Reply to interactive comment by Anonymous referee #2 on "Flood type classification and assessment of their past changes across Europe"

by Yeshewatesfa Hundecha, Juraj Parajka, Alberto Viglione

Assessments of flood change and potential classifications of floods across large regions, such as the continental scale analysis used here, are generally highly valuable contributions to the literature and science. Therefore, I believe the topic of the manuscript is relevant to HESS. However, the manuscript itself needs considerable revision to address some confusing parts of the methods, analysis, and presentation of the results. For this reason, I recommend the manuscript be reconsidered after major revisions and re-review.

We thank the reviewer for the thorough review of the manuscript and thoughtful comments. We will attempt to clarify the issues raised.

Major comments:

1. Explanation of the Data and State Variables:

a. Section 2.1 could be better communicated as a table listing the variables and the data source. The text can then be used to explain any additional pieces of information of particular note with the data, such as the fact that landuse and soils come from different sources (p. 4, lines 20-21). Also, the acronyms are not introduced before they are used.

We thank the reviewer for the suggestion. In response to this comment, we will add a table that will summarize the datasets used in the study together with source links and references.

b. I assume that the "hydrological model" being referred to on p. 4, line 17 is E-HYPE but that model is not named anywhere in this section. In Section 2.3, the model is discussed in more detail but this information is buried under a heading that indicates only the hydrologic and state variables are discussed. Consider renaming Section 2.3 and adding additional subsections to explicitly discuss the model, its input variables, and how the model is used in the analysis. These important pieces of the methods are currently not clear.

Thank you for the comment. In response to this suggestion, we will rename Section 2.3. to "Hydrological modeling and flood event characteristics" and add an additional subsection that will describe the E-HYPE model and its implementation in the present work.

2. The clustering/classification by flood type is a main contribution of the manuscript. However, in Sections 2.4 and 2.5, the description of how the authors arrived at their flood classes needs substantial improvement. Here are examples of where I found these sections highly confusing:

Reviewer #1 has also raised points related to description of the clustering methodology. We will thus revise/describe the methodology in a more detailed way.

a. Section 2.4 discusses the clustering of flood events and yet there are not details explaining how the clustering was quantitatively carried. No information is given as to how the stations were "clustered in space" p. 8, line 4. What variable was used to cluster? What method was used to cluster? Section 2.4 actually appears to be discussing how sites that may be exhibiting correlation in flood events due to their proximity to one another were filtered out - not anything about clustering of flood events.

Grouping of the stations was performed based on peak events at the individual stations. Stations that are grouped into one spatially clustered event could be different for different spatially clustered events. A set of criteria were used to spatially group events at different stations. We defined flood events such that the events can potentially have impact. We set a minimum threshold of the 2 years flood at each station as an approximation to the bankfull flow to delineate the flood extent and put additional criterion that a potentially impact causing flood level occurs at least at one location. This is defined as the 5 years flood in our work. For each peak event exceeding the 5 year flood at each station, nearby stations with peaks exceeding the 2 year flood were searched for. Whether the stations fulfilling this criterion are grouped to form a spatially clustered event is decided based on the temporal lag of the peaks at the stations and the spatial distance between the catchments draining to the stations, as discussed in Section 2.4.

b. Section 2.5 opens with a definition of 4 classes of flood events. This would indicate that the flood classes were determined a proiri and not by a formal clustering method. If the groups arise from a clustering algorithm, then I would consider them results and not appropriate to be placed in the Data and Methods section. As one reads further down however, there is information about a clustering method utilized but that includes the "hydrological and hydro-meterological variables defined from the E_HYPE model" (p. 9, lines 23-24). In my comments above, I do not think enough information has been given about how the model variables are used in a clustering approach.

Yes, the classes were defined a priori. We have discussed this issue in a reply to a similar comment by reviewer #1 and we repeat it here. The clustering technique is employed to guide us into grouping the events in such a way that the resulting cluster groups have the desired distribution of the event hydrological and hydrometeorological characteristics based on our definition of the event types. We could have clustered the events based on all the event characteristics and tried to infer the classes from the characteristics of the resulting cluster groups. However, the resulting groups may not easily be defined in terms of the commonly employed flood process types. Therefore, we started by first defining the main flood generation mechanisms that can be identified based on the data we have. For some of the variables, it is difficult to define a clear border between the different mechanisms. For instance, what should be the rainfall amount to distinguish between snowmelt and rain-on-snow events? We performed clustering of the events based on their characteristics until we got groups whose statistical distributions of the event characteristics reasonably well describe the flood processes we defined rather than defining thresholds subjectively and grouping events based on such thresholds. This involves using different combinations of event characteristics in the clustering algorithm.

c. If the classes resulted from the application of a classification algorithm (in this case, the k-means algorithm), no evidence is given as to how the classification tree was pruned and how these classes were assigned a common behavior such as "short-rain floods." From the description of the flood types, it seems "short-rain floods" are defined as "a flood event caused by rainfall of duration less than one day" (p. 9, line 11). How was this definition arrived at - by looking at classified events to determine common properties or was this a pre-determined definition applied to the flood events.

