Review of the manuscript “Verifying the relationships among the variabilities of summer rainfall extremes over Japan in the d4PDF climate ensemble, Pacific Sea surface temperature, and monsoon activity” by Shao-Yi Lee, Sicheng He and Tetsuya Takemi.

The authors have hypothesized an eventual relationship between Summer (June-July) monsoon rainfall extremes in Japan (at the hour, day and pentad timescales) and some North-Pacific SST seasonal anomaly modes (ENSO; PDO, pseudo-trend), which has been shown to be non-existent or non-statistically significant. This is concomitant, as well, with unclear or non-existent relationship (according to authors and literature) between ENSO and average monsoon rain in Japan. To represent extremes, the authors have used both rain-gauges and the dynamically downscaled D4PDF dataset. In order to aggregate stations and model grid-points, a rather complex cluster analysis is performed, though not using standard geostatistical techniques. The representativeness of the D4PDF dataset to simulate rainfall extremes is somehow ambiguous. For instance, no systematic comparison is made between the distribution of observed extremes and the distribution of simulated extremes. It was done only indirectly through the comparison between extreme rainfall values and climatic indices.
Authors have performed a lot of work, trying many methodological possibilities (e.g. clustering) to present the inexistence of direct control of the extremes by the analyzed climatic indices.
Despite that non-result, the authors have explored (only at the discussion section) the relationship between rainfall extremes and indices of the summer monsoon (Baiu). Moreover, the modulation of monsson indice statistics (average and standard deviation) by the SST Pacific modes has been studied. Those relationships seem robust, new and interesting.
Giving the above considerations, the manuscript can be publishable after concretizing two main tasks (major points):

1) Rewrite some parts of the main manuscript.
2) Explore further the results linking extremes with monsoon and Pacific indices providing physical reasons.

Below are presented the ordered list of corrections asked:

1) Line 15. Are the modes sorted by decreasing explained variance? Clarify. In line 18 ‘higher modes’ refer to the previous order of modes?

2) Line 17-20. Rewrite the phrase of lines 17-20 in a much clear way by splitting it into two sentences.

3) Line 95 Clarify the simulation period of the ‘100-member historical-warming (HPB) climate ensemble’ as well as the temporal and spatial resolution of the referred GCM.

4) Lines 99-100. No systematic comparison is made between the distribution of observed extremes and the distribution of simulated extremes by the 4DPDF dataset. A synthetic study of the representativeness of 4DPDF extremes shall be made.

5) Line 142 and wherever needed: Change ‘Principle Component Analysis’ to ‘Principal Component Analysis’, throughout all the manuscript.

6) Line 144. The linear long trend of SST was not removed to get the anomalous SST. It was obtained farther as a mixed mode (TREND+) or pseudo-trend. Authors could be more direct

by correlating the linear SST trend (for instance averaged throughout all Japan) with extremes. Comment on that.

7) Line 146 Authors have performed extended-PCA (extended Principal Component Analysis) and not a simple PCA. This must be referred explicitly and for clarity, since an extended vector merging 5 delayed (5 trimesters centered in JJA) of spatially distributed values have been taken, from which the covariance and its eigen-decomposition was computed. This is like a MSSA (Multi Singular Spectrum Analysis) with embedding dimension 5 and trimestral sampling.

8) Line 168 The spatial clustering of temporal extremes, particularly the rainfall (e.g. Ma et al. 2020 and references therein) has been studied by different authors, namely by using geostatistical techniques. The reference to those works must be included in the manuscript. The considered metric (distance) to cluster rain-gauges and grid points, used in cluster analysis HDBSCAN depends uniquely on the temporal similarity (Spearman rank correlation) between time-series, being equivalent to the F-madogram. However, for spatially distributed data, a geometrical term, weighing the point-wise distance (e.g. Euclidean) must be added to the statistical distance. The omission of the geometrical term leads to fragmented, topologically complex clusters (even not simply connected, i.e. with 'holes'). This apparently was remedied by adopting ad-hoc clustering rules. The authors are asked to comment on that.

Yingzhao Ma, Mengqian Lu, Cameron Bracken, Haonan Chen, 2020. Spatially coherent clusters of summer precipitation extremes in the Tibetan Plateau: Where is the moisture from?, Atmospheric Research, Volume 237, 104841, ISSN 0169-8095,https://doi.org/10.1016/j.atmosres.2020.104841.

