

Author response to Anonymous Referee #2 (RC2)

Manuscript egusphere-2026-4111: "Cross-scale characteristics of extreme precipitation events under climate change" by Jingkun Yang, Benjamin Poschlod and Felix Ament

We thank Referee #2 for the review and for acknowledging the value of the cross-scale perspective and of the two dGEV frames. Referee comments (from the PDF attached to RC2) are shown in grey boxes, followed by our response and the planned revision.

Main planned revisions

- **Model evaluation:** new subsection summarising the evaluation of the COSMO-CLM simulations for extreme precipitation (Rybka et al., 2023).
- **Event selection:** xWEI computed for an enlarged pool of candidate events, so that the top-200 events are selected by xWEI.
- **Robustness:** maps of the event locations, analysis of three sub-regions, seasonal breakdown, alternative area weighting.
- **Driving data and scenario:** comparison of the MIROC5- and ERA5-driven simulations for their common period 1979–2000; addition of global warming levels; results framed as conditional on model chain and scenario.
- **Uncertainty:** dGEV parameter distributions, return period cap, leave-one-year-out test, fixed shape parameter, bootstrap confidence intervals.
- **Interpretation:** clearer definition of the "second-order" effect, extended limitations, corrections.
- **Correction:** During the revision process we identified a sign-convention mismatch in the GEV shape parameter between the R package used for fitting and the Python routine used for return-period calculation. We will correct this in the revised manuscript and recompute the according return periods, WEI and xWEI values.

This reply describes the planned revision; line numbers refer to the preprint.

General comment

Referee comment RC2.0

The authors analyse the effect of climate change on the cross-scale extremity of precipitation events using the COSMO-CLM convection permitting model and the xWEI. They introduce two different analyses based on different fits of the dGEV, which aim to account for evolving rainfall characteristics. Generally, I think that it makes sense to assess extreme precipitation events across scales with metrics such as the xWEI, especially if the aim is to understand hydrological impacts. While it is a good idea to split the analysis into two different dGEV fits, the study has several shortcomings, which are listed below. In my opinion, the study, in its current version, does not meet NHESS standards.

Thank you. We believe that the planned revisions, in particular the revised event selection, the regional and seasonal analyses, the comparison of the driving data and the uncertainty quantification, address the shortcomings listed below.

Major comments

Referee comment RC2.1

The study only uses one dataset. COSMO-CLM, as you and Rybka et al. (2023) state (line 94 ff.), is biased: The performance of the convection-permitting simulation in terms of spatial distribution and intensity is in agreement with both observational datasets for extreme precipitation with durations greater than 12 hours. Higher return periods for daily extreme precipitation show larger differences, but still providing good

estimates. Referring to rainfall durations below 12 hours, an overestimation of the convection-permitting simulation is evident. In addition to that, the convection-permitting simulation produces a broader distribution for hourly and 6-hourly extreme precipitation referring to a higher spatial variability over Germany (Rybka et al., 2023).

The study should use an ensemble of CPMs to diminish uncertainties in the analysis. It is not common to base a study about climate change on only one simulation.

We agree that a CPM ensemble would be preferable. However, our approach requires multi-decadal convection-permitting simulations with hourly output that cover all of Germany for historical and future periods. To our knowledge, the COSMO simulations performed by DWD are the only openly available data set that fulfils this; other CPM ensembles cover only southern Germany and are not openly available (Hundhausen et al., 2023, 2025) or are driven by reanalysis only (Sieck et al., 2026). The model biases are addressed in our response to Referee #1 (RC1.2). We will state clearly that the results are conditional on one model chain and scenario, quantify the sampling uncertainty by bootstrapping, and emphasise the methodological contribution, which can be applied to future CPM ensembles.

Planned changes in the revised manuscript

- Data and Discussion: single-model limitation and data availability.
- Comparison with the ERA5-driven simulation for 1971–2000 (RC3.2).
- Bootstrap confidence intervals.

Referee comment RC2.2

Selection of extreme events: The authors select extreme events using the CatRaRE framework introduced by Lengfeld et al. (2021). Extreme events are selected based on the WEI. This procedure might lead to missing events that are more extreme across spatio-temporal scales.

