Response to Reviewers

EGUSPHERE-2025-3455: “Hourly Precipitation Patterns and Extremization over Italy using Convection-Permitting Reanalysis Data”

[answers in blue]

Reviewer #2: accepted as it is, but:

However, I still have some concerns regarding the Rx1hour threshold and the precipitation structures. If 11% of summer events exceed the threshold. How does this threshold compare to other more used thresholds in the literature? Eg. Rx1day, and various percentiles? The threshold of above the mean av the Rx1hour during the 37 years is intuitively very high, although justified in the reply. However, if it needs to be this high to be extreme, how does this relate to other studies mentioned in the manuscript that is also classified as extreme (for example line 94, the 90th percentile is stated as extreme). This seems inconsistent to the reader. The manuscript would be improved if the metrics were more relatable, since the structural-based approach is rather new. Some added explanations and relatability in the discussion section would improve the manuscript.

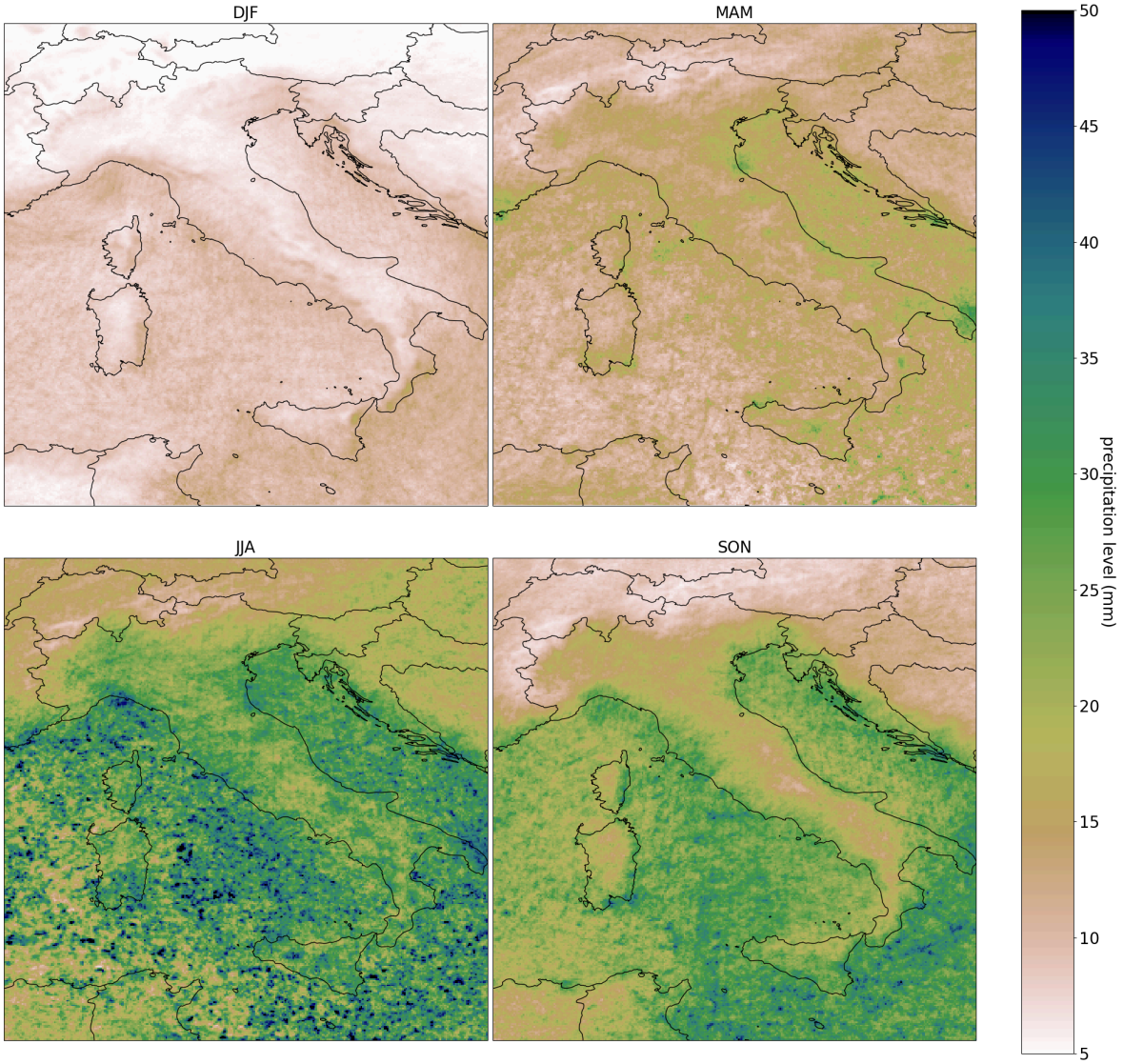
We thank Reviewer #2 for the positive assessment of the manuscript and for the constructive comments regarding the definition of extreme precipitation thresholds. These remarks are very helpful in improving the clarity and interpretability of the study.

