have misunderstood any of your points, we would be grateful for your further guidance.

Best wishes!