General comments

In their manuscript *"A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe. Part 1: Event description and analysis"*, Susanna Mohr and colleagues provide an overview of the disastrous July 2021 flood event.

The authors have done an impressive job in collecting and compiling information on various aspects of the July 2021 flood, and the sheer effort behind this needs to be appreciated.

Having said that, the objectives and specific research questions related to this study remain unclear to me. In ll. 74 ff. of the introduction, the authors state that *"[...] the objective of this two-part study is a multi-disciplinary assessment of the entire process chain of the July 2021 flood in central Europe - from causes to impacts to historical classification and climatological context [...] While Part 1 focuses on the description of the event across various disciplines (meteorological, hydrological, hydro-morphological, economic) [...]"*. This is not a research question, and nowhere in the paper, the authors specified their idea of a *"multi-disciplinary assessment"*. In ll. 754-755 of the conclusions, they state that *"this paper examined the complex interactions among meteorological, hydrological, hydraulic, and geomorphological processes and mechanisms that led to the extraordinary flood [...]"*. I find this statement difficult to confirm: instead, the manuscript largely remains a *description* from different sub-disciplines (meteorology, hydrology, hydro-geomorphology, impacts/damages), listed one after another, but mostly not related to each other in terms of an *analysis*. This is a challenge that many papers have to cope with when they aim to provide a holistic assessment of an extreme event - I noticed a similar referee comment on the paper of Caldas-Alvarez et al. (2022) about the Berlin 2017 event in this same special issue ([link](#)).

While it is, in the direct aftermath of such a disaster, valid and required to focus on the rapid compilation of information and data, and to make such compilations available to decision makers, the research community, and the general public, I wonder whether we should have, by now, reached a phase in which the research community should find more well-defined modes of event analysis. Then again, I am aware that putting together such a manuscript requires a lot of time, and that the processes for this probably had already started in 2021.

My overall recommendation is that the authors take a step back and re-evaluate the purpose of this paper. What is your scientific objective, aside from compiling as much as we know about the event? I think it would help the paper very much to, after a brief synopsis of the event, identify maybe two or three important and *specific* research questions (e.g. with regard to specific interactions), and tell the story along these questions - instead of just listing one discipline after the other. Many aspects of the paper are quite interesting in itself, but maybe not required for a holistic view? Do they have to be a part of this study? For example, I found it difficult to understand why we need the section on the "synoptic overview and atmospheric characteristics". And the predictability of heavy rainfall (weather forecast analysis) would be relevant only if the authors had actually investigated the failure of the early warning chain and the resulting implications on the impacts in terms of damage and loss of lives. Also the interaction between discharge dynamics, hydro-morphodynamics and impacts is not yet sufficiently elaborated.

I understand that it is challenging to revise the paper along these lines, but I am confident that the authors will find an adequate way. Is it helpful to recommend that the paper should be much shorter? It almost took me three days just to work my way through it. I think the length of the paper could be easily reduced by at least a third and I hope this will help separating what is relevant from what is not.

One more comment on the issue of multi-part papers: Surely it is up to the editorial team to assess whether a multi-part publication is warranted. Personally, I have never really understood the need for multi-part papers. A paper should be self-sustained, and *evaluated* as an individual piece of scientific work. This is also the basis for this review. Of course, papers can and should refer to each other and build upon each other, but in my view, that does not require an explicit multi-part approach. In the present context, the multi-part approach supports the impression that PART1 is more about compiling "everything we know" instead of asking and addressing well-defined research questions.

Specific comments

Is this a research article or a review?

To be honest, I am also not sure which type of manuscript I am dealing with. Over large parts, the manuscript more resembles a review paper instead of a research article. It struck me that there is no formal "results" section (instead "event description and analysis"). This is not a fundamental issue in itself; however, it again points to the fact that there are no specific research questions and hence no specific results to address these. Furthermore, the content in section 3 ("event description and analysis") is largely not based on the data and methods which have been described in section 2 ("data and methods").

I would like to discuss this in detail: Section 2.1 (data) documents precipitation data, atmospheric model data, river gauge observations (not the data/methods on reconstruction of water levels/discharge), Sentinel1/2 data (for inundation mapping), and traffic data (reports on road and railway disruptions). Section 2.2 (methods) documents the trajectory analysis (for moisture source analysis), extreme value statistics for precipitation, and the computation of the antecedent moisture index. Together, section 2 incompletely addresses the methods and data that were used to put together section 3; in section 3.2, for instance, reconstructed water levels/discharges play an important large role; section 3.3 (hydro-morphodynamic processes) is almost entirely unrelated to data and methods documented in section 2; and section 4 (impacts and consequences), too, is based on methods and data sets (e.g. aerial/media footage, loss models, insurance data) which were not mentioned in section 2 (except the Sentinel data and the reports on traffic disruptions).

