We thank reviewer 1 for their careful and constructive reviews. The comments will help us substantially improve the clarity of our methods and the framing of our contribution. Below, the original review comments are shown in **bold** and our responses to each comment use normal text. Line numbers refer to the reviewed preprint; specific revised text will be added in the tracked-changes manuscript if the editor is in support of a revised manuscript.

Reviewer 2

This manuscript provides a valuable large-sample analysis of hydrological transitions between drought and flood conditions across the conterminous United States. The use of 3,219 USGS streamgages and multiple event definitions is a major strength, and the paper makes a useful contribution by comparing both drought-to-flood and flood-to-drought transitions. The finding that DtF transitions are generally faster and more common than FtD transitions appears robust and important. The analysis of basin characteristics and the comparison with meteorological whiplash are also valuable. I recommend that the authors clarify several points and consider softening some interpretations.

We thank the reviewer for the supportive assessment and for their constructive suggestions.

First, I think it would be good to define what an impactful transition is. Since it is based on large changes in streamflow percentiles, the word "impactful" may sound like the authors are referring to real societal, ecological, or infrastructure impacts.

We agree the term risks implying documented real-world impacts. We will define "impactful" precisely at first use as a transition occurring within 90 days that involves at least a 50-percentile shift in fixed-percentile space, and will clarify that this is a *hydrologic-magnitude* criterion indexing large movement through the flow-duration curve, a proxy for potential management relevance, not a measure of realized societal, ecological, or infrastructure impact. We will add a caveat to this effect and have considered alternative wording (e.g., "high-magnitude"). We will retain "impactful" for consistency with the defined criterion but will flag its meaning explicitly wherever it is used.

Second, the comparison between hydrological and meteorological whiplash is interesting, but the two datasets use different time scales. Hydrological transitions are based on daily streamflow, while meteorological transitions are based on seasonal events. Because of this, some mismatch is expected. The statement that hydrological and meteorological whiplash "rarely coincide" may be a little too strong. It would be safer to say that they show limited overlap under the definitions and timing criteria used in this study.

We agree and will soften this language throughout (abstract, results, discussion, and conclusions). We will say that hydrological and meteorological whiplash show *limited overlap under the definitions, spatial-overlap thresholds, and seasonal alignment criteria used in this study*, and we will explicitly attribute part of the low overlap to the timescale mismatch between event-scale hydrological and seasonal meteorological transitions (see also Reviewer 1, comment 4a).

Third, FtD transitions are much less common than DtF transitions. The authors should clarify whether the conclusions about rapid FtD transitions are based on enough events, or whether they may be affected by a small number of sites or cases. Adding event counts, uncertainty ranges, or a minimum-event-count rule would make this part stronger.

Fast FtD transitions are less common than fast DtF transitions; however, there are at least as many FtD transitions as DtF transitions overall (including transitions longer than 90 days), in fact somewhat more in our analysis (Fig. S5), and they occur at essentially all sites across all 12 regions. We therefore do not believe the DtF/FtD asymmetry is an artifact of a lack of FtD transition events. We will state this more prominently for clarity and add event counts to instances where FtD transitions are made from only impactful or rapid events.

Fourth, the paper uses ≤90 days for sub-seasonal transitions, ≤30 days for short-term transitions, and ≤14 days for rapid transitions. These thresholds are useful, but the authors should explain why the same thresholds are suitable for all regions and basin types. For example, a 30-day transition may mean something different in a small urban basin compared with a snowmelt basin or a large regulated river.

We agree these fixed thresholds carry different physical meaning across basin types. We adopt them for cross-site comparability and consistency with prior work (Anderson et al., 2025), and we will state this rationale explicitly while acknowledging the limitation. We note that our per-basin-type and per-region analyses (Figs. 3, 6, S6–S8) partly address the concern by showing how transition-time distributions themselves differ across settings, so readers can interpret the fixed thresholds against each type's characteristic timescale. We will add text framing the thresholds as comparable reference bands rather than physically equivalent across all basins.

