Review of: **More frequent flash flood events and extreme precipitation favouring atmospheric conditions in temperate regions of Europe** by Judith Meyer et al.

Ruben Imhoff

Ruben.Imhoff@deltares.nl

January 10, 2022

Summary

The authors present an approach to link an increasing flash flood occurrence in central western Europe to a hypothesized increase in extreme precipitation events and favoring atmospheric conditions for these extreme rainfall events using 20 years of DWD RADOLAN data and 40 years of ERA5 reanalysis data. In a three-step approach, the authors show and claim that (based on the presented data) the number of flash floods has increased in the period 1981 – 2020. The same does not hold for the number of extreme rainfall events, but most of the selected favoring atmospheric convective conditions for both event groups show a trend that supports the set hypotheses.

I would like to thank the authors for a well-written and well-structured manuscript that was a pleasure to read. The presented three-step approach to test the hypotheses is easy to follow and is worth keeping. Nevertheless, I still have quite some questions regarding the experimental setup, event selection, underlying data and potential reasons why hypothesis 2 (the increase of extreme rainfall events) is not supported by the data. This would also make it easier to validate whether the methods are sufficient and appropriate to support the conclusions. Although the authors shortly touch upon this in their discussion section, it would put the conclusions in perspective if we have a better idea of the uncertainty of the found trends, based on the available data and choices made in the methods. In the following sections, I will go more into detail on these topics.

General comments

Dataset length

The authors used RADOLAN data from 2001 – 2020 and ERA5 reanalysis data from 1981 – 2020. I am concerned whether that is a long enough record to make climate-related conclusions? Especially the radar dataset, which only covers 20 years, seems too short to make climate trend-related conclusions. I do see the advantage of the high space-time resolution of radar for such an analysis, and it makes me happy to see it used, but the database length seems not sufficient yet. Although I find it hard to say what the minimum number of years should be in the dataset, I think the work needs at least a more extensive written support for the use of the dataset and the uncertainty that gives in the results.

Regarding the trend found in the data, especially based on the rainfall analysis for the 20 years of RADOLAN data: what does the trend look like if you take out the extreme years 2016 and 2018? I.e., are the trends we see a result of recent extremes?

I wonder if it would make more sense to look back from observed flash floods and extract the ERA5 reanalysis data for these times and locations, instead of partially picking events based on the shorter RADOLAN dataset?

Flood database

The first thing I was wondering is how certain the authors are about the increase in the number of reports from 1981 until 2020. Lines 241 – 242 “While barely any events were reported before 2006, two remarkable years are 2016 and 2018, when flash floods occurred particularly often in the study area (23 and 20 occurrences respectively).” Is there a chance that the number of reports also significantly increased over that period? Although I do believe that there is an increase, it may be good to support it by actual discharge time series of the catchments in the study area.

This also directly lead to how the authors have defined a flash flood. This was not directly to clear to me when reading the manuscript. In addition, is a flash flood that occurred on a certain day counted double if it occurred in a different location on the same day? It would be biased to base the frequency of occurrence on such a double counting, while it actually says something about the intensity and spatial extent of the flash flood (and rainfall events). This is also highly relevant, but not the objective of this study.

Concluding, would it be good to take a step back and (1) define what a flash flood is in this study, and (2) search for the events in time that correspond to this definition backed-up by both the literature study and discharge time series? I am aware of the amount of extra work this asks for, but it would make the conclusions stronger.

Specific comments

Lines 32 – 33: I would make this sentence a bit longer (to increase readability): E.g. “Flash floods, generally originating from severe convective storm fed by deep moist convection, rank among the most destructive hazards and result in economic losses, damage to infrastructure and high mortality rates (refs).” Or something similar, of course.

Lines 84 – 86 “This generally occurs in case of very weak pressure/geopotential gradients when the mean wind speed and the bulk shear between the surface and the lower to mid troposphere are weak.”: True, but what about orography causing or enhancing this?

