Authors' response to Reviewer 1

[hess-2021-628-RC1]

We thank the reviewer for taking the time for the evaluation of our manuscript (hess-2021-628). Below we address the reviewer's comments (full text) indented by arrows and coloured in blue.

Their dataset and methods adopted in analysis are seriously flawed. The three hypotheses that they raised in the manuscript cannot be validated based on the existing analytical framework.

→ *From the comments made by the referees we understand that there is a need for clarification on both the definition of flash floods and the way how the related database has been built. Here below, we will provide additional elements that shall eventually also be included in the revised version of the manuscript.*

Aside from the technical issues, a key problem is that throughout the manuscript the authors do not specially define what is exactly a "flash flood" (in their perspective). We all know flash floods can be different from other types of riverine floods in various ways. However, it is never proper to simply classify floods during the summer months as flash floods (as distinguished from the winter floods). Without clarification of the basic concept, some of the sentences seem logistically biased. For instance, "The development of flash floods relies on long-lasting, extreme precipitation" (Line 108). This is not true, since extreme rainfall does not have to be "long-lasting" to generate a flash flood, although it is true for a subset of flash floods (not vice versa).

→ *We agree that multiple flash flood definitions have been proposed in literature. A non-exhaustive list of references defines flash floods as:*
- *A subset of pluvial floods (Owen et al., 2018),*
  - *exhibiting different characters than river floods (WMO, 2017)*
  - *runoff rates often exceeding by far those of other flood types, due to the rapid response of a catchment to intense rainfall (Borga et al., 2010)*
  - *often occurring on steep slopes (Van Campenhout et al., 2015)*
  - *with high flow velocities (Van Campenhout et al., 2015)*
  - *composed of less than 30% of solid material (Kron, 2011)*
- *Being caused by intense rainfall (Owen et al., 2018),*
  - *with event durations < 34 h (Marchi et al., 2010)*
  - *short-lived storms with high intensities < 24h (Gaume et al., 2009)*
  - *high intensity rainfall, mainly of convective origin and affecting small areas (Borga et al., 2010)*
  - *rainfall totals with return periods exceeding 50 yrs (Marchi et al., 2010)*
  - *rainfall totals exceeding 100 mm over a few hours (Gaume et al., 2009)*
  - *mainly of convective origin, spatially confined and often orographically enhanced (Gaume et al., 2009)*
- *Being characterised by short response times (Marchi et al., 2010)*
  - *generally, less than 6h (Marchi et al., 2010)*
  - *with rapidly rising and falling limbs of the hydrograph (Owen et al., 2018)*
  - *occurring within minutes to several hours, depending on the region (WMO, 2017)*
- *Remaining very difficult to forecast (WMO, 2017; Owen et al., 2018)*

        - *Flash floods developing at spatial and temporal scales that conventional hydro-meteorological observation systems are not able to monitor (HYDRATE, 2008)*
    - *Occurring in rather small catchments (WMO, 2017; Owen et al., 2018)*
        - *< 1000 km² (Marchi et al., 2010)*
    - *< 500 km² - varying from tenths to a few hundreds of km² (Gaume et al., 2009)*

→ *The flash flood definition criteria used in the references listed above consist mainly of metrics referring to the intensity of events observed via hydro-meteorological monitoring networks. When occurring in areas not covered by a monitoring instrument, extreme rainfall-runoff events may eventually remain undetected with this approach. Therefore, we propose here an alternative method - tapping into multiple sources of extreme hydro-meteorological events, combining scientific papers, agency reports, insurance inventories, personal communications, and newspapers. Moreover, we connected the flash flood database to the RADOLAN data matching it with the onset of a precipitation exceeding 40 mm/h (as defined by the DWD, leading to extreme weather warnings) within the grid cell of the flash flood. We will further improve this approach by considering a wider area around the occurrence location for identifying the corresponding precipitation event. The precipitation database should account for unnoticed flash floods.*

→ *We eventually relied on this multi-source approach for: (i) identifying extreme convective precipitation events in summer, triggering floods with considerable stream power, erosive force and impact potential to (inundated) infrastructures (line 172f); (ii) accounting for the high spatial heterogeneity that characterizes extreme hydro-meteorological events during the summer season - as opposed to inundations occurring in large river floodplains, caused by advective precipitation events (albeit not exclusively) - mainly during winter months.*

→ *Note that the vast majority of events identified through our query remained bound to catchments smaller than 120 km^2. Large summer floods, triggered by prolonged rainfall over extended areas and that may have occurred on larger rivers (e.g., Moselle, Rhine), were disregarded in our study.*