Finally, Lines 220–226 should better explain why these specific drought and flood definitions were chosen. The authors later say that the selected definitions help balance FtD and DtF transitions, but that reason alone is not fully convincing. It would help to also explain why these definitions make hydrological sense.

We will expand the justification. Beyond balancing event counts, the study definition was chosen because (i) the 10th-percentile variable threshold with a 15-day minimum duration captures sustained low-flow conditions consistent with hydrological (not merely meteorological) drought while avoiding the very short "dry periods" flagged by Reviewer 1, (ii) the 5-day inter-event pooling reflects typical short recessions between low-flow spells, (iii) and the POT-1-event/year flood definition yields a physically interpretable, site-adaptive flood frequency comparable to the drought rate and avoids the fixed-percentile bias against long snowmelt peaks (documented in Section 4.1). We will present these hydrological rationales alongside the balancing argument.

Detailed suggestions

Figures

All figure sizes and font sizes could be bigger.

We will increase figure and font sizes throughout for legibility.

Some subplot labels are before the description, some are not — make caption formatting consistent.

We will standardize all captions to place the panel label before its description.

Ensure all legends don't overlap with other parts.

Legends will be repositioned to avoid overlap in all figures.

Figure captions should be more clearly separated from the main text (e.g., lines 491–499 could be mistaken for main text).

We will set captions in a distinct style (bold "Figure N:" lead-in, and where possible reduced/italic type) so they are clearly separated from body text.

Figure 3: legends are not well aligned with subplots and overlap labels/elements, especially panels c/h, d/i, e/j.

We will realign each legend with its corresponding row and remove overlaps for the c/h, d/i, and e/j pairs.

Figure 6: difficult to read; numbers/decimals on the boxplots are hard to see and may need repositioning.

We will reposition and enlarge the overlaid values in Fig. 6 (and reduce their number where they crowd the boxes) to improve readability.

Abstract

Line 16: Define CONUS when it first appears.

We will spell out CONUS (conterminous United States) at first use in the abstract.

Lines 15, 18, 20: verbs "analyzed," "test," "identify" use mixed tense — make consistent.

We will make verb tense consistent (past tense for what we did) throughout the abstract and manuscript.

Line 24: "the line" should be defined more clearly (e.g., where P ≈ PET, or P/PET ≈ 1).

We will define it on first use as "the line from North Dakota to Texas where precipitation approximately equals potential evapotranspiration (P/PET ≈ 1)."

Introduction

The Introduction feels somewhat long — consider subheaders.

We will add subheaders to guide the reader (e.g., motivation and impacts; large-sample and regional studies; the FtD direction; methodological considerations; objectives) and will tighten repetitive passages.

Terms "wet-to-dry," "rapid wet-to-dry shifts," and "rapid FtD and DtF transitions" should be defined and distinguished.

We will define these terms explicitly and distinguish meteorological "wet-to-dry"/"dry-to-wet" transitions from hydrological "FtD"/"DtF" transitions at first use, consistent with the definitions added for Reviewer 1, comment 1.

Lines 69–70 and 88–90 discuss similar ideas on hydrological intensification/volatility — shorten or combine.

We will combine these into a single, more concise statement to reduce repetition.

Methods

Line 198: specify where these streamgages are located when they first appear.

We will state that the gages span the conterminous United States and will note their spatial distribution (Fig. 1a) at first mention.

Line 234: "conterminous" — "contiguous United States" may be more understandable.

We will define CONUS as the conterminous (contiguous) United States at first use and retain the standard term thereafter, so both readerships are accommodated.

Line 242: "a fraction of 0.1" seems unnecessary — simplify.

We will simplify this — we will state "less than 10% of the drought volume" without the parenthetical fraction (the criterion is also now defined per Reviewer 1, comment 2b).

Line 286: "HCDN; Lins, 2012" may read as if HCDN is part of the citation — revise (e.g., "the Hydro-Climatic Data Network (HCDN), as described by Lins (2012)").

We will revise this as suggested to "the Hydro-Climatic Data Network (HCDN; described by Lins, 2012)."