I really liked to read this paper on the compilation of global flood hazard maps using 'physically based' model cascades since I am currently involved in a comparable endeavour even though on a much smaller scale. Still there are some issues I would like to comment on and recommend to improve understandability.

We thank Kai for his review and are grateful for his kind words.

Without question, uncertainties and limitations are an important aspects to address. Given the numerous sources of uncertainty involved in the application of a cascade of different models and underlying data I think it is important to name these sources in detail and to discuss their potential implications on the results. Likewise the relative contribution of the different uncertainty sources should be ranked at least in a qualitative way. Otherwise the uncertainty bounds included in Fig. 4 and table 2 (which only refer to the uncertainty stemming from the extreme value statistics, right?) run the risk to draw an exessively optimistic picture of the performance of the overall approach.

We fully agree with the reviewer and will make sources of uncertainty more clear. The uncertainty bounds are definitely too optimistic and we will make this even more clear in the revised manuscript. We cannot provide any ranking regarding their importance, simply because we do not know and it will depend largely on the spatial and temporal domain. We do not have sufficient observations and information to do so. However, this is an important discussion and we will raise it in the revised manuscript including a table describing potential sources of uncertainty, examples and a qualitative potential ranking.

Further the labelling of figures 5 and 6 is confusing because the colour palette does not include the colour white. Presumably these cells are below the minimum thresholdof 5% as mentioned in the text (page 6625 line 6).

Correct - will be changed

The comparison to the benchmark data is discussed using several performance scores. In the text it is stated (page 6628 line 2) that ETS is above 0 for all return periods, i.e. 2 to 500 years. However the corresponding figures 7 and 8 are limited to return periods of 75 and 80 years. Is this due to a limitation in the benchmark data?

Again correct and will be clarified. Thank you for this comment!

The use of the term rating curve on (page 6632 line 13) should be supplemented with a note on the data base or methodology of how these 'rating curves' have been derived. This is to avoid misconceptions with the term rating curve at a gauge.

Thank you for pointing out this issue. In fact the term "rating curves" was misused in this context. It will be replaced by a clear explanation of the transformation between river water storage and river water level.