<u>General comments</u>

The authors make use of a high-resolution reanalysis dataset that represents a considerable effort in terms of length (37 years), spatial coverage (all of Italy), and resolution (4 km). This combination makes it a valuable resource for a national-scale study of changes in extreme precipitation. The availability of hourly rainfall data also makes the study relevant for analyzing the seasonality and spatial variability of hourly precipitation in Italy.

I consider this work useful for the scientific community because it brings together results that were previously limited to local or regional scales into a study covering the whole of Italy. This allows for a better understanding of the spatial variability of extreme precipitation event (EPE) characteristics and their changes, while showing consistency with earlier regional studies.

That said, I find that the Abstract and Introduction could be more explicit regarding the specific scientific contributions and aims of the work. The chosen methodology is interesting in that it introduces a structure-based perspective on hourly precipitation; however, some methodological aspects would benefit from clearer explanation or reconsideration.

Referring to hourly precipitation structures as *"events"* is, in my view, misleading. Since their temporal evolution (duration, displacement, deformation) is not considered, nor the total rainfall volume produced by multi-hour storm systems, the term *"event"* may create confusion and should be replaced with a more precise designation. Finally, the section addressing non-extreme precipitation should be better introduced, with its purpose and relevance clarified.

Overall, in my opinion, the manuscript is suitable for publication after minor revisions. These include adjustments in specific vocabulary, clearer argumentation, additional explanations to help reader understanding, and refinements in the methodology.

<u>Specific comments</u>

1. The 4 km resolution of the reanalysis is at the limit of what is usually called convection-permitting. This should be made explicit in the text, especially since the dataset paper uses "high-resolution" instead. I would keep the term "convection-permitting" for clarity but suggest adding a note that it is at the edge of the definition.

2. Since about half of the figures in the Results section concern hourly rain structures and not EPEs, it is important to either (i) include non-extreme hourly rain discussion in the abstract as well as adapt the title accordingly, or (ii) introduce non-extreme hourly rain results as a necessary step before moving to extremes. Otherwise, the part on extremes, which is announced to the reader, takes too long to arrive.

3. The Introduction and Abstract should state more clearly the purpose of the work and its scientific contribution.

4. The HOPE-X dataset is a collection of hourly rain structures, not "events" in the usual sense of storm systems or local rain events. For instance, a precipitation system lasting two hours can be counted as two "events," even though it is the same system. Similarly, a moving storm can be counted as several "events" in different areas as it displaces. For this reason, the use of the term "events" is inappropriate. The title is also misleading, since the study is not, in my opinion, event-based. The same applies to the dataset name HOPE-X. This does not invalidate the study, but the terminology issue is central and should be revised. I strongly recommend using the term "hourly rain structures." If the term "event" is retained, it should not appear in the title, abstract, or in any part of the text where the definition is not clearly introduced yet (line 164: "Hereafter, the term 'event' denotes the precipitation structures identified using this method"). In that case, it should also be stated that the term "event" is used for readability purposes only.

5. Following the previous comment, in maps such as Figure 6 the number of "events" (N) can reach ~300 per season in a 0.5×0.5 window. This N reflects a mix of the number of hours with a rain structure and the number of surrounding structures at each hour, rather than distinct events. For example, an hour with many small rain structures and an hour with one large rain structure could produce the same hourly rainfall total in that 0.5×0.5 area, yet N would differ greatly. It should therefore either (i) be clarified in the text what the analysis of N is intended to represent if it is something other than the amount of rain, (ii) reconsider the counting methodology, or (iii) include additional analysis showing that N reflects the actual volume of hourly rain.

