REFEREE'S COMMENTS ON egusphere-2025-1590

<u>RC#1</u>

R: Thank you for the opportunity to review this short-form manuscript. The authors design a set of straightforward experiments that include testing the efficacy of using 0- to 15-day antecedent soil moisture information from a modeled global reanalysis data product, in conjunction with rainfall data, to identify the triggering conditions for landslides using machine learning. The Results and Conclusion are intuitive in that antecedent soil moisture improves model performance, with the benefit decreasing somewhat with increased lag. This deprecation in model performance seems minor for a lag that is equivalent to the latency of the modeled soil moisture data product (~5 days). Although I appreciate the streamlined presentation of this study, I think it would be helpful for readers to see more text related to (1) the kind of landslides this study is relevant to, (2) why the spatiotemporal resolution of the modeled soil moisture data product is appropriate for the landslide type(s) considered here, and (3) a deeper interpretation of the Results. Regarding #3 - What are the rainfall depth/duration characteristics and the antecedent soil moisture levels that the best-performing model settles on? And do these characteristics make sense relative to the landslide type(s) and/or any previously published regional thresholds? The objective of this study is crystal clear, but the authors may consider questions like these to expand the relevance of their work for the broader scientific community.
Sincerely,
Matthew A. Thomas

A: Thank you for your general appreciation of our manuscript and your valuable comments. Here our replies to the issues raised:

1) Regarding the type of landslides our study is relevant to, we have limited the information in this manuscript as they are basically the same of our previous cited study (Distefano et al., 2023). Nevertheless, given also similar requests by other referees, more details may be added in the "Study area and data section". Text to add may be as follows:
"*The landslide database distinguishes between two categories of events: single landslide events (SLEs) and areal landslide events (ALEs). Both types were incorporated in the reconstruction of the most comprehensive set of landslide occurrences. SLEs generally provide more precise temporal and spatial information on the failure, whereas ALEs consist of multiple landslides within a defined area and are typically associated with lower spatial resolution—often limited to administrative units such as municipalities. Despite the potential uncertainties in the location and data for both SLEs and ALEs, all events were retained for the threshold analysis, as excluding them would have led to an insufficient sample size for training, validating, and testing the neural network models. More specifically, only those landslides linked to rainfall as the triggering factor were considered. This includes events where the mobilized material was consistent with rainfall-induced landslides, while other types (e.g., rockfalls) were excluded. The final dataset consisted of 207 rainfall-related landslides. The CTRL-T software facilitated the reconstruction of triggering conditions for 144 of these events. Among them, movement type information is unavailable for 126 events (87.5%). Of the remaining cases, 10 were identified as areal rockfalls (6.9%), and both flows and slides accounted for 4 events each (5.6%). Regarding temporal resolution, for 103 landslides only the date of occurrence was known. For the rest, more specific information—such as the hour or time of day (e.g., morning, afternoon, or evening)—was available. In cases with daily resolution, the failure was assumed to have occurred at the end of the day, while in events with more detailed timing, the landslide was assigned to the time of peak rainfall.*"
2) Regarding the spatiotemporal resolution of the modeled soil moisture, the following details will be added to the revised manuscript in the section "Case Study and data":
"Regarding temporal resolution, we recognize the importance of temporal resolution in the context of landslide modeling. However, considering the typical precision and uncertainty associated with landslide timing information in most available datasets, we believe that an hourly temporal resolution is already sufficiently detailed for capturing the relevant triggering conditions in most cases. In fact, finer resolutions may not necessarily lead to a meaningful improvement in model performance, given the inherent limitations in the temporal accuracy of landslide occurrence records. With reference to spatial resolution, each cell of the ERA5 Land grid covers an area of about 9x9 = 81 km^2, which is consistent with the average density of available rain-

gauges, that corresponds to 1 rain-gauge per 84 km^2. This can provide a measure of a sufficient resolution for applying our proposed approach."

3) Regarding a deeper interpretation of the results, i.e. more details on the rainfall depth/duration characteristics and the antecedent soil moisture levels that the best-performing model settles on, and whether these characteristics make sense relative to the landslide type(s) and/or any previously published regional thresholds, please consider the following. As highlighted by Distefano (2023) (see Fig. 7), the rainfall events used for the development of the neural networks exhibit durations ranging from 1 hour to 10 days and event rainfall totals up to 300 mm. These values define the applicability domain of the trained models. Nonetheless, it is worth noting that the same study reported even better predictive performance for a climatically distinct region in Norway, where event durations spanned from 24 hours to 40 days and rainfall accumulations reached approximately 400 mm. This suggests that the methodology may maintain robust performance even under differing climatic conditions, particularly when using lagged hydrometeorological inputs. Additionally, the soil moisture values included in the analysis range between 0.15 and 0.45, effectively covering the porosity spectrum of most soil types. This supports the generalizability of the proposed approach across varying environmental settings. These aspects will be covered in Discussion section of the revised manuscript.

Other Notes:

R: LN 20: Comma needed in "4862"?

A: Will be done

R: LN 30: May consider highlighting that ANNs have also proven successful for forecasting subsurface hydrologic response for landslide-prone hillslopes. https://doi.org/10.1029/2020GL088731

A: Thanks for indicating the linked study. The following lines may be added in the Introduction Section:
"Orland et al. (2020) developed a deep learning model that utilized soil moisture, pore pressure, and rainfall monitoring data from landslide-prone hillslopes in the USA to predict both the timing and magnitude of hydrologic responses at various soil depths. Their findings suggest that machine learning offers an accurate and computationally efficient alternative to empirical approaches and physical models for landslide hazard warning. Other studies have also confirmed the effectiveness of machine learning in different contexts, such as Distefano et al. (2022, 2023) [...]"

R: LN 34-41: Is it worth mentioning that virtually all of these kinds of rainfall and soil moisture products (e.g., NASA GPM and SMAP) have some kind of latency?

A: We fully agree that latency is a common feature of these kinds of rainfall and soil moisture products. For instance, the NASA GPM (Global Precipitation Measurement) mission provides different data streams with varying latencies between the observation time and the publication varying from 4 hours to 3.5 months (Huffman et al., 2019). Similarly, NASA's SMAP (Soil Moisture Active Passive) soil moisture retrievals typically have a latency of 1–3 days, with latency increasing for higher-level or gap-filled products (Dashtian et al. 2024). Thus, in the revised manuscript, we will highlight that the latency issue affects a broader range of Earth observation datasets, and that the approach presented in this study is not only relevant to ERA5-Land reanalysis dataset, but an issue common to other sources of data that may be of interest for landslide prediction studies.

REFERENCES: Huffman, G.J.; Bolvin, D.; Braithwaite, D.; Hsu, K.; Joyce, R.; Kidd, C.; Nelkin, E.; Sorooshian, S.; Tan, J.; Xie, P. NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG)). Algorithm Theoretical Basis Document (ATBD) Version 06; NASA/GSFC: Greenbelt, MD, USA, 2019; Volume 30

Dashtian H, Young MH, Young BE, McKinney T, Rateb AM, Niyogi D, Kumar S V. (2024) A framework to nowcast soil moisture with NASA SMAP level 4 data using in-situ measurements and deep learning. J Hydrol Reg Stud 56:102020. https://doi.org/10.1016/J.EJRH.2024.102020

R: LN 57-58: "On the other side," may be unnecessary text.

A: Will be done

R: LN 133-134: What are the implications for assuming landslide timing as the end of the day?

A: We acknowledge that assuming landslide occurrence at the end of the day introduces temporal uncertainty, which may affect the estimation of rainfall conditions leading to the event. As discussed by Peres et al. (2018), even small shifts in the assumed landslide timing—such as reporting the event earlier or later than it occurred—can significantly affect the derived rainfall thresholds. Nonetheless, it is not possible to overcome this lack of precise timing, and the best we could do is to use the available information consistently throughout all the analyses we have conducted. A comment will be added in the Discussion section of the revised manuscript.