Thank you a lot for this hint. We fully agree. Following Voit and Heistermann (2022), who computed the xWEI for the 100 CatRaRE events with the highest WEI, we pre-selected the 200 events with the highest object-based extremity $E_{T,A}$ per period and then ranked them by xWEI. Events that are extreme across scales but have a lower $E_{T,A}$ can thus be missed; based on your comment, we calculated the rank correlation between $E_{T,A}$ and xWEI within the top-200 events. Spearman only shows moderate values in the range of 0.3–0.5. In the revision, we will compute the xWEI for a much larger candidate pool (e.g., the 2000 events with the highest $E_{T,A}$), select the top-200 by xWEI, document the convergence of the selection and update all results. The procedure will be described explicitly in the Methods. We hypothesize that your idea will also fix the lower end of the top-200 xWEI values, which show a relatively strong decline in the current preprint. When calculating the xWEI for a large number of events and selecting the top-200 based on the xWEI, we expect a more linear behaviour of the xWEI values at ranks 150 to 200.

Planned changes in the revised manuscript

- Methods: explicit description of detection, pre-selection and selection.
- xWEI for an enlarged candidate pool; all results updated with the new top-200.

Referee comment RC2.3

The study lacks depth and novelty. It seems that with recent AI coding clients this study could be realized without much effort. The methods, namely the framework to select extreme events, the calculation of the xWEI and the dGEV, are established, and at least the code for the xWEI and the dGEV is publicly available. For this reason, I believe it would be appropriate that the study is extended to an ensemble of CPMs.

We agree that the components (event detection, xWEI, dGEV) are established; we use them unchanged to ensure comparability with radar-based studies (Lengfeld et al., 2021, 2023; Voit and Heistermann, 2022). The use of a single model chain is due to data availability (RC2.1), not to the effort involved. The novelty lies in (1) the first application of the xWEI to multi-decadal convection-permitting climate simulations, (2) the comparison of a fixed and an adapted dGEV as frames of reference under non-stationarity, (3) the shift in the composition of the most extreme events towards more persistent and extensive events, which grid-cell-based return levels do not reveal, and (4) the lessons for transferring the radar-based framework to climate model projections (e.g., return period cap, in-sample versus out-of-sample fits). We will state this more clearly in the Introduction and Conclusions.

Planned changes in the revised manuscript

- Introduction and Conclusions: clearer statement of the novelty.

Referee comment RC2.4

Limitations are not discussed enough: Fitting the dGEV to 30 years of data leads to a high uncertainty. 30 years are not representative for a selection of extreme events. This should be stated clearly in the limitations. Where did you cap the return periods? Did you check the fits that resulted from the dGEV (e.g. the shape parameter)?

We agree and will extend the limitations (details in RC1.9). Pooling across durations stabilises the dGEV, but the shape parameter remains uncertain for 30 annual maxima, and Lengfeld et al. (2023) showed that WEI and xWEI change by 17–28 % depending on whether the event year is part of the fit. Return periods are capped at 1000 years, following Müller and Kaspar (2014) and Voit and Heistermann (2022). We will show the dGEV parameters, including the shape parameter, for all periods, report the share of capped grid cells and the sensitivity to the cap, test a fixed shape parameter and leave-one-year-out fits, and provide bootstrap confidence intervals.

Planned changes in the revised manuscript

- Methods and Limitations: return period cap; uncertainty of 30-year fits (Lengfeld et al., 2023).
- Supplement: dGEV parameters and share of capped cells;
- Uncertainty assessment: test fixed-shape and perform leave-one-year-out tests; generate bootstrap confidence intervals.

Referee comment RC2.5

Climate change might affect different regions of Germany differently. However, the xWEI is only comparable within the same spatial domain because it is based on return periods that are adapted to local conditions. Using Germany as a whole as the spatial domain might conceal important signals of change.

We agree. Since return periods are local and the window size is the same for all events, an xWEI value describes the same rarity and extent in any region, which allows a Germany-wide ranking; still, the

ranking may be dominated by the regions with the largest changes. As also requested by Referee #1 (RC1.3), we will repeat the analysis for three large sub-regions and compare the regional changes for both dGEV frames, also with the regional changes in return levels reported by Rybka et al. (2023).

Planned changes in the revised manuscript

- Sub-region analysis (main text summary, Supplement figures).

Minor comments

Referee comment RC2.m1

line 114: The statement that Voit and Heistermann (2022) used the object-detection algorithm of Lengfeld et al. (2021) appears to be incorrect. As stated later in the manuscript (l. 193), a 200 x 200 km bounding box was used there for the calculation of the xWEI.