Altogether, I understand that the variety and mass of methods and data that section 3 is based upon is almost impossible to describe in section 2. The reason for that is that much of the content shown in section 3 is not really based on the application of data and methods in the context of study, but rather a compilation and synthesis from other recent studies about the July 2021 event, namely Fekete and Sandholz (2021), Dietze et al. (2022), Thieken et al. (2022), Schäfer et al. (2021), Apel et al. (2022), BM (2022), Roggenkamp and Hergert (2022).

Hence, I would like to ask the authors to clarify, from the beginning, how they combine the original analysis of data and methods with the results of other studies in the context of the paper. Furthermore, I suggest that the authors publish the data which they used for this study in a single dedicated and documented data set to accompany this manuscript as an asset, even if parts of this data are available elsewhere (like the precipitation data or the Sentinel data). That way, they would not only make the original contribution of this study more transparent, but it would also be a valuable service to the research community. I understand that this will require some coordination with the source institutions, e.g. for the river gauge observations or the railway disruption data. Still, I think it'll be worth the effort.

Text is reproducing figures and tables

Very often, the content of figures and tables is reproduced/reiterated in the main text. I think the authors should trust more in the information content of their figures, and if they don't, figures should be made more concise.

Introduction and conclusions sections do not sufficiently frame the study

The introduction reflects the general issue of this paper: instead of using the introduction to systematically develop and justify specific research questions, it appears to string together previous studies in a rather unrelated fashion.

Similarly, the section "discussion and conclusions" (four pages!) provides a long list of statements, and for many of them, it is not really clear how they are based on the results presented in this study and how they relate to a study objective.

Other comments

Fig. 1: Please make sure that the border of Luxembourg is visible behind the river Sauer; I am not sure why the catchments are hatched - it does not improve the readability of the map. Furthermore, please show the catchment boundaries for all river catchments discussed in this paper (see section 3.2).

L. 90: While we see an area in the map, it remains unclear what signifies the "study area". Which parts of the area are actually *studied*? Is it the area in which specific precipitation totals were exceeded? Is it a combination of catchments, and if yes, which?

Ll. 93-100: I am not sure how helpful this paragraph is for the audience. If specific districts or municipalities are important in a spatial context, they should be included in a map. If the main map is too crowded or its scale to small, you can provide another inset or sub-plot which e.g. focuses on the administrative structure within e.g. the Ahr and the Erft catchments.

Ll. 125-130: Why is a weather forecast model used to analyse geopotential patterns and precipitable water? Why not use an analysis? The section header says "Weather forecast and analysis data", but the section does not describe any analysis data.

Ll. 146: Again, you mention the study area with regard to the selection of river gauges. Please explicitly specify, e.g. in Fig. 1, what you consider as the study area. You also specify

a lot of selection criteria; but why did you not just use, in the context of your study, all river gauge data that were available to you?

Ll. 153 ff.: I think the reconstruction of water level or discharge is an important methodological feature and should hence be described or adequately referenced.

Ll. 180 ff.: You are using ERA5, but you did not specify ERA5 as data used in this study in section 2.1.

L. 195: you say *"for example, [...] daily precipitation totals"*. Which variables, apart from that, were subject to extreme value statistics?

L. 239: *"very high values of total precipitable water of more than 40 kg m⁻² were reached, which occur only very rarely."* - how rarely?

Ll. 256 ff.: The authors end the section with the sentence *"Heavy precipitation associated with quasi-stationary low pressure systems, their fronts, or convective systems located on the western flank of persistent blocking systems is common in Europe during summertime, so the large-scale situation is not unusual"*. So what do we actually learn from the section *"Synoptic overview and atmospheric characteristics"*? How does it help us to understand how the event unfolded?

Ll. 262: How do you know the presence of embedded convection?

L. 269: please quantify what you mean by *"major part"*.

Fig. 3: Why not zoom all maps into the study area? I do not find it helpful to show all of Germany. For the API, it would also help to show return periods in order to appreciate whether/how the API was anywhere near unusual for this event.

Fig. 4: for the sake of comparability, please use the same x-axis for subplots a and b; I assume the black dashed line (observed reference from RADOLAN) is the spatial average - but over which region? Please specify in the figure caption.

Ll. 319 ff.: At some point, you need to provide a bit of context on the interpretation of the EFI - also quantitatively: How frequent is an EFI exceedance > 0.8 in the region, hence how well does it signify the potential for such a singular event?

Ll. 343: *"twice the climatological mean"* is not very helpful, in my opinion, as the reader is not informed about the statistical distribution of API values in the area. It would be more informative to provide a (extreme value) statistical assessment in terms of frequency.

