Comment on nhess-2022-225
Sergiy Vorogushyn (Referee)


Review of the manuscript by Ludwig et al. submitted to NHESS

The presented manuscript characterizes the extreme flood event in July 2021 in western Germany from meteorological, hydrological and geomorphological perspectives. The manuscript is a companion manuscript to Mohr et al. (2022) and focuses on putting the above-mentioned event characteristics into historical perspective. Furthermore, modelling experiments following pseudo-global warming storyline approach were performed with a climate model at convection permitting resolution for scenarios within -2K to +4(+3)K range. The resulting ensemble was used to run a hydrological catchment model to estimate the impact on flood flows. Additionally, the hydro-morphological changes and their impact on flood hazard are discussed.

The manuscript is well-written and mostly well structured. However, it is not concise and some excursions can be significantly shortened as proposed below. The analysis of event severity in meteorological and hydrological terms is not only nicely placed into historical context, but due to selection and analysis of various regions, the spatio-temporal perspective of the July 2021 is well elucidated. The use of LAERTES-EU ensemble and two complementing approaches (PGW and scenario ensemble) to assess the effect of climate change on meteorological and hydrological hazard strongly enrich the presentation. I consider the manuscript to be a valuable contribution to the analyses of 2021 flood event and a good fit to NHESS journal. Said that, the current presentation needs significant revision to address several mostly structural issues and a number of minor issues related to formulations. At several occasions clarity is lacking. I therefore suggest major revision.

With kind regards,

Sergiy Vorogushyn
**Major comments:**

L105ff: I feel research questions III and IV are very similar and refer to the potential characteristics of precipitation under future climate conditions and their implications for flood events. Consider merging these two questions.

L170-174: At this stage it is not quite clear what exactly you intend to do, i.e. what you mean by “broader classification”. Some more specific goal setting would be helpful.

L274-275: The definition of HPEcrit should be more precise. Do you first identify the exceedance of 50-year return period on the KOSTRA 8x8 km grid and then count cells with exceedance if they form a contiguous area above 1000 km²? Do you compute return periods at 1x1 km HYRAS-DE resolution and compare to 8x8 km² KOSTRA or do you regrid HYRAS-DE to the KOSTRA resolution? Specify details in the manuscript.

Related to the previous comment: Is the area (A) in Table S2 the contiguous area of HYRAS-DE grid cells exceeding 50-year return period for 24 hour rainfall in comparison to KOSTRA? If so, specify this in caption.

L303-305: First, not all 26 but only 20 events are displayed in Fig. 2 and Fig. S1. I am not sure if you have enough evidence to claim that if heavy rainfall occurs in eastern and southeastern Germany during summer it is associated with Vb weather patterns. This is indeed documented in the literature for some of the presented events, like 2002, but I am not aware of such established association with Vb for all of the listed events. This statement should be relaxed and maybe shortened.

L313-329: this description can be significantly shortened, particularly by focusing on the July 2021 event in relation to a few other events in terms of precipitation location, area, intensities etc. L325-329: this somewhat arbitrary distinction into two types of rainfall field is difficult since the data only covers Germany and not riparian countries. I would omit this.

Having read the analysis in section 3.1.1 I question the introduction and use of PSI in the context of this study. It does not bring much additional information, but causes quite some confusion in the results. Yes, it has persistence as additional characteristic, but results in the end in a quite different ranking of events. Since the purpose of the study is
not to compare different event indices (and there are a few others out there, like WEI (Müller & Kaspar, 2014) and xWEI (Voit & Heistermann, 2022)), but to put the July 2021 into the historical context, I would suggest to focus only on the HPEcrit index and omit PSI. It would improve the clarity of the manuscript. Figure 2 should then be redesigned for HPEcrit index.

Chapter 3.3 needs considerable revision. It starts with putting the July 2021 event into the historical perspective but then loses the focus. I would suggest to focus on general changes in the landscape including (1) urbanization, (2) construction of bridges and transport of infrastructure, (3) land use & agricultural practice and support them with a few examples.

Section L510-529 can be omitted.

You can use the examples from local places (Altenahr, Schuld, Dernau), but these should be introduced and shown on a map. It is not possible for a reader not being acquainted with the local geography to understand these details. Either refer to figures PART1 or introduce those localities here. In overall, I would significantly shorten the description of those examples of geomorphological long-term changes.

L548-551: This finding is not surprising if you prescribe CC-scaling for specific humidity of the initial and boundary data (L230-232), is it? Then one should rather discuss the limitation of this setup and not directly confirm it with observational evidence. L553 -561: This paragraph should be formulated more precisely. Are you talking here of the average July 2021 precipitation in a specific region? I am then puzzled, how can the return period of this precipitation in a colder climate become smaller (5 years) compared to the warmer climate (20 years). L558: there is no return period of the control run, but of some precipitation amount. Overall, it is difficult to keep track of numbers and their changes for specific regions, please, consider bringing these numbers and regions into one table. Also, the hydrological results (L580-586) can be introduce into this table.

L586: Can the conclusion about the amplification of hydrological response be supported by the model analysis? You can derive runoff coefficients, share of quick(overland) flow, concentration times. Of course, these are modelled indicators which may or may not reflect the reality, but at least it will explain the model response.

