Review for NHESS submission, "Brief Communication: Drought Likelihood for East Africa" by Yang and Huntingford

General comments

This study by Yang and Huntingford uses a combination of historical data from a single reanalysis product (ERA-Interim), combined with models from the CMIP5 suite, to quantify changes in drought risk over the East Africa region under a high-warming scenario (RCP8.5). The study does not include validation analysis to determine whether the models being considered can in fact provide a reasonable representation of the observed climate for the region of interest – this is particularly problematic given the long-recognised absence of high-quality observational records over these regions. There have also been several other studies analysing changes to East African drought in recent years, which the authors have failed to acknowledge. Finally, there is little to no treatment of statistical uncertainty estimates accompanying the results, which is fundamentally misleading to the reader.

Therefore, my recommendation is that ***this article should be rejected***. In accordance with the NHESS review criteria, my assessment is as follows:

Scientific Significance: 4 (poor)

Scientific Quality: 4 (poor)

Presentation Quality: 3 (fair)

Specific comments

1) Why was ERA-Interim reanalysis chosen? Has there been any sensitivity analyses using other reanalysis products, or comparison with observational rainfall products for common time periods? I would like to see some demonstration as to whether the results would vary if NCEP2, JRA or 20th Century Reanalysis products were used instead, as well as some consideration of actual observational data (such as TRMM, CRU-TS, or individual station-based records).
The authors are referred to the following papers for reference:

- Angelil et al (2016, *Weather and Climate Extremes*, https://doi.org/10.1016/j.wace.2016.07.001)
- Sillmann et al (2013, *JGR Atmospheres*, doi:10.1002/jgrd.50203)

2) Why is there no broader consideration of the suitability of the climate models used, beyond a bias correction based on climatological mean precipitation? This approach to bias correction has been widely considered inadequate in the context of analysing precipitation extremes, and especially so for within an attribution context. For example, if the width of the distribution of precipitation is unrealistically narrow, than even a 'realistic' shift in the mean of the distribution would result in an overestimation of the increased likelihood of future low-precipitation extremes.
The authors are referred to the following papers for reference:

- Sippel et al (2016, *Earth System Dynamics*, doi:10.5194/esd-7-71-2016)
- Jeon et al (2016, *Weather and Climate Extremes,* https://doi.org/10.1016/j.wace.2016.02.001)
- Angelil et al (2017, *Journal of Climate*, https://doi.org/10.1175/JCLI-D-16-0077.1)

3) The authors don't mention the wide variety of papers already in the peer-review literature that have evaluated changes to East African drought under a warming climate, and in my opinion, this manuscript does not add sufficient value beyond these previous studies to warrant publication in a high-quality journal like *NHESS*. Moreover, the authors fail to consider the many other contributions beyond low precipitation which

can contribute to a severe drought. In this context, an exploration of more robust drought metrics would be helpful.
The most relevant previous studies for the authors include:

- Lott et al (2013, *Geophysical Research Letters*, 10.1002/grl.50235)
- Lott et al (2014, *Environmental Research Letters,* https://doi.org/10.1088/1748-9326/9/10/104017)
- Funk et al (2015, *Bulletin of the American Meteorological Society*, DOI: 10.1175/BAMS-D-15-00106.1)
- Marthews et al (2015, *Bulletin of the American Meteorological Society*, DOI:10.1175/BAMS-D-15-00115.1)
- Funk et al (2016, *Bulletin of the American Meteorological Society,* DOI:10.1175/BAMS-D-16-0167.1)
- Climate Central (2017, https://wwa.climatecentral.org/analyses/kenya-drought-2016/)

4) The absence of uncertainty estimates in Figure 1c is troubling. Most attribution studies provide, at the very least, bootstrapping estimates of uncertainty. I suspect that if error bars did accompany the probability changes in the inset panel, the changes for 2001-2031 and 2035-2065 would be statistically insignificant relative to 1861-1891. Further, it is not clear as to whether the future probabilities of witnessing less than 46mm month^{-1} have been calculated using raw model data, or the excessively-smoothed PDF constructions.

5) The use of 1861-1891 is a misleading representation of 'pre-industrial' as it includes the influences of the Krakatoa volcanic eruption, and associated cooling effects. I strongly recommend changing the baseline period, perhaps to 1861-1880 in accordance with the IPCC's definition.

Technical Comments:

Specific comments:

P1, L9: Are these small increases statistically significant?

P1, L10: East Africa, not East African.

P1, L10: It's not really an analysis 'merging' ERA-Interim data with CMIP5 data. Instead you are using ERA-Interim as a basis for bias-correcting the model data.

P1, L14: '… shows that *during* August to October'?

P1, L20: RCP8.5 is not designed as a 'business-as-usual' scenario. It is in fact an acceleration of the present-day rates of increase in radiative forcing. The closest to 'business-as-usual' would be RCP6.0. I suggest the authors modify this particular phrase.

P1, L20-L23: This is poorly phrased. So you are bias-correcting based on climatological mean precipitation for ERA-Interim?

P1, L27-28: I suspect a change from 5.6 to 5.8% is not statistically significant in any way, and repeating these calculations with even just two or three models removed could lead to a completely different answer. What is the range in answers of this probabilistic increase for each individual model?

P1, L29: 'stretched left-tails' is a poor description, and I suggest changing this.

P2, L2-7: This is a very poor concluding paragraph. You do not summarise the key results of the study, but instead offer reasons why significance testing is needed (reliability of different models), before highlighting all of the reasons why monthly-mean precipitation deficits may in fact be a poor proxy for drought impacts over East Africa.