The threshold Rx1day refers to a longer accumulation period and therefore leads to numerically larger values (in mm/day), while also capturing partially different physical phenomena. Daily maxima are not necessarily associated with short-lived convective events, which are instead the specific focus of the hourly precipitation analysis presented in this study.

Concerning percentile-based definitions of extremes, as reported in the literature review (e.g. the 90th percentile you mentioned), we acknowledge that there is no unique or universally accepted definition of “extreme” precipitation. Different definitions emphasise different aspects of the precipitation distribution and can all be legitimately referred to as extremes, despite being conceptually different and not directly comparable. Percentile-based thresholds characterise the upper tail of the full distribution. In contrast, thresholds based on annual maxima focus exclusively on the most extreme values reached over a time window (a year).

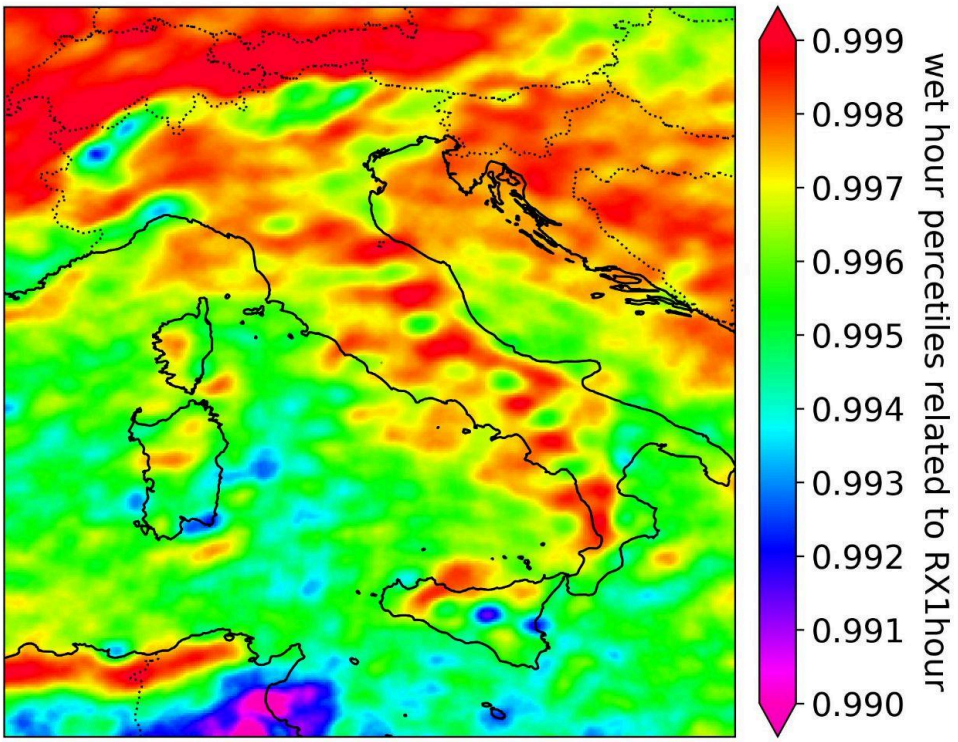
As an illustrative example, we can consider the 99th percentiles of wet-hour precipitation (hours with precipitation > 1 mm) for the four seasons (Figure below, without spatial smoothing). These statistics show that, particularly in winter, percentile-based thresholds are dominated by the large number of hours with weak precipitation, so that the few intense events contribute little to the resulting values. In contrast, a threshold derived from annual hourly maxima isolates only the most extreme value at a given location. This cuts out most of the lower values, but makes this kind of threshold conceptually different from percentile-based ones.