L611-612: Why is it so? Can you speculate about a possible reason?
Chapter 4.3 feels like a foreign body. It is not really linked to the rest of the manuscript and contains very general discussion not linked to any concrete findings. I suggest to remove it completely, but use some (few) thoughts and literature references in redesigned Chapter 3.3 focusing on geomorphological perspective. But the focus on historical changes should be kept in order to be consistent with the original scope of the manuscript.

Discussion and conclusion chapter require considerable revision. It is partly very general and does not distill main findings. L684-721: the attempt to provide answers to the originally formulated research questions does not work in my view. The questions are not narrow enough to be answered in a few sentences, so that they can be understood without further information. A more detailed information is provided below, but it is decoupled from the above answers. I suggest to dissolve this structure and provide the concise discussion of the four questions/issues and related findings in plain text. It would have 4 (or 3) paragraphs (if you merge QIII and QIV) summarizing the discussion.

**Minor comments:**

L3: Part 1 – substitute by the proper reference to the (discussion) paper.

L5 flood hazard

L7: return values and periods – what is the difference between the two?

L11: hazard assessment of flood risk – reformulate

L29: remove ‘widespread’

L31: please, specify that death toll of more than 180 refers to Germany only. The description above was focusing also on non-German part. Also, afterwards economic losses including neighboring countries are mentioned.

L44-54 can be omitted.

L58: ‘in terms of peak discharge’. Actually, the 1910 event peak was about the half of that in 1804 and 2021. I would therefore not speak of 1910 being comparable to 2021.
L59: you mean here official estimates by LfU RLP. This should be mentioned.

L60: Vorogushyn et al. (2022) – meanwhile published.

L114: three rectangular geographical domains

L147ff: It should be noted that all values are rough estimates and are not directly measures due to failure of gauging stations.

L166: However, the estimations of return periods across different federal states are inconsistent – Is this what you want to say?

L258-263: Can this be omitted?

L205 & L228: Here you are talking about the PGW up to +3K. In the abstract up to +4K – correct.

L247: +2K – GWL2 or +3K – GWL3?

L273: specify here that you focus on daily precipitation totals.

L298: is illustrated

L299: Omit last sentence.

L317: awkward sentence – reformulate

L376: do you mean “extent” or rather the location of the precipitation field within CReg?

L397: You can easily fit linear models and show how strong the linear relationships are.
What is “approximately linear”?

L399: I cannot confirm that only 5 GRDC events exceed the 2021GD. How do you determine this? If you fit a linear model to 2021GD data and even if you shift it to the most outlying point retaining the slope (so, that the other red dots are below), would it be still exceeded by only five GRDC events? Looks like not. But it should be shown. L400-405: these peaks/gauges can be marked in Figure 6.

In Figure 6 I miss the peak for Altenahr and Müsch in the 2021GD dataset, ~1000 m3/s and ~500 m3/s, respectively. The red dot at ~500 m3/s is not Müsch but likely Kordel at Kyll with 816 km2 basin area. I still count 10 red points. If two are missing, what are the other two?

L418: a few other return periods

Figure 7: What are orange dots? These are not annual maximum peaks, but a few peaks above a threshold? Please, explain the caption. Hence, the statement in L431-432 is not correct. The HQ100 estimate of 241 m3/s is based on the AMS in the period 1947-2016.


L436-437: gauge recordings did not miss the historical peaks, because no recordings were carried out prior to 1946 - reformulate.

L439: Henrichs with one ‘n’.

L443ff: I suggest a slight reformulation here related to the work of Vorogushyn et al. (2022), where I was a co-author. (1) It was exactly the purpose of the study to show the limitation of the extreme value statistics without considering historical floods, and not to estimate an unrealistic return period of 10^8. (2) It should be noted that Vorogushyn et al. (2022) did not use 2021 event for fitting the distributions, while Henrichs (2022) did. So, the estimates are not fully comparable.

L468-472: is there photographic evidence or any type of documentation and analysis?
L477: this statement requires a reference.

L563-566: This can be omitted.

L624: This statement is not precise, unless areal dimensions are not specified. Since thick lines in Figure 12 have different slope, the relative change will be different. Instead of the discussion in L624-629 I would suggest to discuss the mechanisms of increasing precipitation clusters in the model, and support this by literature references analysis observational evidence if exists. The notion that for large clusters model boundaries provide represent a limitation is valuable.

L717: not setup, but weather pattern.

L716-718: this comes out of blue. At any point in this manuscript it was mentioned that July 2021 was due to Vb weather pattern or its variation.

L723-724: this statement referred to 2013 flood. It gets messy here (L719-723). Please, discuss rather your concrete findings on the uniqueness of precipitation related to Question I here. E.g. one important message: the event was quite exceptional for the region, but not unique in terms of intensities and extent for a larger area in Germany and so on. In this respect, the discussion of the hydrological perspective (L724-734) is a perfect blue print.

L746: +4K or +3K?

L750-754: the formulations should be reformulated more precisely. It is not clear what you mean: The increase in precipitation corresponding to increase in return period from 5 to 10 years... (which precipitation, in which scenario? Increase compared to what?). Doubling of the return period from 5 to 10 years actually means reduction of probability of occurrence of this amount and not doubling. It depends on what you mean in relation to what.

L770: precipitation is not a task – reformulate.

L777-779: Where were photos and chronicles used in the presented manuscript to reduce the uncertainty in HQ100?
References:


