Dear Reviewer,

We sincerely thank you for these detailed and constructive comments. In response, we have carried out a set of new analyses, which are reported in this letter, including landslide inventory declaration, a multi-resolution scale-sensitivity assessment, robustness and reproducibility using stability-based edge selection over repeated spatial partitions; and a SHAP-based pathway attribution scheme that replaces the earlier heuristic fusion to generate the contribution of the identified pathway. We also agree that parts of the reasoning, logical structure and wording require reorganization, and we specify below how each will be revised. We note that, as the journal does not permit submission of a revised manuscript at the current stage, we will systematically revise the manuscript in the next stage by incorporating your suggestions and will provide further clarification and additional information where appropriate.

Q1: Landslide inventory issue. It is recommended that additional information be provided, including landslide type, occurrence time and triggering factors. In that case, you can use the DAG method to explore direct or indirect evidence relationships in each type of landslide. If sufficiently reliable type- or trigger-specific information is available, it would be valuable to examine whether the inferred dependency structures differ among rainfall-induced, snowmelt-related, earthquake-induced, or other landslide categories. Such an analysis could provide stronger evidence for the proposed direct and indirect dependency pathways.

Reply: Thank you for your constructive suggestions. Landslide inventory attributes such as landslide occurrence time, size/type, and triggering factors are important for interpreting the direct and indirect dependencies among the pre-conditioning factors. In response to this comment, we have re-examined the inventory, which contains 1,198 landslide records. The attribute fields — landslide type, size class, recorded date of occurrence and triggering factor — are complete for all records; Table 1 lists the attribute fields of the inventory to document this completeness. Figure 1 presents the statistical characteristics of the inventory. In terms of size, the inventory is dominated by small- and medium-sized landslides. In terms of timing, landslide occurrence is concentrated in two periods: spring (March–May), associated with snowmelt, and July–October, associated with the rainy season; the recorded events span multiple years. The triggering factors are dominated by rainfall, earthquakes and snowmelt, together with their combined effects. More importantly, the inventory has undergone independent expert review and has been formally released on the official data-sharing platform of the Third Xinjiang Scientific Expedition Program of the Ministry of Science and Technology of China (https://cstr.cn/33110.11.XIEG.XJSEDATA.2024.00100302). This detailed inventory information can be used as external evidence to assess the consistency of the directed spatial dependency pathways inferred by the DAG and to infer possible propagation processes in regional landslide development.

Table 1. Selected attributes of the landslide inventory.

ID	Elevation difference (m)	Aspect (°)	Landslide length (m)	Landslide width (m)	Area (m²)	Occurrence time	Longitude	Latitude	Scale	Triggering factor
1	35	101.17	100	30	3402.49	2007.9	83° 9' 59.136" E	43° 53' 41.746" N	Small	Rainfall and earthquake
2	20	186.37	80	50	3265.21	2007.9	83° 11' 44.127" E	43° 53' 56.440" N	Small	Rainfall and earthquake
3	160	236.19	770	360	260501	2007.9	81° 24' 8.584" E	43° 14' 47.322" N	Medium	Rainfall and earthquake
4	165	229.45	950	210	251962	2007.9	81° 24' 11.973" E	43° 14' 36.964" N	Medium	Rainfall and earthquake
5	85	261.57	290	90	26190.7	2007.9	81° 24' 14.360" E	43° 15' 4.118" N	Small	Rainfall and earthquake
…	…	…	…	…	…	…	…	…	…	…
1198	140	30.02	376	518	127826	2005.3	83° 16′1.200″E	43° 18′42.707″N	Medium	Rainfall and engineering slope cutting

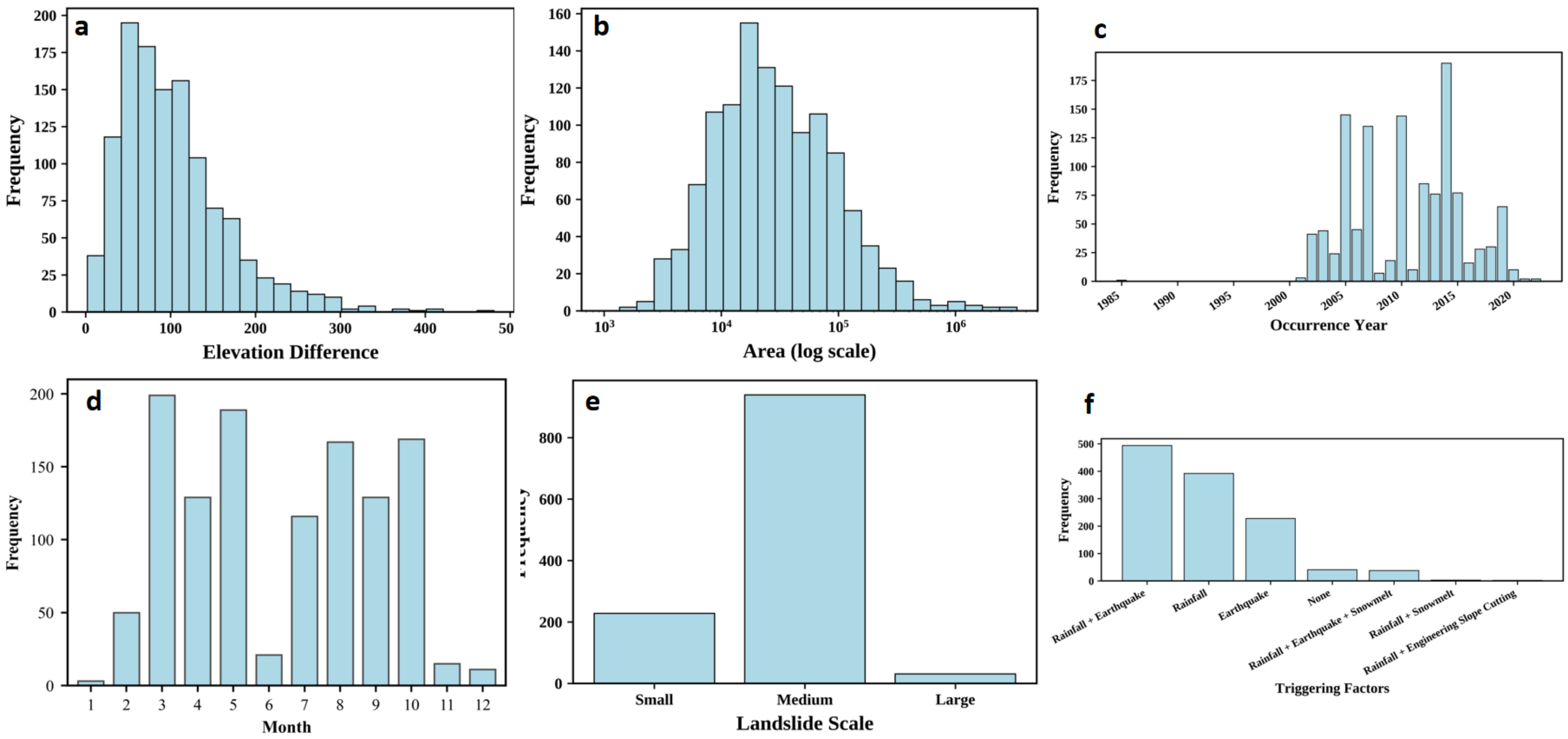

Figure 1. Statistical characteristics of the landslide inventory: (a) elevation difference, (b) area, (c) occurrence year, (d) occurrence month, (e) landslide scale, and (f) triggering factors.

Q2: The rationale for selecting variables such as NDVI, NDWI, annual and monthly mean temperature, precipitation, and snow depth should be further strengthened. For example, why NDVI represents root-reinforcement/interception, why NDWI proxies antecedent soil saturation, why snow depth matters for freeze-thaw, that will improve physical justification.

Reply: We thank the reviewer for this suggestion. In the revised manuscript we will add the links between these environmental factors and landslide occurrence in the study area, and we state explicitly that the remote sensing and climatic variables are used here as proxies for the relevant physical processes rather than as direct measurements of the corresponding physical quantities.