Seasonal levels associated to 99th percentile (wet hours) (1986-2022)



This does not imply that high percentiles are an incorrect way to define extremes; rather, within the context of a structure-based approach, we found a threshold based on annual maxima to be more appropriate and consistent with extreme value theory (e.g. Coles et al., 2001). Most importantly, the Rx1hour threshold is applied not at the single-grid-point level but to spatial clusters defining precipitation events. An event is classified as extreme if at least one point within the cluster (its maximum) exceeds the local threshold, effectively enlarging the set of events identified as extreme.

If the reviewer inquiry instead concerns the relationship between the mean Rx1hour threshold and percentile-based thresholds, we have prepared an additional figure (shown below) illustrating which percentile (computed among wet hours with precipitation > 1 mm) corresponds to the mean Rx1hour value (in mm/h). As noted, this percentile generally exceeds the 99th percentile and is often close to the 99.9th percentile. Nevertheless, a substantial number of events are selected because the threshold is applied to spatial clusters rather than to individual grid points.



Following the reviewer's suggestion, we have added a clarifying statement in the Discussion section (lines 437-439).

Reviewer #3: Minor Revision:

General comments: The manuscript has been improved, with the authors appropriately addressing the reviewers' comments. The use of more precise terminology (e.g., "hourly precipitation patterns" instead of "events", "spatial extent" instead of "size"), a more instinctive choice of acronyms, and several redaction adjustments enhance the readability of the paper and improve the reader's understanding, particularly regarding the methodological aspects. The contextualization of the analysis has also been strengthened. The authors now acknowledge the limitations associated with the 4-km resolution, integrate the nonextreme precipitation analysis more smoothly (through revisions to the title, abstract, and introduction), and clarify the overall purpose of the study. Most of my previous concerns have been satisfactorily answered and incorporated. However, I reiterate that the metric N still presents a caveat. Although the authors have partially addressed this issue, it requires either further verification or more cautious phrasing in the conclusions regarding changes in extreme precipitation events (HPEs), as detailed in the specific comments. For this reason, I recommend minor revision before publication.

We sincerely thank the Reviewer for the careful re-evaluation of the manuscript and for the positive and encouraging assessment of the revisions.

We also thank the Reviewer for reiterating the remaining concern regarding the metric N. We recognize that this aspect still warrants caution and appreciate the opportunity to further clarify and refine our interpretation. Below, we provide point-by-point responses to the specific comments, detailing the additional verifications performed and the corresponding revisions made to the manuscript.

