

Referee #2 Comment 01

The manuscript addresses a relevant and timely topic by combining explainable AI with rainfall-induced landslide prediction. The proposed Dynamic Rainfall Threshold framework appears promising and has the potential to improve model interpretability for operational Landslide Early Warning Systems. However, several methodological aspects require clarification before the scientific contribution can be fully assessed. The manuscript is generally understandable; however, a careful proofreading is recommended to improve the fluency and readability of the text.

We sincerely thank the anonymous Referee 2 for the thorough assessment of our manuscript and the positive feedback regarding the relevance of our work. We thank the reviewer for the constructive suggestions for improvement, particularly regarding the need to clarify the validation methodology. We appreciate the complementary perspective to the comments provided by referee #1 and believe that a revised version of the manuscript will adequately address the points raised by the reviewer.

Referee #2 Comment 02

In general, the validation procedure is not clearly explained in the Methodology section. Details that belong to the methodology are found in the Results section (lines 125-132 and 140-150). I suggest adding a short but clear paragraph in the Methodology section where the validation procedure is addressed.

We agree that explaining the validation procedure more clearly in the methodology section will improve the manuscript significantly. At the same time, this can further improve the readability of the results section. The revised manuscript will thus add a more thorough validation methodology description.

Referee #2 Comment 03

Only 40 events are reserved for validation, which appears rather limited considering the overall database. Please justify this choice and clarify whether the validation subset adequately represents the temporal and spatial variability of the complete dataset.

As addressed in the comment above, a clearer explanation of the validation method should include an explanation of the event selection for validation. The 40 events are for validation as

for example in Mondini et al, 2023 (<https://www.nature.com/articles/s41467-023-38135-y>), whereas the reported scores stem from the test set which contains 10% of all events in the dataset. The validation events have thus been drawn from the complete dataset before constructing training and testing data to demonstrate the working of the model on previously withheld data. We propose to rename this subset of samples the “demonstration” dataset for clarity and will add a more thorough description of the sample generation and splitting in the methodology.

Referee #2 Comment 04

Regarding the comparison among the different models, the methodology does not describe the characteristics chosen for each model :

- *No hyperparameters*
- *No details on optimization procedures*
- *How thresholds were chosen*

Without this information it's difficult to see if the reported superiority of KANs is a consequence of better tuning or not. To ensure a fair comparison, it would be useful to report whether all competing models underwent comparable hyperparameter optimization procedures.

We will expand the methodology section to include information on the grid search conducted to fine-tune the hyperparameters. This will include the hyperparameter ranges, optimisation details (optimiser used, learning rates and early stopping criteria as well as loss function). Generally, the hyperparameter grid searches were designed to be as fair as possible with regards to the differences in model architectures to allow for a fair comparison of the models (see also Referee #1 Comment 04).

Referee #2 Comment 05

The paragraph 2.2 Network Explanations is well written, however, while the mathematical background of KANs is useful, some of this material is already available in the cited literature. In contrast, additional methodological details on the generation of the proposed Dynamic Rainfall Thresholds (citing the produced figures) would improve reproducibility

The reviewer suggestion partly overlaps with Referee #1 Comment 07. We will address the issue in the revised manuscript by expanding the DRT methodology description. We thank the

reviewer for the suggestion to shorten the mathematical background description. This shortening will make the section after amendments more readable, and we aim thus to implement both recommendations.

Referee #2 Comment 06

The Discussion section is well structured. I appreciate the interpretation of results and the acknowledging of limitations, as well as the outlining of future improvements. However, there are a few aspects that could be improved or discussed: The main scientific issue remains the lack of comparison with conventional rainfall thresholds. Although the authors highlight their potential role in enhancing predictive capabilities; in the end we do not know how they perform. Since the manuscript introduces DRT as an interpretable decision-support tool, it would greatly benefit from a comparison with traditional empirical rainfall thresholds (e.g., intensity-duration or event-duration thresholds). Such a comparison would help readers understand the practical advantages, limitations, and potential complementarity of the proposed framework within existing operational Landslide Early Warning Systems

We thank the reviewer for the positive feedback on the discussion. We acknowledge that the section can be improved by directly comparing the Dynamic Rainfall Thresholds to empirical thresholds. An interpretation comparison combined with a hypothetical Early Warning System scenario will be added to the discussion section in the revised manuscript accordingly.

Referee #2 Comment 07

Generalizability: The authors state “The method is widely applicable as long as landslide inventory data is available”. I suggest softening this statement. In fact, the model has only been trained and tested in Northern Italy. Rainfall-landslide relationships vary considerably due to climatic conditions, geological setting, vegetation, soil properties etc., therefore, I would suggest being a little more cautious.

Generalisability stems from the fact that the model is “portable” as long as data of similar quality is available. We acknowledge however that softening the statement is a valid suggestion and will follow the reviewer's recommendation in the revised manuscript version.

Referee #2 Comment 08

The Discussion could further analyse how the proposed interpretability framework would translate into operational Landslide Early Warning Systems. In addition, the manuscript would benefit from a discussion of the uncertainty associated with the predicted dynamic thresholds or model outputs, particularly if the framework is intended for operational applications.

We appreciate the suggestion to include a discussion point regarding the integration of DRTs into an operational Early Warning System and will elaborate on this point in the discussion of the revised manuscript.

Referee #2 Comment 09

In the introduction section: A deeper discussion on the limitations of current approaches and an explanation on why KANs are supposed to provide a substantial advantage beyond interpretability.

We agree that elaborating on the advantage of KAN based networks in the introduction will improve the manuscript. We will add more context accordingly in the revised manuscript.

Referee #2 Comment 10

A clearer introduction should be given to DRTs which seem to represent one of the main scientific contributions of the paper.

As the DRTs are a methodological contribution and not found in this form in the literature, we propose to improve the methodology section where the DRT method is described. A more thorough description will further improve the manuscript as suggested above.

Referee #2 Comment 11

The novelty of the work is currently presented across several parts of the Introduction. I suggest adding a short concluding paragraph summarising the main scientific contributions.

We thank the reviewer for this constructive feedback and will add a summary paragraph to the introduction section of the revised manuscript to briefly highlight the main contributions.

Referee #2 Comment 12

Please, carefully check the Figures and their citation in the text:

- *Figure 3 and 4 seem to have been exchanged,*
- *Is Figure 3b missing?*
- *Line 154, is citing Figure 3 correct?*

We carefully reviewed the figures and their citations throughout the manuscript and will implement corrections accordingly. The revised manuscript will correct the figure references including the missing visual reference in figure 3b.