Lines 88 – 109: I think this paragraph can be shortened. The authors give an extensive overview of proxy parameters used in literature. This is appreciated, but it is, in my opinion, a bit too long and distracting from the main message in the introduction. Perhaps give a couple of examples and then come to the main point of the paragraph.

Lines 116 – 118 “In addition, relative humidity levels decrease at low levels of the atmosphere, connected to rising temperatures, which also reduces the number of thunderstorms (Taszarek et al., 2021a).”: Although I am not an expert on this topic, I can image that with higher temperatures evapotranspiration also increases, which leads to higher moisture contents again (besides the fact that the air can contain more moisture at higher temperatures). As said, I am not an expert on this, but I think the statement at least calls for more references.

Line 142 “May to August”: Doesn't that leave out some potential late-summer storms in September?

Lines 157 – 159: Can you add some more information about the RADOLAN product? E.g., what kind of radars used, adjusted with rain gauges and how? Hence, how 'good' or reliable is this dataset? Were there any changes in the radar product during the 20 years that also results in different estimations over the years?

Lines 160 – 161 “an extended rain gauge adjustment with supplementary local rain gauges”: How many rain gauges were used, what time step was used and what kind of adjustment have the authors applied?

Lines 161 – 162 “We extracted the events for the database from the radar database by identifying 1x1 km grid cells with precipitation amounts ≥ 40 mm h^{-1}”: But you do not have RADOLAN coverage in the full study area? Or is the study area constrained to the area covered by the RADOLAN observations?

Lines 172 – 174: Is this database giving all the floods for the study domain and which catchments does it contain?

Lines 177 – 178 “The maximum hourly precipitation value was considered the trigger for the flash flood and atmospheric parameters were extracted from the identified grid cell and time.”: What about the cells around this grid cell, as their parameters may also have influenced the rainfall that fell there?

Lines 178 – 180: How did you find the flash floods here and the rainfall intensities, as this is outside the RADOLAN data coverage? In addition, do you have time series of the catchments, which could already indicate the presence and timing of a flash flood?

Line 203 “extremely rare in Central Europe”: just out of interest (and perhaps worth mentioning), how rare is it (quantified)?

Line 206 “700 hPa”: Why have the authors chosen to pick the 700 hPa level?

Line 215 “soil moisture (Swvl) [m3 m-3] at depths of 0-7 cm, 7-28 cm, and 28-100 cm from ERA5”: Why have the authors chosen for these three depths and would it make sense to average them in some way, as they will be (cor)related to each other?

Lines 222 – 223 “Therefore, we chose upper or lower boundaries including 75 % of extreme events.”: Do the authors mean the events IQR of the extreme events or did I understand it incorrectly?

Line 248 “Between 2001 and 2020, we observed a slight increase in the number of events per year (Figure 3a).”: But not a significant one, right?

Lines 266 – 267 “Moisture conditions during extreme precipitation and flash flood events were found to be mostly within the upper percentiles of the overall simulated values.”: That is also what you expect seeing the Clausius-Clapeyron (CC) relation and in fact even the 2CC relation for extreme precipitation. It probably deserves mentioning that, including some references (e.g. Lenderink & Van Meijgaard, 2008; Mishra et al., 2012; Manola et al., 2018; Wasko et al., 2018; Dahm et al., 2019).

Lines 269 – 270 “All moisture parameters, and especially RH tend to be even higher during flash flood events compared to general extreme precipitation events (Figure 4d-f).”: As clearly not all heavy rainfall events lead to flash flood events, can you also give some event statistics (earlier in the manuscript) between the two groups? What were average rainfall intensities in both groups, does the duration differ, does the size of the rainfall storms differ, etc.? This will give an idea why we see differences between the two groups.

