Manuscript Modifications: Point-by-point Responses

Dear Reviewers and Editors,

Thank you very much for allowing us to revise our manuscript further. We would like to express our appreciation to you for your valuable comments and suggestions regarding our manuscript. We have made revisions following your comments and suggestions, and the revised contents are marked using the “Track Changes” function of Microsoft. You can view all changes using the “Display for Review” function of Microsoft Word. The line number corresponds to the revised manuscript without changes marked. We have tried our best to correct all grammatical mistakes and statement errors in the manuscript. Please see our point-by-point responses to the Editors’ and Reviewers’ comments below.

Reviewer #2

The study tackles the relevant problem of regional debris-flow susceptibility. I think the authors chose adequate methods (Bayesian statistical methods) and an interesting combination of features representing hydro-geomorphological factors for debris-flow triggering. However, I found the methods to be intransparent in important acspects of the study, such that I cannot assess the reproducibility or plausibility of the results. I list my major concerns below.

Comment 1: What debris flow data was used for training, namely to estimate the occurrence probability in P(Ci) in Eq. 1? You mention that you use “debris flow survey sites” (L123) but there is no reference or description of how you obtain the data. Also Fig. 2 on the study implementation doesn’t mention any use of observational debris flow data to train the model. In L115, I see one citation that may refer to such data (Yu and Tang, 2016), but the full reference is missing. Fig. 5 indicates, that debris -flow fans were identified, but how exactly and how do you differentiate debris-flow fans from alluvial fans?

Response 1: In the revised manuscript, we have provided a more detailed explanation of this section and included the corresponding MATLAB code. The survey data were obtained from the Ministry of Natural Resources of the People’s Republic of China and are publicly accessible via the Resource and Environmental Science Data Platform of the Chinese Academy of Sciences (https://www.resdc.cn/). Regarding references, we have added relevant literature to support and contextualize our methods. Our research is designed to identify a greater number of potential risk points by training on a limited set of fundamental survey locations, aiming to approximate a full-scale 1:1 mapping of debris flow susceptibility. This represents the practical significance of our work. Based on topographic maps and remote sensing imagery (see Fig. 1-2), we do not rely on all survey points as training data. Instead, we optimize data quality by selectively screening points using a range of attribute indicators, thereby maintaining the model error within a reasonable margin. This methodological refinement is now emphasized in the revised text. Although the scale of our survey map may not be exhaustive, the dataset remains objective and robust. In Fig. 5, the presence of debris flow fans serves as a geomorphic indicator to validate the threshold used for defining high-energy valleys. Specifically, the application of an average stream power gradient threshold of 1×10^{-4} W/m² allowed us to effectively distinguish high-energy valleys, and the observed debris flow fans (Fig. 5f-1 and 5f-2) empirically confirm the suitability of this classification criterion. We have added corresponding annotations in the revised manuscript (see lines 13–14 of Section 3.1 “Data and Preprocessing”).