Voit and Heistermann (2022) took their events from CatRaRE, which is based on the detection of Lengfeld et al. (2021), and computed WEI and xWEI in a 200 × 200 km box. As no catalogue exists for the model data, we apply the detection of Lengfeld et al. (2021) ourselves. We will rephrase L114–116.

Referee comment RC2.m2

Figure 4: Please make the markers in the legend larger so the colors are easier to distinguish.

We will enlarge the markers and use distinct colours and symbols in Figures 4 and 6.

Referee comment RC2.m3

line 244: The most pronounced change occurs only once (Figure 6), namely between the past (1971–2000) and the subsequent periods. Could this be a data artifact?

The step also occurs within the MIROC5-driven chain (about +15 % in the mean xWEI for 2031–2060 and 2071–2100 relative to 1971–2000), so it is not only due to the change of the driving data. We will test whether it is related to differences in the dGEV fits between the periods, to the return period cap or to sampling variability, and use the comparison of the MIROC5- and ERA5-driven simulations for 1979–2000 (RC3.2). We will revise the interpretation if it turns out to be an artefact.

Referee comment RC2.m4

line 297 ff.: Rybka et al. (2023) confirmed that COSMO-CLM at convection-permitting scale projects stronger precipitation intensities for future scenarios, particularly on sub-daily timescales. Doesn't this finding contradict your results?

Thank you for pointing this out, our sentence misrepresents Rybka et al. (2023). They find similar relative increases (about 30 %) for 3-h and 24-h return levels in this convection-permitting simulation, i.e. no stronger sub-daily intensification, unlike models with parameterised convection. This is consistent with our finding that short-duration events do not gain prominence among the top-200. We will correct the sentence.

References cited in this response

- Hundhausen, M., Feldmann, H., Laube, N., and Pinto, J. G.: Future heat extremes and impacts in a convection-permitting climate ensemble over Germany, *Nat. Hazards Earth Syst. Sci.*, 23, 2873–2893, <https://doi.org/10.5194/nhess-23-2873-2023>, 2023.
- Marie Hundhausen, Hayley J. Fowler, Hendrik Feldmann, Joaquim G. Pinto, Sub-hourly precipitation and rainstorm event profiles in a convection-permitting multi-GCM ensemble, *Weather and Climate Extremes*, 2025, 100764, <https://doi.org/10.1016/j.wace.2025.100764>.
- Lengfeld, K., Walawender, E., Winterrath, T., and Becker, A.: CatRaRE: A Catalogue of radar-based heavy rainfall events in Germany derived from 20 years of data, *Meteorol. Z.*, 30, 469–487, <https://doi.org/10.1127/metz/2021/1088>, 2021.
- Lengfeld, K., Voit, P., Kaspar, F., and Heistermann, M.: Brief communication: On the extremeness of the July 2021 precipitation event in western Germany, *Nat. Hazards Earth Syst. Sci.*, 23, 1227–1232, <https://doi.org/10.5194/nhess-23-1227-2023>, 2023.
- Müller, M. and Kaspar, M.: Event-adjusted evaluation of weather and climate extremes, *Nat. Hazards Earth Syst. Sci.*, 14, 473–483, <https://doi.org/10.5194/nhess-14-473-2014>, 2014.
- Rybka, H., Haller, M., Brienens, S., Brauch, J., Früh, B., Junghänel, T., Lengfeld, K., Walter, A., and Winterrath, T.: Convection-permitting climate simulations with COSMO-CLM for Germany: Analysis of present and future daily and sub-daily extreme precipitation, *Meteorol. Z.*, 32, 91–111, <https://doi.org/10.1127/metz/2022/1147>, 2023.
- Sieck, K., Pinto, J. G., Geyer, B., Keuler, K., Beier, C., Braun, C., Ehmele, F., Feldmann, H., Frisius, T., Heinrich, P., Hundhausen, M., Petrik, R., and Trachte, K.: NUKLEUS – A First Kilometre Scale Multi-model Climate Ensemble for Germany: Evaluation, *EGUsphere* [preprint], <https://doi.org/10.5194/egusphere-2026-1024>, 2026.
- Voit, P. and Heistermann, M.: A new index to quantify the extremeness of precipitation across scales, *Nat. Hazards Earth Syst. Sci.*, 22, 2791–2805, <https://doi.org/10.5194/nhess-22-2791-2022>, 2022.