Reply on RC1

Jana Ulrich, Felix S. Fauer and Henning W. Rust

August 20, 2021

Dear Reviewer,

Thank you very much for your informative and helpful comments. We address them point by point below.

General comments

The general point concerns the small block size of one month for the extreme value analyses. Even if the GEV is fitting well as shown by the authors the basic assumption of sufficient large n for the validity of Fisher–Tippett–Gnedenko theorem might be questionable also considering the fact that in certain months of some years no extremes might be observed. This problem becomes especially relevant given the results from which months the maxima from different durations originate with respect to the effective block size. Especially the latter is much larger than one month for long durations. I would suggest to discuss this a little further, may be also considering future research regarding specific analyses periods/ seasonal sub-divisions depending on the duration.

Answer

You are correct that a detailed analysis of the monthly block size is needed, however this has been addressed in previous studies (Maraun et al., 2009; Rust et al., 2009; Fischer et al., 2018). The choice of block size typically involves finding a balance between rapid convergence to the GEV distribution and a sufficiently large number of block maxima for parameter estimation.
Since the aim of our study is to resolve the seasonal variations, we need to choose a sub-annual block size. We used q-q plots to investigate whether the monthly maxima can be well described by the d-GEV model. We present the q-q plots for the example station Bever-Talsperre again below in Figure 1. The plots show that the d-GEV distribution describes the monthly maxima of the different durations sufficiently well. To investigate the influence of the choice of a larger block size, we additionally provide the q-q plots for a block size of 2 and 3 months, respectively, in Figure 2. Figures 1 and 2 illustrate that with increasing block size information about the seasonal variations is reduced, while an increasing block size does not lead to a significant improvement of the model fit.
We agree with your concern that extreme values may not occur in certain months of the year, depending on the region under investigation. Indeed, it could be problematic to consider monthly maxima if there is no precipitation occurring at all during certain months. Nevertheless, this is not the case in Germany. You are right that Figure 4 in the manuscript indicates that for certain months the exceedance probabilities of the annual maximum are very low. But this should not be interpreted as a sign that in these months no extreme precipitation values occur. Rather, it means that the precipitation intensities, which are extreme in these months, are comparatively low with respect to the whole year. To illustrate that the GEV distribution describes the monthly maxima of individual durations sufficiently well and that a larger block size does not lead to a significant improvement in the model fit even when modeling individual durations, we additionally provide the q-q plots for two selected durations in Figures 3-6. Consideration of the remaining durations led to the same results.

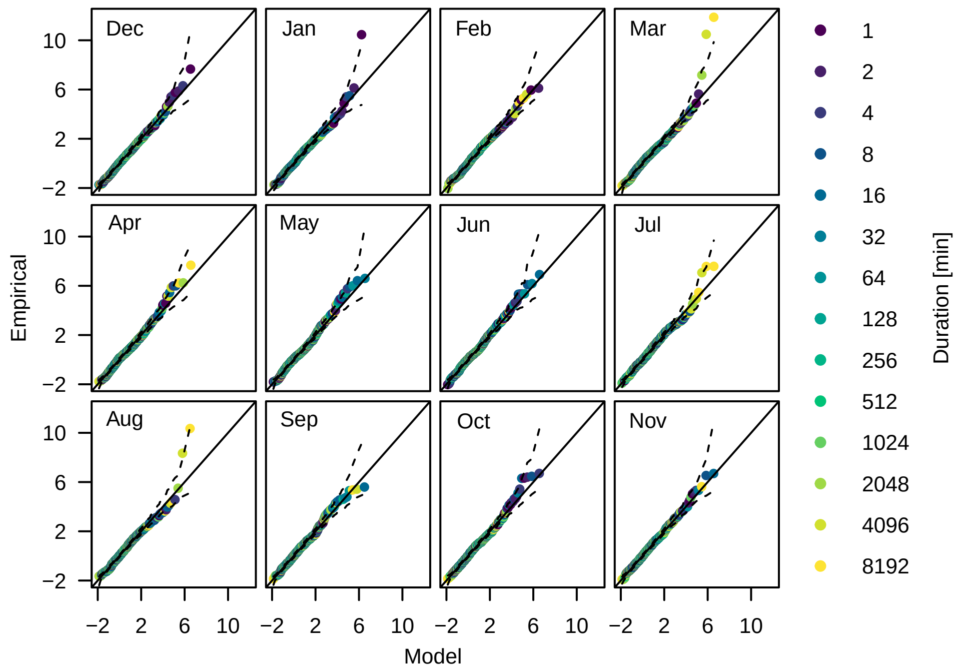

Figure 1: Diagnostic q-q plots of the d-GEV model for each month at station Bever-Talsperre as provided in Fig. A1 of the manuscript . The observations and the modeled quantiles are transformed to standard Gumbel $G(\mu = 0, \sigma = 1, \xi = 0)$ to remove the duration dependency. Dashed lines represent 95% confidence intervals.

