Interactive Discussion: Author Response to Referee #2

Present and future trends of extreme short-term rainfall events in Germany, by downscaling convective environments of ERA5 and a CMIP6 ensemble

Gerd Bürger and Maik Heistermann
EGUSphere, doi:10.5194/egusphere-2025-3584

RC: ***Reviewer Comment***, AR: *Author Response*, ☐ Manuscript text

Dear referee,

thank you for taking the time to review this manuscript as well as for the positive feedback.

We agree to most of your comments, and will address them in a revised manuscript accordingly. Please find a point-by-point response to your comments below.

Thanks again for your efforts!

Kind regards,
Gerd Bürger and Maik Heistermann

1. Comments and responses

1.1. Major comments

RC: ***The manuscript is very well organized and structured and clearly written. The scientific questions are stated clearly, and the method is for the most part well presented and explained. Although, since the details of the methodology were described in a previously published paper, I felt the need to read that one as well (which was also an interesting read!). The topic is interesting and the study represents, in my opinion, an interesting illustration of how statistical tools can be used to derive trends and changes in local phenomena, such as extreme precipitation, based on large-scale fields which are represented with greater confidence in coarse scale climate models. Below follows specific, mostly minor, comments and suggested changes. If these are appropriately addressed, I would find this manuscript ready for publication.***

AR: We truly appreciate this positive comment.

RC: ***L13-14: "with an estimated damage...". Suggestion to change to "with an estimated cost due to flood damages..." or similar.***

AR: We changed the sentence to "As a result, flash-floods are for Germany the most costly type of extreme events, with an estimated cost of more than 70 billion EUR due to flood damages for the last two decades."

RC: ***L15: "a warmer atmosphere" → "a warmer Earth atmosphere" (since it may not apply to all atmospheres, even though the moisture-holding capacity increases).***

AR: As a warmer atmosphere can hold more water on every planet, specifically referring to Earth here may distract the reader. But for clarity we change the phrase to 'can hold more water'.

RC: ***L18-20: Is this sentence referring to the use of very high-resolution convection-permitting models? If so, I suggest the sentence be modified slightly since these models do not concern the implementation of convection processes, but rather enabling the removal or turning off of convection parameterizations and treating (deep) convection explicitly.***

AR: We changed the sentence as follows:

> To quantify future impacts, state-of-the-art climate models increasingly replace the numerically cheaper form of convection parameterization by an explicit implementation of the very complex physics of (deep) convection into their thermodynamic routines.

RC: ***L24: I'm not sure I fully agree with this, or maybe I misunderstand. I'm under the impression that conventional statistical downscaling approaches have not been used historically, since they haven't shown capability of generating realistic high-resolution convective rainfall fields with correct spatio-temporal patterns and extreme values. Hence, the promising avenue of ML methods...***

AR: We generally share the reviewer's critical view on the statistical downscaling of high-resolution convective rainfall fields. But corresponding fields from dynamical downscaling are not 'correct' either. Both have their pros and cons. We have augmented our original sentence (although it had been quite neutral in this respect) with a corresponding note of caution, as follows:

> More conventionally, a variety of empirical schemes known as 'statistical downscaling' have been employed, with varying success, to obtain realistic heavy local rainfall patterns that are consistent with the large-scale atmosphere; for a more recent overview see Devadarshini et al. (2025).

RC: ***L36-38: Do you plan to extend the analysis using the new CMIP6/EURO-CORDEX suite when these downscalings have been published? Would be interesting to see the impacts on the classification of using input data with significantly higher spatial resolution.***

AR: We agree that this would be an interesting analysis, especially with respect to obtaining better *cape* and *cin* estimates.

RC: ***L56: Even though the reader most likely understands what P00 represents, it's worthwhile to define it explicitly in the text.***

AR: We now have "[...] denoted as P_{00} (all events) [...]"

RC: ***L64: overcome → reach (?)***

AR: Good suggestion, has been implemented.

RC: ***L79: "same resolution" → "same temporal resolution" (if you refer to the temporal resolution)***

AR: The suggestion has been implemented.

RC: ***L127: What is the main reason for the reduced or non-existing skill for the P99 severity level? Sampling size? Would be good if you could elaborate on this further. Is it still worthwhile to continue consider P99 given the small skill?***

AR: We can only speculate about the reason behind the small skill for the P_{99} severity level, but of course, the small sampling size is an obvious hypothesis. Note that this also affects P_{90}. We have augmented the text accordingly, as follows:

> The decreasing scores from P_{00} to P_{99}, with the latter having almost no positive skill left, are in line with the decreasing sample size, or increasing rareness, of the classified events. This may have a number of reasons: very extreme events may not be represented as well in the ERA5 predictor fields; they may contain stronger stochastic elements; and as a simple arithmetic effect, misclassifications have a larger influence when numbers are small.

In the revised manuscript, we have also emphasized that the results for P_{99} have to be interpreted with particular care. Still, the ANOVA analysis shows that neither the severity level (SEV) nor the classification method (DSC) are a major source of uncertainty in our trend analysis.

