

*First, we want to thank both reviewers for their comments and suggestions. On their basis, we propose several changes in the manuscript. Generally, we intend to adjust our methodology to the method of Uhlemann et al. (2010) and Schröter et al. (2015) for our study area, compare different discharge limits and present a set of extreme floods with their main spatial and temporal characteristics. Below, we describe the proposed changes in more detail.*

#### **Referee #1 comments:**

Referee: The paper addresses the topic of identification of flood events in western and central Europe based on multiple series of mean daily discharge for a period of 61 years. Building on the method of Uhlemann et al. (2010) the authors aim to highlight the influence of different parameter choices on the the flood severity index which in turn influences the event identification. Extending the analysis of Uhlemann et al. 2010 to a larger study area, i.e. identifying a true central European event set of large flood events is a valid objective. However, the paper lacks scientific rigour and presents little novelty on the event severity assessment. I therefore do not recommend the paper for further publication in HESS. In the following, I will outline my key criticism and encourage the authors to commence with their research on the important topic of understanding flood event frequency, severity and causes in central Europe. Uhlemann present a thorough sensitivity analysis of the severity index already and present the impact of different thresholds and input data on the resulting event set. The work presented by Gvoždíková only addresses the sensitivity of two parameters: subcatchment area and flood discharge limit (which is a threshold of flow expressed as the ratio of the peak flow at a gauge against the mean annual maximum flow of the entire time series at that gauge). The selection of the parameters and the chosen range in which they are being tested is not supported by argument. I.e., what is the hypothesis for defining the three variants of the thresholds of what is called the discharge limit ( $Q_s/Q_{ma}$ ;  $Q_s/Q_{ma} > 1.2$ ;  $Q_s/Q_{ma} > 1.5$ )? Likewise, choosing either the subcatchment area, its root or logarithm as range for testing the impact of the spatial weight on the severity index is arbitrary.

*Authors: Thank you for your comment that our main objective – evaluation of flood extremes without limits of state borders – is valid. Our motivation for modifications of the Uhlemann et al. (2010) index was the future comparison with precipitation extremes where the affected area is a necessary parameter. Nevertheless, we accept your comments regarding the methodology, see our next response. The threshold will be set to return levels in the reconstructed paper.*

Referee: The event sets identified are limited in a first step to 80 events, and then, for comparing results of the different variants of the index, to 30. Why is that so?

*Authors: The first selection was done in order to eliminate less extensive floods. The set of 30 extreme floods is finally presented as we wanted to select one flood per two years on average.*

Referee: My strongest criticism is with the principle choice of subcatchment area as spatial weighting factor to account for the relative contribution of a peak recorded at a river gauge to the overall flood severity. This is a classical regionalisation problem in hydrology. Subcatchment area however fails to address this problem and introduces a severe spatial bias into the analysis. Unlike precipitation, for which area indices are well suited, floods are not a space-filling phenomena. In particular, peak flows recorded at downstream gauges of the large streams Rhine, Elbe, to some degree Danube, Weser, and Meuse are in most cases not caused by inflow from their intermittent catchments but are a result of the flood wave propagating from further upstream. Also, when choosing subcatchment area, the density of the gauge network and particularly the uneven distribution of gauges in the river network becomes relatively more important in the severity index calculation and needs sensitivity testing. The original

severity index presented by Uhlemann 2010, and also the application of the index in Schröter et al. 2015, provide a method for regionalisation of peak flows to the river network rather than to the subcatchments. I strongly recommend the authors to review their method for computing the spatial extent of the flood events in any further study.

*Authors: We fully accept your comment and intend to recalculate our results with respect to the methodology by Uhlemann et al. (2010).*

Referee: In summary, the conclusions reached on the best suited variant of the severity index and resulting event set need thorough reworking. In fact, I think, that the sensitivity study provided in Uhlemann et al 2010 provides all the necessary findings to allow for a fairly straight forward adoption of the severity index to the context of identifying flood events in central Europe. In the final paragraph of the paper the authors highlight that they want to commence with an analysis of the hydro-meteorological causes of large flood events in central Europe. I think, this is where the innovation will come and I highly encourage the authors to proceed on this avenue. The assessment of the severity of events and consequently the identification of the relevant flood events in the region can be natural part of any paper submitted on this.

*Authors: Yes, analysis of hydrometeorological causes of central European floods will be the next step of our research. Though after accepting the Uhlemann's methodology the paper will not be innovative from the methodological viewpoint, we are still convinced that after basic changes in the methods, the results are valuable for publishing in HESS. The focus of the reconstructed paper will shift from the methodological issue to the analysis of the set of flood events. However, the discussion of the role of threshold needs to remain a substantial part of the paper because, as demonstrated in our Fig. 6, the thresholds play an important role in evaluation of floods within the large central European region.*

Referee: On the aspect of analyzing severe floods in W/Central Europe in their frequency and severity and also in their spatial-temporal patterns and potential changes of these + attribution of these changes to causes: Reading the title I had expected to see the Odra basin included in the study. This basin forms the eastern boundary of the very wide transitional zone between atlantic and continental influences on flood genesis and at present I expect that in particular some of the summer flood events are insufficiently represented in the event set(s). Also, extending the event analysis to the most recent period, e.g. 2015, would add value to any change detection and finally attribution. In principle, I think Central Europe is the better description for the area under study.