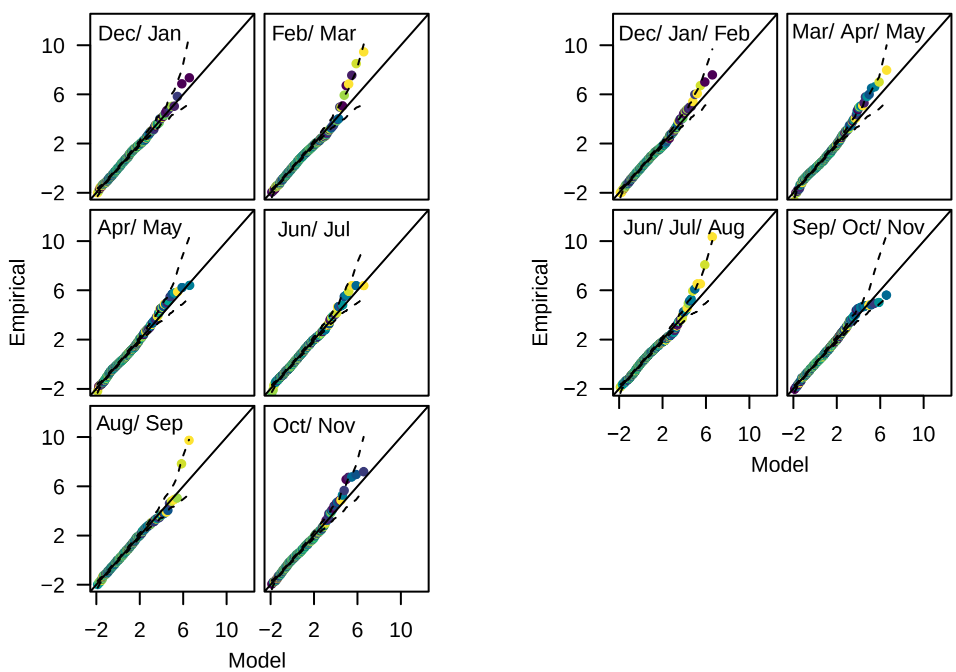

Figure 2: Diagnostic q-q plots similar to Figure 1, but for a block size of 2 months (left) and 3 months (right).

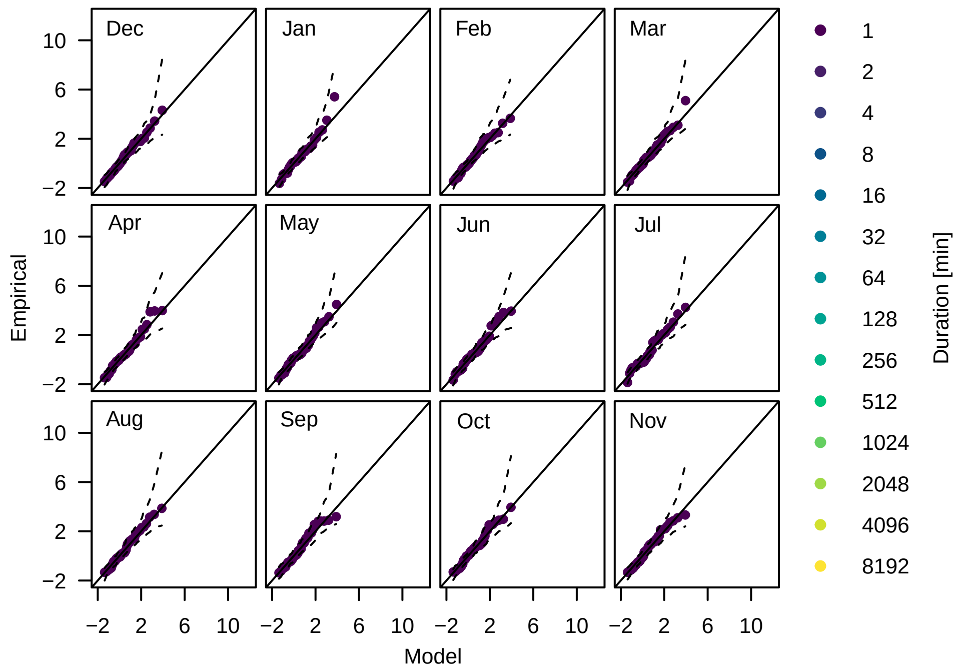

Figure 3: Diagnostic q-q plots of the GEV model for a duration of 1 minute for each month at station Bever-Talsperre. Again, the observations and the modeled quantiles are transformed to standard Gumbel $G(\mu = 0, \sigma = 1, \xi = 0)$.

