REVIEWER #1

First, we would like to thank the reviewer for her/his insightful comments, which have greatly contributed to improving the text. In making corrections, we have tried to follow as closely as possible the suggestions made.

The manuscript by Ludwig et al. is part II of two papers discussing an exceptional flood event in 2021 in Central Europe. Part II's goals are manyfold: it discusses precipitation and discharge records and valley morphology in historical context and potential future flood events in storyline and in projection approaches. The paper is interesting to read. But it is also very long, sometimes a bit difficult to read, and it is not obvious what is the new take-home message. The manuscript seems to have taken advantage of perspectives, methods, and simulations available at the Karlsruhe Institute of Technology and merged them into one text. This is my main concern. The manuscript should better explain

(a) why precipitation, discharge, and valley morphology are discussed in one paper and what is to be learned from this (it discusses these aspects more or less independently),

A: These are all important points/ingredients to better understand the resulting flood event in 2021. One of the main aims of this study is to put the event in a historical context and to decipher the connections between extreme precipitation and discharge and the role of the specific morphology in the Ahr valley. Based on our analysis, we found that although precipitation was strong, even stronger events could occur based on the LAERTES-EU ensemble. The discharge of course depends much on the exact location of the precipitation, particularly when considering very narrow valleys like the Ahr and its relatively small catchment area. Although the precipitation was not record breaking during this event, discharges where at the upper limit of the observations, also fostered by the valley morphology. Thus, a more severe precipitation event (whether under present or future climate conditions) at the same location could potentially lead to a much more severe discharge (flood) event.

(b) why the future climate is discussed with PGW and with high-resolution projections (KIT-KLIWA), but the discussion of advantages and disadvantages concerning telling something about floods in Central Europe is lacking ("KIT-KLIWA ... confirms the CC-scaling" is about all; are the PGW experiments needed if anyway only the well-known C-rates are applied to obs. data in runoff simulations?), and

A: In the PGW experiments, we only investigated the influence of thermodynamic changes on the specific 2021 event. With the KIT-KLIWA ensemble, we consider the full climate change signal under specific global warming levels, including changes in atmospheric dynamics that are neglected by the PGW approach. Thus, the results are also a proof of concept, that such PGW studies might be an alternative way to discuss the impact of global warming on extreme (precipitation) events. However, we will modify the text in terms of a discussion of advantages and disadvantages of the two different approaches.

(c) what can be learned for other catchments worldwide (concerning climate change, land use, methods or whatever - why a scientific paper and not a report?).

A: First of all, the main goal of the two companion papers PART1 (Mohr et al., 2022) and PART2 (this manuscript) is to provide a scientifically sound and comprehensive account of the 2021 event in all its facets (see PART1). But it also goes clearly beyond a mere report of the event by putting it in PART2 into the perspective of (i) historical events, and (ii) climate change. We think this justifies submission as a research paper. Nevertheless, it is a valid question to ask – going beyond the target region and the 2021 event itself - what can be learned for other catchments worldwide. We suggest there are the following three points related to the estimation of hydrological extreme values in a changing climate: (i) Beware of extrapolating gauged-based extreme values too far (e.g., estimating 10,000-year floods from less than 100 years of gauge observations), (ii) cc-scaling takes place, and should be taken into account, (iii) cc-scaling might be further amplified by the non-linear rainfall-runoff transformation. All of these points are already discussed in the discussion and conclusions section 5, but for clarity we suggest adding to the same section in a revised version of the manuscript a sentence summarizing the generalizable hydrological implications of our study.

Mohr, S., Ehret, U., Kunz, M., Ludwig, P., Caldas-Alvarez, A., Daniell, J. E., Ehmele, F., Feldmann, H., Franca, M. J., Gattke, C., Hundhausen, M., Knippertz, P., Küpfer, K., Mühr, B., Pinto, J. G., Quinting, J., Schäfer, A. M., Scheibel, M., Seidel, F., and Wisotzky, C.: A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe. Part 1: Event description and analysis, Nat. Hazards Earth Syst. Sci. Discuss. [preprint], https://doi.org/10.5194/nhess-2022-137, in review, 2022.

In some places the wording should be clarified, or minor mistakes corrected:

line 16: "scales to first order"? Does it mean it is between 1/10 and 10 times the CC-rate?

A: We removed “*first order*” as precipitation scales with CC-rate

l32: "Europe in the last half-century" - last 50 years?

A: Changed as suggested

l35-39/40 could be deleted.

A: Deleted as suggested

l55: perhaps "A key aspect is a deeper analysis of the 2021 flood event taking a ... perspective"?

A: Changed to: “*A key aspect is a deeper analysis of the 2021 flood event based on a long-term climatological perspective*.”

l73: "GCMs (usually 100 to 200 km)" should be more specific. HiResMIP or paleMIP models have very different grid spacings.

