nhess-2020-192: An analysis on temporal scaling behavior of extreme rainfall of Germany based on radar precipitation QPE data (Pöschmann et al.)

Reply to the comments from Referee #1:

We are thankful for referee #1's constructive and very detailed comments and suggestions that helped us to improve the original draft. We have provided responses to all comments (in blue) and updated the manuscript according to the suggestions. All line numbers given in our responses refer to the original version of the manuscript in order to avoid confusion.

Comments:

Introduction

Extreme rain events may cause severe damage and fatalities and they are of principle interest as they mark the physical limits. This manuscript describes the depth-duration relationship as observed in Germany by the DWD radars (RADOLAN product YW, adjusted by rain gauges) for the period from 2001 to 2016. Although neither the principle method nor the data are new, the application of this method to these data is.

The manuscript offers a new sight to Germanys operational radar data and thus has a chance to be published. Nevertheless, it suffers from several drawbacks and misteks making it partly difficult to read and it contains several minor errors that should be removed before publishing.

Major issues

1. Radar observations cover the complete area of Germany. This advantage is discussed in the paper. Rain gauges may miss the most intense events. Nevertheless, the data quality of radar measurements is spatially variable, depending on the orography and distance from the next radar. Furthermore, radar provides precipitation measurements at a different scale (here 5 min./1 kmˆ2) than rain gauges (commonly 1 min./200 cmˆ2 in Germany). The paper lacks a discussion on data quality. Especially the shorter extremes might be impacted by ground clutter (in case of 5 min. extremes even from wind turbines or airplanes). With increasing distance the area of each range bin increases, reducing the frequency of extreme values. For a self-contained publication the authors have to describe and discuss these effects. How do they impact the results? Is the spatial distribution of extreme rainfall caused by the precipitation process or by the method of observation? It should be noted that the radar measurement consist of only one sweep at the lowest undisturbed elevation angle. Scan pattern were variable during the years. (So called precipitation scan.) Additionally, the authors need to add a short description on the data processing from the measurement to offline quality control and the adjustment with RADOLAN.

Response: The data quality has indeed been a big concern during the data processing and evaluation of the results. Our biggest concern had been a potentially high number of outliers due to radar errors as have been observed with the RADOLAN products, which would result in too high values for rainfall maxima. In this study, the focus is on the post-processed data of RADKLIM. The RADKLIM data should be significantly and consistently improved compared to RADOLAN. Therefore, we included more details on the data quality of RADKLIM.

Lines 81 – 84 were edited as follows:

“Since the quality enhancement of RADOLAN is ongoing without post-correcting previous data, the so–called radar climatology project of the DWD, RADolanKLIMatologie (RADKLIM,Winterrath et al., 2017) has consistently reanalyzed the complete radar data archive set since 2001 to attain homogeneity of the data that were processed through different algorithms. Compared to RADOLAN, RADKLIM has implemented additional correction algorithms that lead to much more plausible spatial distribution of precipitation totals including fewer typical radar artefacts, improved representation of orography as well as efficient correction of range-dependent path-integrated attenuation at longer time scales (Kreklow et al., 2019). Whereas RADOLAN is not suited for climatological applications and aggregated precipitation statistics, RADKLIM is a promising data set for these kinds of applications. The RADKLIM data is available ..“

Lines 91 – 94 were edited as follows:

“The YW product covers the area composed of 1100 x 900 pixels with the spatial resolution of 1 km (improved compared to former version of RADOLAN). Remaining weaknesses of RADKLIM (as outlined in Kreklow et al. (2019)) are a higher number of missing values (compare below) than for RADOLAN as well as an overall negative bias causing a rather “underestimation” of high intensity rainfall due to spatial averaging and rainfall-induced attenuation of the radar beam.”

In addition, we combined sections 2.1.1 and 2.1.2 into 2.1. (by removing their subtitles). Furthermore, we added text in section 3.1., explaining the possible underestimation of sub-hourly values.

The following sentence was added at line 151 (center):

“As mentioned in the data quality description, it is possible that these sub-hourly values do not represent the “real” German-wide extremes for 2001-2016 since very short durations are specially effected by averaging effects of the radar processing.”

The uncertainty of radar-based products remains a challenge for the coming years and the product quality and characteristics obviously have an effect on the distribution and intensity of extreme values of rainfall. However, a detailed discussion of the validity of the data set is not the scope of this study, rather to discuss how this dataset reflects well-documented features of rainfall extremes from the “pre-radar world”. RADKLIM is currently the best consistent representation of rainfall for the whole of Germany (according to DWD). Thus, we purposely used the dataset despite unknown (as always) deficiencies. As the results appear to be plausible and consistent, the discussion deals with the RADKLIM extremes as observed extremes. Generally said, almost half a million pixels will provide a good representation of the “true” characteristics.

Following sentence was added at the end of the conclusion:

“Also, the known "underestimation" of rainfall extremes by RADKLIM-YW and the potential impact on the results needs further evaluation.”

2. Figure 3 shows the "total within-clusters sum of squares". This term is not defined in the paper. Please, describe the procedure to determine the shown curve so it can be comprehended and interpreted.

Response: We changed the passage in order to make it more comprehensible.