→ *The table here below relates to (i) the information sources used in our study, (ii) the criteria retained for identifying flash flood events, and (iii) the number of events that had eventually been identified. Note that the time period covered by our multi-source query spans from 1981 to 2020.*

	Criteria	*No. from this source*
Scientific papers	*keywords: "flash flood"*	*4*
Reports (LFU, LUA, Ministère de l'Ecologie, du Développement durable et de l'Energie)	*keywords: "Sturzflut", "Hochwasser", "crue éclair", "crue subite", "inondation", during summer months, excluding big rivers (Moselle, Rhine)*	*Göppert: 21* *Johst: 7* *Pfister: 4* *Ministère de l'Ecologie, du Développement durable et de l'Energie: 1*
Insurance reports	*keywords: "inondations" in combination with thunderstorms, heavy/extreme precipitation, summer.* *All floods listed have cause mentionable damage*	*1 (insurances only report major events, i.e. Braunsbach)*
Personal communication	*local hydrological events caused by heavy convective rainfall that have led to damage and made it to the awareness of the collector*	*23*
Newspaper	*keywords: "Sturzfluten", "Inondations", in summer, after thunderstorm/ intense rainfall*	*France 3: 16* *Trier: 1*

DWD Radolan data	Precipitation intensity per grid cell exceeding 40mm/hour → We will have a closer look in the revised version, to be able to better identify P for FF, that should have been covered by RADOLAN.	In 45 (of 83) cases a P event could be related to a FF event. 5 were before the start of the RADOLAN, and < 9 outside the area covered by RADOLAN.

→ *As shown by the manifold definitions provided in literature, a precise and clear definition of flash floods remains challenging. For our study, we considered extreme pluvial floods, as reported in scientific papers, agency & insurance reports, personal communications, and newspapers. Note that based on the available RADOLAN weather radar dataset we could eventually attempt to quantify the precipitation amount, intensity, and duration for some of the reported events in a backward approach. However, this would not improve the quality of the database required for our study – targeting in the first place the building of a comprehensive set of extreme (summer) rainfall-runoff events. The response time between the precipitation peak and the runoff peak occurred within only a few hours for each event. Since both precipitation and discharge data were only available for a subset of events, we opted for building a database reporting on the sole occurrence of extreme summer rainfall-runoff events (as per the criteria listed above). Note that the catchments in which the retained events occurred were all "small" - spanning from individual slopes (where major surface runoff had been reported) to catchments up to the size of a bit more than 120 km² (e.g., Ernz Blanche river, Starzel river). For all reported events, stream power and inundated water levels were strong enough to create substantial impact and damage (e.g., the displacement of large objects, such as cars).*

→ *We will revise the introduction of our manuscript paragraphs to better reflect a) on the challenges inherent to the definition of flash floods, as well as on b) the criteria that we have eventually retained for selecting events for our summer extreme rainfall-runoff event database. Basically, we are looking at extreme, pluvial, small-scale floods with a high impact that we call flash floods.*

Another concern of mine is that the flash flood database is not consistent in space and time. Any trend analysis based on the dataset would not be able to generate true insights into the real world. The authors also admit that the database is non-exhaustive. I would suggest the authors to demonstrate their efforts in making the database at least consistent in time. Otherwise, people would argue whether the significant trend is due to sampling biases or not. This corresponds to their first hypothesis (Line 404-405).

→ *We admit that we cannot guarantee an equally consistent database for the entire time period. While CCR has inundation data available online since 1989, Wald + Corbe collected data more systematically since 2006 and France 3 only since 2012. Moreover, we have checked the database by the European Severe Storm Lab for heavy precipitation, that also shows biases in space and time. As long as holding on to the database creation based on reports, this bias will inherently remain. The advantage of this approach is, however, to obtain a sample of extreme pluvial floods much larger, as it can be inferred from gauge catchments alone.*

→ *The cleanest option we can suggest is to take back this first part of the hypothesis and focus on the identification of their atmospheric parameters and calculate the trend based on them as a proxy.*

In addition, the authors use cumulative statistics to quantify the occurrences of flash floods for each year. Since floods cluster in space and time, the authors need to be aware of the issue of repeated counting. This is relevant to their second hypothesis where they evaluate trend in the occurrences of extreme rainfall. It would be biased to count the number of grids with rain rate exceeding certain thresholds. The statistics thus reflect the combined effect of intensity and spatial coverage of rainfall, not changes in the frequency.

