

Geosci. Model Dev. Discuss., referee comment RC2  
<https://doi.org/10.5194/gmd-2021-225-RC2>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on gmd-2021-225

Anna Kiriliouk (Referee)

---

Referee comment on "Variability and extremes: statistical validation of the Alfred Wegener Institute Earth System Model (AWI-ESM)" by Justus Contzen et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-225-RC2>, 2021

---

The authors present a method for climate model validation which focuses on extremes rather than on classical measures such as means and covariances. Second, they apply a clustering algorithm to detect regions in which extremes occur simultaneously. The paper is well written, the authors clearly expose the different steps and methods used in their approach. However, I have some methodological remarks, mainly concerning the use of extreme-value theory.

General comments:

- The introduction does not seem to cite other papers that study model validation focused on extremes, yet I think it is already an active research topic. For an example, see for instance

*Timmermans, B., Wehner, M., Cooley, D., O'Brien, T., & Krishnan, H. (2019). An evaluation of the consistency of extremes in gridded precipitation data sets. Climate dynamics, 52(11), 6651-6670.*

- Page 2, lines 35-36: please mention that there are two popular approaches, block maxima and peaks-over-thresholds.
- Page 4, lines 110-111: I would remove "especially for the yearly maximum of daily average precipitation" since the GEV has been used extensively for many types of data.
- Page 5, lines 131-132: this sentence is rather unclear, could you maybe reformulate?
- Page 5, lines 145-146: please note that the lower bound of the extremal index corresponds to perfect dependence (comonotonicity), which is more general than a Pearson correlation of 1 (a Pearson correlation of 1 implies comonotonicity but not vice

versa).

- Page 6, equation (8): does it make sense to « weight » the three marginal parameters equally? Maybe the shape parameter could play a bigger role than the mean and scale parameters?
- Page 9, line 210: it is not necessary to use bootstrap-based confidence intervals for the PWM estimators, since their asymptotic covariance is known and has a simple expression; see

*Hosking, J. R. and J. R. Wallis (1987). Parameter and quantile estimation for the generalized Pareto distribution. Technometrics 29(3), 339–349.*

*Ribereau, P., P. Naveau, and A. Guillou (2011). A note of caution when interpreting parameters of the distribution of excesses. Advances in Water Resources 34(10), 1215–1221.*

- Page 13, Fig 7: the graphs are difficult to compare because of the many clusters and colors. Could you please resume the main differences in the text?
- Page 13, line 250: ``While parametric copula families are applicable only to a very limited extent in high dimensions...'' I disagree, there are many possibilities to model high-dimensional data using parametric copulas, for example, through vine copulas.
- Page 14, line 259: aren't the spatially stationary method well suited to model the clusters as identified previously?

Typographical errors:

- Page 4, line 114: resultung -> resulting
- Page 9, line 224: Fig 4 should be Fig 7?