

Response to Reviewers' Comments

Dear Editor and Reviewers,

We sincerely thank the editor and reviewers for their careful evaluation and constructive comments. We have revised the manuscript accordingly, including the detailed annotations provided by Reviewer #1 in the PDF. Relevant text, figures, captions, and methodological descriptions were modified where needed, while some redundant or out-of-scope annotated text was removed to make the manuscript more concise and focused. Below, we provide a point-by-point response to the comments and describe the corresponding revisions.

Response to Reviewer #1:

General comment:

This is a really interesting paper reviewing threshold behaviors and runoff generation mechanisms for many catchments around the world and linking them to catchment properties. The analysis, and in particular the figures, are very insightful. However, the text is not always very clear and precise and the conclusions seem to go beyond the scope of the study. Below, I will give 3 examples, but there are more instances in the manuscript. These may seem like nuances, but they are important differences, especially in the abstract and conclusions:

Response:

We sincerely thank the reviewer for the careful, constructive, and insightful comments on our manuscript. We are encouraged that the reviewer finds the analysis and figures valuable. We fully agree that the previous version of the manuscript was not always sufficiently precise in separating (i) the observation of threshold-like runoff behavior, (ii) the interpretation of dominant runoff generation mechanisms in the original studies, and (iii) our higher-level connectivity-based synthesis. We also agree that several statements in the Abstract, Discussion, and Conclusions were too strong for the scope of a literature synthesis, and that the manuscript needed clearer positioning relative to previous review papers.

In response, we have substantially revised the manuscript by:

- (1) clarifying the distinction between observed threshold behavior and process/connectivity interpretation;
- (2) softening claims that previously implied quantitative performance comparisons not carried out in this review;
- (3) refining statements on threshold comparability and transferability across catchments;
- (4) expanding the Introduction and Discussion to position this study relative to previous review papers;
- (5) clarifying terminology and standardization rules for state variables and threshold indicators;
- (6) improving transparency regarding the literature search, statistical unit, and analysis workflow;
- (7) revising figures and captions for clarity; and

(8) shortening the Discussion and removing duplicated or repetitive text.

Comment 1:

There is a persistent mix-up in the text on the presence of a threshold (based on data) and the interpretation of either a hydrological process or connectivity. This distinction needs to be made much more carefully and clearly throughout the text. For some suggestions, see the annotated pdf.

Response 1:

We thank the reviewer for this important comment and agree that the previous version did not consistently distinguish between the observation of threshold-like runoff behavior and the interpretation of that behavior in terms of runoff processes or hydrologic connectivity. We have revised the manuscript throughout to make this distinction explicit.

In the revised manuscript, we now distinguish three levels of interpretation: (1) threshold-like runoff behavior, which refers to nonlinear or abrupt runoff responses reported in the source studies; (2) dominant runoff generation mechanisms, which refer to process interpretations made by the original authors based on hydrometric, soil-moisture, groundwater, tracer, or other field evidence; and (3) hydrologic connectivity transitions, which represent the conceptual synthesis developed in this review rather than a directly observed process in every case.

Following this distinction, we revised the Title/Key Points, Abstract, Methods, Results, Discussion, and Conclusions. In particular, we changed statements that previously implied that thresholds are directly equivalent to connectivity transitions into more cautious wording, such as “threshold-like responses can be interpreted within a connectivity-based framework” or “reported thresholds are often associated with changes in runoff mechanisms or flow-path connectivity.” We also clarified in the Methods that mechanism classifications were based on the interpretations of the original studies and that our connectivity framework is a synthesis of those reported observations and interpretations, not a reclassification of each dataset.

Specifically, we added a methodological clarification in Section 2.3, revised the Abstract and Conclusions to avoid treating connectivity transitions as directly observed in all cases, changed “direct observations of mechanism transitions” to “reported evidence for mechanism transitions” in Section 3.4, and revised the Discussion to present hydrologic connectivity as a conceptual synthesis framework rather than as the definition of threshold behavior itself.

Comment 2:

On L425 it is said that “A key finding is that rainfall alone is insufficient to explain nonlinear runoff responses”. And on L586 “Across climates and landscapes, antecedent state variables consistently provide stronger explanatory power for nonlinear runoff responses than rainfall-based descriptors”. While these statements may be true, this study didn’t test nor compare the explanatory power of rainfall and other factors in explaining the threshold response for different datasets. Thus, this is not a finding of this review and someone who only reads the abstract and conclusions may get the wrong impression of the manuscript and its findings based on the quoted sentences. A key finding of this

review is that few studies use rainfall alone to explain nonlinear runoff responses and that more studies used antecedent state variables. It is likely that more people report thresholds with antecedent wetness conditions because they perform better, but it may also be because they do this to compare their results with other published results (there are indeed cases where both rainfall and rainfall plus storage explain almost the same amount of variation in the data). Furthermore, the choice of a certain explanatory variable to describe the threshold may depend on the type of data that are available (only rainfall, soil moisture, or also groundwater levels) – which is not really discussed in the manuscript.

Response 2:

We thank the reviewer for this important comment and fully agree that the previous wording overstated what can be inferred from the present review. Our synthesis compiled the threshold indicators reported in the literature and their associations with runoff mechanisms, but it did not quantitatively test or compare the explanatory power of rainfall-based versus state-based indicators across datasets.

In response, we have revised the Abstract, Discussion, and Conclusions to replace statements implying explanatory superiority (e.g., “rainfall alone is insufficient” and “state variables provide stronger explanatory power”) with wording that more accurately reflects the scope of the review. We now state that antecedent-state and composite indicators were more frequently reported than rainfall-only metrics in the reviewed studies.

We also added text clarifying that this reporting pattern may reflect several factors in addition to hydrological relevance, including differences in data availability, measurement design, and the desire to maintain comparability with previous studies. Accordingly, the revised manuscript now presents these patterns as reporting tendencies and literature-based interpretations rather than as a formal ranking of indicator performance.

Comment 3:

L579 – 582 and elsewhere in the manuscript “Our findings demonstrate that runoff thresholds are not fixed hydrologic constants, but emergent properties resulting from the interactions between rainfall forcing, evolving catchment states, and relatively stable structural constraints”. I agree with the later part of the sentence but why would the threshold not be a constant or emergent property for a specific catchment? Again, the study didn’t test if it is constant or if thresholds have changed over time or seasonally, so these types of statements seem to be too broad and beyond the scope of the study. Of course, it is not a constant for all catchments, but no one would expect that.

Response 3:

We thank the reviewer for this important comment and agree that the original wording was too broad. Our review did not test whether runoff thresholds are temporally constant or seasonally variable within individual catchments. We therefore agree that it was not appropriate to state without qualification that runoff thresholds are “not fixed hydrologic constants.”

Our intended point was narrower: across catchments, the reviewed literature does not support a

single universal threshold metric or directly transferable numeric threshold value. We have revised the relevant statements in the Conclusions and Discussion accordingly. The revised text now emphasizes cross-catchment non-transferability rather than implying temporal non-stationarity within individual catchments. We also added clarifying text stating explicitly that the present review does not evaluate threshold stationarity through time within a given catchment.

Specifically, we replaced the phrase “runoff thresholds are not fixed hydrologic constants” with wording referring to the absence of a universally transferable threshold metric across catchments, and we added a scope statement clarifying that threshold constancy within individual catchments was not assessed in this review.

Comment 4:

Important review papers on threshold responses, like Ross et al (<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020WR027498>), Ali et al. (<https://onlinelibrary.wiley.com/doi/10.1002/hyp.10527>) and McDonnell et al (<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020WR027514>) are not mentioned or discussed and there is no discussion on how this study complements these previous studies in the introduction, nor any comparison of similar findings or mention of discrepancies in the discussion. Similarly, the recent review papers highlighting the variation in runoff processes, such as Penna (<https://www.nature.com/articles/s44221-025-00547-z>) and McMillan et al (<https://www.nature.com/articles/s44221-025-00407-w>) and <https://onlinelibrary.wiley.com/doi/full/10.1002/hyp.14537>) should be highlighted and discussed. That other reviews are available doesn't mean that this work is not useful. Instead, it should be made clear how this review complements the previous reviews (i.e., the global scale and the linkage to runoff generation processes and catchment characteristics in this review) and where findings are different.