RC: ***L134: cf. 2.3 → cf. Sec. 2.3***

AR: Thanks, this was corrected.

RC: ***L144: The "base rate" expression is used frequently through the manuscript, however, it hasn't been clearly defined. Is it the frequency of occurrence of a certain class; either observed, based on ERA5-fields or simulated in GCM historical simulation? Please clarify.***

AR: The base rate is, as the referee assumed, the frequency of occurrence of a certain class, based on the observations (CatRaRE). We have added a corresponding definition near Tab. 2:

> The relative frequency of each of these 6 CatRaRE-based classes, in the following referred to as the *base rate* of the class, are displayed in Table 2.

and changed the caption accordingly.

RC: ***L156-158: Remarkable difference between the two GCMs, not only in trends but also in the inter-annual variability. Could you also add the ERA5-driven classification in the same figure (Fig. 7)? Perhaps just the variability (standard deviation?), so you will get an indication of how well the GCM represent this aspect (I guess you cannot expect the trends to be similar to ERA5-classification). How does the variability in these GCMs compare with the other GCMs?***

AR: The differences are remarkable indeed. If you inspect the SI, similar differences are seen between, for example, CMCC-ESM2 and CNRM-CM6-1. Adding even more curves to the already quite populated Fig. 7, however, would in our view blur the figure too much. Instead, Fig. S4 can be taken as an ERA5-reference for the SW region, with interannual variability being somewhat between the two GCMs. Note that standard deviation is composed of the trend signal and the detrended data, so simply displaying the annual data themselves is informative enough in our view.

RC: ***L165: Regarding the different model behavior; have you looked at the temperature and humidity estimates and trends from the GCMs, both over the very local target area and on the larger (European?) scales? Could this give a clue on the different behavior?***

AR: No, we have not looked into these variables, and yes, it can provide important clues. The reason must eventually be found in the different convective environments and, hence, in the temperature and humidity profiles. We have checked spatial resolution, but that was inconclusive. A closer look into these profiles is definitely indicated, as it may reveal crucial differences in the GCMs. But that is beyond the scope of the current study.

RC: ***L166: "Examples analogous to 7 are Figs. S5-S8" → "Examples analogous to Fig. 7 are presented in Figs. S5-S8"***

AR: Thanks, was corrected.

RC: ***L172-173: Have you analyzed specifically the changes in intensity and frequency of convective rainfall events, or do you infer this somehow from Fig. 8? The normalized cp in Fig. 8 is derived from cp intensity, right, not frequency?***

AR: This was inferred directly from Fig. 8, based on the ideas of Chen et al. 2020 (e.g.: higher *cin* –> lower *cp* frequency). And no, the normalized *cp* in Fig. 8 is convective precipitation itself, as provided by the GCMs.

RC: ***L174: What do you mean, with respect to cin/cape changes, by "...in line with the law of clausius-clapeyron and the global warming narrative"? Decrease in rainfall frequency but at the same time more intense rainfall when it occurs? Consider rephrasing.***

AR: We have rephrased the sentence as follows:

> Growing precipitation intensity is in line with the law of Clausius-Clapeyron and the global warming narrative; *cin* trends appear to depend on more factors, such as land vs. ocean, see Chen et al. (2020), Myhre et al. (2019), and also section S3 of the SI.

RC: ***L176: "More examples are Figs. S9-S12" → "More examples are presented in Figs. S9-S12"***

AR: Was corrected.

RC: ***L176: Also, you have mainly discussed results for SW region in this section. It would be nice to include some information on the results over the other sub-regions as well (even though you refer to the SI). Do you see the same model behavior or are there any other conclusive results or take-home messages?***

AR: We have added the following:

> The above results pertain to the SW region. Corresponding results do not change much for the other regions, which also reflects the message of the last section.

RC: ***Figure 5: Why are the (stacked) bars plotted from top (prob 1) – down (prob 0)? Wouldn't it be more intuitive to have them originate from prob 0? Or am I misreading this?***

AR: Indeed, the bars are stacked from bottom (0) to top (1) (with "class 0" being gray) in the plot, but reversed in the legend. This should be fixed in the revised version. Thanks for spotting this!

RC: ***Figure 7: As mentioned in comment above, it would be nice to have some metric of the spread (variability) from ERA5-driven classification in the plots as well. Perhaps same in Fig. 8 (if possible)?***

AR: This was in fact considered, but we decided against it because the figure is already heavily populated. We think variability is clearly visible from the annual dots themselves. See comment above.

RC: ***Figure 10: In the figure text/caption, it would be good to explain again what the "base rate" represents.***

AR: See our comment above.

RC: ***Figures 11, 12: A minor thing; these figures are quite large (in printed manuscript version), while the content of both is rather minimalistic. You could easily make these smaller and put them together in one figure, or even merge into one panel with semi-transparent (and colored! Unless the economy is an issue :))***

AR: We are planning to merge both plots into one for the revised version.