Specifically, vegetation increases the shear strength of shallow soils through root anchoring and modifies the moisture state of the slope through transpiration and canopy interception. NDVI is a retrieval index of vegetation cover at the regional scale and can therefore be used to reflect indirectly the spatial variation of root reinforcement, transpiration and interception effects, without directly quantifying the shear contribution of roots (Carlson and Ripley, 1997; Bordoloi and Ng, 2020). NDWI is sensitive to the liquid water content of vegetation and of the land surface, and is used here as a remote sensing proxy for regional moisture conditions and antecedent wetness, rather than as a direct measurement of soil saturation or pore-water pressure (Gu et al., 2008; Sánchez-Ruiz et al., 2014).

The hydro-climatic factors characterize the long-term hydro-climatic and thermal background conditions. Monthly mean precipitation (July–October) is used to characterize the spatial variation of moisture conditions between seasons (Iverson, 2000; Gariano and Guzzetti, 2016), and monthly mean temperature (March–May) reflects seasonal temperature variation and differences in the freeze–thaw environment. Snow depth characterizes the snow storage and the seasonal hydrothermal conditions of different areas. Previous studies have shown that increased snow depth combined with rising temperature is an important factor promoting snowmelt-induced landslides in this region (Zhang, 2005; Zhao et al., 2023).

On the basis of these mechanisms, NDVI, NDWI, precipitation, temperature and snow depth are used in this study as proxy variables for vegetation–hydrological processes, antecedent wetness, and seasonal hydrothermal climatic conditions, respectively, in order to characterize the spatial variation of slope stability conditions under different environmental settings. We further elaborate the rationale for selecting these factors in the revised manuscript, so as to avoid equating remote sensing or climatic indices with direct physical quantities such as root strength, pore-water pressure or freezing depth.

Q3: We found you have partially addressed already via the ablation and data-sparsity experiments (Figures 10–11), further discussion is recommended regarding the feasibility, robustness, and transferability of the model, including the stability and reliability of identified spatial dependency under different data partitions and parameter settings.

Reply: We thank the reviewer for this suggestion. The experiments in the original manuscript already provide a preliminary assessment of the feasibility, robustness and transferability of the model from several perspectives.

First, regarding feasibility, this study does not rely solely on purely data-driven structure learning. Instead, expert knowledge is combined with data-driven DAG learning, and current understanding of landslide formation mechanisms in the study area is used to constrain the dependency structure, thereby reducing connections that lack physical meaning.

Second, regarding robustness and transferability, the original manuscript already examined the stability of the model under different data conditions and model configurations through module-ablation experiments, data-ablation experiments and repeated independent training runs. In the revised manuscript we further examine the stability of the spatial dependency structure itself, using two metrics:

(1) I-DAG structural similarity. The Jaccard similarity of the directed edge sets is used to measure the degree of overlap between the I-DAG obtained in each repeated experiment and the baseline structure, calculated as:

$$J(E_i, E_0) = \frac{|E_i \cap E_0|}{|E_i \cup E_0|} \tag{1}$$

where E_0 denotes the set of directed edges of the I-DAG obtained by the baseline model, and E_i denotes the set of directed edges obtained in the i-th repeated experiment, or under a different parameter setting or a different spatial partition.

(2) Pathway-ranking stability. The Spearman rank correlation is used to compare the consistency of the importance rankings of the main dependency pathways across experiments, calculated as:

$$\rho_s = \frac{\sum_{j=1}^{m}\left(R_{ij} - \bar{R}_i\right)\left(R_{0j} - \bar{R}_0\right)}{\sqrt{\sum_{j=1}^{m}\left(R_{ij} - \bar{R}_i\right)^2}\sqrt{\sum_{j=1}^{m}\left(R_{0j} - \bar{R}_0\right)^2}} \tag{2}$$

where R_{ij} and R_{0j} denote the rank of the j-th dependency pathway in the i-th experiment and in the baseline experiment, respectively.

In addition, we carry out the sensitivity analysis for the key model parameters. Keeping all other settings unchanged, the main structure-learning and network-training parameters described in the Methods section were each varied around their baseline values, and the I-DAG structural similarity and the Spearman correlation were recomputed. The results are reported in Table 2. They show that the DAG is essentially stable across the tested parameter settings, with Jaccard similarity between 0.79 and 1 and ρs between 0.91 and 1. In the revised manuscript we further analyse the stability of the DAG and determine the parameter settings to be adopted.

Table 2. Preliminary parameter-sensitivity results.

