

This manuscript explores the spatial and temporal variability of heavy precipitation events across Germany using high-resolution data from various rain gauges. By applying a block maxima sampling, the authors analyse annual maximum precipitation totals for durations ranging from five minutes to seven days. They apply simple statistical metrics and tests to assess variability and trends. The findings indicate that while short-duration events are scattered and driven by local convective processes, longer-duration events are more influenced by topography and large-scale atmospheric patterns. Although positive trends in intensity are reported in short-duration (e.g. hourly) events, a large fraction of trends are not significant. For long-duration rainfall (e.g. 7 days) also negative trends are reported, where the significance depends on the length of the analysed time series.

The study is affected by major drawbacks:

- The manuscript lacks a scientific discussion section
- 5-min rainfall intensities show jumps, which are presented as a scientific finding. In turn, they should be described in the data description, where also the reason (change of sensors) should be reported. Can those jumps be corrected? Otherwise, I'd suggest removing these 5-min-results from the manuscript, especially in the trend analysis.
- Statistical tests seem not to be adjusted for multiple testing correcting the false discovery rate (FDR; see e.g. Wilks, 2016: <https://doi.org/10.1175/BAMS-D-15-00267.1>).
Trend analysis for long durations might benefit from seasonal maxima, as the behaviour of decreasing/increasing maxima might be hidden in changes of seasonal maxima. In some parts of Germany, long duration maxima dominate during summer (e.g. southern Germany), while in other parts long duration maxima dominate during winter (e.g. north-western Germany). Both seasons might show different trends (drier summers, wetter winters).
- The manuscript would benefit from a more consistent usage of the term "variability," which currently conflates several distinct concepts. Specific instances of inconsistent usage include:
 - Spatial variability: The term is used to describe the geographic distribution of AMPT medians (Line 191)
 - CV is used to describe the relative interannual fluctuations of AMPT (Line 197) and "also the event-to-event variability at each station"
 - Temporal variability: This refers to long-term monotonic trends (Line 161)
 - Scale-dependent "visual" variability: In lines 175–178, the authors refer to a "visually apparent" greater "overall variability" in long-duration events due to the widening of the interquartile range (box plots), while simultaneously clarifying that this does not represent an "actual increase in their variability". Describing the absolute range of variability for cumulative rainfall is somewhat misleading; if the y-axis of Figure 3 would show intensities instead of cumulative rainfall, the visual notion of variability would probably be inverse.
 - One important concept of variability (internal climate variability; ICV) is not fully captured in the manuscript. Partly, event-to-event variability is governed by ICV. Later, in lines 366-375, "decadal fluctuations" implicitly describe the realization of ICV on the decadal scale. It is important to understand that the ICV is the driver for year-to-year temporal variability but also spatial variability,

especially for smaller-scale convective events governed by the chaotic behaviour of the climate system.

While these different concepts can and should be applied in the manuscript, they should be clearly distinguished and also understood.

- The authors discuss possible reasons for the reported spatial patterns (convection for short-duration AMPTs, large-scale drivers for long-duration AMPTs, “decadal fluctuations” of short-duration AMPTs “explained by changes in the behaviour of large-scale circulation patterns, potentially reinforced by an intensified convective response to anthropogenic warming”. A thorough statistical analysis of these assumptions could add value (and scientific novelty) to the manuscript making use of the high-resolution data set.
- In sum, the degree of scientific novelty is low to moderate; the data set is valuable, however, this study is not a data description paper and the data cannot be accessed openly. Trends are still not significant (and might be even less after correcting the FDR); the from my perspective biggest novelty would be the finding that short-duration (5 min, 60 min) AMPTs show similar intensities across Germany, i.e. they are not as much governed by topography. In turn, Shehu et al. (2023: <https://doi.org/10.5194/hess-27-1109-2023>) show a considerable topography-influenced gradient for 1-year return levels of 5 min and 60 min rainfall (Fig. 13 in their article). I am not sure if your findings contradict (based on similar data sources; then you have to carefully explain why) or if the visualization of Fig. 4 leads to a biased perception of low differences, as the colorbar extends to 0mm. Hence, all circle markers are in relatively dark greenish-blueish color making them hard to distinguish leading to a low-difference-perception.

Due to these major points I'd suggest to reject the current version of the study but to encourage the authors to re-submit if they are willing to address these major points. At this point, I won't add any minor comments on figures, references, wording, etc.