Response 4:

We thank the reviewer for this important comment and fully agree that the previous version of the manuscript did not sufficiently position the study relative to earlier review and synthesis papers. Although some of these papers were cited in passing, we did not clearly explain in the Introduction how the present review complements them, nor did we compare our findings with them in the Discussion.

In response, we revised the manuscript in two targeted ways. First, we added a concise positioning paragraph in the Introduction that now explicitly discusses Ross et al. (2021), Ali et al. (2015), and McDonnell et al. (2021) as key threshold- and connectivity-related antecedents, and also situates our study relative to the broader global syntheses of McMillan et al. (2025), McMillan (2022), and Penna (2026). Second, to avoid overextending the Discussion, we did not create a new subsection; instead, we added one short comparative paragraph at the end of Section 4.2 clarifying where our synthesis is consistent with these earlier studies and what the present review adds.

Specifically, we now clarify that Ross et al. (2021) and Ali et al. (2015) focused on threshold detection and intersite comparison, McDonnell et al. (2021) provided a conceptual fill-and-spill

framework for grouping runoff-process observations, McMillan et al. (2025) synthesized dominant hydrologic process patterns and their controls globally, McMillan (2022) addressed process taxonomy and terminology, and Penna (2026) synthesized global controls on runoff processes in forested catchments. We then state more clearly that the distinct contribution of the present review is to connect threshold indicators with reported runoff generation mechanisms, wetness-dependent mechanism transitions, and recurring soil–geology associations across diverse experimental catchments.

Comment 5:

Finally, it is unclear what the difference is between antecedent soil moisture condition, antecedent soil moisture, and soil water content. Are these not all referring to the same thing? And what is the difference between coupled soil–groundwater, storage state, and storage? It will be very useful to define these terms more clearly in the methods sections where the annotations are also described. Even if there are small differences in these terms perhaps, they can still be grouped so that Figure 2 contains fewer terms?

Response 5:

We thank the reviewer for pointing out that several terms used in Fig. 2 were overlapping and insufficiently defined. We agree that presenting antecedent moisture condition, antecedent soil water content, and soil water content as separate categories could be confusing. We also agree that the distinction among coupled soil–groundwater state, storage state, and storage needed to be clarified.

In the revised manuscript, we clarified the distinction between the middle and right columns of Fig. 2. The middle column now represents harmonized hydrological state domains, whereas the right column represents the specific threshold indicators reported in the source studies. To reduce unnecessary complexity, we grouped antecedent soil water content and soil water content under “soil water content,” because both refer to explicitly reported soil water conditions. We retained “antecedent moisture condition” as a broader category for pre-event wetness descriptors, including antecedent precipitation, antecedent precipitation index (API), antecedent baseflow, or qualitative wet/dry catchment-state descriptions.

We also revised the storage-related terminology. “Coupled soil–groundwater” and “storage state” were combined into a broader category, “integrated storage state,” which includes catchment storage, storage deficit, relative storage, and explicitly coupled soil–moisture–groundwater conditions. In contrast, storage-related terms in the indicator column refer to the specific quantitative storage metric or storage threshold used to identify nonlinear runoff responses. We revised the Methods section and Fig. 2 caption accordingly to make these distinctions clear and to reduce the number of overlapping terms in the figure.

Specific comments:

Comment 6:

Methods and results: It is not clear if percentages are given per catchment or per study. The search

terms probably led to multiple studies for some catchments and some studies that report results for multiple catchments. How were these treated in the statistics?

Response 6:

We thank the reviewer for identifying this ambiguity. We agree that the previous version did not clearly state whether percentages were calculated per publication or per catchment, nor how repeated catchments or multi-catchment studies were handled.

In the revised manuscript, we now clarify that the primary unit of synthesis is the catchment, not the publication. Unless otherwise stated, all percentages in the Results are calculated relative to the number of catchments included in the corresponding analysis. When one publication reported multiple catchments, each catchment was treated as a separate record because threshold indicators, runoff mechanisms, and catchment attributes could differ among catchments. Conversely, when multiple publications reported results for the same catchment, the information was consolidated into a single catchment record to avoid double counting. Additional publications for the same catchment were used only to supplement information on threshold indicators, state variables, or soil–geology characteristics, and did not increase the statistical weight of that catchment.

We have added this clarification to the Methods section and revised the Results and figure captions to state the relevant denominators explicitly. In particular, we replaced ambiguous wording such as “studies” with “catchments” where the statistics are catchment-based, and clarified that percentages in Figures 3–5 are calculated using catchments as the unit of analysis. For the soil–geology networks, we now specify that link thickness represents the number of independent catchments, not the number of publications.

Comment 7:

Methods L140: how many "hits" did your search terms get? Did you screen all of them? Or is the 200 a selection? To what time period did you limit your search?

Response 7:

We thank the reviewer for pointing out that the literature-search workflow was not sufficiently clear. We have revised Section 2.1 to report the number of records returned by each database, the screening procedure, and the time period covered by the search.

In the revised manuscript, we now state that the database search covered publications from database inception to 2024 and was conducted using Web of Science Core Collection and Scopus. The search returned 745 records in total, including 344 records from Web of Science and 401 records from Scopus. After removing 292 duplicate records, 453 unique records remained. All 453 unique records were screened by title and abstract, followed by targeted full-text assessment when title and abstract information was insufficient.

Based on this screening, 153 studies met the inclusion criteria for extracting catchment-scale threshold and runoff-generation information. These studies provided information for 176 experimental catchments. After excluding 39 model-only catchments without direct field observations, the final synthesis dataset contained 138 experimental catchments. We have added

these details to Section 2.1 to make the search and screening process more transparent and reproducible.

Comment 8:

Methods: Information on what software and packages were used to create the different diagrams and do the analyses needs to be added.

Response 8:

Thank you for this helpful suggestion. We agree that the previous Methods section did not provide sufficient information on the software and packages used for data processing, analysis, and figure generation.

In the revised manuscript, we added a new subsection entitled “Software and reproducibility” to the Methods section. This subsection now specifies the software environment and packages used for database processing, frequency calculations, transition-matrix construction, permutation testing, Sankey diagrams, bar charts, heatmaps, co-occurrence networks, mapping, and the conceptual figure.

Comment 9:

Figure 2: This is a really nice figure but spell out the acronyms in the caption. Also it would be useful to include a table (or figure) with the % of studies that used what metric (and possibly then split this also by runoff mechanism).

Response 9:

Thank you. We have revised the Figure 2 caption to spell out all acronyms and improve readability.

Comment 10:

Section 3.3 and elsewhere – it is important to clearly state that studies that found a certain dominant runoff generation mechanism mainly took place in a certain climate - not that the runoff mechanism does not also have to be dominant in that climate. This is important as we know that research catchments are not representative of the entire land surface (as shown in Figure 4b by the dominance of forested catchments and an exclusion of catchments in croplands or (sub)urban areas, even though these cover a large part of the global land surface). This may seem like a nuance but is in my opinion an important one. See specific suggestions in the annotated pdf.

Response 10:

We thank the reviewer for this important clarification and agree that the previous wording could be interpreted as implying that certain runoff mechanisms are globally restricted to, or necessarily dominant in, specific climates or vegetation classes. This was not our intention. Our analysis describes the distribution of reported mechanisms within the reviewed experimental-catchment

literature, not the global prevalence of runoff mechanisms across the land surface.

