Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/hess-2021-334-RC1, 2021

Comment on hess-2021-334

Rasmus Benestad (Referee)

Referee comment on "Flexible and consistent quantile estimation for intensity–duration–frequency curves" by Felix S. Fauer et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-334-RC1, 2021

The paper 'Flexible and Consistent Quantile Estimation for Intensity-Duration-Frequency Curves' by Fauer et al. provides a timely discussion about intensity-duration-frequency (IDF) analysis of sub-daily rainfall intensities. They present and discuss a flexible method for calculating IDF curves based on a new parametric approach and evaluate three different options.

Due to the small samples of the rainfall from clouds which rain gauges with a diameter of centimetres represent, I wonder if it's justifiable to combine data from stations with a distance below 250 m into a single station data set. It's probably OK for the statistical properties, but maybe not so for the raw data - especially on short time intervals (too much random sampling noise?). At least, there ought to be a test of the robustness of the results to this assumption.

Perhaps it would make the flow of the text better if there was some connection between the return value (probability) and the quantile? I thought the part on Quantile Skill Index (e.q. 13) wasn't as easy to follow as the preceding sections (how does it link to the preceding discussion on the GEV and the estimation of the parameters?).

The description of the bootstrapping was a bit difficult to follow - perhaps explain it more carefully or add an illustration?

Very brief catches (e.g. minutes) of rainfall with rain gauges are expected to be subject to a large degree of sampling uncertainty, aren't they? (depending on the number of rain drops falling onto the cross section representing the measurement). Maybe this also can explain some discrepancies at the extreme short end of the scale?

One reason why annual maxima of different durations do not follow the same scaling process could be that different rain-producing meteorological phenomena have different temporal and spatial scales. If the rainfall can be considered as a 'by-product' of different processes and conditions (e.g. convection, weather fronts, cyclones, and derechoes), then different statistics may perhaps show the true situation? But I'm still struggling to understand what the skill estimates really say.

A new and relevant paper DOI: 10.1088/1748-9326/abd4ab suggests a simple formula for expressing IDF curves even for sites with limited data. This formula is based on more 'physical' parameters (wet-day mean precipitation and wet-day frequency), rather than the stronger reliance on the statistical/mathematical theory behind GEV. It would be interesting to compare the results presented here with this formula. It also fits in the comparison of different ways to parametrize IDF curves. At least, it could be included in the introduction and the discussion of different ways of calculating IDFs.

Appendices: when describing what calculations and processing was done in this analysis, it's more elegant to use past tense rather than present tense (my subjective opinion). But mixing past and present tense makes the text inconsistent and a 'clumsy' read. Also check the references therein ('??').