Parameter	Value	Edge Jaccard	ρs
GAT heads	2	1	1
GAT heads	4	0.92	0.99
GAT heads	8	0.97	0.91
Learning rate	0.0005	1	1
Learning rate	0.001	0.82	0.95
Learning rate	0.005	0.85	0.94
Hidden units	32	1	1
Hidden units	64	0.79	0.97
Hidden units	128	0.85	0.99

Q4: Random splitting on spatially autocorrelated environmental data will inflate reported accuracy/AUC, since train and test pixels can be near-neighbors. It is recommended that spatial cross-validation be incorporated to reduce spatial dependence between training and testing samples and to evaluate model performance across multiple independent training-testing splits. This would provide a more objective assessment of the model's generalization ability.

Reply: We thank the reviewer for this suggestion. To address this issue, we have added a 5-fold spatial cross-validation (SCV) in the revised manuscript, following the recommendations of Brenning (2012). As shown in Figure 2, the landslide samples were first partitioned into five spatial subregions, F1–F5, by applying k-means clustering (k = 5) to their spatial coordinates. One complete spatial fold was held out as an independent test region and the remaining four folds were used for model training, with the five regions rotated in turn.

For each SCV training run, the corresponding I-DAG and dependency-pathway importance were re-derived. Table 3 reports the results: the Accuracy, AUC and F1 score across the spatial folds range from 0.77 to 0.86, from 0.78 to 0.84 and from 0.76 to 0.83, respectively. Although the three metrics fluctuate to some extent between folds, no consistent or pronounced decline in performance is observed. Compared with the random partitioning used in the original manuscript, this approach reduces the effect of spatial autocorrelation caused by spatially adjacent samples entering both the training and the test sets, and provides a more rigorous test of the model's generalisation ability and spatial robustness in areas not used for training.

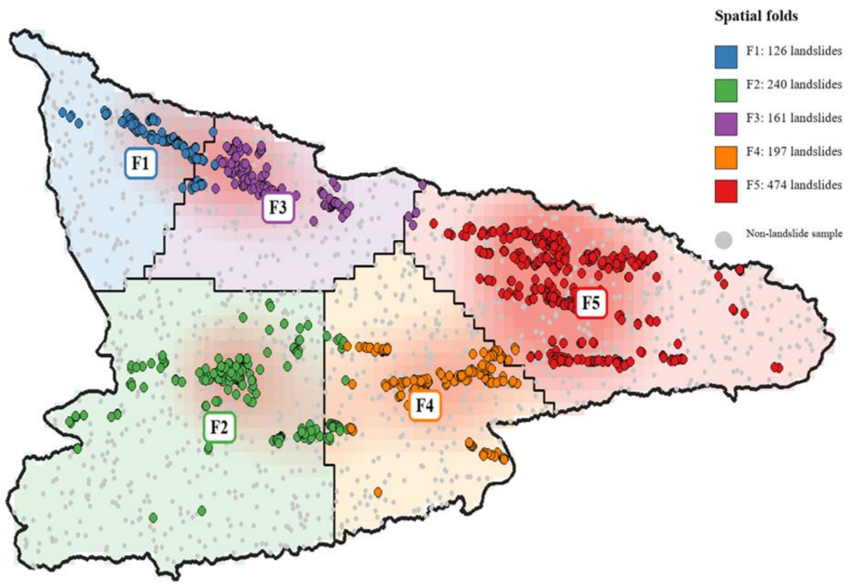

Figure 2. Representative repetition of the 20 repeated five-fold k-means spatial cross-validation. The background shows landslide-sample density overlaid on nearest-centroid k-means spatial partitions (k = 5) clipped to the study-area boundary, and the points show the landslide samples assigned to each spatial fold.

Table 3. Performance of spatial cross-validation (pre-experiment).

Spatial fold	Accuracy	AUC	F1
F1	0.85	0.84	0.82
F2	0.82	0.80	0.82
F3	0.86	0.78	0.76
F4	0.77	0.81	0.81
F5	0.79	0.79	0.83

Q5: Derivation and validation of pathway-specific susceptibility maps. Extracting dominant dependency pathways and generating pathway-specific susceptibility maps are among the most distinctive contributions of this study. Please add explanatory context about the construction and interpretation of these maps.