In response, we revised Section 3.3 and related figure captions to consistently use wording such as “among the reviewed catchments,” “within the compiled dataset,” and “were more frequently reported,” rather than wording that could imply global occurrence or absence. We also revised statements such as “SSF-dominated catchments are primarily found in humid climates” and “GWF-dominated catchments are restricted to humid and semi-humid climates” to clarify that these patterns refer only to the catchments included in our review.

We further added a methodological clarification that aridity and vegetation classes are used to describe the composition of the reviewed literature and to support cautious interpretation, not to infer the global distribution of runoff mechanisms. In the Discussion, we added a short limitation statement noting that humid and forested experimental catchments are overrepresented, whereas cropland, dryland, urban, and suburban systems are underrepresented or absent. We believe these revisions better reflect the sample-based nature of the synthesis and avoid overgeneralizing beyond the reviewed dataset.

Comment 11:

Section 3.3 (and 3.5) is interesting and very nice, but these results are not discussed in the discussion at all. How do these results compare to the recent studies of McMillan et al and Penna?

Response 11:

We thank the reviewer for this helpful comment. We agree that the results in Sections 3.3 and 3.5 were under-discussed in the previous version, particularly in relation to recent global syntheses by McMillan et al. and Penna.

In the revised manuscript, we added a concise comparative paragraph to the Discussion, while avoiding a long new subsection in response to the reviewer’s broader concern that the Discussion was already too long. This new paragraph explicitly compares our aridity–vegetation patterns and soil–geology fingerprints with the findings of McMillan et al. (2025) and Penna (2026). We now clarify that our results are broadly consistent with these studies in showing that runoff processes are shaped by environmental context, including climate, soils, vegetation, geology, geomorphology, and hydrological state.

At the same time, we emphasize that the present review differs in focus. McMillan et al. synthesized global patterns in observed hydrologic processes, and Penna synthesized controls on runoff processes in forested catchments. Our study adds a threshold-focused perspective by linking reported threshold indicators with dominant runoff mechanisms, wetness-dependent mechanism transitions, and recurring soil–geology associations across experimental catchments. We also clarified that the aridity–vegetation patterns and soil–geology fingerprints in our review should be interpreted as patterns within the reviewed threshold literature, not as global estimates of runoff-process prevalence.

Comment 12:

Figure 8: I understand the reason for the inclusion, but shallow groundwater is not a geological structure.

Response 12:

We agree with the reviewer. Shallow groundwater is more appropriately described as a hydrogeological condition rather than a geological structure. In the revised manuscript, we changed the label of Fig. 8d from “Geological structure” to “Geological and hydrogeological features.” We also revised the figure caption to clarify that descriptors such as faults, fractured or weathered bedrock, impermeable layers, and rock outcrops are treated as geological or structural features, whereas shallow groundwater level is treated as a hydrogeological condition.

We further revised the corresponding text in Section 3.5 to state that GWF is associated with shallow groundwater levels as a hydrogeological condition and with structurally conductive geological settings such as faults and fractured formations.

Comment 13:

L456: What is meant with “perform”? This is not a model study where one can compare simulations and observations, nor are the “performance” of different metrics in explaining the thresholds tested. Instead, the review shows that some metrics are less frequently used (possibly because they describe the data less well, or because the data were not available, or simply because the use of other metrics makes comparisons with the literature easier).

Response 13:

We thank the reviewer for this clarification and agree that the word “perform” was inappropriate in the context of this review. We did not quantitatively compare the explanatory or predictive performance of different threshold indicators across datasets, nor did we conduct a model-performance evaluation.

We have therefore revised this sentence and related wording in Section 4.1. Instead of stating that state-based indicators “perform less consistently” in infiltration-excess-dominated systems, we now state that these indicators were less frequently reported in catchments interpreted as infiltration-excess-dominated, whereas rainfall intensity was more commonly emphasized as the immediate triggering factor. We also added clarification that this pattern should be interpreted as a reporting tendency in the reviewed literature, not as evidence of lower explanatory performance.

In addition, we now explicitly acknowledge that the choice of threshold descriptor may reflect data availability, monitoring design, and comparability with previous studies, as well as hydrological relevance.

Comment 14:

L469-472: There are some non-intuitive results here that need to be explained or discussed better. Why would coarse material (which generally has a high conductivity) lead to HOF? Why would a limited storage capacity lead to HOF, rather than SOF? Why would low conductivity soils lead to

SOF (via near surface saturation) but not also to HOF due to the low conductivity? This requires some discussion (and perhaps even some speculation).

Response 14:

We thank the reviewer for this insightful comment. We agree that the original explanation was too deterministic and that some of the reported associations in Fig. 8 require more careful interpretation. The soil–geology networks identify recurrent co-occurrence patterns in the reviewed literature, but they do not establish one-to-one causal relationships between individual descriptors and runoff mechanisms.

We have therefore revised Section 4.2 to clarify the context dependence of these associations. For HOF, we now state that the association with rock outcrops, shallow soils, gravel, and sandy loam should not be interpreted as evidence that coarse materials inherently have low hydraulic conductivity. Instead, shallow soils and rock outcrops may indicate thin or discontinuous soil cover, limited effective near-surface storage, and exposed or poorly infiltrating surfaces, whereas gravelly or sandy loam materials may only promote HOF under specific surface conditions or high rainfall intensities. We also note that the occurrence of deep soils among some HOF cases indicates that HOF cannot be explained solely by limited storage capacity.

We also revised the discussion of SOF. We now clarify that fine-textured soils, peat, low-permeability layers, and shallow groundwater can favor SOF when they promote perched saturation, near-surface water accumulation, or expansion of connected saturated areas. However, low hydraulic conductivity can also contribute to HOF when it limits surface infiltration. Thus, the mechanism associated with low permeability depends on where the permeability limitation occurs in the soil profile, the antecedent wetness state, topographic position, and rainfall intensity.

Comment 15:

Figure 9: I like the figure and think that it could be very useful for discussions and teaching. However, I don't understand the y-axis and it is not explained in the text. Why is connectivity considered to be higher for groundwater flow than for subsurface stormflow and why would it be higher for SSF than OF? What is this based on and what is the evidence for that? And what is the arrow for soil depth supposed to reflect? Explain better in the caption.

Response 15:

We thank the reviewer for this helpful comment and are pleased that the figure was considered potentially useful for discussion and teaching. We agree that the previous version did not explain the y-axis sufficiently and could be interpreted as implying a universal ranking of overall hydrologic connectivity among runoff mechanisms. This was not our intention.

In the revised manuscript, we clarified that the y-axis refers specifically to the conceptual vertical integration or depth extent of activated hillslope–channel flow paths contributing to event runoff, rather than to total hydrologic connectivity in all directions. Under this definition, HOF is placed lower because it mainly involves surface or near-surface pathways; SOF partly overlaps with HOF but extends higher because it commonly involves near-surface saturation or connected saturated

areas; SSF is placed higher because it represents activation of subsurface pathways such as perched flow, macropore flow, or flow along restrictive layers or soil–bedrock interfaces; and GWF is placed highest only for the groundwater-dominated cases considered in this review, where shallow groundwater rise or shallow water-table connectivity contributes to event runoff and hillslope–riparian–channel coupling. We also explicitly state that this ordering is conceptual and should not be interpreted as a universal ranking of overall hydrologic connectivity, because overland flow can be highly connected laterally in some landscapes.

We also revised the explanation of the soil-depth arrow. The revised caption now clarifies that soil depth and effective storage can have a non-monotonic influence. Very shallow or discontinuous soils and rock outcrops may favor HOF because of limited effective storage, whereas intermediate to deeper profiles may support SOF, SSF, or GWF when saturation, restrictive layers, or shallow groundwater promote connected flow paths. Conversely, extremely deep and dry profiles may again favor HOF under extreme rainfall because subsurface or groundwater connectivity thresholds are difficult to reach. We revised the figure caption and added a short explanatory paragraph in Section 4.3 to make these assumptions explicit.