*Authors: We agree that Odra river should be studied together with other central European rivers. In our original paper, Odra river basin as well as several last years were not included because of lack of available data. The situation changed nowadays so we will be able to extend the study area and partly also the study period according to your comment.*

### **Referee #2 comments:**

Referee: The study presented in this paper aims at redefining flood extremity indices over a large region between Western and Central Europe over the period 1950–2010. The approach followed consists of designing flood extremity indices by combining discharge values and the spatial extent of floods. Several versions of such indices were tested, with different weightings of the threshold value of discharge or area parameter for considering a flood event. The topic is suitable for publication in HESS but major revisions would be necessary in my opinion before the paper be published. General: - The paper lacks a discussion on the consistency of the choices to be made for designing the extremity

indices (determination of  $Q_s$  and the threshold  $Q_s/Q_{ma}$ ). There is almost no discussion about this point which constitutes the basis of the whole approach. Also, it seems from the discussion/conclusion section that the main difference between this work and previous other ones upon which the present study builds relies on the choice of the threshold selected for discharge: 1) this emphasizes even more the importance of strengthening the discussion on criteria for choosing the best suited  $Q_s/Q_{ma}$  threshold, 2) it questions the value-added of including an area parameter in the approach (the authors themselves state that extremity indices are not very sensitive in changes in the area parameter: if so, then this approach is very similar to previous ones?).

*Authors: Thank you for your comment. As we said, we propose to change discharge limits and use return periods for determining the event sets. Also, the subcatchment area will be replaced by the length of river sections of certain order.*

Referee: Conducting a more detailed study on the determinism of the occurrence of flood events seems important in order to relate the extremity indices defined to more concrete or practical hydrological/hydrometeorological processes (in this sense it is surprising that the role of ground water is never even mentioned), it should be addressed here and would certainly constitute the value-added to other previous studies such as those of Uhlemann et al., etc. As a first step, the authors could try to relate the interannual variability and trends of extremity indices to some climate indices for instance.

*Authors: The climatology of extreme flood events is just the first step of the research. The distribution of extreme flood events can be compared e.g. to some drought indices. The role of groundwater will be mentioned within the introduction section.*

Referee: Specific comments: - Title: Something like "large spatial extent floods" or "extensive floods" (as used in the introduction) could be included in the title to be more specific as it is an important aspect, and would prevent from using "trans-basin" which indeed could be misleading?

*Authors: We agree, that the term "extensive floods" could be included in the title.*

Referee: P.4, line 8: I don't get why only the downstream sub-catchment area is considered when an upstream gauging station is available. The downstream station is still representative of flow occurring over the whole upstream area anyway unless the upstream part of flow is subtracted.

*Authors: You are right. One possibility would be to subtract the upstream part of flow. On the other hand, the actual discharge at certain station cannot be ignored, as it is related to severity of flood in that place. However, on the basis of other comments, we propose to use the length of river sections of certain order instead of subcatchment areas.*

Referee: P.4-5, "Methods": I do not recommend using the word "significant" in this context, as this does not refer to any statistical meaning here. I think the rationale for using this method to determine a time series of "significant" discharge values lacks explanation.

*Authors: We agree and remove the word from the text. Moreover, when we use return periods as discharge limits, the identification of what we call "significant" mean daily discharges will be released.*

Referee: As well, I am concerned by the approach for determining a significant flood event: the choice of the length of the time window needs a little more explanation. As is, it looks like the method suffers from a lack of either statistical or deterministic basis, and the definition of a flood event seems to be too much data- and operator-dependent and not enough transposable (see for instance "After analyzing all of the data series, we chose a time window [...]"). The fact that the time window had to be extended for one river, or that an additional rule had to be included to prevent merging events that have different atmospheric origins is also problematic: is an automatic split of flood events in two parts when they are separated by 5 days enough to conclude to different atmospheric causes ?

*Authors: The time window of 10/12 days was used because of time of propagation of flood waves downstream. For example, during the event of March 1979, it took 11 days from first detected peak to the 10-year discharge value observed at Ketzin station on Havel river. To simplify it, we propose to use single time window of 12 days before and after the observed 10-year discharge. If two or more flood waves occur, it is clearly visible in time series - two peaks greater than 10-year discharge appear at several stations and the distance between these peaks at each station was at least 5 days. E.g. it was the case of August 2002 flood, when two flood waves were detected, which corresponds to other studies about this event.*

Referee: P.5, line 19: does the separation date between the cold and warm halves of the year also hold for other regions than the Czech Republic?

*Authors: We decided for this separation mainly because of flood in 1998, which started on October 29. This flood clearly belongs to cold half year floods due to its meteorological causes. But I think that we could use more classic division when the cold half year will start in November and this particular flood can be listed in cold half year as the major part of the event occur in November.*