Reply: We thank the reviewer for this suggestion. In the original manuscript, the pathway-specific susceptibility maps were obtained by fusing the global landslide susceptibility (Global LSP), the pathway importance derived from the GAT, and the spatial correlation between pathway importance and global susceptibility. Specifically, we first computed the pathway importance of each dependency chain at every grid cell, then used the Spearman correlation coefficient to determine whether the direction of that pathway was consistent with the spatial pattern of global susceptibility, and finally combined the two through a weighted fusion governed by the preset empirical parameters α and γ. Although this approach could to some extent reflect the spatial differences among dependency chains and their promoting or inhibiting effects, when following the reviewer's suggestion to set out the construction and interpretation of these maps more clearly, we recognised that the method involves several empirical fusion steps and preset parameters, which make the correspondence between pathway contribution and the final model prediction insufficiently objective and increase the interpretive complexity of the pathway-specific maps.

In the revised manuscript we have therefore changed to a SHAP-based pathway attribution approach. This follows the idea, introduced in Shapley Flow, of treating pathways as the objects of attribution, and takes each complete dependency chain Cc as the unit of pathway-level attribution (Wang et al., 2021). For any combination of pathways $S \subseteq C$, the model retains only the DAG information-propagation relationships corresponding to that combination and computes the resulting prediction $vx(S)$. The contribution of the c-th pathway at grid cell x is defined as the Shapley average of the marginal changes in prediction produced when that pathway is added to the different pathway combinations:

$$\phi_c(x) = \sum_{S \subseteq C \setminus \{C_c\}} \frac{|S|!(K-|S|-1)!}{K!} \left[v_x(S \cup \{C_c\}) - v_x(S) \right]. \qquad (3)$$

Thus, $\varphi_{c(x)}$ represents the SHAP value of a complete dependency chain with respect to the final landslide susceptibility prediction. A positive SHAP value indicates that the pathway increases the predicted susceptibility at that location, i.e. a promoting effect, whereas a negative SHAP value indicates an inhibiting effect. The SHAP values of all grid cells are then remapped according to their spatial position, yielding the pathway-specific contribution map for that dependency chain.

Finally, we validated these maps spatially using landslide samples with a clearly documented triggering background. For example, areas of high positive contribution of the precipitation-related pathways were compared with rainfall-triggered landslide points, and the earthquake-related pathways were compared with earthquake-related landslide points. The results show that landslide points of the corresponding type are generally concentrated in areas where the corresponding pathway has a high positive contribution, indicating that the spatial distribution of the derived pathway contributions is spatially consistent with the actual triggering background and geologically reasonable (Figure 3).