Comment 16:

Discussion: The discussion is long and at times extends beyond the topic and findings of the review and analyses. Try to shorten it and to stay closer to the core results of the manuscript.

Response 16:

We thank the reviewer for this helpful comment. We agree that the previous Discussion was too long and, in several places, extended beyond the direct findings of the review. In the revised manuscript, we substantially shortened and refocused the Discussion to keep the interpretation closer to the core results of the synthesis.

Specifically, we removed repeated explanations of rainfall versus antecedent-state controls, shortened broad conceptual discussion on hydrologic connectivity, and deleted or condensed statements that were not directly supported by the analyses. We also reduced overlap between the Discussion and figure captions, especially for the conceptual framework and co-occurrence-network interpretation.

At the same time, we retained concise discussion needed to address key reviewer concerns, including: (i) the distinction between reported threshold-like behavior and process/connectivity interpretation; (ii) the difference between cross-catchment environmental patterns and within-catchment event-scale mechanism transitions; (iii) the interpretation of soil–geology fingerprints as reported co-occurrence patterns rather than deterministic controls; and (iv) the role of structural and functional connectivity as a conceptual framework rather than a directly measured index.

The revised Discussion now focuses on three main points: how threshold indicators are reported across mechanisms, how reported event-scale mechanism transitions are interpreted under increasing wetness, and how recurring soil–geology–hydrogeology settings provide structural context for runoff generation. We believe these revisions make the Discussion shorter, more focused, and more closely aligned with the scope and evidence of the review.

Comment 17:

General comments on the text: Several blocks of text are given at two locations in the manuscript (i.e., copied at two places). Please remove the duplicates (see annotated pdf). Other parts of text are not necessary as they should be part of the captions and can be removed (see annotated pdf for suggestions).

Response 17:

We thank the reviewer for pointing this out. We agree that the previous manuscript contained duplicated text and that some figure-explanation material was unnecessarily repeated in the main text rather than being confined to figure captions.

In the revised manuscript, we carefully checked the text against the annotated PDF and removed duplicated paragraphs. In particular, we removed the repeated Introduction block describing the study gaps and objectives. We also streamlined the Results sections by moving figure-reading details, such as node size, flow-band width, self-links, link thickness, and node meanings, to the figure captions where appropriate. The main text now focuses more directly on the key results and their interpretation, while the captions provide the necessary information for reading the figures.

We also shortened several repetitive parts of the Discussion and removed wording that repeated the conceptual contribution without adding new interpretation. These changes improve readability and help keep the manuscript closer to the core results of the review and analyses.

Response to Reviewer #2:**General comments:**

This manuscript presents an ambitious synthesis of runoff threshold behavior across diverse catchments, offering a connectivity-based framework for understanding nonlinear rainfall-runoff responses. The authors have compiled an impressive dataset of 138 experimental catchments and conducted systematic analyses of threshold indicators, dominant runoff mechanisms, transition pathways, and soil-geology fingerprints. The central finding as stated by the authors—that runoff thresholds represent state-dependent hydrologic connectivity transitions rather than fixed rainfall exceedances—is conceptually valuable and advances the ongoing discourse on runoff generation processes.

However, I think the key finding of the paper is its clear demonstration that as catchments become wetter, the runoff generation zone progressively shifts from the soil surface or near-surface layers into the weathered bedrock layer. This indicates that, despite the dominant role of tectonic setting and lithology, climate still controls weathering and pedogenesis, thereby strongly influencing runoff generation. However, the paper still needs to further deepen our understanding of how climate, vegetation, soil, and rock weathering control runoff generation mechanisms across different catchments.

Therefore, the manuscript requires substantial revision to address several critical issues. The conceptual distinction between hydrologic connectivity across catchments (structural vs. functional

connectivity) needs clearer articulation. The classification of dominant runoff mechanisms lacks operational clarity, and the methods section (Section 2) is insufficiently detailed for reproducibility. The interpretation of mechanism "transitions" within catchments versus across climatic gradients fundamentally different processes, requiring separate treatment. Or I suggest that they only focus on the research of mechanism transitions across catchments. Additionally, the section of Data and Methods need a more detailed elaboration, and the presentation of results requires more transparent explanation of metrics and analytical procedures. The specific comments are presented following to guide the authors in strengthening the manuscript's conceptual framework, methodological rigor, and interpretive clarity.

Response 1:

We thank the reviewer for the thoughtful and constructive assessment of the manuscript and for recognizing the value of the compiled dataset and the connectivity-based synthesis. We agree that the previous version needed clearer conceptual framing, more transparent methodological description, and more cautious interpretation of mechanism transitions and cross-catchment patterns.

We have revised the manuscript accordingly. First, we clarified the distinction between structural and functional connectivity in the conceptual framing of the manuscript. Structural connectivity is now described as the relatively persistent catchment architecture that constrains potential flow pathways, including soil depth, soil texture, permeability contrasts, lithology, weathered or fractured bedrock, topography, and drainage organization. Functional connectivity refers to the event-scale activation of these pathways as rainfall interacts with antecedent wetness, storage, soil moisture, and groundwater level. This clarification helps separate longer-term catchment controls from event-scale state-dependent runoff activation.

Second, we improved the operational wording for dominant runoff mechanism classification. We now spell out Hortonian overland flow (HOF), saturation-excess overland flow (SOF), subsurface stormflow (SSF), and groundwater flow (GWF) at first use, and clarify that the mechanism classes were extracted from the process interpretations reported in the original studies, rather than inferred solely from climate, vegetation, soil, or geological descriptors. We also clarified that mixed categories were assigned only when the source studies explicitly reported co-dominant mechanisms or event-dependent switching, and that GWF refers here to event-scale shallow groundwater responses within weathered bedrock or fractured regolith, not to regional groundwater baseflow.

Third, we revised the text to distinguish within-catchment mechanism transitions from across-catchment environmental patterns. The transition analysis is now described as applying only to the subset of catchments where the original studies explicitly reported changes in dominant runoff mechanisms during events or across event sequences under increasing wetness. In contrast, differences among climate, vegetation, and soil–geology settings are now described as cross-catchment distributions or associations in the reviewed dataset, not as temporal transitions. This revision avoids conflating event-scale mechanism reorganization within catchments with differences among catchments along climatic or physiographic gradients.

Fourth, we refined the interpretation of the wetness-related transition patterns. We appreciate the reviewer's observation that wetter conditions may be associated with progressively deeper or more integrated runoff pathways, including pathways in weathered bedrock in some catchments. However, we revised the manuscript to avoid implying that all catchments follow a universal vertical

progression from surface runoff into weathered bedrock. Instead, we now describe the evidence more cautiously as a reported tendency, within the transition subset, for surface-dominated responses to reorganize toward saturation-, subsurface-, or shallow-groundwater-influenced pathways as wetness increases.

Fifth, we strengthened the discussion of climate, vegetation, soil development, and rock weathering as coupled controls on runoff generation. We now explain that climate influences long-term vegetation, weathering, pedogenesis, soil depth, and storage architecture, which together shape structural connectivity, while event rainfall and antecedent wetness regulate whether these potential pathways become functionally connected during storms. This revision better links the threshold and mechanism-transition results to broader controls on runoff generation across catchments.

Finally, we improved the methodological transparency of Section 2. We added details on the literature search and screening workflow, the unit of analysis, treatment of repeated catchments and multi-catchment studies, threshold-indicator standardization, mechanism classification, transition-matrix construction, soil–geology descriptor extraction, software used for analysis and visualization, and statistical procedures. Together, these revisions clarify the conceptual basis, improve reproducibility, and ensure that the interpretation remains closely tied to the evidence provided by the reviewed studies.

Specific comments:

Comment 2:

Lines 13-14: What is the recurrent soil geology combination? Does it refer to the repeated occurrence of similar soil geological and stratigraphic structures? From the following text, the author naturally believes that key band structures vary in different regions, thus forming structural fingerprints that regulate connectivity development. However, though readers are aware that they are different, and we want to know what characteristics the structures in different regions have, which can lead to different mechanisms of production flow connectivity.

