Authors' Response to Reviews of

Impact-based temporal clustering of multiple meteorological hazard types in southwestern Germany

Katharina Küpfer, Alexandre Tuel, Michael Kunz
Nat. Hazards Earth Syst. Sci., https://doi.org/10.5194/egusphere-2024-2803

RC: ***Reviewers' Comment***, AR: Authors' Response, ☐ Manuscript Text

Reviewer #1

General comments

RC: ***The paper analyzes the relevant question for risk assessment of the temporal clustering of events. The study is based on insurance loss data covering south-west Germany over the period 1986-2023. The hazard types associated to the losses are flood, storm and hail and they are associated to meteorological data. The hazards flood and storm are further separated into different phenomena since they can derive from different weather conditions: fluvial, pluvial or mixed floods on one hand and large-scale storms and convective gusts on the other hand, which makes sense. Then the methodology is described, and the results analyzed and discussed.***

AR: We would like to thank Sylvie Parey for reviewing the manuscript and their valuable comments.

Specific comments

RC: ***While seasonality is handled both in the loss data and in some hazards' characterization, it is not considered when identifying the major loss events. In my opinion, because seasonality is clearly identified in the loss distribution, it should be considered in the characterization, otherwise the percentile is computed in mixing different types of losses, which can be misleading. Therefore, I would suggest that the discussion devoted to the loss distribution analysis in section 4 is moved before the major loss events identification in section 2, justifying that the identification should be made on a seasonal basis. Then the same methodologies can be applied for each main season of occurrence and further summarized at the annual scale if necessary.***

AR: We thank Sylvie Parey for raising this point, which we understand. As mentioned, the loss distribution shows strong seasonal patterns depending on the hazard type. However, by defining events depending on their seasonality, we would miss events and clusters that occur across the limits of seasons (e g., starting in late August and continuing until September, see Fig. 5). A seasonal event identification would also overly emphasize the relevance of, e. g., convective gusts during December–February. Conversely, if they were not considered at all because of lower losses or a low number of events, medium loss events of that type and season would be missed (e. g. fluvial floods during May–August). Actually, this annual definition of events allows us to see the seasonal pattern identified in Fig. 3. Off-season events or events across seasons can become relevant, for example, if they are also associated with clustering (see Fig. 5b). Events need to be defined before the seasonal analysis, as they are defined as meteorological event categories and not as seasonal categories.

Another reason for identifying events before the meteorological characterization is that the methods we use to differentiate between hazards require a subset of extremes (e. g., convective gusts are identified by those events that do not occur within a large pressure gradient). Given the small-scale nature of these convective events, we are limited in our choice of methods of assigning meteorological categories.

Furthermore, if we moved the loss distribution analysis (Sect. 4.1) before the event identification (Sect. 2.2), the loss distribution analysis would consider daily losses instead of aggregated multi-day events. This would mean that events which actually occurred within the same system would be considered as separate events (see final comment for reviewer #1).

Finally, we agree that the percentile is computed mixing different types of losses, however these losses are registered as one category identified by the insurance company (storm vs. flood vs. hail event). Therefore, we would like to keep the order of the analysis and sections, but for further clarification, we have included a sentence in the manuscript after L163:

> Although all hazards follow strong seasonal patterns (see Sect. 4.2), percentiles are computed on an annual basis. This is because events and clusters can span over the limits of seasons, and because with an annual definition, the seasonal pattern within the loss data is kept.

RC: ***The identified clustered hazards are physically relevant, which is reassuring, but one may wonder whether such an analysis was really necessary to derive the results. An interesting question regarding these events is the role of decadal variability, which is hard to infer with less than 40 years of observations. The identified clustering may be explained both by the fact that the clustered hazards derive from the same weather situation and by the fact that those weather situations occur more frequently during certain decades compared to others. This should be considered when analyzing trends too.***

AR: We thank Sylvie Parey for raising these points. We would like to split our answer in three parts.

1. **Relevance of clustering analysis vs. seasonal patterns:** Our results on the degree of clustering, as measured by Ripley's K, are somewhat surprising: We find significant clustering compared to random samples for the combination of extremes during 1986–2023, but not for all single hazards (Figs. 6–8). This cannot be identified from the mere occurrence of hazards (e. g. Fig. 3). To make this clearer, we have rearranged parts of the Abstract (L5-10):

   > Results show that clustering is significant only for certain hazard types compared to a random time series. However, clustering is robust for a combination of multiple hazard types, namely hail, mixed or pluvial floods and storms. This particular combination of hazard types is also associated with higher losses compared to their isolated occurrence. Clusters of damaging hazards occur mainly during May–August and depend on the method of defining independent events (Peaks-over-Threshold with flexible lengths vs. Hours Clause with fixed lengths) and their resulting duration.