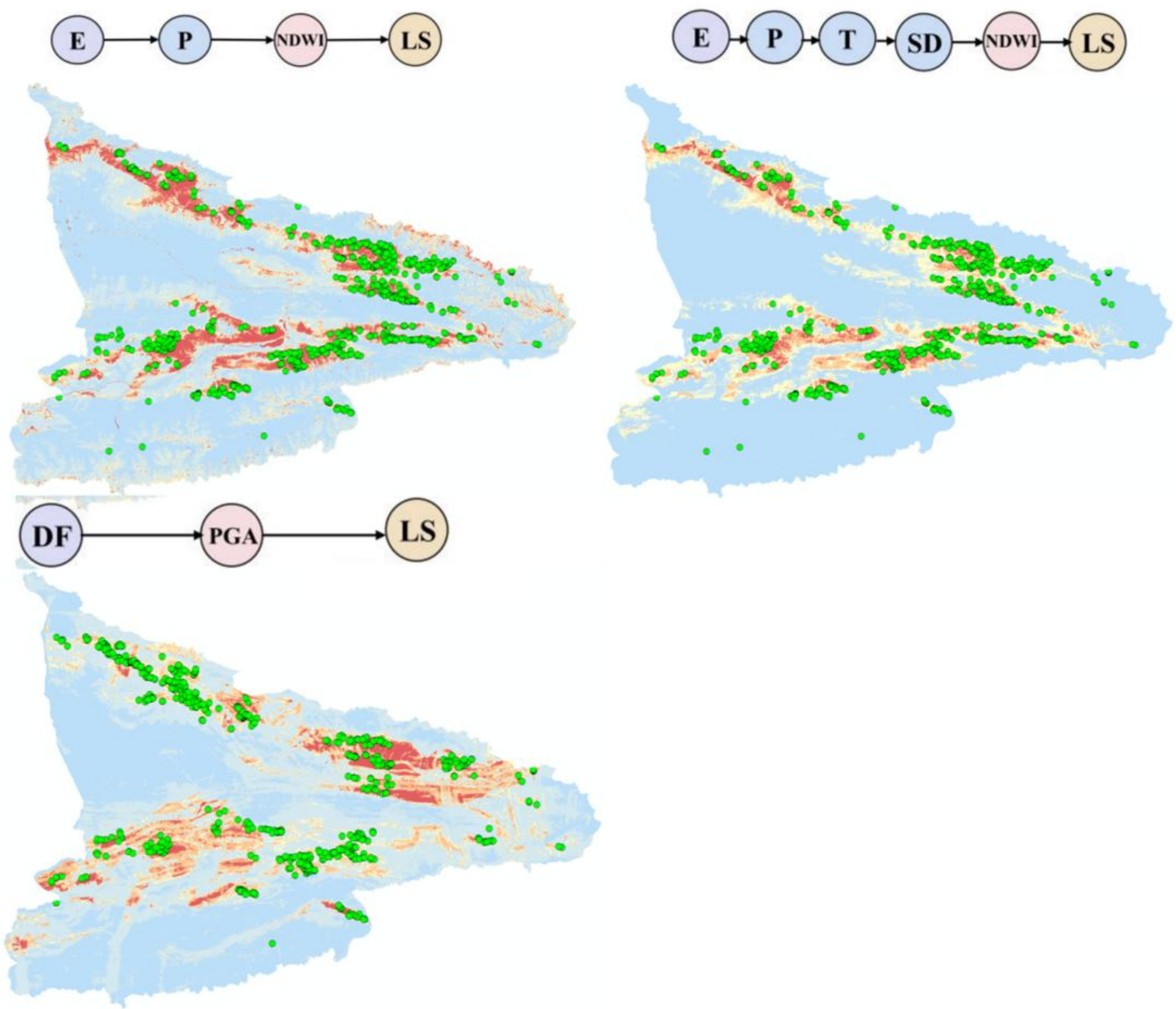

Figure 3. Relationship between pathway-specific susceptibility results and observed landslide subsets with corresponding trigger information.

References

Bordoloi, S., and Ng, C. W. W., 2020. The effects of vegetation traits and their stability functions in bio-engineered slopes: A perspective review. Engineering Geology, 275, 105742.

Brenning, A., 2012. Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: The R package sperrorest. In 2012 IEEE International Geoscience and Remote Sensing Symposium, 5372-5375.

Carlson, T. N., and Ripley, D. A., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sensing of Environment, 62, 241-252.

Gariano, S. L., and Guzzetti, F., 2016. Landslides in a changing climate. Earth-Science Reviews, 162, 227-252.

Gu, Y., Hunt, E., Wardlow, B., Basara, J. B., Brown, J. F., and Verdin, J. P., 2008. Evaluation of MODIS NDVI and NDWI for vegetation drought monitoring using Oklahoma Mesonet soil moisture data. Geophysical Research Letters, 35.

Iverson, R. M., 2000. Landslide triggering by rain infiltration. Water Resources Research, 36, 1897-1910.

Sánchez-Ruiz, S., Piles, M., Sánchez, N., Martínez-Fernández, J., Vall-llossera, M., and Camps, A., 2014. Combining SMOS with visible and near/shortwave/thermal infrared satellite data for high resolution soil moisture estimates. Journal of Hydrology, 516, 273-283.

Wang, J., Wiens, J., and Lundberg, S., 2021. Shapley Flow: A graph-based approach to interpreting model predictions. Proceedings of the 24th International Conference on Artificial Intelligence and Statistics, PMLR 130, 721-729.

Zhang, T., 2005. Influence of the seasonal snow cover on the ground thermal regime: An overview. Reviews of Geophysics, 43.

Zhao, L., Peng, J., Ma, P., et al., 2023. Microstructure response to shear strength deterioration in loess after freeze-thaw cycles. Engineering Geology, 323, 107229.