- → *We agree that clustering can be an issue, due to the dependency of floods on event and pre-event conditions. Extreme events that are highly variable in time and space, such as flash floods and hailstorms, are telling examples in this respect - as shown for example by Changnon (1984) for hailstorms in the American Midwest. The flash flood clustering effect is also shown in Figure 2b of our study - showing the temporal occurrence of floods. We counted every single event, also when they occurred in the vicinity of another flood, or within a few consecutive days. The connecting element of flash floods are the related meso-scale atmospheric systems and/or similar pre-event conditions. In 2016 and 2018, for example, most flash flood events happened within 2 weeks across our study area. Multiple events may be linked (e.g. Piper et al, 2016), but they eventually also cause more damage and fatalities than a single flash flood event that may have occurred in only one isolated catchment. Therefore, we argue that counting all flash floods is important information, as multiple events should be weighted stronger than isolated events.*
- → *What was pointed out by the reviewer, is that "The statistics thus reflect the combined effect of intensity and spatial coverage of rainfall, not changes in the frequency." This is an important part of what we try to show. We had already aggregated the small radar grid cells to the resolution of ERA5 grid cells. This was to extract the atmospheric parameters in the next step. In the revised version we offer to cluster the radar grid cells by events for the second sub-hypothesis and count the number of grid cells contributing to an event, as we do not want to lose the spatial information and extent of the data. If high intensities of precipitation are present over a larger spatial area, the chance of a flash flood to occur should increase. We moreover suggest adjusting the wording and move from "more frequent precipitation events" to a "more frequent occurrence of precipitation intensities potentially generating flash floods".*

Changnon, S., 1984: Temporal and spatial variations in hail in the upper Great Plains and Midwest. J. Climate Appl. Meteor., 23, 1531-1541.

Lastly, I did not see significant increases in the proxy parameters for flash flood potential. This is mainly a concern with Fig. 5. Increases in moisture content are kind of expected according to the Clausius-Clapeyron relationships, but other than that, the other two proxy parameters show negligible significance (especially for DLS). In addition, flash floods are tied to comprehensive combinations of atmospheric conditions. By examining the trend in individual component of the comprehensive conditions as the authors did here offer limited insights into the real changes in flood potential. The threshold values are also chosen in a subjective way that needs further justification.

- → *It is true that many factors are involved in the development of a flash flood, which also need to interact in the right proportions. If individual parameters (or indices) would adequately describe their formation, then flash flood forecasts would most certainly already be routinely done. However, it is not our aim to comprehensively describe and mechanistically relate the occurrence of flash floods to individual parameters and/or indices of atmospheric conditions. We will clarify this aspect in our manuscript.*
- → *We confirm that the trends identified for DLS and CAPE are weak and largely insignificant. This was reported in the manuscript (for DLS, line 297-300 & line 413-415 / for CAPE, line 289-292). We will further develop on this finding in the revised version of our manuscript. The expectation of the Clausius-Clapeyron relationship is a 7%/K scaling in water vapour pressure if relative humidity remains invariant under climate change. While it is a frequent assumption in most climate change scenarios, it is after all an assumption.*
- → *In order to account for the combined occurrence of flash flood relevant atmospheric parameters, we added the subchapter 3.5, Fig. 6. This simple approach using low thresholds already excludes some of the occurred events. Any effort to specify this would not do justice to the variety of extreme rainfall events. Moreover, we considered using a GLM, but rejected the idea, as the parameters are not independent from one another.*
- → *The thresholds were originally chosen in a range relevant in literature. These values seem however often too high to include most detected events. Therefore, we adjusted the approach, while lowering the thresholds to include 75% of all our identified precipitation events. These choices are described*

in line 221-224 and 281-284, and a discussion of it is added in line 369-376. We will develop this aspect in more detail in the revised version of the manuscript.

I would not go into any further details about the presentation of the manuscript. Some of the sections (like Introduction, Discussion) need to be shortened and merged. These issues are relatively less important compared to the aforementioned concerns of mine.

→ *We had a close look at our manuscript again and also see the need for shortening some sections of the manuscript. We will take this into account in the revised version, whereby we - as previously stated - will also develop some aspects more clearly and reduce the ambiguities that may have prevailed in some statements of the initial version of our manuscript. However, we do consider a certain level of detail necessary, since this manuscript covers an interdisciplinary range of topics (flash floods, extreme precipitation, and atmospheric conditions) and that it might be read by specialists with one or the other background. The long introduction was initially meant to give valuable context about both the meteorological and hydrological aspects dealt with in the manuscript*.