Response 2:

We thank the reviewer for pointing out that the phrase “recurrent soil–geology combinations” was insufficiently clear in the Key Points. We agree that the original wording was too abstract and could be interpreted as implying deterministic control.

In the revised manuscript, we replaced this Key Point with more specific and cautious wording. We now refer to “recurring soil–geology associations” rather than “transferable structural fingerprints that govern threshold behavior.”

We further clarified in Section 3.5 that “structural fingerprints” refer to recurrent co-occurrence patterns between reported runoff mechanisms and standardized soil, geological, and hydrogeological descriptors, not to deterministic soil–geology controls. In the Discussion, we now explain how different structural settings may condition connectivity development: shallow rocky or discontinuous soils may favor rapid surface runoff, fine-textured or low-permeability soils and shallow groundwater may favor saturation-excess runoff, and permeability contrasts, preferential flow pathways, fractured bedrock, or weathered regolith may facilitate subsurface or shallow-groundwater connectivity.

Comment 3:

Lines 33-40: The hydrological connectivity here includes both structural and functional connectivity. I suggest the author to delve deeper into the discussion of structural connectivity, i.e., how structural fingerprints of soil–geology combinations affect state-dependent connectivity which is functional connectivity from its literal meaning.

Response 3:

We thank the reviewer for this helpful suggestion. We agree that the previous wording did not sufficiently distinguish structural connectivity from functional connectivity, especially in the Abstract. We have revised the relevant text to clarify that the soil–geology fingerprints represent structural connectivity, that is, the relatively persistent potential flow-path architecture created by soil depth, texture, permeability contrasts, lithology, weathered or fractured bedrock, and hydrogeological setting.

We also now clarify that functional connectivity refers to the event-scale activation of these potential pathways as rainfall interacts with antecedent wetness, storage, soil moisture, and groundwater level. In the revised Abstract and Discussion, we state that structural fingerprints condition state-dependent functional connectivity by influencing where water is stored, where saturation or preferential flow can develop, and which pathways become activated during storms. We also revised the wording to avoid implying that soil–geology structures alone determine runoff mechanisms; instead, they condition how evolving catchment states are translated into functional connectivity.

Comment 4:

Lines 49-51: Traditionally, we also do not consider production flow to be linear, and the expression is not very rigorous.

Response 4:

We thank the reviewer for this clarification and agree that the original wording was not sufficiently rigorous. We did not intend to imply that traditional runoff-generation concepts assume linear behavior. We have therefore revised this sentence to avoid the phrase “challenges traditional interpretations” and to clarify that threshold-like runoff responses have often been interpreted as the activation of previously inactive flow pathways or hydrologic connectivity under sufficiently wet catchment states.

Comment 5:

Lines 62-63: In fact, hydrologic connectivity framework is proposed for explaining why similar storms generate different runoff responses due to different catchment state of connectivity. I do not think it is a good question in context. Rewrite this paragraph.

Response 5:

We thank the reviewer for this helpful clarification. We agree that the previous wording was not appropriate because hydrologic connectivity frameworks already provide a conceptual explanation

for why similar storms can produce different runoff responses depending on catchment wetness and connectivity state. We have therefore removed the original question and rewritten the paragraph.

Comment 6:

Lines 109-110: it is quite similar to above question. The readers may be more concerned about how much humidity increases in different climate basins can lead to abrupt changes in connectivity, resulting in so-called mechanism transitions. In addition, the language expression is not direct and difficult to understand. It is recommended to use simple language to express it.

Response 6:

We thank the reviewer for this helpful suggestion. We agree that the original third research question was too similar to the preceding question and that the wording was unnecessarily abstract. In particular, the phrase “progressively strengthening hydrologic connectivity” could be difficult to interpret and might imply a stronger or more quantitative assessment of connectivity than was conducted in this review.

We have therefore rewritten the research questions in simpler and more direct language. Specifically, we merged the previous questions on mechanism transitions and connectivity development into a single question that asks which runoff mechanisms are reported to become active as catchments become wetter, and what wetness, storage, soil-moisture, or groundwater conditions are linked to abrupt changes in runoff connectivity. This revision better reflects the scope of the review and focuses on the reported state conditions associated with mechanism shifts, rather than implying a universal or quantitatively measured progression in connectivity.

Comment 7:

Lines 104-112: In my opinion, none of these four questions are very clear and can be said to be expected. For example, for the second question, the answer is in a sense affirmative, but the transition of the runoff generation mechanism should also be different. In extremely arid and sparsely vegetated loess watersheds, the range of changes in runoff mechanisms may be limited despite increased humidity. In addition, the author uses the concept of mechanism transitions here, and I believe they should elaborate on what mechanism transitions are in the introduction? Traditionally, we believe that a location, such as a slope, has a fixed mechanism for runoff generation. If the slope does not produce runoff or is in the early stages of runoff generation, it is only due to differences in its internal hydrological connectivity that it is not suitable for runoff mechanism transitions. Personally, I think this is the case.

Response 7:

We thank the reviewer for this helpful comment. We agree that the original research questions were too abstract and that the term “mechanism transitions” required clearer definition. We also agree that increased wetness does not necessarily lead to the same type or range of runoff-mechanism change in all catchments.

“Mechanism shifts.” refers to reported changes in the dominant active runoff pathway during events or across event sequences as catchment wetness increases. It does not imply that the structural

runoff-generation potential of a hillslope or catchment changes. Rather, it describes the wetness-dependent activation, deactivation, or reorganization of flow pathways as interpreted in the source studies. In the revised Introduction, we therefore rewrote the research questions in simpler and more direct language.

Comment 8:

Lines 145-146: Studies focusing solely on linear rainfall-runoff relationships....., without event-scale threshold interpretation, were excluded. Why is there a linear relationship between precipitation and runoff in some watersheds, and shouldn't it all be nonlinear?

Response 8:

We thank the reviewer for pointing out this ambiguity. We agree that the original wording could be misread as implying that runoff generation in some catchments is inherently linear. This was not our intention.

In the revised manuscript, we clarified that “linear rainfall–runoff relationships” refers to the analytical treatment or reported interpretation in the source study, not to the intrinsic hydrological behavior of the catchment. Some studies analyzed rainfall–runoff response only as an approximately linear input–output relationship or as a long-term water-balance problem, without identifying event-scale thresholds, regime shifts, or nonlinear runoff responses. Because the present review specifically focuses on reported event-scale threshold behavior and its interpretation in relation to runoff generation mechanisms, such studies did not provide the information needed for our synthesis and were therefore excluded.

We revised the Methods text accordingly to avoid implying that runoff generation processes are assumed to be linear in those catchments.

Comment 9:

Lines 154-186: Though, according to the authors definition, i.e., Classification reflects the mechanism that primarily controls stormflow during events exhibiting nonlinear or threshold behavior, however, for me, this classification definition is still very vague, not direct and clear enough. Hence, I strongly suggest that the authors present more clear definition and classification standard. Moreover, a conceptual diagram maybe needed to describe how HOF, SOF, SSF, GWF dominate or combine to shape stormflow of a headwater area.

Response 9:

We thank the reviewer for this helpful comment and agree that the previous description of the runoff-mechanism classification was not sufficiently operational. Although HOF, SOF, SSF, and GWF are standard terms in catchment hydrology, the manuscript needed to explain more clearly how these labels were harmonized across source studies and how combined mechanisms were treated.

In the revised manuscript, we rewrote Section 2.2 to provide clearer classification criteria. We now define HOF as source-study interpretations of infiltration-excess or Hortonian overland flow; SOF as saturation-excess overland flow, return flow, or runoff from saturated or near-saturated source

areas; SSF as event-scale lateral subsurface stormflow through soil, regolith, preferential pathways, perched zones, or restrictive interfaces such as the soil–bedrock boundary; and GWF as event-scale shallow groundwater responses within weathered bedrock or fractured regolith contributing to stormflow, not regional groundwater baseflow.