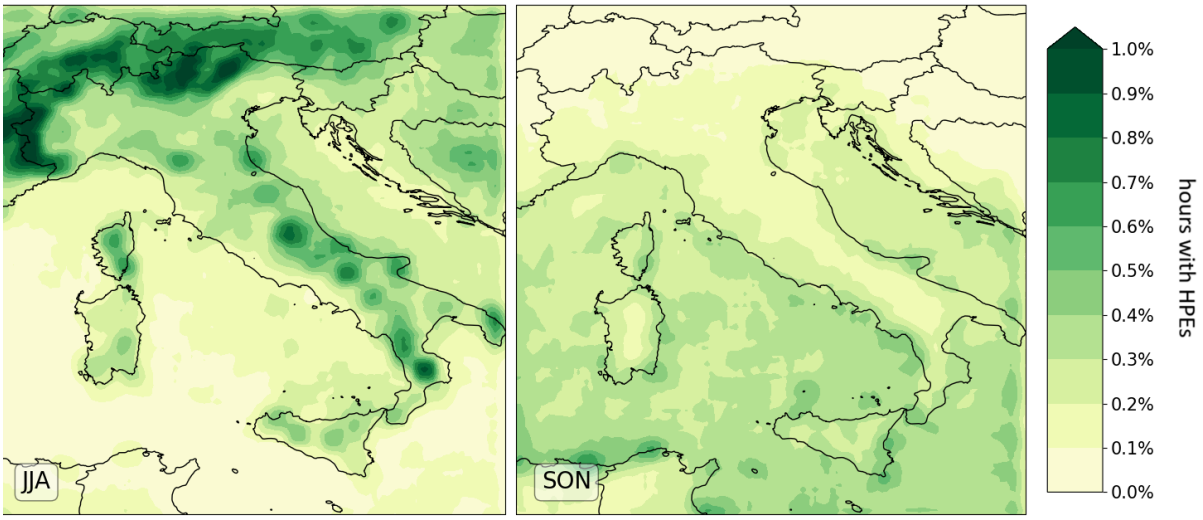
We believe that these additional clarifications and adjustments further strengthen the robustness and transparency of the study, and we thank the Reviewer again for the constructive feedback that contributed to improving the manuscript.

Specific comments:

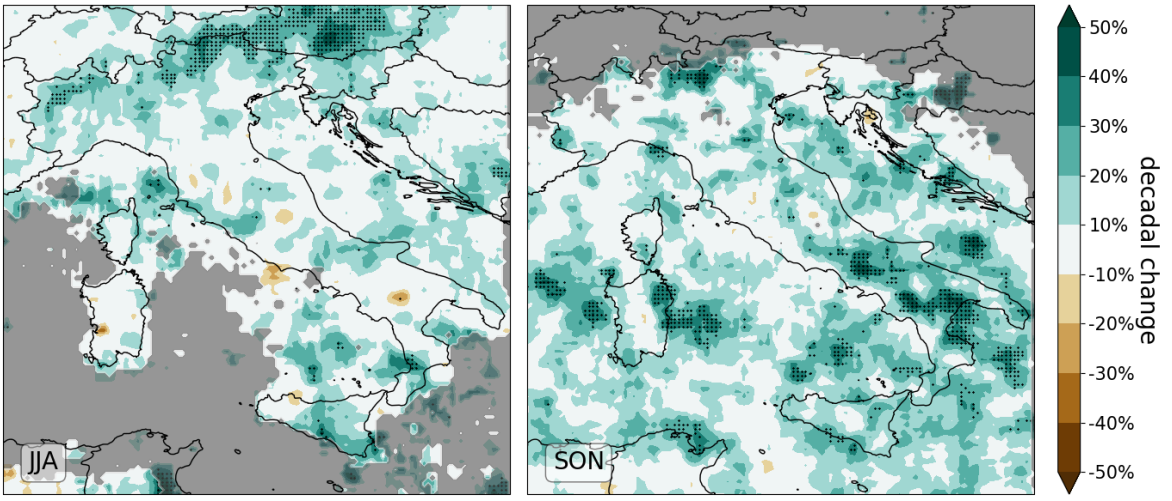
1. Regarding the N metric: My initial concern was that N could reflect both the hourly spatial density of precipitation structures and their frequency. The authors addressed this by adding, in the supplementary material (Fig. S1), an analysis of the frequency of wet hours, which reproduces the behaviour of N. These satisfactorily address the issue, but only for hourly precipitation structures, not for HPEs and not for HPE changes derived from N. I leave to the authors the choice to either (i) extend the wet-hour analysis to HPEs and to changes in HPEs, or (ii) explicitly state the limitations of using N to interpret HPEs behaviour (See point 2). In addition, in the sentence at l.204–206 ("Alternative metrics to N, such as the frequency of wet hours…"), it should be clarified what this alternative metric is intended to test. I also do not agree that N is more "tangible" than the wet-hour metric, but this may be personal.