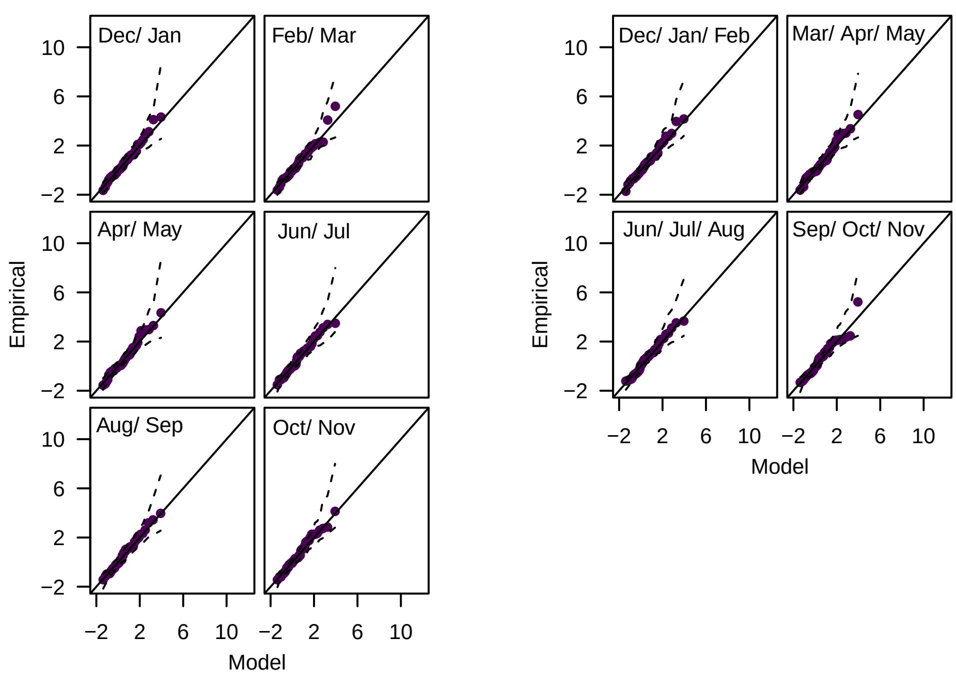

Figure 4: Diagnostic q-q plots similar to Figure 3, but for a block size of 2 months (left) and 3 months (right).

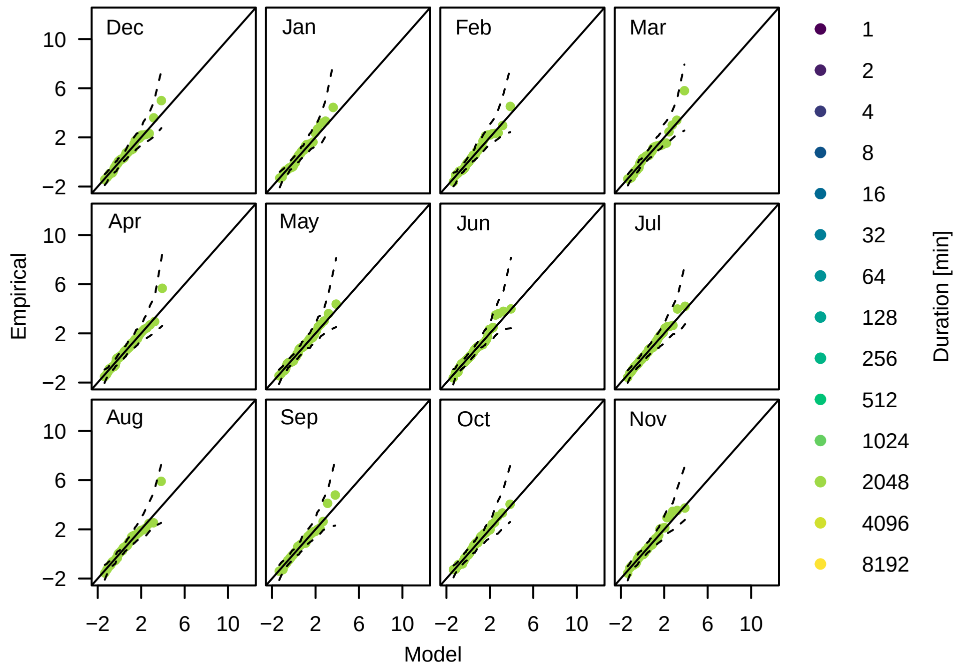

Figure 5: Diagnostic q-q plots of the GEV model for a duration of 2048 minutes ($\approx$ 34 hours) for each month at station Bever-Talsperre.