A: In the revised versions, we will be more specific here, adding information of the range of grid spacing for different purposes.

l94: PGW is also imprinted on the lateral boundary forcings, I assume?

A: '*lateral boundary*' forcing has been added

l171: "Therefore, ..." - please explain

A: This sentence will be deleted, as it is more a less a doubling of information already provided before.

l210 vs l747: GWL of present-day is 0.46 K or 1.09 K

A: The 0.46K corresponds to the GWL of the reference period (1971-2000) in comparison with pre-industrial conditions based on different observational datasets (Vautard et al., 2014). As the regional model simulations only start after 1950, the period 1971-2000 is considered as the reference period, thus already featuring a GWL of 0.46K with respect to the pre-industrial period. This means that climate change signals of the RCM simulations under GWL2 (with respect to the pre-industrial period) correspond to a +1.54K global warming with respect to the reference period (1971-2000). The 1.09K is the current (2011-2020) GWL based on observation with respect to the pre-industrial period (IPCC, 2021). To avoid any confusion, we deleted the 0.46K in line 210 and only refer to the method applied by Teichmann et al. (2018).

Vautard R. et al. (2014) The European climate under a 2 °C global warming. Environ Res Lett 9:034006

l265: It should be better explained why you can assume that the LAERTES-EU data provides independent 12500 years of data?

A: As presented in the early years of operational numerical weather forecasts by, for example, Lorenz (1982) or Dalcher and Kalnay (1987), and recently by Fuging et al (2019), the intrinsic predictability of the atmosphere has a limit of about 12-15 days due to the chaotic nature of the atmospheric system. The largest part of LAERTES-EU (namely data blocks 2 and 4) is a so-called forecast ensemble which means, each simulation is initialized ones and then run free for 10 years. Therefore, it can be treated the same way as a numerical weather forecast but on a 10-year (decadal) scale. Furthermore, the simulations were initialized on a day in November but within LAERTES-EU only data from 1 January are used. So, we already have a spin-up phase of about 2 months which is not considered in the ensemble. As mentioned above, after about 2 weeks there is hardly any dependency between the single members of the ensemble anymore, so using a 2-month spin-up, we can assume that LAERTES-EU provides independent 12,500 years of data. We will add a comment to the manuscript related to this topic.

Lorenz (1982): Atmospheric predictability experiments with a large numerical model. Tellus A, 34, 505–513, https://doi.org/10.3402/tellusa.v34i6.10836.

Dalcher and Kalnay (1987): Error growth and predictability in operational ECMWF forecasts. Tellus A, 39, 474–491, https://doi.org/10.3402/tellusa.v39i5.11774

Fuqing et al. (2019): What Is the Predictability Limit of Midlatitude Weather? J. Atmos. Sci., 76 (4):1077–1091, 10.1175/JAS-D-18-0269.1

l298: "in total" - delete?

A: Deleted as suggested

Fig 3: bubble 13 cannot be seen.

A: Bubble 13 is very tiny; we will find a way to make it better recognizable. Furthermore, as suggested by Reviewer 2, we will change from PSI to HPEcrit related quantities in Fig.3.

l372: "However ..." - this is a statement for the conclusion

A: We agree with the reviewer, that this statement should also appear in the conclusion. This will be done in the revised version.

l655: "The second type ... scaring ... changing the boundaries ...". Please, reformulate.

A: The sentence was in fact convoluted, we reformulated as suggested.

l703: "precipitation intensities" or do you mean the daily mean amounts discussed above?

A: Yes, we are referring to the daily mean amount discussed above. This will be changed in the revised version.

l722: "... found no observed ..." - observed but not found?

A: We agree, this sentence reads strange. We will reformulate it in the revised version.

l772: "basically" - reformulate, please

A: Changed to "*physically*"

l786: here you refer to the spatial information in the PGWs not used in the hydrological modelling? In what respect shall the new simulations help? In understanding the 2021 event?

A: Basic idea behind the PGW approach is to extract the pure thermodynamical effect of climate change on extreme precipitation by forcing the atmospheric dynamics to follow the present-day state. However, due to the model characteristics, parameterizations, etc., there are little deviations in the spatial structure of the precipitation field which would further propagate into the hydrological model. The intention of the presented analysis was to extract the pure hydrological response to changed atmospheric conditions. Therefore, we estimated the scaling factors for -1K and +2K and applied it on the "best-available" precipitation data for which the flood peak and course were represented best. Hence, conclusions can be drawn in how far climate change already influenced such peak discharges and to which extent it will change in the future. We agree that the last sentence is misleading and changed them accordingly. Furthermore, we reformulate the beginning of Sect. 4.1.2 (see also replies to Reviewer 2).