Lines 274 – 277: This also says a lot about the initial catchment wetness prior to a flash flood event. As stated earlier by the authors, the wetter, the quicker a flash flood can occur. Now, from these results, I do not directly see a significant difference between the three groups. Only the 'P' group has somewhat lower initial soil moisture values, which gives the impression that heavy, convective rainfall does more often occur during drier periods. Something which corresponds a bit to the summer weather patterns in Northwest Europe. It also suggests that initial soil moisture conditions were on average not different from other days in the studied periods, so the flash floods are mostly a result of the weather system and not initial conditions here.

In addition, perhaps it is interesting to show the soil moisture as a relative scale (so % of the capacity).

Line 291 "These findings were particularly significant for the northern part of the study area (Figure 5b).": Any idea why in the north?

Lines 332 – 335: How is the trend if you take out 2016 and 2018?

Lines 361 – 362 "Regarding low wind speeds and weak bulk shear, we found slightly increasing but barely significant trends.": But you did find a significant trend for LLS, right?

Lines 380 – 382: This might also be related to the finite gauge-adjusted radar dataset of 20 years.

Lines 407 – 409 "Future analyses could incorporate the intra-annual temporal distribution of extreme precipitation events. Perhaps, formerly evenly distributed rainfall events tend to occur more condensed within a few days.": This is something you could already focus on in this study, by also looking at longer event durations. So what if you don't only look at 1-h accumulations, but also 6-h, 24-h, etc?

Lines 420 – 422 "In addition to the hypothesis, we found mostly higher upper (0-7 cm) and lower (28-100 cm) layer soil moisture during flash flood events compared to general extreme precipitation events.": This did not seem that significant in the results.

Figure 1c: I suggest to put here an actual DEM with a higher resolution, which makes the mountain ranges and the differences between them clearer.

Figure 4: Would the differences (which are clearly present!) become clearer when you take the P and FF events out of the all group?

Figure 5:

1. An idea for the figure, make the color scale discrete instead of continuous, then it is easier to distinguish the actual values.
2. In addition, the slope is in [unit] per year. So, don't forget to give the unit.
3. To get an idea of the timeseries underneath, could the authors provide for one pixel the timeseries + trend?

Technical corrections

Line 25: We tend not to confirm a hypothesis, but it is supported by the results.

Line 112 "point out, that CAPE": the comma can be removed.

Lines 240 – 241 “0.382 events per year (p-value 0.039)”: These values are not exactly the same as in Fig. 2a.

Line 247 “precipitation sums >40 mm h-1”: you can remove “h-1”, as it states precipitation sums and not intensities.

Line 404 and 418 “confirmed”: supported by the found data.

Conclusions section: as many people only quickly read through the abstract and conclusions, make sure to write out abbreviations (once again) in the conclusions.

Figure 3c: “intensities” in y label should be “sums”.

Figure A1: K-Index increase is also clearly significant in the north. Perhaps worth mentioning in the results.

References

Dahm, R., Bhardwaj, A., Sperna Weiland, F., Corzo, G. & Bouwer, L.M. (2019). A Temperature-Scaling Approach for Projecting Changes in Short Duration Rainfall Extremes from GCM Data. Water, 11, 313. https://doi.org/10.3390/w11020313.

Lenderink, G. & Van Meijgaard, E (2008). Increase in hourly precipitation extremes beyond expectations from temperature changes. Nat. Geosci., 1, 511–514. https://doi.org/10.1038/ngeo262.

Mishra, V., Wallace, J.M. & Lettenmaier, D.P. (2012). Relationship between hourly extreme precipitation and local air temperature in the United States. Geophysical Research Letters, 39, L16403. https://doi.org/10.1029/2012GL052790.

Manola, I., van den Hurk, B., De Moel, H. & Aerts, J.C.J.H. (2018). Future extreme precipitation intensities based on a historic event, Hydrol. Earth Syst. Sci., 22, 3777–3788. https://doi.org/10.5194/hess-22-3777-2018.

Wasko, C., Lu, W.T. & Mehrotra, R (2018). Relationship of extreme precipitation, dry-bulb temperature, and dew point temperature across Australia. Environ. Res. Lett., 13, 074031. https://doi.org/10.1088/1748-9326/aad135.