We thank the Reviewer for the detailed comment and for the constructive suggestions regarding the interpretation of the metric *N*. We fully acknowledge the potential ambiguity of *N*, as it may reflect both the hourly spatial density of precipitation structures and their frequency.

Following the Reviewer's recommendation, we extended the frequency analysis also to HPEs:



This additional analysis shows that, for extreme precipitation, the spatial and seasonal patterns of the frequency of occurrence are highly consistent with those obtained using *N (see Figure 10 of the manuscript),* with no substantial differences that would suggest a dominant role of frequency changes over spatial density. The same conclusion applies to the temporal changes, as trends in HPE wet-hour frequency closely mirror those derived from *N*:



As these additional figures do not provide further independent insight beyond what is already conveyed by *N*, we opted not to include them in the Supplementary Material. However, we now explicitly state in the Method section 2.5 (lines 239-240) that the behaviour illustrated in Fig. S1 for hourly precipitation structures also applies to HPEs and to their temporal trends.

We have also revised the sentence at lines 204-207, specifying that the purpose of the alternative metric based on wet-hour frequency is precisely to disentangle the respective contributions of the hourly spatial density of precipitation structures and their frequency of occurrence.

2. Regarding interpretation of changes in N: In l.384–386, the authors write: “This suggests that changes over time are more likely associated with the frequency of HPEs rather than their intensity or spatial extent. It may also reflect the lower noise sensitivity of N compared to other indicators.” This is an important statement, but in my view the methodology is not sufficient to conclude that frequency changes dominate over intensity changes in HPEs. As discussed in point 1, an increase in N could indeed reflect (i) more hours containing extremes, but it could also arise from (ii) an increase in number of structures within the 0.5 × 0.5 window during the same hour.

We thank the Reviewer for this important observation and for highlighting the need for caution in interpreting changes in N. We agree that, in principle, an increase in N may arise from (i) an increase in the number of hours containing HPEs and (ii) an increase in the number of HPEs occurring within the same spatial window during a single hour.

However, case (i) result the dominant mechanism, since the analysis of wet-hour frequency for HPEs shows spatial, seasonal, and temporal patterns that closely mirror those of N, including comparable trends, indicating that changes in N primarily reflect an increasing number of hours containing extreme precipitation. Moreover, we believe that an increase in the number of HPEs exceeding the threshold within the same spatial window and at the same hourly time step (i.e., simultaneous events) does not represent a methodological issue in itself. When more than one HPE is counted within a window at a given hour, these correspond to spatially disjoint structures and therefore to distinct precipitation events, rather than to multiple counts of the same event. As such, their contribution to N reflects a genuine increase in the number of extreme precipitation structures, not an artificial inflation of the metric.

> The point about noise sensitivity is relevant and should be explicitly incorporated into the interpretation in that same phrase: N is less affected by noise, which makes statistical significance easier to achieve compared to peak-int and mean-int. However, both peak-int and mean-int also show increases in the selected regions (though not significant), and there are high chances that an intensification is responsible for the increase in the number of extremes. For these reasons, this interpretation should be rewritten with clearer limitations, as a deeper investigation would be needed to affirm that frequency is primarily affected compared to intensity in my opinion.

We have revised the sentence at lines 388-391 to explicitly incorporate this interpretation, clarifying that an intensification of extremes—while not statistically significant at the hourly scale—can lead to an increased number of HPEs (more HPSSs exceeding the threshold). This results in a clearer and more robust signal in N, allowing statistically significant trends in HPE occurrence to emerge.