Ll. 345 ff.: *"In the southern parts of the Eifel, the Ardennes in the north-west, and in the north-east of the study area in the Wupper region, generally less than 10 mm of soil water storage were still available for infiltration. In the remaining regions, free soil water storage was larger, but still below average, ranging mainly between 10 and 30, sometimes 75 mm."* On which basis are these statements made? The API does not allow for such an assessment (by the way, since the API is only computed from precipitation, it is basically a meteorological quantity, not a hydrological one). More importantly, in the context of your study and its stated objective, it would be crucial to assess the extent to which drier soils

(e.g. average soil moisture in July) could have been able to retain substantial amounts of water and hence significantly reduce the hydrological response.

Ll. 368 ff.: Apart from the fact that peak flow was exceptionally high with a steep rise, I am not really sure what to learn from the entire section about "Ahr, Kyll, and Prüm river basins". I think it would be a perfect opportunity to analyse why quite a similar rainfall in the headwaters lead to different event amplitudes in the different basins. These lines (ll. 368 ff.) aim at that, but I have the impression that the statements are rather based on speculation.

Fig. 5 and ll. 373 ff.: How were the time series reconstructed (dashed lines)? I understand that peak water level and hence peak flow can be reconstructed from debris lines etc., but how about the recession? How can we explain the high water level at gauge Schönau after the peak, while discharge had already recessed to normal levels?

Ll. 456 ff. / section 3.3: many statements in section 3.3 are not identifiably based on scientific data - either observation or models -, but seem to be speculative, hear-say, or partly reproducing textbook knowledge. The actual scientific contribution of this section remains unclear to me, especially in comparison to the paper of Dietze et al. (2022) which is widely cited by the present study.

Fig. 9: I think that the scale of this figure is not really helpful for appreciating the inundated area.

Ll. 589 / section 4.1: I think the issue of rapid inundation mapping is an important research topic, but I do not see how it is relevant in the post-event analysis almost one year after the fact.

Ll. 630 ff. / section 4.2: The same applies to section 4.2: Why is the rapid damage estimation process important? In the context of this study, shouldn't the best possible estimates of damages and damage processes be used? Furthermore, section 3.3 emphasised very much the role of debris/sediment/hydro-morphodynamics for damage processes, but I do not find this issue in section 4.2. Isn't this damage process chain one of the key properties of the July 2021 event?

Ll. 700 ff. / section 4.3: This section is mainly based on the number of traffic reports and the number of affected railroad sections, irrespective of the severity of disruption. Doesn't that limit the meaningfulness and the comparability to the 2013 event, given that it should matter whether a road/bridge is just disrupted or whether it actually disappeared? Maybe the persistence of the disruption for specific sections/lines could be interpreted as a proxy for severity, but there should be a reference from other events to compare to.

Fig. 11b: In total, I find subplot b not very informative, particularly since most of the specific road and railway numbers are not referenced in the main text. Wouldn't a simple plot of affected sections/lines over time be more informative and concise?

Ll. 744 ff.: *"The July 2021 flood in western Germany and neighbouring regions was one of the five most severe and expensive natural catastrophes in Europe in the last half century"* - what is the source of that statement? How important is the consideration of EFI in the context of your study, given the apparently high skill of the ICON models?

L. 763: *"However, because of the high political relevance, the failure of the warning chain was not discussed here."* I understand, but why bring it up then? Could in this case the topic of predictability also be dropped from the manuscript?

L. 809: *"Remote sensing of rapidly available imagery from social media, television, and news media [...]"* - I don't understand what that means.

Ll. 817 ff.: *"[...] our still insufficient knowledge must be further improved through more dedicated research"* - is it really necessary to spell out the obvious?.

Ll. 823 ff.: Is this item a new conclusion drawn from this study? Wasn't this already pointed out by Roggenkamp and Herget (2022)?

Ll. 849 ff.: Sure, but how??

Ll. 853 ff: How is this a conclusion from the present study?

Ll. 876 ff.: Please see my above comment: I think it would be really helpful to publish the data along with this manuscript.

Technical comments

L. 160: Please replace "flood plain" by "inundated areas"

L. 189: were used

L. 191: 95th percentile

L. 210: suitability instead usability

L. 224: second "decade"?

Ll. 272 ff.: *"The long-term average for the month of July at this station is 69 mm (1981-2010); thus, in just a few hours, the rainfall added up to more than twice the usual monthly precipitation."* - This kind of statement is unnecessary.

L. 307: as early as

L. 355: replace same by similar

L. 580: replace "the hardest affected location" by "the location affected most severely"

Fig. 10: It impossible to distinguish EMSR517_v1 from EMSR517_v3

Fig. 11: the resolution is very low.

L. 762: delete "the occurrence"

L. 831: an extraordinary

L. 837: the term "re-forecast" sounds weird - isn't "hindcast" the correct term?