

## **Response to RC3:**

### **S. Sippel et al., 2016**

The authors examine the robustness to choices made in the analysis of a recent analysis of observed trends in precipitation in dry regions around the world. In general I quite like this, as results of studies are usually interpreted beyond the specific experiment design of the analysis, and so this paper performs the important task of determining the extent to which it is possible in the case of observed precipitation trends in dry regions. However, I think there are a couple of additional aspects to this that the authors have not considered, as well as one important syntactic issue, that I believe need to be addressed before publication.

#### **Thanks for the positive evaluation!**

First, the motivation you frequently mention is for informing adaption decisions. For that motivation, though, it is not clear to me that what you do in terms of normalising to the full period is necessarily any more appropriate than what Donat et alii (2016) did. Many decisions are based on climatic or hydrologic data from a specific time period, for instance in the case of international treaties allocating water on an international river. Thus adaptation decisions need to be made with respect to divergence from that reference baseline (ignoring non-climate stuff). So while e.g. you may be correct that there has been no actual trend in precipitation totals, say, that does not necessarily mean there has not been a trend as measured by stipulated monitoring procedures used in many decision-making settings. Cast another way, we have the same problem in dealing with future climate change: projections are based on, say, the full historical period you use, but that does not include the future itself. I expect you are not arguing that we cannot make useful projections of future climate change simply because we have not monitored the future yet. In this context, I laud your effort because you highlight the sensitivity to this point, but I think it is important – and entirely consistent with your consideration of robustness – to emphasise that there is not necessarily a single "correct" answer.

**Thanks for this point - indeed we are not arguing that no useful policy choices can be made just because the future is not part of the reference period. This would be somewhat nonsensical and clearly overstate the problem. However, for the latter problem, please note that it is not just an issue of random variation of the reference period - it is indeed a systematic bias that is 1) positive outside the reference period, and 2) scales with the relative change in the time series. The expectation value of the bias is a function of the reference period length, and the mean:standard deviation ratio (or location:scale ratio and shape parameter in a GEV). Please**

have a look at the analytical approximation (attached pdf-file), which we intend to include in the Appendix or Supplement of a revised manuscript.

In many practical settings the bias might be small (e.g. if the ref. period is long enough, or the mean:standard deviation ratio is high), but in other cases it might be relevant for example because the trends in the quantity of interest might be small. In our particular case, we have chosen to normalize to the full period because this avoids the bias. However, in a hypothetical case where one would be bound to a given baseline period (for instance in your example where an international treaty would define climate change or its impacts relative to a specific baseline), one could still take our analytical results and estimate the expected magnitude of the biases relative to the observed trend. Thanks again for this hint, and we will emphasize in the revised manuscript that there is no correct answer for the choice of the reference period, but with either avoiding a "reference - non-reference breakpoint" or by analytically adjusting for it using e.g. our approximation (which also works if there are in-stationarities in the time series outside of the reference period).

Second, in terms of all of the discussion about what constitutes a "dry" region, the most striking thing to me is that none of the definitions of dry regions you consider include what I think of as prototypical dry regions: the Sahara, the Saudi Peninsula, Central Asia (particulaly for Rx1D), southwestern Africa (other than South Africa), western Australia, northern Mexico (for Rx1D), nor the driest areas of South America (for Rx1D). The reason for this of course is monitoring coverage, but given the absence of all of these regions (the Sahara!) I do not think these results can plausibly be considered as being indicative of how precipitation is actually changing over the world's dry regions. Again, I consider this a point about robustness that is entirely consistent with your paper, but it most certainly needs to be acknowledged/noted/highlighted.

**Yes, we agree (please see also additional analyses in the arid data-scarce regions and more detailed reply to Reviewer #2 who made a similar comment). In a revised manuscript, we will this point in the text, Abstract, and Conclusions, and also move Appendix A (Figure 0) in the main manuscript to ensure that even a casual reader will encounter the aridity and dryness maps, including the gaps in spatial coverage.**

Third, on the syntactic side, while the title refers to precipitation and it appears to be precipitation you are actually analysing, within the text this is generally referred to as "rainfall". Please clarify which you are examining, as these are certainly not identical for annual totals (and, if

defined in certain ways, for heavy extremes) in many of the regions you examine.

**Thanks for this point. "Rainfall" and "rainfall extremes" are indeed used erroneously and we will change it to precipitation in a revised manuscript.**

Specific comments:

page 1, line 1 The title says you are examining precipitation extremes and annual totals, but here you indicate it is rainfall. Which is it? It seems to be a precipitation dataset you are using, so it looks like the usage of "rainfall" is wrong?

**Yes it is, please see above.**

page 2, lines 24-25 If they are being underestimated, then it sounds like the errors are not completely cancelling, right?

**Not completely, correct. Will be rephrased in a revised manuscript for clarification.**

page 2, line 25 "These results": Which, Donat's or yours?

**Actually both, but we'll clarify.**

page 2, lines 25-26 I think such an assertive statement concerning the decision-making processes utilised in dry regions requires some supporting evidence, e.g. to other research on decision-making in those regions.

**That is true. We have rephrased our statement to reflect that an accurate quantification of change in precipitation characteristics (which includes monitoring, etc.) is important because it is simply a prerequisite to be able to make climate change adaptation decisions.**

page 3, line 9 "in both time periods" -> "over the combined periods"

**thanks.**

page 4, lines 21-22 This is a case where if you are considering rainfall, and not precipitation, then indeed North-East Siberia is rather dry.

**This is true, but we have changed the discussion to discuss precipitation; and discuss "dryness" in terms of "low precipitation" (precipitation alone), "low annual-maximum rainfall", and "dry in terms of water availability", i.e. supply and demand (arid climate).**

page 4, lines 25-26 I do not believe that Fischer and Knutti (2015) studied the decision-making processes used by those

involved in responding to climate change, and in particular what they considered "relevant" information for informing those processes.

**Yes, true. We have rephrased to better reflect what we mean and removed the reference to the paper mentioned.**

page 6 "Figure 0" should have a different identifier.

**Yes, and the figure will be moved into main manuscript.**

page 6, caption Can you confirm that for only different between columns for the lowest two rows is the mask? I found this caption confusing, for instance with the distinction between the columns being introduced only halfway through. Subtitles on each panel could help.

**The difference between the columns for the lowest two rows is the mask and the data availability of HadEX2 for PRCPTOT (left) and Rx1d (right). We'll clarify the caption.**

Figure 2, caption line 3 By "red lines" do you mean yellow? Tables 1 and 2 What does "Period Inc. (%)" mean?

**Thanks for reading thoroughly! Red should be orange, and "Period Inc." means the relative period changes between the first and second period (i.e. 1951-1980 vs. 1981-2010). Will be corrected/clarified in a revised manuscript.**

Tables 1 and 2 Why do the "Sample size" values differ? Aren't all the trends calculated over the same number of years?

**By "sample size" we mean the number of grid cells over which the spatial averages are taken. We'll clarify.**

Table 2 There is one trend values listed as "<0.000". Why do you not give the numerical value for a negative trend? This one is interesting, because it is the only significant negative trend.

**Oh, sorry. The trend is actually almost exactly 0 and not significant. Will denote this one as "0.000".**