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 #1

Comment 1: The Authors did a good job in addressing my review comments one by one and in adapting the manuscript. Of my original comments, there are two remaining where I require a clarification and an adjustment of the manuscript. These are given below. Other issues are solved from my point of view.

Response 1: Thank you for your positive feedback and the constructive suggestions for improvement. We have made the revisions in accordance with your recommendations and have added the appropriate references.

Comment 2: **"**Comment 3: I think the method in the manuscript, which neglects other factors of importance, could be partially responsible for their own relatively low accuracy. Factors such as vegetation, lithology and soil transmissivity (also mentioned by the authors for classical approaches, line 68) are what come to mind. I think deliberately neglecting these factors bends the aim of the manuscript from an overall debris flow hazard indicator to introducing a specific source-sink process-based method. This is still innovative and interesting, but I think the authors should mention their choices in this regard more explicitly at the end of the introduction and in the methods.

Response 3: Our fundamental approach is to constrain the design of the machine learning scheme using the basic principles of watershed erosion and transport. During the research framework design process, we did not overlook the role of vegetation. The reason for not deliberately incorporating vegetation data is that current DEM data products are generally based on InSAR satellite observation technology, which does not filter out the elevation affected by vegetation. The calculation process of the geomorphological connectivity index (IC value) is based on this type of DEM, and thus the resulting IC values naturally include spatial variations in surface connectivity caused by vegetation. We also did not consider using geological maps to describe lithology, as descriptions based on geological maps are typically qualitative, which is not conducive to a quantitative assessment process. In fact, regardless of lithology, loose surface soils and weathered layers are the key contributors to debris flow formation. Therefore, we introduced the erodibility factor, K, from the Universal Soil Loss Equation. This indicator reflects the degree to which the surface is prone to erosion and is only related to the properties of the soil or weathered layer itself. It is a quantitative metric with clear physical meaning, which facilitates a more rigorous quantitative assessment. To aid in reader understanding, we have added relevant explanations (the modified text can be found in lines 8-10 of "3.1 Data and Preprocessing"). Subsequent Response reviewer 3: Correct me if I'm misinterpreting here, but it reads as if you treat canopy height as an addition to the DEM. This is not how I think vegetation should be included. Vegetation has a complex interaction with soil hydrology and geotechnics. Itis not 'additional elevation'. My advice would be to somehow reflect vegetation presence in your model as independent variable or incorporate it in the erodibility/connectivity. Another option would be to ignore it of course, as it might not be a focus of the study. I think you should also mention that the DEM you use is in fact a Digital Surface Model."

Response 2: The approach to fine-scale simulation typically aims to consider as many factors as possible to minimize errors. However, models built with this philosophy often face challenges in scaling to larger spatial assessments due to their computational intensity. We do not intentionally overlook any factors; rather, starting from general physical processes of surface dynamics, we aim to identify the most critical factors. This approach allows us to achieve high-quality results while greatly reducing data processing demands, making it suitable for large-scale evaluations at the scale of thousands to tens of thousands of square kilometers.

The topographic data used in this study were not subjected to "vegetation elevation correction." Consequently, the surface connectivity derived from this data inherently includes the influence of vegetation, which we treat as a parameter for surface roughness. Furthermore, the water-soil coupling process is highly complex, with vegetation playing an integral role, often acting as an obstacle in the "source-sink" process. The dynamics between water and soil are more central to discussions on the driving environmental processes. In addition, the primary influence of vegetation on the formation of mountain floods and debris flows occurs on slopes, while our indicator specifically targets the valley floor, where the role of vegetation is relatively indirect. As such, vegetation was not singled out in the initial model. In the revised manuscript, we have supplemented the discussion on the role of vegetation while also emphasizing the focus of this study.

Comment 3:

Comment 26: Figure 11. How is the relative importance calculated?

Response 26: In machine learning, if a change in the value of a particular factor leads to a more significant change in the dependent variable, then the relative importance of that factor is higher. This can be understood through a simpler example. For instance, 9 in multiple linear regression, each independent variable in the results corresponds to a significance level p-value. The smaller the p-value, the more significant the factor, and thus, the importance of the factors can be ranked based on the significance of the p- value.

Response reviewer: Mention this method explicitly, or with a reference.

Response 3: The relative importance of a factor can be determined by calculating the difference in the log-likelihood ratios for different factors. The mutual information measure permits analysis with both continuous and categorical variables and has been widely adopted in the literature; we therefore select this metric (Blanquero et al., 2021). It quantifies the information about variable *X* contained in variable *Y*, defined formally as:

$$I(X,Y) = \iint P(x,y)\log\left(\frac{P(x,y)}{P(x)P(y)}\right)dxdy \quad (6)$$

In addition, we have provided clear references with explicit annotations in the revised manuscript (Blanquero et al., 2021).

Reviewer #2

Comment 1:

In my opinion the paper provides innovative methodology to model debris flow susceptibility. Overall, the paper is good shape with concise language and descriptions. However, I found it sometimes very hard to read and understand the figures. There is a lot of information on the figures with small fonts, I suggest increasing the size of the fonts in the figures for better readability. Additionally, I wonder if some subplots could be removed in order to focus more specifically on key aspects of the figures. Furthermore, I found the method and result section about extreme precipitation difficult to understand. Specifically, I did not understand how the computed severity relates to the Standardized Precipitation Index (SPI) of Table 1? And how is heavy precipitation (used in Fig 8) defined? Can you better explain these sections?

Response 1: The observed image degradation likely resulted from lossy compression during file format conversion; the original vector graphics have been restored in the manuscript. Additionally, regarding the severity of extreme precipitation, we have included a clear definition formula (18) in the Methods section, which represents the average SPI value during the duration of extreme precipitation events. Furthermore, Table 1 outlines the classification standards for precipitation intensity levels, including the categorization for extreme drought conditions. In addition, we have redrawn some of the illustrations (Fig. 8).

Comment 2:

Specific comments:

• Table 1 is not referenced in the text.

• Sometimes spaces are missing in the text. On following lines, I found missing spaces. But please

carefully check the entire manuscript for further missed or double spaces. Lines:

41, 103, 182, 185, 217, 229, 456

- Line 159: I suggest writing 1.6 * 104 N/m3 as 16,000 N/m3
- Figure 5 misses the index f-1
- Figure 8 misses a description of subfigures b, e, h in the caption
- Line 299 state the year of the date (2024)
- The sentence of lines 333-335 misses a references
- The sentences about climate of lines 354 – 358 also need references.

Response 2:

We greatly appreciate your meticulous attention to detail, which has significantly enhanced the quality of our work. In response to your suggestions, we have made the following revisions:

Added labels for Table 1.

Conducted a thorough review to eliminate duplicate spaces.

Corrected the notation of 1.6 * 10^4 N/m^3 to 16,000 N/m^3.

Revised the figure legend for Figure 5f-1.

Updated the figure legend for Figure 8 to include descriptions of subfigures b, e, and h.

Incorporated references to relevant literature on news events.

Included additional references concerning climate change.

Blanquero, R., Carrizosa, E., Ramírez-Cobo, P., Sillero-Denamiel, M.R., 2021. Variable selection for Naïve Bayes classification. Computers & Operations Research, 135, 105456.