2. **Clustered hazards potentially deriving from the same weather situation:** To avoid that the identified clustering derives from the same weather situation, we performed a declustering with the Peaks-over-Threshold and Hours Clause methods (see Sect. 2.2). This declustering aggregates daily losses to multi-day events which ideally derive from the same weather situation. We have already included this in a detailed description in Sect. 2.2 (L163ff), to which we would like to add a sentence after L165 to make it clearer:

> Furthermore, as a prerequisite for applying extreme value statistics, the events are required to be independent. Toward this end, and to avoid clustering on the timescale of synoptic systems (around 5 days), clustering on the timescale of a few days needs to be removed (Wilks, 2006). This is called (runs) declustering (Coles, 2001) and means, in our case, that the daily data are aggregated to events with a length of either one or several days. Thereby, we avoid that events from the same synoptic cause appear as distinct events and lead to artificial clustering.

We acknowledge that in case of persistent patterns, several multi-day events may occur during the same weather situation. To test for robustness, we will therefore include a sensitivity test using the Peaks-over-Threshold method (POT) with $r = 3$, i. e., independent events defined when they are separated by less than 3 days. With this increase of days between events, we reduce the probability of events from the same weather situation to be wrongly considered as separate events. We expect little change in our results.

Furthermore, the average duration of an event is much higher using the HC method compared to POT (see L436f). This longer event duration, and therefore, also higher probability of independence, does not reduce the degree of clustering (see Figs. 6-8). This contradicts the argument of events being clustered due to the same physical causes. We would like to include this argument after L441:

> This furthermore proves that the (multi-day) events identified are not deriving from the same weather systems, since the degree of clustering is higher with longer durations (HC method) than with short ones (POT).

Lastly, a large part of the events during May–August (where our focus lies) is of convective nature, which is a short-lived phenomenon.

3. **Weather situation occurrence over decades:** We agree with Sylvie Parey that it would be interesting to analyse decadal variability, but also agree the time period is too short to perform an in-depth analysis. First, we would like to refer to Fig. 4 and the corresponding section for our discussion on inter-annual variability of event occurrence and to Fig. 11 and the corresponding section for our discussion on trends. To complement this discussion, we would like to include an analysis of the large-scale circulation by evaluating the event occurrence in relation to the North Atlantic Oscillation (NAO). We would like to discuss this in Sect. 5.5 and add two Figures in a supplement. As shown by Fig. 1 at the end of this document, events occurred most frequently during a negative NAO in May–August from 1986–2023. Particularly in recent years (since 2010), the NAO has been mainly negative during May–August. For events during December–February (see Fig. 2 at the end of this document), the opposite is the case. Synoptic storm events, which are most frequent during this season, occur predominantly within a positive NAO, which has become more frequent in recent years. This provides an additional explanation for the upward trend evaluated in Fig. 11.

0.1. Technical corrections

RC: ***Technical corrections: Line 18: "Damage by those hazard": hazards; Line 98: the closing bracket should be removed after 36 000 km2; Line 110: "onlyf" is written instead of "only"; Line 126: "Given the the different environmental conditions" 2 instances of "the", one should be removed; line 472: "(see Fig 8" the closing bracket is missing; line 570: "It should furthermore not be neglected is that there is a stochastic element": "is" should be removed***

AR: We thank Sylvie Parey for the technical corrections; we addressed and implemented all of them in the main text.