We also clarified that mechanism classes were extracted from the process interpretations reported in the original studies and were not inferred solely from climate, vegetation, soil, or geological descriptors. Dominance now refers to the mechanism interpreted as the primary contributor to stormflow during threshold-like events, but does not imply exclusive operation of a single process. Mixed categories were assigned only when source studies explicitly reported co-dominant mechanisms or event-dependent switching, whereas ordered wetness-dependent mechanism shifts were treated separately in the transition analysis.

To further improve clarity, we also added a simple conceptual schematic in the Supplementary Information illustrating how HOF, SOF, SSF, and GWF may dominate or combine to shape stormflow in a headwater catchment.

Comment 10:

Lines 218-245: What are the sensitivity tests (Line 243).

Response 10:

We thank the reviewer for noting that the sensitivity procedure was not clearly described. We revised the wording from the vague phrase “sensitivity tests” to “threshold-sensitivity checks” and now explicitly state that we reconstructed the co-occurrence networks using stricter minimum co-occurrence thresholds of three, four, and five independent catchments. These checks were intended to evaluate whether the main mechanism–descriptor associations depended on the chosen threshold for retaining links. The stricter thresholds reduced the number of retained links, as expected, but the main mechanism-specific patterns remained qualitatively similar.

Comment 11:

Section 2 Data and Methods is not detailed, and many methods were not provided. For instance, how do you obtain the result of figure 2 and the similar structural clusters (mentioned in section 2.5)? Moreover, the meaning of indicators (in figure 2), e.g., Pi and API, should be provided.

Response 11:

We thank the reviewer for pointing out that the Data and Methods section did not provide sufficient detail for reproducing several analyses. We agree that the procedures used to generate Fig. 2 and the threshold-sensitivity checks in Section 2.5 needed clearer explanation, and that the abbreviations used for threshold indicators should be defined.

In the revised Methods, we added a more explicit description of how Fig. 2 was generated. We now state that each catchment record was coded for the reported dominant runoff mechanism, the relevant state or antecedent-condition descriptor, and the threshold indicator used to identify or describe the nonlinear runoff response. We also clarify how reported indicators were standardized

into common categories, including event precipitation amount (P), precipitation intensity (Pi), antecedent precipitation index (API), soil water content or soil moisture (SWC), groundwater level or water-table depth (GWL), storage-related indicators, and composite rainfall–state indicators. We revised the Fig. 2 caption and added a supplementary table defining these abbreviations.

We also revised Section 2.5 to clarify the sensitivity procedure for the soil–geology–hydrogeology co-occurrence networks. Instead of the vague phrase “sensitivity tests,” we now describe these as threshold-sensitivity checks. Specifically, the main network retained mechanism–descriptor links reported in at least two independent catchments, and we reconstructed the networks using stricter minimum co-occurrence thresholds of three, four, and five independent catchments. These stricter thresholds reduced the number of retained links but preserved the main mechanism-specific associations qualitatively. We added this explanation to the Methods and summarized the sensitivity results in the Supplementary Information.

Comment 12:

Lines 269-270: how many studies without specified threshold indicator?

Response 12:

Thank you for pointing this out. We agree that the caption should state the number of cases included in the “None reported” category. In the revised Fig. 2 caption, we now clarify that “None reported” refers to reviewed catchments for which the source studies described nonlinear or threshold-like runoff responses but did not specify the corresponding state variable or threshold indicator. Specifically, 29 catchments did not report an explicit state variable, and 41 catchments did not report an explicit threshold indicator. We also revised the wording from “studies” to “catchments” to be consistent with the unit of analysis used in the Results.

Comment 13:

Section 3.2 Frequency and proportional contribution of dominant runoff generation mechanisms is not important, and not recommended as a separate section, suggested to be included in section 3.1.

Response 13:

We thank the reviewer for this suggestion and agree that the frequency and proportional contribution of dominant runoff mechanisms did not need to be presented as a separate Results subsection. In the revised manuscript, we merged the previous Section 3.2 into Section 3.1 so that the mechanism-frequency information now serves as background for interpreting the reported state variables and threshold indicators.

The revised Section 3.1 is now titled “Frequency of reported mechanisms and associated threshold indicators.” It first summarizes the frequency and proportional contribution of reported dominant mechanisms across the 138 reviewed catchments, and then presents the state variables and threshold indicators associated with those mechanism classes. The subsequent Results sections have been renumbered accordingly.

Comment 14:

Line 302: how do you classify four aridity classes, i.e., humid, semi-humid, semi-arid, and arid? As it is not a standard pattern in Köppen climate classification, Provide the standards in section 2.

Response 14:

We thank the reviewer for noting that the four classes were not sufficiently defined. We agree that humid, semi-humid, semi-arid, and arid should not be presented as Köppen climate classes. In the revised manuscript, we now describe them as operational climatic wetness categories used to summarize the reviewed catchment sample.

We revised Section 2.4 to provide the classification criteria. Catchments were assigned based on the climate description reported in the source study, supported by mean annual precipitation and Köppen climate information where available. We now clarify that catchments described as humid, perhumid, tropical humid, or temperate humid were classified as humid; sub-humid or semi-humid as semi-humid; semi-arid or dry steppe as semi-arid; and arid or desert as arid.

We revised Section 3.3 and the figure caption to avoid implying that these categories are standard Köppen classes.

Comment 15:

Figure 5 presents well that as the climate in the watershed becomes more humid, the runoff yielding interfaces gradually penetrate from the soil surface or surface layer to the weathered bedrock layer. This indicates that despite the dominant role of structure and rock lithology, climate still controls weathering and soil formation, thereby strongly affecting runoff. I suggest the author strengthen the discourse on this important discovery.

Response 15:

We thank the reviewer for this insightful interpretation of Fig. 5. We agree that the climatic composition of the reported mechanisms deserves stronger discussion, particularly regarding how climate may influence vegetation, weathering, soil development, regolith structure, and shallow groundwater storage, thereby affecting runoff-generation pathways.

In the revised manuscript, we strengthened the discussion of this result. We now explain that Fig. 5 suggests a climate-related shift in the reported depth domain of active runoff pathways: surface-dominated HOF is relatively more represented in drier subsets, whereas SSF, SOF–SSF, and GWF are more strongly represented in humid or semi-humid catchments, where deeper soil wetting, regolith storage, shallow groundwater rise, or weathered-bedrock connectivity are more commonly reported.

At the same time, we revised the language cautiously to avoid over-interpreting Fig. 5 as proving a universal climatic sequence. We now state that this pattern is consistent with the idea that climate influences runoff generation not only through rainfall and antecedent wetness, but also through long-term controls on vegetation, weathering intensity, pedogenesis, soil depth, regolith architecture, and groundwater storage. However, we also emphasize that lithology, tectonic setting, soil structure, topography, and vegetation mediate this relationship, and that the observed pattern should be interpreted as a reporting tendency within the reviewed catchment dataset rather than as a universal

progression from HOF to GWF.

Comment 16:

Line 343: Does node size refer to the bar height? And the rings represent the catchments where dominant mechanism remain unchanged?

Response 16:

We thank the reviewer for pointing out this ambiguity. We agree that the previous wording did not make clear what “node size” referred to, and that the self-returning bands could be confused with other graphical elements.

In the revised manuscript, we clarified the Fig. 6 caption and changed “node size” to “rectangular node height.” The caption now states that the height of each rectangular node is proportional to the number of catchments in the transition subset with that initial mechanism. We also clarified that colored self-loop bands surrounding or returning to the same node indicate cases where the reported dominant mechanism remained unchanged under wetter conditions.