Please refer to the reply to the previous comment.

d. Following on this, p. 9, lines 30-32 note that after the classification was complete, "manual adjustment" was used to move events around from group to another if they "happened to end up in a group which doesn't reasonably represent them." The authors need to provide objective criteria here as to how this was assessed. Since the remaining part of the manuscript centers around this classification, how can a reader be ensured the results are not biased by these initial adjustments? What was the point then of using a classification algorithm in the first place?

As we discussed in the reply to previous comments, the clustering technique was used as a guide to enable us classify events by avoiding subjective thresholds. The final cluster groups we arrived at have the desired statistical characteristics of the event characteristics based on the definitions but some of the individual events in a certain group may have event characteristics that are counterintuitive to the

way the events are defined. For instance, an event with no snowmelt may end up in a group that represents snowmelt or rain-on-snow event. That is why we had to examine the events in each group and move around events accordingly. We used simple logical rules to move events around. We will describe these rules in more detail in the revised manuscript.

3. Because it is unclear how the flood classes were arrived at, the novelty of this work is not apparent. It would be more useful to pose the manuscript as a testing of several hypothesis about flood generating mechanisms or classes using the current state of the literature as support rather than general objective statements such as those found in p. 4, lines 1-4. I am not clear even as to whether the flood typing classes are a contribution because I am not sure if they are determined from the data or imposed by the authors to perform subsequent analysis.

Based on our replies to the previous comments, we hope that our objective is clearer now. Our objective is to classify past flood events into flood types that are commonly discussed in the literature based on the characteristics of the meteorological drivers and hydrometeorological states across Europe and study the regional differences in the dominant flood generation mechanisms and the temporal trends. We didn't intend to introduce a novel methodology for flood type classification.

4. Figures 3-7: These figures should be stand-alone. Referring back to previous captions decreases the readability and interpretation for the reader. I think it would also be helpful to show boxplots next to each map of the flood events grouped by region to show the distribution of the flood event types.

Thank you for the suggestion. We will make the captions standalone in each figure and add distributions of the flood types in Figures 5-7.

5. Figure 9: This figure is not well-explained and need clarification. Was a regional Kendall test used to obtain the significance values? If so, this is not cited or defined in the methods section. The conclusions made based on this analysis (p. 20, lines 1-8) do not reference any specific figure or evidence for these statements. This needs to be remedied.

The Kendall test was performed on the regional counts of the different types of flood events and this was mentioned in section 2.6. The results of the trend test are presented in Table 3 and Figure 9 shows, for each region, the types of flood events that have shown significant trend. The conclusions on page 20 are based on figures shown in Table 3. We will make reference to the table.

6. Section 4: It is very difficult to follow some of the statements made in Section 4 (p. 21, lines 6-13 for example) because Figure 8 is so difficult to understand. I do commend the authors on using the discussion to pull together the literature on flood change and typing from smaller regions within Europe and describe how those studies fit with these results. I recognize synthesizing these results into the text was not a simple task.

Thanks for the comment. From line 8 on, it is actually a new paragraph and it is meant to discuss the characteristics of the different event types in terms of the meteorological and hydrometeorlogical characteristics. We will use subsections to enhance readability of the section.

Minor comments:

p. 4, line 15: The authors note a variable data period without mentioning how actually variable the periods are. Please note at least the minimum data period allowed.

The data obtained at the different stations are of different length. Some date as far back in time as 1812. The analysis in this work is based on a common period 1961-2010 and stations with at least 90% data over this period were used. We will add this information to the manuscript.

p. 2, lines 22-24: This sentence is quite confusing as to what is meant here. The use of “different” twice creates most of the confusion.

Thanks for the comment. We will reformulate the statement to “Studies have suggested that flood regime has been changing in Europe over the last decades, although the change pattern has been found to be regionally different”.

Figure 8: Show the flood-type names instead of numerical values. This helps the reader to better understand the relation between the variables and the flood types.

Numbers were used not congest the figure area with texts as the names do not fit in an optimal way. We will show the flood types in a short form.