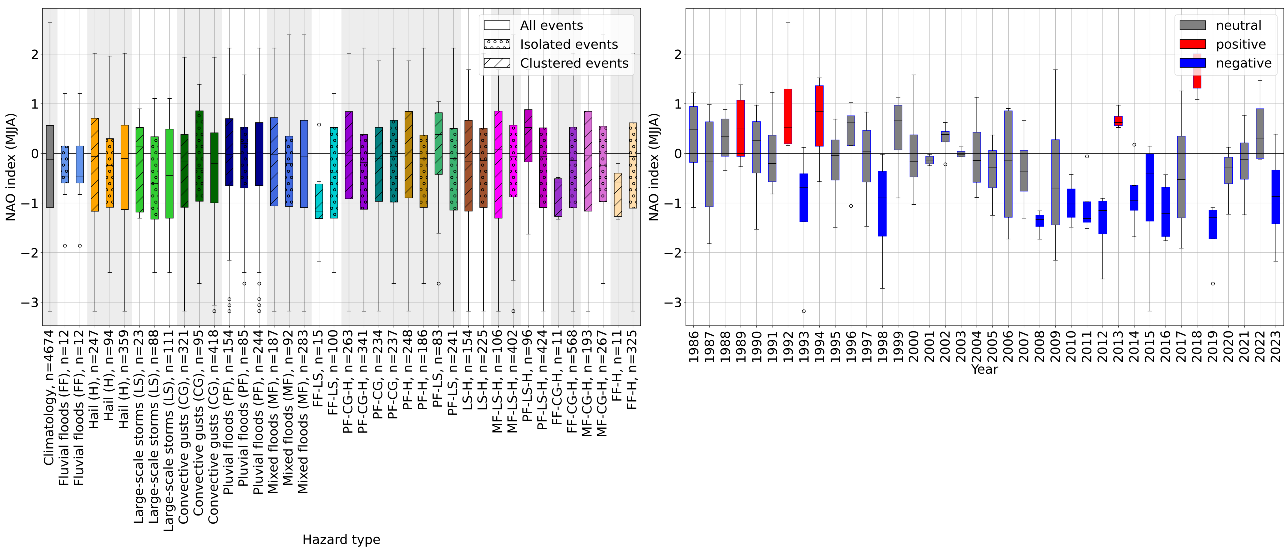

Figure 1: May–August: (a) Distribution of monthly NAO values during 1986–2023, depending on the event type (colors), isolated occurrence or occurrence in clusters (hatches), (b) monthly mean NAO values from 1986–2023. Positive NAO values are detected when with mean values > 0.5 and max values > 0.75. Negative values relate to mean values < -0.5 and max values < -0.75. Neutral years are all years neither classified as positive nor as negative.

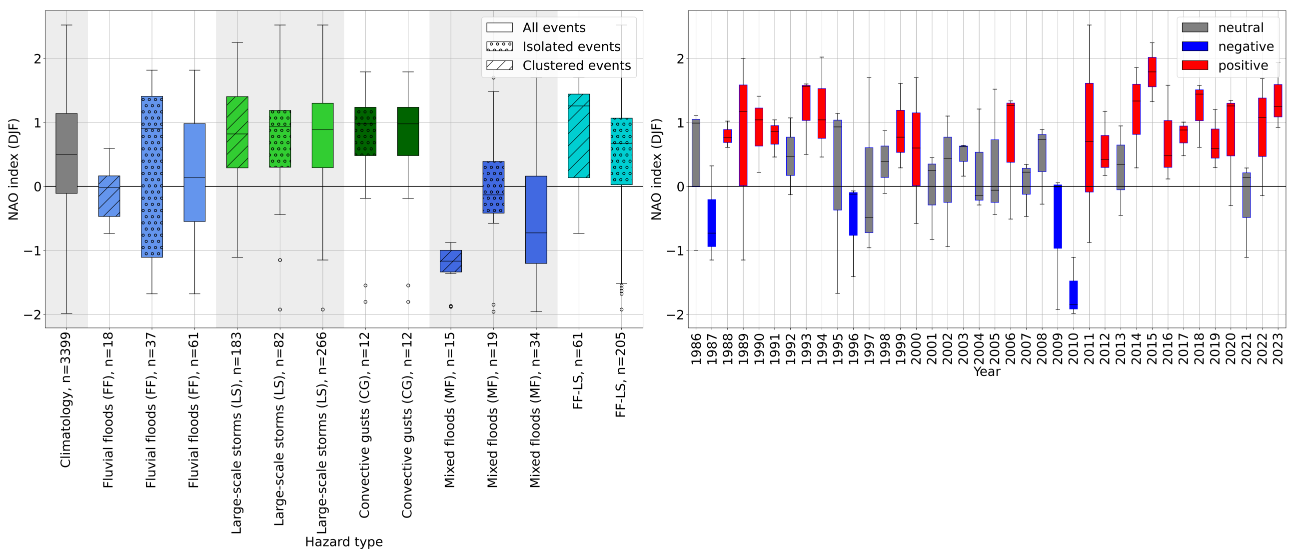

Figure 2: December–February: (a) Distribution of monthly NAO values during 1986–2023, depending on the event type (colors), isolated occurrence or occurrence in clusters (hatches), (b) monthly mean NAO values from 1986–2023. Positive NAO values are detected when with mean values > 0.5 and max values > 0.75. Negative values relate to mean values < -0.5 and max values < -0.75. Neutral years are all years neither classified as positive nor as negative.