Review on "Statistical characteristics of extreme daily precipitation during 1501 BCE-1849 CE in the CESM

June 17, 2021

The manuscript is clear, well-written, and provides a new approach to the understanding of extreme precipitation, and I recommend it for publication with minor revisions.

The authors introduce a methodology to analyze the characteristics of extreme precipitation in the past 3000 years, using peak-over-threshold extreme value analysis and the identification of modes of natural variability using EOF decompositions in distinct geographical areas. The adequacy of the model was first verified over the recent observational period against reanalysis data, before focusing on simulations of the pre-industrial era, with and without external forcing; and the statistical analysis then allows to link the characteristics of regional extreme events to internal variability and different types of forcing.

1 Minor comments

- L27: could cite Allen and Ingram (2002) for the different drivers of mean and extreme precipitation.
- L28 (rich gets richer): could cite Chou et al. (2009); Chou and Neelin (2004)
- L30 "is constrained by the available atmospheric moisture": to be more precise, it is constrained by the maximum available moisture content at a given temperature ($q_v^\star$ is what Clausius-Clapeyron knows about, not q_v), but global extremes occur in regions close to saturation, so it does make much difference for global extremes. It could make a difference for regional extremes, in regions where the values of rain percentiles are rather constrained by the relative occurence of different precipitation regimes.
- L72: you mention some model limitations later, but at this stage what is coming to mind is that you can add references for that fact that in CMIP5, the rich-gets-richer mechanism breaks in the torpics (Chadwick et al., 2013) and references for the drizzling bias (e.g. Stephens and Hu, 2010)
- L84: interesting
- L110: would there be any interest in having several ensemble members for the robustness of attribution of extremes to different modes of variability?
- L120: Is there any uncertainty reported on the forcing data, that could propagate to uncertainties in extreme rainfall characteristics?
- L147, section 3.1: extremely clear, well-explained
- L168 "each cluster is composed of consecutive days of extremes": which threshold is used to separate the clusters, how is it chosen, and is it spatially uniform?

- L181: I am bit confused, it is not clear to me how the data is sampled for the calculation of a given non-stationary GPD. For stationary GPDs, I thought you were using the entire time series at each location and computed the shape, scale parameters for that whole distribution over time. For non-stationary distributions, it seems that you somehow fit σ_0 and σ_1 for the entire dataset, but how do you get the evolution of the tail distribution over time for each (x, y) point? Do you compute the rain distribution and the GPD for data sampled on moving windows, then obtain a $\sigma(t)$ and regress it over time to get σ_0 and σ_1?
- L244 (Mann-Kendall trend test): I don't know this test, would you have a reference? Is it standard practice?
- L246 (Mann-Whitney U-test): I don't know this test, would you have a reference? Is it standard practice?
- L269: very interesting
- L284 "the de-clustering method *is* (typo) only around 40% of the initial numbers of extremes": same comment as above, does that depend on the threshold used to define clusters? Is it uniform?
- L286: interesting comparison of the extremal index between tropical and extra-tropical regions. It seems there could be a correspondence with the convective-organization viewpoint, saying that organization is more likely to occur at high SSTs, and which would be consistent with clusters mainly occurring in the tropics.
- L350-351 "distinguishing *the* regions": do you mean that 'the POT analysis allows to separate distinct and coherent contiguous regions for similar types of distributions', or do you mean that the regions that are exhibited somehow map onto known regimes, ie the regions on Fig. 5b?
- Figure 2: What would the distributions look like with a logarithmic y-axis? Maybe that would allow to better illustrate when there is a finite upper-bound. Is it hard to see as it is shown now.
- Figure 7, middle panel title: "mean above 99th threshold" would me more explicit.
- Appendix A, L761-762: several variables are used and I assume they are normalized with mean and variance before analysis so that the RMSE can be compared numerically. But is there a reference RMSE value, a threshold above/below which the error is large/small, or is it just for relative comparison across modes?

References

M. R. Allen and W. J. Ingram. Constraints on Future Changes in Climate and the Hydrologic Cycle. *Nature*, 419(6903):224–232, sep 2002. ISSN 0028-0836. doi: 10.1038/nature01092. URL `http://www.ncbi.nlm.nih.gov/pubmed/12226677`.

Robin Chadwick, Ian Boutle, and Gill Martin. Spatial patterns of precipitation change in CMIP5: Why the rich do not get richer in the tropics. *Journal of Climate*, 26(11):3803–3822, 2013. ISSN 08948755. doi: 10.1175/JCLI-D-12-00543.1.

Chia Chou and J David Neelin. Mechanisms of Global Warming Impacts on Regional Tropical Precipitation. *Journal of Climate*, 17:2688–2701, 2004.

Chia Chou, J David Neelin, Chao-An Chen, and Jien-Yi Tu. Evaluating the "Rich-Get-Richer" Mechanism in Tropical Precipitation Change under Global Warming. *Journal of Climate*, 22(8):1982–2005, apr 2009. ISSN 0894-8755. doi: 10.1175/2008JCLI2471.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/2008JCLI2471.1`.

Graeme L Stephens and Yongxiang Hu. Are Climate-Related Changes to the Character of Global-Mean Precipitation Predictable? *Environmental Research Letters*, 5(2):025209, apr 2010. ISSN 1748-9326. doi: 10.1088/1748-9326/5/2/025209. URL `http://stacks.iop.org/1748-9326/5/i=2/a=025209?key=crossref.d5277cd96f73aefba73aad997e4af21a`.