Fig. 1 Distribution of debris flow disaster sites

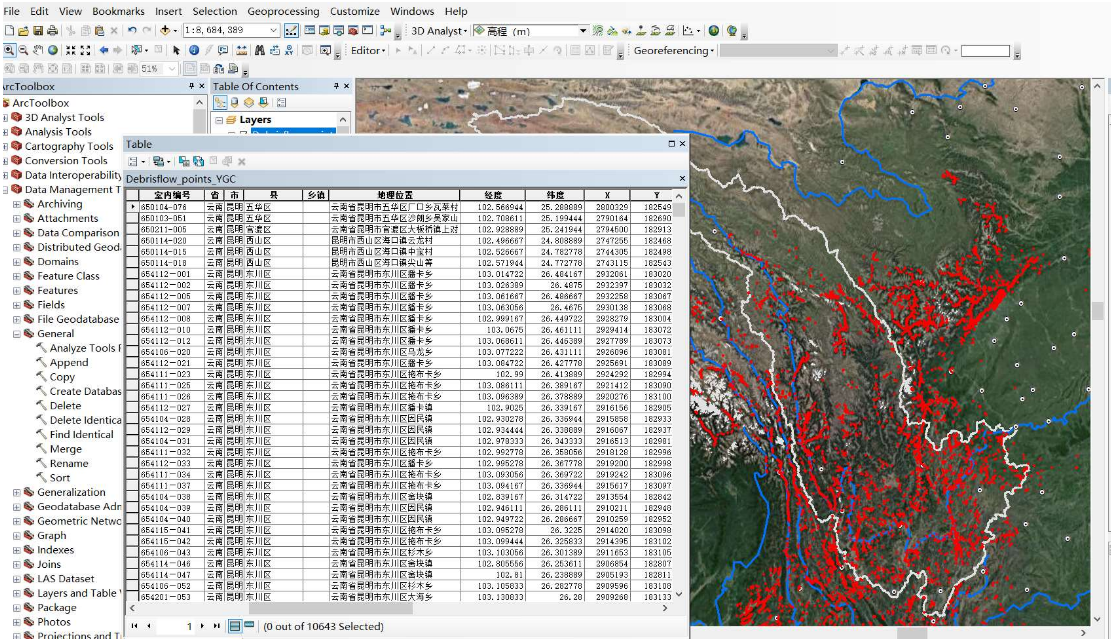

Fig. 2 Vector data

Comment 2: Any information on model training and testing is missing, except for the showcasing the model for one event

Response 2: Regarding the modeling process, it shares similarities with neural network approaches in that it operates as a "black-box" model. The core mechanism involves constructing probability distribution functions to generate probabilistic predictions. In the revised manuscript, we have elaborated further on the underlying principles of the model to enhance transparency and interpretability. These additions can be found in the revised text, specifically in lines 9–34 of Section 3.2 "Modeling Approach".

Comment 3: There is no uncertainty assessment or discussion of model limitation

Response 3: In the revised manuscript, we have added a discussion on the limitations of our modeling approach. These limitations mainly lie in two areas: First, the model assumes strict independence among input variables. This constraint implies that any inherent correlation between variables—especially those sharing a common physical mechanism—must be intentionally avoided during the construction of the indicator system. As a result, the selection of indicators is inherently restricted in both type and number, limiting the comprehensiveness of the parameter set. Second, while the model achieves reasonably accurate spatial predictions, it does not incorporate temporal information. It is important to note, however, that the parameters used in this study are derived from geomorphic processes and are intended to capture key dynamic elements of landscape evolution. These parameters possess clear physical meanings, and thus the identified high-probability zones represent locations of confirmed hazard potential, despite the absence of precise timing. We have incorporated these points into both the "Modeling Approach" and "Conclusion" sections to more fully acknowledge the methodological limitations.

Comment 4: The data availability statement states that datasets are being made available, but there is no link. Anyway, more important would be the data to reproduce the results and this would include the debris flow observations

Response 4: We have uploaded the dataset and explicitly documented the source of the survey data.

Comment 5: The conclusions are largely a copy of the abstract. Both should be rewritten such that they are complementary (e.g., more focus on research question and methods in abstract and more focus on conclusions, implications, outlook in conclusions)

Response 5: We have revised and rewritten both the abstract and the conclusion sections.

Comment 6: Specific comments, ~L70-77 : I cannot follow the critique on previously used indicators for DF susceptibility. It may be that the risk is highest in the valley bottom, the source area characteristics govern susceptibility. Can you specify what current methods exactly are missing and what you do differently? Contradictory to your argument on the importance of valley bottom characteristics, I would assume that the factors you report in L77 (stream power, surface erosion, etc) characterize source are rather than sink area.

Response 6: Our original expression in the manuscript may have lacked clarity. In response to your suggestion, we have revised this section to provide a more precise explanation. Our intention was to emphasize the need for a more targeted indicator system—one that specifically focuses on the linear river channels at the bottom of gullies. This system is designed to directly capture the physical and energetic characteristics of valley floors. In contrast, the broader hillslope source areas generally exhibit more subdued material transport processes. By accurately characterizing the matter–energy dynamics at the gully base, we aim to enhance the predictive power of the model, particularly in terms of spatial localization accuracy.

Comment 7: L105: if you use ERA5, higher resolutions that daily are available to my knowledge. Could you justify why you don't use these? Sub-daily rainfall is commonly much more useful than daily for debris-flow triggering

Response 7: While the ERA5 dataset provides precipitation records at hourly resolution, our research is primarily concerned with improving the spatial accuracy of debris flow susceptibility predictions, rather than modeling the triggering process of individual debris flows at the gully scale. Given the relatively broad spatiotemporal scale of our analysis, daily-scale precipitation data are sufficient to address the scientific questions posed in this study.