Comment 17:

In section 3.4, transitions of dominant runoff generation mechanisms with increasing soil wetness during rainstorm events and transitions across different catchments with increasing annual rainfall are quite different. I suggest the author to conduct relevant analyses separately. I can easily understand that as the humidity level increases, such as annual precipitation, the watershed runoff mechanism transitions from HOF to SOF. However, I don't understand why there is a similar transition within a watershed as rainfall events continue, HOF and SOF or SSF are essentially completely different mechanism of runoff generation, which is determined by vegetation-soils system. Hence, what is the underlying mechanism?

Response 17:

We thank the reviewer for this important clarification. We agree that transitions within a catchment during storm events and differences among catchments along climatic or annual-precipitation gradients are fundamentally different. We would like to clarify that the transition analysis in this study refers only to within-catchment, event-scale or event-sequence changes reported in the source studies. We did not infer mechanism transitions from differences among catchments with different annual precipitation or climatic wetness classes.

In the revised manuscript, we therefore clarified the wording throughout Section 3.3. We now state explicitly that the transition analysis is based only on 23 experimental catchments for which source studies reported changes or persistence in the dominant active runoff mechanism within the same catchment as antecedent wetness, storage, soil moisture, or groundwater level increased. Cross-catchment climatic patterns, such as those shown in Fig. 5, are now described only as differences in the climatic composition of reported mechanisms, not as mechanism transitions.

We also clarified the meaning of “mechanism shifts.” We do not mean that the fixed vegetation–soil–geology structure of a hillslope changes during a storm. Rather, the term refers to wetness–

dependent changes in the dominant active runoff pathway. The structural system defines the potential runoff pathways, while event-scale wetness controls which pathways become functionally connected. Thus, reported shifts from HOF toward SOF or SSF can be interpreted as activation of saturation-excess or subsurface pathways as storage fills, saturated areas expand, perched water develops, or preferential and soil–bedrock interface flow paths become connected.

To avoid ambiguity, we revised the section title and wording from general “mechanism transitions” to “within-catchment, event-scale shifts in dominant active runoff mechanisms,” and we added explanatory text distinguishing this analysis from cross-catchment climatic composition analyses.

Comment 18:

Line 384: what is Link thickness in figure 8? How do you calculate the frequency of repeated co-occurrence? Explain what co-occurrence is. Lack of necessary introduction makes it difficult for people to keep up with the author.

Response 18:

We thank the reviewer for pointing out that the meaning of co-occurrence and link thickness in Fig. 8 was not sufficiently clear. We agree that the previous explanation did not provide enough information for readers to follow how the network was constructed.

In the revised Methods section, we now explicitly define co-occurrence as the presence of a given runoff generation mechanism and a standardized soil–geology–hydrogeology descriptor within the same independent catchment record. For each mechanism–descriptor pair, the co-occurrence frequency was calculated as the number of independent catchments in which both the mechanism and descriptor were reported. Repeated mentions within the same catchment and multiple publications describing the same catchment were counted only once.

We also revised the Fig. 8 caption to clarify that colored nodes represent runoff mechanisms, gray nodes represent standardized descriptors, and link thickness represents the number of independent catchments supporting each mechanism–descriptor co-occurrence. We further state that link thickness represents reporting frequency across independent catchments, not the number of publications and not causal strength.

Comment 19:

The results in Section 3.5 and figure 8 doesn't seem to be very convincing, for example, in small watersheds of limestone or granite, groundwater runoff can usually contribute to flood runoff processes, which seems to be missing from the author's analysis. In addition, the keywords listed in Figure 8d are not very accurate and cannot effectively represent the geological structure that affects runoff. It is recommended to merge Figure 8d and Figure 8c.

Response 19:

We thank the reviewer for this helpful comment. We agree that the original label “Geological structure” was not sufficiently accurate, especially because the panel included descriptors such as shallow groundwater level that are hydrogeological conditions rather than geological structures. In

the revised manuscript, we changed the label of this panel to “geological structural and hydrogeological features” and revised the caption to clarify the meaning of these descriptors.

We also considered the suggestion to merge the lithology and geological/hydrogeological panels. We retained them as separate panels because they represent related but distinct descriptor types: lithology describes rock type, whereas geological structural and hydrogeological features describe how the rock mass or subsurface system stores, transmits, and connects water, for example through fractures, faults, weathering, impermeable layers, rock outcrops, or shallow groundwater conditions. We revised the caption to state explicitly that panels 8c and 8d should be interpreted together as complementary descriptors of subsurface architecture.

We agree that groundwater or shallow bedrock flow can contribute to flood runoff in small limestone or granite catchments, especially where karst features, fractures, weathered zones, shallow groundwater, or hillslope–stream connectivity are present. We have therefore revised the text to clarify that Fig. 8 shows reported mechanism–descriptor co-occurrence patterns, not all possible hydrological contributions within a given lithology. The absence of a strong GWF link to limestone or granite should not be interpreted as evidence that groundwater contributions are absent from those lithologies; rather, it reflects whether event-scale groundwater flow was explicitly reported as a dominant or combined mechanism in enough independent catchments to be retained in the co-occurrence network.

Comment 20:

Lines 443-444: I totally agree that threshold are not fixed, catchment-invariant constant within each catchment. Many dynamic factors impact rainfall-runoff threshold, for instance rainfall intensity referring to Zhang et al. (2024). However, this factor has little impact on the threshold especially in humid climates, so the water storage capacity of a watershed is usually assumed to be a constant value, which explains why the water storage capacity parameter is set to a constant value in almost all hydrological models (e.g., the Xinanjiang model). While using the relationship curve of rainfall-runoff for determining the threshold and runoff, we always take storage state or antecedent rainfall into account. For example, the widely used method API is a Rainfall-Runoff Empirical Correlation Method.

Response 20:

We thank the reviewer for this helpful clarification. We agree that the original wording could be interpreted too broadly. In particular, we did not intend to imply that catchment storage capacity or structural storage architecture cannot be treated as a relatively stable catchment property in conceptual hydrological models. Nor did we intend to suggest that traditional rainfall–runoff empirical approaches ignore antecedent conditions; methods based on API or storage state explicitly account for such effects.

In response, we removed the statement that runoff thresholds should not be viewed as fixed, catchment-invariant constants and replaced it with a more precise clarification. We now distinguish between relatively stable catchment storage-capacity parameters and event-scale runoff thresholds or threshold indicators. The former may often be treated as catchment-specific structural properties, whereas the rainfall amount or intensity required to trigger nonlinear runoff depends on antecedent

storage deficit, soil moisture, groundwater level, API, and event rainfall characteristics.

Comment 21:

Lines 447-448: I also totally agree that it is hydrologic connectivity that determines runoff generation. However, hydrologic connectivity including function and structure connectivity (Zhang et al., 2026) is quite difficult to be quantified. Hence, it is largely not a practical or directly available index.

Response 21:

We thank the reviewer for this important clarification. We agree that hydrologic connectivity, including both structural and functional connectivity, is difficult to quantify consistently and is not usually available as a direct practical index across the studies reviewed. Our intention was not to imply that connectivity itself can be directly measured or compared across all catchments using a single metric.

We have therefore revised the sentence. Instead of stating that cross-catchment comparability is found in the “connectivity transition” represented by a threshold, we now clarify that hydrologic connectivity is used as a conceptual organizing framework. Cross-catchment comparison is better sought in the combination of reported threshold indicators, catchment-state proxies, and runoff-mechanism changes that can be interpreted in relation to connectivity development.

We also added text noting that, in practical applications, connectivity-related threshold behavior is usually inferred from observable proxies such as antecedent precipitation, soil moisture, catchment storage, groundwater level, saturated-area expansion, or changes in runoff response, rather than from a directly quantified connectivity index. This revision better reflects the difficulty of measuring connectivity while preserving its value as a conceptual framework for interpreting nonlinear runoff responses.