6. The use of seasonal thresholds to adapt to the intensity scales of each season is well justified and explored. However, the choice to let the seasonal threshold vary spatially with the local 50th percentile of 1 mm+ hourly rains requires stronger justification. As a result of threshold methodology, identical rain structures may be included in HOPE-X in one region but not in another. While such area-relative thresholds are understandable when defining extremes (e.g. EPEs) due to the definition of rarity being potentially region-based, they are harder to justify for

non-extreme rain structures. Choosing fixed threshold across the domain for each season would be more correct. If the authors choose to go with a justification of the current methodological choice of non-homogeneous threshold, consider also that in areas with very little rainfall, the threshold based on wet hours 1mm+ may also be computed on a very small and potentially unrepresentative sample, so this would need to be tested.

7. Potential artifacts due to the threshold methodology chosen discussed in previous point: It is not clear whether some of the patterns in the maps (e.g. the higher JJA values over Ravenna in Fig. 8 AvIn and Fig. 9 PkIn) are genuine results or artifacts of the methodology filtering out structures with higher thresholds (higher thresholds over Ravenna in Fig. 1). Line 257: Are the "hotspots" results of higher local thresholds? Or an actual signal? Same for "isolated spots" line 278.

8. The role of the "minimum enclosing ellipse" is unclear. Its mention in the title of Fig. 2 and in the text (line 169) is confusing: are the variables in Table 1 computed on the ellipse itself or on the rain object it encloses? Is the ellipse only used for visualization of selected structures, or does it play a role in extracting the properties of the rain objects? This part should be clarified for better understanding.

9. Line 169: for readers unfamiliar with Wernli et al. (2008), it would be helpful if the purpose of using this methodology were explained.

10. Following the previous points, it is unclear whether *axis_maj* refers to the major axis of (i) the rain structure itself or (ii) the enclosing ellipse. In both cases, the phrasing in line 264 ("summer events are generally smaller") is problematic, as are later formulations in the SpS analysis regarding size. If (i), and the aim is to characterize the size of the structures, the "area" variable would be more appropriate. For example, a thin elongated structure may be characterized as large with SpS even though its area of coverage is small, while a rounder structure with the same SpS value could cover a much larger area, yet still be classified as the same "size" when using SpS. If (ii), the analysis would instead describe the size of influence of the structure rather than its actual size. It should therefore be clarified what SpS is intended to capture, and statements about "size" should be made with caution. Replacing the SpS analysis with an analysis of the area variable would make the interpretation more straightforward.

<u>Technical comments</u>

1. Line 10: "the most extreme component." The word *component* is unclear here. I suggest rephrasing to: *"The most extreme rain events (EPEs)…"*

2. Line 111: "event-based approach." As discussed earlier, this would be better described as an approach using hourly rain spatial structures. I also suggest adapting the phrase to be less generic and more explicit about how clustering helps address the limitations:
   "In light of these limitations, an object-based approach using a clustering technique was adopted to capture coherent hourly precipitation structures and reduce sensitivity to small-scale discrepancies between simulations and observations."

3. Line 146: "The event-detection … precipitation events…" The vocabulary of *event* throughout the text before line 164 is misleading, since line 164 specifies:
   "Hereafter, the term 'event' denotes the precipitation structures identified using this method." Event terminology should not appear earlier without this clarification.

4. Line 296: "is conducted on a subset of EPEs." A clearer phrasing would be: *"a subset of the HOPE-X dataset."* (But HOPE-X should be renamed)

5. Lines 321–322: The phrase *"seasonal redistribution is likely driven by the persistence of summer-like convective activity into early autumn"* is confusing, since the focus is on differences between summer and autumn. This cannot be justified by describing autumn activity as "summer-like." A phenomenon-based explanation could instead highlight that the persistence of warm sea surface conditions beneath a cooler atmosphere creates instability favorable to convection. Autumnal convection is not "summer-like". This is also reflected in Fig. 12, which shows differences in the "size" characteristics of hourly rain structures between summer and autumn.

6. Line 354: Instead of *"alternating pattern,"* I suggest: *"spatial heterogeneity in the sign of the signal"* or *"patterns of alternating signs."*