EGUsphere, referee comment RC2
https://doi.org/10.5194/egusphere-2022-391-RC2, 2022

Comment on egusphere-2022-391

Anonymous Referee #2

Referee comment on "A storm-centered multivariate modeling of extreme precipitation frequency based on atmospheric water balance" by Yuan Liu and Daniel B. Wright, EGUsphere, https://doi.org/10.5194/egusphere-2022-391-RC2, 2022

I have read over the paper. I think it is very well written and if I was reviewing this I would only ask for moderate or minor revisions.

The authors developed a storm tracking algorithm (using a combination of existing algorithms in a novel way) and use it to create a dataset of large storm events which they perform frequency analysis on. Despite the novel storm tracking algorithm, performing frequency analysis on a storm dataset (as opposed to gauge records) is not in itself novel. This paper's key contribution comes from the way the authors uses copula based multivariate analysis on atmospheric variables from ERA5 to develop a way to stochastically generate annual maxima series representative of the observed storm catalogues.

Major comments

My first major concern with the paper is that the authors do not make enough attempts to validate their method or at least compare it to external data sources.

The only comparison to other methods they make is with a GEV fitted to storm annual maxima in Figure 6-7. I believe there is also opportunity to compare DAD curves in Figure 8 to an external data sources. The authors mention the relationship between DAD curves

and ARFs, so any ARF information available for the Mississippi basin could be used to formulate a comparison here.

I appreciate that these comparisons may be difficult to facilitate because of the authors have taken a storm-centred approach while the majority of other datasets are based on gauge-centred data. Still for their approach to be applied outside of research we need to understand how it compares to existing approaches.

My second major concern is about the use of empirical CDFs for all atmospheric variables except the divergence term for which a GEV is fitted. While divergence shows the highest correlation to precipitation I find this insufficient justification for why only this variable is modelled using a GEV. I also note that other terms such as the residual also have non-negligible contributions to rainfall. I would be interested to know if there is any change in results if similar extreme value distributions are used for other atmospheric variables.

Minor comments

I'd prefer the use of spelling gauge to gage, I think it's most common in modern literature

Section 3.1: I believe the authors could draw more attention to their storm search method being a novel combination of existing approaches

Line 170-175: I think the explanation of the 'binary search' is not clear and could be improved

Line 255: The GEV scale parameter must be greater than zero ($\sigma>0$)

Line 354: Should reference Figures A1-2?

Additional - uncertainty in the ERA data is not accounted for in the approach here. Alternate reanalyses do not agree with each other, even for atmospheric moisture - see Moalafhi, D. B., Evans, J. P. & Sharma, A. Influence of reanalysis datasets on dynamically downscaling the recent past. Climate Dynamics 49, 1239-1255 (2017). Some discussion on the impact of this uncertainty and how it could be included in the GEV modeling may be helpful.