Reply to RC2 (reviewer comments in grey)

Review of Wasko et al, "A systematic review of climate change science relevant to Australian design flood estimation"

Overall assessment

This is an exceptionally thorough and well-written review, which I think will be of interest to a wide readership. The authors methodically review the different approaches for design flood estimation, with a specific focus on the Australian context. They apply objective methods to assess the level of consensus of different sources of evidence. The findings are mostly uncontroversial, and some are quite important, such as the fact that no defensible method currently exists for factoring climate change into flood frequency estimates. The summary of changes in rainfall with warming is particularly interesting, as is the table summarising all the findings (Table 4).

We thank the reviewer for their assessment and thoughtful comments. Please find below our reply in blue and proposed modifications in underlined.

Main comments

L.81. The authors state that there has been "little research undertaken on the non-stationarity of other factors affecting the design flood estimate". There has been considerable research on a range of nonstationary factors affecting large floods, so this statement seems a little surprising. Does it matter that the research has focussed on nonstationary factors affecting general floods rather than design floods? Presumably the challenge is more about understanding how to incorporate the nonstationarity in the design estimates, rather than a lack of research on nonstationary factors.

The reviewer is correct. Our intention was to state that most research does not explicitly focus on the impacts of climate change on the inputs into design flood estimation. This sentence would be reworded to explicitly state that although much research has been undertaken into the factors affecting flooding, they are often not directly comparable or translated to the inputs used in design flood estimation using event-based models (the exception to this being peak rainfalls).

L.97-145. Section 2 is a very helpful overview of the main methods, but it provides almost no citations. I think it would be helpful to include at least the key references for each of the methods described. It might also be helpful for the reader to provide a summary table of the strengths/limitations of each method (I leave this decision with the authors).
The paper reviews different methods for design flood estimation based on streamflow and based on rainfall. At various points in the review, the authors seem to suggest that the streamflow-based flood projections are either too complex (L.1040) or too uncertain (L79) for practitioners to apply. I wonder if it might be worth reflecting a little more on the adequacy of different approaches, and the limitations of the simpler precipitation-based approaches (which make up a very large part of the review). They may work well for certain types of flooding (e.g. pluvial floods in an urban context), but is it fair to say they may not work so well for fluvial flooding, where other factors such as changes in groundwater storage need to be taken into account?

In response to the reviewer comments, we would:

(1) Add key references to each of the methods (i.e., continuous simulation, flood frequency analysis, and event-based simulation).

(2) Lines 142-145 would be expanded to discuss the strength/limitations of each method (this would tie in with Section 5.3 where the limitations with respect to climate change are briefly discussed).
(3) To the above point and the reviewer's suggestion, Line 1006-1015 would be expanded to explicitly discuss the adequacy but also limitations of each method in the context of climate change. Particularly at Line 1007 where we state there is a lack of consensus on the best method to adopt under climate change, but do not explain why this is the case.

Pg.17-18. It is clear that a lot of effort went into producing the 'best estimates of central scaling rates' in which the authors independently assigned weights to different studies to arrive at a median estimate and 66% range. The Supplementary table is an impressive exercise in collating and trying to synthesize information from 179 different sources. However, it is a little difficult to see how the summary values in Table 1 were arrived at, since the weights are not provided and the thresholds shown in Table 1 (<1hr or >24hr) don't match those shown in Figure 2 (<6hr, >12hr) (perhaps I missed why Table 1 and Figure 2 differ). It would be helpful to provide more data on how the values were obtained, so it doesn't seem too subjective.

Agreed. A table will be added to the supplementary material of the individual researcher quantitative results and a further summary of the methods of averaging adopted by each researcher would also be provided. In addition, the labelling on Figure 2 will be updated to match Table 1.

Minor comments

L.99 "the primary differences between methods relating to where in the causal chain of flooding the data are obtained, and where the probability model is fitted" could be rephrased for greater clarity, perhaps being more specific.

Agreed. This sentence would be removed and replaced with a brief introduction to the design flood estimation methods to help introduce the text that follows.

L.139. The term "efficacy" in the caption of Figure 1 is a little ambiguous and could be clarified.

Agreed. Efficacy would be replaced with words that emphasize that in the highlighted AEP ranges the methods show the most utility.

L.172/183 "average effect size" – please specify the effect size of what.

This will be replaced with "magnitude of extreme rainfall depth change" or words to that effect.

L.177 "with variation to storm duration .. and location preserved" could be rephrased for clarity.

Agreed. These factors were considered as additional variables, and the text will be amended to make this clear.

L.186 "was weighted by"?

This typo will be fixed as per the reviewer's suggestion.

L.206-211. Recent work has shown that groundwater is a more important driver of flooding than either antecedent soil moisture or antecedent extreme rainfall (see work by W.Berghuijs); this may be worth mentioning.

The text will be amended to add the reviewer's suggestion.

L.237-240 "even the use of physically-based covariates is problematic as the covariates should capture the differing processes": please clarify/elaborate. Also, which of the "statistical associations may not remain constant with climate change"?- it is worth being more explicit.

Here we wished to state that although the covariate may be physically linked to the processes of climate change, if the covariates does not explicitly model the physical relationship driving the change, then extrapolation may result in incorrect estimates. And even if this is physical relationship is captured, there is still no guarantee that the governing processes will be the same in the future. The wording will be rephrased to make this clear.

L.432 "their application to the future period remains untested": the phrasing is a little odd; application to the future is always untested. Do you mean their predictive ability (to predict out of sample) is untested?

This is correct. The suggested rephrasing will be adopted.

L.957-961. I am not sure how helpful it is to tell the reader that the papers were assessed independently and through weighting of evidence, if the outcomes of those analyses are not presented. It's a bit like asking the reader to simply *trust* that the analysis is objective.

Agreed. As per the response above, a table will be added to the supplementary material of the individual researcher quantitative results and a further summary of the methods of averaging adopted by each researcher would also be provided. In addition, the labelling on Figure 2 will be updated to match Table 1. This section would then reference the additional results table.

L.1068 "the impact on rare floods diminished" could be rephrased for clarity.

This will be rephrased to state that the impact on rarer floods is lesser.

Figures

Figure 1. Nice summary figure. The labels of the x-axis could explained in the caption. The "S" shape of the curve could also be explained.

Figure 2. Again, very interesting (and novel) summary figure. I don't think 2xCC and 3xCC are defined.

The omissions referred to by the reviewer will be added.