9) Line 175. Authors use the procedure 'percentiles were calculated for each individual ensemble member, then the ensemble mean was taken'. This is an alternative to take quantiles of a super-sample (collecting all sub-samples). Please comment on that.

10) Line 184 Say here the value of CC significant at alpha=0.05.

11) Line 202-205. Authors say: 'At each timestep, the maximum rainfall in each set was selected. For example, in a cluster of three points, this would be the point with the highest rainfall and the other two points would be discarded. This was based on the concept that a set included locations which experienced rainfall from the same event, and the maximum rainfall of that event was sampled by an observer that could "see" the entire location rather than only one specific point.' This procedure seems ad-hoc and not well supported statistically. The maximum of the extremes within a cluster is taken, instead for instance the cluster average of maxima. What is the representativeness of the maximum (among the extremes) within the cluster?

12) Line 220. Authors say: 'Since there were over a hundred rain-gauges, we would like to group them into regions that persisted across the three time resolutions'. The obtained clusters depend on the time resolution (hour, day or pentad), what seems expected, since the associated meteorological systems have different spatio-temporal scales. Authors shall provide a clear justification for their choice. For instance, if instead we had taken the time-

scales: hour, 6 hours, 3 days, would the clustering output be the same? Ideally, a uniform (or time-scale integrated) criterium should be adopted. Comment on that.

13) Line 220-221. Authors have performed cluster analysis for each of the three analyzed extremes and then they have grouped the respective clusters into common regions. Explain the rationale of that.

14) Line 241 Authors use both the symbol R and CC for the Spearman correlation. Choose only one symbol.

15) Lines 275 The seasonal anomalous Jlat, Flat, QU and QV were obtained as residuals from a sinusoidal and logistic fit of daily values. The result can be biased due to choice of the fit. Why not computing explicitly the daily-basis seasonal cycle from averages along the 61 days with the same Julian day (or within a smoothing window of 2-3 days)?

16) Line 285 Emphasize again that extended-PCA is applied, not PCA, to help the reading. The extended-PCA (Ext-PCA) analysis in the North-Pacific, shown by authors, have omitted other relevant modes of SST Pacific variability like the Northern Pacific Gyre Oscillation (NPGO) (DiLorenzo et al. 2008) that can be also candidates as drivers of extremes. Moreover, the Ext-PCA looks for variability all along the year (all trimesters) and not uniquely focused in the summer season. For instance, a PCA of SST in (JJA) would capture uniquely the difference between summers, eventually producing more suited modes driving the summer rain in Japan. Comment on that.

Di Lorenzo E., Schneider N., Cobb K. M., Chhak, K, Franks P. J. S., Miller A. J., McWilliams J. C., Bograd S. J., Arango H., Curchister E., Powell T. M. and P. Rivere, 2008: North Pacific Gyre Oscillation links ocean climate and ecosystem change. *Geophys. Res. Lett.*, 35, L08607, doi:10.1029/2007GL032838.

17) Line 323 The so-called TREND+ mode seems to contain interannual variability. It is not clear how much is attributed to the trend and global warming. It seems to be a mixed mode including the linear trend (maybe nor the largest part), the PDV (or more commonly the PDO: Pacific Decadal Oscillation) and the NPGO.

18) Line 476. The relationship between monsoon parameters and rainfall extremes is done uniquely for daily extremes. Explain the omission of the other scales.

19) Lines 481-483 Authors say: 'The spatial patterns of correlation were similar between d4PDF and rain-gauges for the μFLat, μJLat, and μQV, showing correlation along the Sea of Japan coast and anti-correlation along the Pacific coast (Figs 8e, 8g, 9e)'. This statement is based on very few rain-gauges, while in some others the agreement is neutral (weak correlations). For other parameters the similarity is quite inexistent, so any apparent similarity seems to be unfair.

20) Line 503. Authors refer to 'scores of the Pacific SST'. Do you mean PCs (ENSO+, ENSO-, TREND+ etc.)? Rewrite and clarify. A Table with correlations between PCs and monsoon indices should be presented with analysis of significance and discussed afterwards. No physical explanation is provided for the correlations found at the end of the Discussion Sec. 4. Improvement needed.