Technical comments:

1. L. 248-249 “Before focusing on the extremes, an analysis of the overall patterns of HPSSs across the dataset is presented, providing the necessary context for the interpretation of subsequent results on extremes”. It is not straightforward that the non-extreme precipitation context is “necessary” for extreme precipitation analysis. Please clarify or rephrase.

We thank the Reviewer for this comment and agree that the term “necessary” may be too strong. The analysis of non-extreme precipitation patterns is not strictly required for the study of extremes; rather, it is intended to be supportive and informative, helping to better contextualize and interpret the results on extreme precipitation. We have revised the sentence at lines 250-251 accordingly to reflect this distinction more clearly.

2. L.307-308 Regarding Figure 8 “In autumn, slightly higher intensities, ranging from 4 to 5 mm/h, cover most of the country, while lower values persist only in the Prealpine and Alpine regions.” Change “only” to “mostly”, as values under 4 mm/h are not exclusive to the Prealpine and Alpine regions.

We thank the Reviewer for this suggestion. The text at line 311has been revised accordingly, replacing “only” with “mostly” to accurately reflect that values below 4 mm/h are not exclusive to the Prealpine and Alpine regions.

3. At several points the manuscript refers to ‘bias’ and ‘inconsistencies’ in the dataset. Please clarify that these refer to findings from previous work, since the dataset is not evaluated in this study:

   (i) l.320-323 “However, while MeanInt and PeakInt seasonal maps appropriately reflect higher values during the autumn and summer seasons, they also display certain inconsistencies. In particular, some areas exhibit an overrepresentation of convective activity during summer, which may not fully align with observed patterns. This issue will be examined in greater detail in the Discussion section 4.” Please clarify what is meant by “inconsistencies”. My understanding is that these refer to inconsistencies in the sense of biases relative to observations. If this is the case, please specify which observational dataset they are compared to and cite the corresponding references. If it is not the case, give more precision on what you mean by inconsistencies.

We thank the Reviewer for this observation. We agree that “inconsistencies” was not precise, and we have revised the text at lines 324-326 to explicitly refer to biases relative to observations, citing previous work where the dataset was compared with observational data.

   (ii) l.402-403 “Then, as described in Section 2.1, a precipitation overestimation bias is present in summer. These localized wet biases are likely due to overly active explicit convection in the model, as shown in Figure 8.” Precise what is shown in Figure 8, as it is not a bias map. Rephrase the sentence 402-403 so that it is more explicit that you are referring to bibliography here.

We thank the Reviewer for this comment and agree that Figure 8 does not represent a bias map, but rather shows the spatial patterns of HPSSs seasonal mean intensities. We have revised the sentence at lines 407-409 to clarify that the overestimation in summer precipitation refers to findings reported in previous studies, rather than to Figure 8 directly.

4. Discussion l.413-414: “In principle, such biases could have masked decreasing trends in those areas; however, the overall spatial pattern suggests that this scenario is highly unlikely.” I do not understand how the overall spatial pattern suggests this scenario unlikely. Also be careful as you are mixing bias on annual precipitation percentages with percentages of trends in HPEs. Are they comparable? What if the bias in annual precipitation specifically comes from a misrepresentation of extremes?

We thank the Reviewer for this careful observation. The reason for that statement is that Figure 12 shows that the majority of significant and non-significant points display an increasing tendency of HPEs occurrences. Based on this overall spatial pattern, it appears unlikely that extensive areas could instead exhibit an opposite signal that is subsequently masked by model biases. Also, we agree that the comparison between biases in annual precipitation percentages and percentages of trends in HPEs involves distinct quantities and is therefore not strictly comparable numerically. Here, the intention was qualitative, to provide an idea of the magnitude of potential impacts and to illustrate that, even if some bias could influence trends, this effect is likely to be small. We have revised the text at lines 418-420 to clarify this interpretation and to emphasize the qualitative nature of the comparison, acknowledging that the two quantities refer to different temporal intervals and precipitation characteristics.