

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC2
<https://doi.org/10.5194/nhess-2021-66-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on nhess-2021-66

Anonymous Referee #2

Referee comment on "Using high-resolution regional climate models to estimate return levels of daily extreme precipitation over Bavaria" by Benjamin Poschlod, Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2021-66-RC2>, 2021

The author estimates 10-years return levels with the Generalized Extreme Value (GEV) distribution from three different Regional Climate Models (RCMs), namely the Canadian Regional Climate Model version 5 (CRCM5) at 12 km resolution, the Weather and Forecasting Research model (WRF) at 5 km resolution and the WRF model at 1.5 km resolution, showing that the finer spatial resolution of the WRF-5km with respect to the 12 km CRCM5 reduces the bias of GEV. Moreover, he investigates uncertainties due to the use of three different extreme value models (GEV with fixed shape parameter, Generalized Pareto (GP) and Metastatistical Extreme Value (MEV) distributions) to estimate the 10-year return period quantiles using the WRF model with the finest (1.5 km) resolution. Through this analysis he concludes that GEV and GP distributions are equivalently biased ($\sim +1\%$), while MEV tends to underestimation ($\sim -6\%$) and that high-resolution RCMs provide promising results for the estimation of spatially homogeneous rainfall return levels.

The study is interesting and shows potential for evaluating extremes in a changing climate, the manuscript is well written and easy to follow. I have some major comments, and minor comments follow.

Major comments:

1) The use of the high-resolution products (REGNIE, RADOLAN, SPARTACUS) would avoid to homogenize the gauge precipitation values and would make possible a more accurate validation of the RCMs with the finest resolution. Why not considering them?

2) Why only return level of 10 years? I understand the concern of the author that 30 years of data are few for estimating higher quantiles, but return periods higher than 10

(e.g., 100) years are more relevant for engineering applications/(re)insurance purposes and the challenge is indeed to estimate them with the availability of short time series. How would the estimation of higher return levels compare e.g. with the official ones from KOSTRA? As the manuscript is presented now, the conclusion stated in the abstract "it follows that high-resolution regional climate models are suitable for generating spatially homogenous rainfall return level products" is not fully supported by the analysis, since only the 10-years return levels have been evaluated.

3) The study area is characterized by some high-elevated regions affected by orographic precipitation. I'm wondering if using all the values as "ordinary events" in the MEV might not respect the independence hypothesis required by the MEV framework. See for example Marra et al. (2018) and Miniussi et al. (2020) for some discussion on temporal correlation.

4) Why using a GEV distribution with a constant shape parameter and not, for example, a Gumbel? Previous studies (e.g., Grieser et al. (2007)) have shown that the Gumbel distribution is a good model for precipitation in the Bavarian area, and its location parameter has a strong correlation with altitude, while its scale parameter has a noisy pattern (except for the Bavarian Alps). Moreover, you say that the shape parameter based on all the three RCM setups is centered around a value close to 0.114, in line with the one recommended by Papalexiou and Koutsoyiannis (2013): is this really a fair comparison, as these shape parameter values are already affected by estimation uncertainty?

Minor comments.

Section 3.

L225: Another title for section 3.3 would be more appropriate

L226-227: please add a couple of words about the adjustment, so that the reader understands it directly from here without the need to go looking at the reference.

L239: you state that "the location and scale parameter are governed by the topography". From Figure 3 one can notice that the spatial pattern of the location parameter is somehow coherent with topography, but the noise for the scale parameter does not make its pattern straightforward to understand. Maybe also the colors scale is not helping.

L240: why a chaotic pattern for the shape parameter? Is it related to the uncertainty that one can get due to the limited series available to estimate it?

L259: you mention you made a "goodness of fit" (despite its limitation in prediction) for the GEV and the GP distributions. Have you made a similar analysis also for the Weibull distribution?

L264: in L253-255 you mention that for sample sizes > 50 estimation via ML is recommended. Why then using PWM for the Weibull distribution in the MEV framework?

Section 4.

L287 and 310 (captions of Figures 4 and 6): "difference calculated as climate model return level minus observational return level" -> difference between the return level from the climate model and the observational one. Why using of the absolute error instead of the relative error?

L454-457: in Zorretto et al. (2016) the analysis has been made by means of a cross-validation approach, so that the sample used for parameter calibration is independent from the one used for testing the performance of GEV and MEV distributions. When GEV is fitted and tested on the same sample (unless the sample is shorter, i.e. 10-20 years, when issues in the parameter estimation –especially for the shape parameter- might arise), it usually outperforms MEV, but it is not flexible in prediction.

A curiosity: are you considering or discarding snow events?

Supplementary material.

FigS2: how are the 95% confidence intervals computed?

You have the example for the Munich grid cell, and only for GEV-LMOM and GP models, why not for GEV-ML and MEV? Moreover, a comprehensive validation of all the extreme value models for the whole area would add value to the analysis.

FigS5-S6: now the REGNIE product is shown; why not showing the observation-based product used in the analysis? It would be also useful to evaluate differences among the products (even if for some events only).

References

Grieser, J., Staeger, T., & Schonwiese, C. D. (2007). Estimates and uncertainties of return periods of extreme daily precipitation in Germany. *Meteorologische Zeitschrift*, 16(5), 553–564. <https://doi.org/10.1127/0941-2948/2007/0235>

Marra, F., Nikolopoulos, E. I., Anagnostou, E. N., & Morin, E. (2018). Metastatistical Extreme Value analysis of hourly rainfall from short records: Estimation of high quantiles and impact of measurement errors. *Adv. Wat. Res.*, 117, 27–39. <https://doi.org/10.1016/j.advwatres.2018.05.001>

Miniussi, A., Villarini, G., & Marani, M. (2020). Analyses Through the Metastatistical Extreme Value Distribution Identify Contributions of Tropical Cyclones to Rainfall Extremes in the Eastern United States. *Geophysical Research Letters*, 47(7). <https://doi.org/10.1029/2020GL087238>