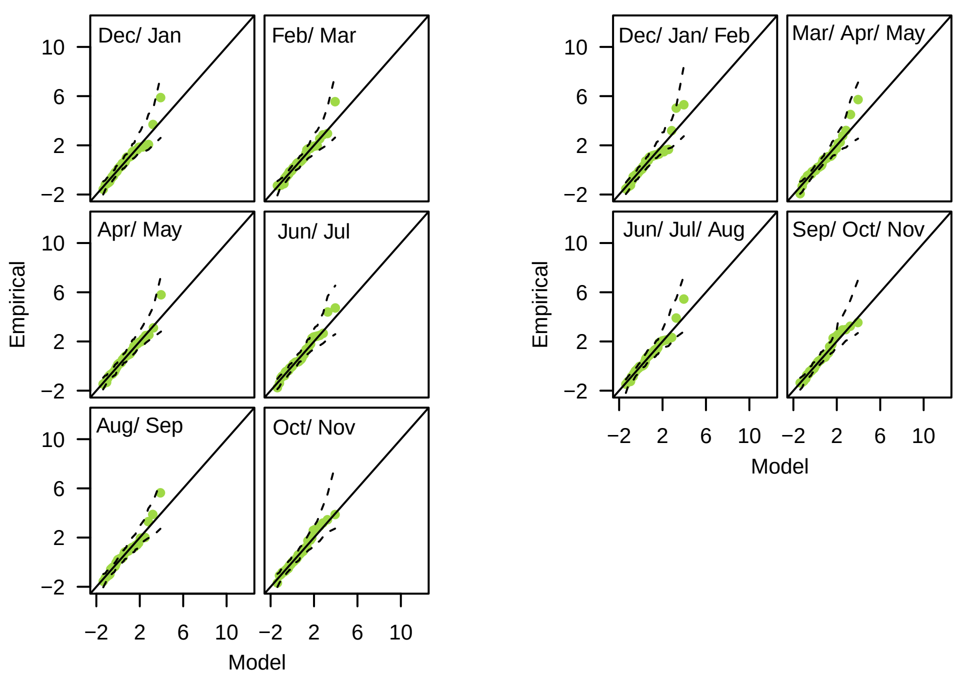

Figure 6: Diagnostic q-q plots similar to Figure 5, but for a block size of 2 months (left) and 3 months (right).

References

- Fischer, M., Rust, H. W., and Ulbrich, U.: Seasonal Cycle in German Daily Precipitation Extremes, Meteorol. Z., 27, 3–13, https://doi.org/10.1127/metz/2017/0845, 2018.
- Maraun, D., Rust, H. W., and Osborn, T. J.: The annual cycle of heavy precipitation across the United Kingdom: a model based on extreme value statistics, Int. J. of Climatol., 29, 1731–1744, https://doi.org/10.1002/joc.1811, 2009.
- Rust, H., Maraun, D., and Osborn, T.: Modelling seasonality in extreme precipitation, Eur. Phys. J. Spec. Top., 174, 99–111, https://doi.org/10.1140/epjst/e2009-01093-7, 2009.

Specific comments

1.

Please indicate the specific GEV parameters on the figures of relate a) – e) to the specific parameters in the caption.

Answer: In Figure 2 in the manuscript, the d-GEV parameters are indicated at the y-axis of each of the plots (a)-(e), but we can also add this informtaion to the caption.

2.

Lines 233ff: Please add some more explanation how the bootstrap is carried out.

Answer: Thank you for pointing out that the bootstrapping procedure is not sufficiently explained. We will improve the explanation.

3.

Equation 16: The equation is not completely clear to me. The variable u is not explicitly defined. However, if I assume $u = (o_n - q_p)$ then it should read e.g. $\rho_p = pu$ and not $\rho_p(u) = pu$?

Answer: You are right, the explanation of the Quantile Score is very short. We will add the information that $\rho_p(u)$ is the check loss function where for the Quantile Score the arguments is $u = o_n - q_p$.

4.

Lines 425ff: The uncertainties are estimated with different methods, the Fisher information matrix and the bootstrap method. Are these results comparable? Why not using bootstrap for all uncertainty assessments?

Answer: Thank you for pointing this out. You are correct that it is better for reasons of comparability to estimate all uncertainties using the same method. We will change this and estimate all uncertainties via the boostrap method in Figure 7.