Comment 22:

Lines 456-457: in infiltration-excess-dominated systems, hydrologic connectivity determines runoff generation. However, what determines the former? It is soil-vegetation root systems (Gao et al., 2024). As mentioned above, hydrologic connectivity is difficult to quantify.

Response 22:

We agree with the reviewer. The original statement has been removed because it implied an unsupported comparison of indicator performance and overemphasized rainfall intensity. We now clarify that, in infiltration-excess-dominated systems, rainfall intensity may act as an immediate trigger, but the development of connected surface runoff depends strongly on soil–vegetation–root systems and surface conditions, including infiltration capacity, root structure, macroporosity, crusting, compaction, vegetation cover, and microtopographic flow paths. We also revised the text to acknowledge that hydrologic connectivity is difficult to quantify directly; therefore, in this review it is interpreted through observable proxies rather than treated as a directly available index.

Comment 23:

Lines 467-482. You must define what is SOF or SSF firstly in section 2. Otherwise, I could not figure out what the differences between SOF and SSF. As they are very close, and, in both conditions, overland flow will occur to shape flood discharge in creek channels.

Response 23:

We thank the reviewer for pointing out that the distinction between SOF and SSF needed to be made clearer before the soil–geology fingerprint discussion. We agree that both SOF and SSF can contribute to flood discharge in creek channels, but they differ in the dominant pathway by which storm water reaches the channel.

In the revised Section 2.2, we now provide clearer operational definitions for the runoff generation mechanisms used in this synthesis. SOF is assigned when stormflow is interpreted as saturation-excess overland flow, return flow, or runoff from saturated or near-saturated source areas. SSF is assigned when the source study reports event-scale lateral subsurface flow through soil, regolith, preferential pathways, macropores, pipes, perched zones, or along restrictive layers such as the soil–bedrock interface. We also added clarification that SOF and SSF may both contribute to channel flood discharge, but SOF is classified by saturated-area surface runoff, whereas SSF is classified by lateral subsurface transmission before water enters the channel.

Comment 24:

Lines 504-505: we certainly know structural attributes, i.e., structural connectivity control connectivity and then threshold. However, the authors should provide more in-depth discussions about what (e.g., root biomass?) controls structural connectivity?

Response 24:

We thank the reviewer for this helpful suggestion. We agree that the previous wording was too general and did not sufficiently explain what controls structural connectivity. We have revised the relevant Discussion text to clarify that threshold differences reflect the interaction between rainfall forcing, antecedent state, and structural connectivity, rather than structural attributes in an abstract sense.

In the revised text, we now specify that structural connectivity is shaped by relatively persistent catchment properties, including vegetation cover and root systems, soil texture and aggregation, macroporosity, organic matter, surface crusting or compaction, soil depth, permeability contrasts, lithology, weathered or fractured bedrock, and topographic convergence. These properties define potential storage and flow-path architecture, whereas event wetness, soil moisture, groundwater level, storage filling, and rainfall intensity determine which pathways become functionally connected during storms.

We also clarified that the soil–geology fingerprints analyzed in this review capture only part of structural connectivity, mainly soil depth, texture, permeability, lithology, fractures, weathering, and shallow groundwater conditions. Vegetation-root architecture and biological macropore controls are important, but they were only included when explicitly reported in the source studies. We therefore added this point as both an interpretation and a limitation, rather than presenting root biomass as a

systematically analyzed variable.

Comment 25:

Lines 542-548: runoff mechanism transitions across catchments and within a catchment are totally different. The authors should provide discussions respectively.

Response 25:

We thank the reviewer for this important clarification. We agree that mechanism differences across catchments and mechanism transitions within a catchment are fundamentally different and should not be conflated. In the revised manuscript, we clarified that the transition analysis refers only to reported within-catchment, event-scale or event-sequence changes in the dominant runoff generation mechanism as catchment wetness increases. It does not refer to transitions inferred from differences among catchments along climatic or annual-precipitation gradients.

We revised the relevant Discussion text accordingly. Cross-catchment climatic and soil–geology patterns are now discussed as differences in structural setting and reported mechanism occurrence, whereas within-catchment transitions are discussed as event-scale changes in the mechanism dominating stormflow. We also removed wording such as “evolutionary pathways,” which could imply a universal or cross-catchment sequence.

We now explain that within-catchment mechanism transitions do not imply that the vegetation–soil–geology structure or the catchment’s structural runoff potential changes during an event. Instead, they reflect the wetness-dependent activation of different pre-existing flow pathways. As antecedent wetness, storage, soil moisture, or groundwater level increases, saturation expansion, perched subsurface flow, preferential flow, soil–bedrock interface flow, or shallow-groundwater connectivity may become functionally connected and dominate stormflow.

Comment 26:

Lines 549-556: things about rainfall have been mentioned repeatedly in the whole manuscript.

Response 26:

We thank the reviewer for pointing out this repetition. We agree that the role of rainfall as a trigger, together with antecedent wetness and catchment state, had already been discussed earlier in the manuscript. To reduce redundancy and keep the Discussion focused, we deleted the repeated explanatory sentences around Lines 549–556.

In the revised text, we retain only a brief linking statement indicating that rainfall forcing is treated as the event input, whereas antecedent wetness, storage state, and structural connectivity determine which flow pathways become functionally connected.

Comment 27:

In figure 9, I agree that as wetness and connectivity increase, runoff mechanism could transit from HOF to GWF among different climates. However, subsurface flow could occur at very lower soil

moisture level ($SM < \text{Field capacity}$), which could help explain why there are hockey stick rainfall-runoff curves in many catchments (Scaife and Band, 2017; Zhang et al., 2021). Moreover, in arid loess catchment, though with deep soils, HOF is the dominant runoff mechanism. In addition, SSF are easy to transit into SOF in the valley floors in steep humid hillslopes in the fields. Soil thickness or regolith thickness as well as vegetation root biomass may all control runoff mechanism transit. So, the conceptual framework still needs further review and revisions.

Response 27:

We thank the reviewer for this constructive comment. We agree that the conceptual framework needed clearer explanation and that the previous caption could be misread as implying a simple monotonic sequence from HOF to GWF or as treating field capacity as a universal runoff threshold. This was not our intention.

In the revised caption and figure, we clarified that Fig. 9 is a conceptual, non-metric framework for within-catchment, event-scale transitions in the dominant runoff mechanism. We now explicitly state that the red dashed line represents an approximate field-capacity reference, not a universal runoff-initiation threshold. The SSF box spans both sides of this reference to indicate that subsurface stormflow can occur below field capacity where preferential flow paths, macropores, pipes, perched zones, or soil–bedrock interfaces become connected.

We also clarified the non-monotonic role of soil or regolith depth. The circular soil-depth arrow is intended to show that very shallow soils or rock outcrops may favor HOF because of limited effective storage, whereas very deep and dry profiles, such as arid loess catchments, may also remain HOF-dominated during extreme rainfall because subsurface storage and groundwater connectivity thresholds are difficult to reach.

we revised the explanation of dashed arrows. The dashed SSF-to-SOF pathway now explicitly represents cases where lateral subsurface flow exfiltrates in footslopes, valley floors, or riparian zones and contributes to saturation-excess overland flow or return flow. The dashed SSF-to-GWF pathway represents cases where continued wetting connects lateral subsurface flow with shallow groundwater, weathered regolith, or fractured bedrock pathways, resulting in groundwater-influenced stormflow.

These revisions make clear that the framework does not describe a universal or deterministic HOF–SOF–SSF–GWF progression, but rather summarizes reported and plausible event-scale transition pathways conditioned by both wetness state and structural controls.

We sincerely thank the reviewers for their helpful suggestions, which have greatly improved our manuscript. In addition to addressing the reviewers' comments, we have also refined the content and expression throughout the paper for better clarity and coherence. The revised manuscript has been submitted to your esteemed journal, and we look forward to your favorable response.

Yours
Zhen Cui, Fuqiang Tian

Jun 17, 2026