Edited text passage (from line 138):

"If the number of clusters is not predefined, it can be identified by drawing an elbow chart as seen in Fig. 3. For different number of clusters K the measure of the variability of the observations within each cluster (Total within-cluster sum of squares, y-axis) is calculated and the curve should bend like an elbow at the optimal value. Since the algorithm did not suggest a number of clusters, we chose six clusters for a sufficiently detailed analysis since it gave consistent results when repeating the automatic algorithm for several time (each time the algorithm clusters slightly different)."

3. The authors explain (line 151f) "Between 25 min and 16 h, maximum values are calculated for the southeastern edge of Hesse state in May 29th 2016." This is not reproducible. From Figure 4 we can see, the 25 min extreme is already significantly above 100 mm. In the area between 8.96◦ and 9.37◦ east and 50.15◦ and 50.32◦ north (this area should cover the location the maximum) rain amount is below 32 mm. Maximum precipitation on that day is 123.67 mm at 50.54◦ N, 12.61◦ E in Ore mountains, Saxony. This maximum is followed by 122.98 mm at 49.22◦ N, 9.83◦ E, close to Braunsbach, Baden-Württemberg. That rain event caused estimated 100 million Euro damage and three fatalities, as newspapers reported a few days later. - I did not control the further maxima. A table is missing, indicating duration, start time, rain amount, and location for each of the blue dots in Figure 4. Without these data no reproduction of the findings is possible.

Response: We want to thank the referee for carefully checking our results. At the same time, we apologize for our mistake. In Figure 4, locations and rain amounts are a correct representation of what we get from the data, but the identification of dates went very wrong. We changed that, updated the text and added a table indicating the suggested characteristics. From the table the findings are now reproducible. The Binary RADKLIM-YW files are available at the DWD online depository (https://opendata.dwd.de/climate_environment/CDC/grids_germany/5_minutes/radolan/reproc/2017_002/bin/) for the given dates and the locations are identifiable by their WGS84 coordinates.

The text is edited as follows (from line 151):

"Between 25 min and 16 h, maximum values are calculated for a location at the border of Hesse state and Bavaria in August 25th 2006, which has not been documented in public news. The extreme event around September 29th/30th 2003 around Berlin comprised the maximum depth-duration relationship at the duration between 18 h and 2 d."

The following Table 1 is inserted into the manuscript:

Table 1. Rainfall records for different duration from RADKLIM-YW for 2001 - 2016 with corresponding locations.

Duration	Start Date	Start Time (Time Zone: Berlin)	Precipitation Sum (mm)	Location (WG84)
5 min	2009-07-04	2:10 PM	40.94	48.50015° N, 9.35161° E
10 min	2006-07-07	9:30 AM	80.82	51.22436° N, 8.767699° E
15 min	2010-07-12	11:05 PM	105.61	52.79713° N, 12.39296° E
20 min	2002-07-30	5:15 PM	127.32	48.82225° N, 9.577044° E
25 min-16 h	2006-08-25	05:25 AM - 1:25 PM	141.13-230.67	50.21148° N, 9.201292° E
18 h-1 d	2003-09-29	09:05 AM-03:05 PM	258.91-327.45	52.52761° N, 13.5271° E
1.5-2 d	2003-09-28	02:20-9:35 PM	471.67-503.66	52.52761° N, 13.5271° E
3d	2001-04-08	06:50 AM	525.89	53.67822° N, 10.00056° E

Maxima of 25min - 16h as well as from 18h - 2d correspond to the same location and date and are thus summarized.

4. The value of Figure 6 and its interpretation is not clear for me. What elevates the 3921st greatest event (shown) above the 3920th (not shown)? What is meant by "The lower the quantile, the sparser the location of the quantile rainfall occurrence"? What is a sparse location? The location of the dots is totally random, as it is random if an event is the 3920th or 3921st. I do not get the message, the authors want to transport here.

Response: We are sorry that the interpretation of the Figure is unnecessarily difficult. The purpose of placing the Figure like this is related to Figure 5: We want to show that if not taking the maxima of maxima, but certain quantiles of maxima (we chose to take numbers that correspond to 99.999%, 99.99 %, 99.9% and 99%, but obviously other values could have been chosen):

1) No longer seven locations (as in Figure 4) hold all maxima, but that the number of locations increase and also spread over all of Germany.

2) Despite 1), the corresponding depth-duration relationships are straightening out (=getting smoother) and start to reflect rather natural conditions of rainfall (e.g. dominance of Alpine region in Figure 6d) instead of singular “extremes of extremes”.

We simplified the sentences around the irritating sentence (l. 164) that the referee pointed out in order to make it clearer.

Text edition (from l. 164):

“It shows that the number of locations increases the lower the quantile of maximum rainfall is. This suggests the reduction of the influence of one single rainfall event on the depth-duration relationship causing inflection in the curve. Additionally, from a certain degree of quantile (Fig. 6 d) the locations of maximum rainfall contributing to the development of the rainfall-duration relationship seem to happen mainly in the wider Alpine region in South Germany. This suggests that rather natural rainfall mechanisms are dominating the scaling relationship, such as regional characteristics and meteorological conditions (e.g. orographic lifting or leewards effects). Naturally, one would assume that this heterogeneity of the meteorological conditions and rainfall generating mechanisms will reflect rather regional characteristics and will exhibit some irregular scaling behavior. Contrary to this conjecture, the curves in Fig. 5 (99.9% and 99%) show a rather smooth scaling behavior” [end of subsection 3.2]

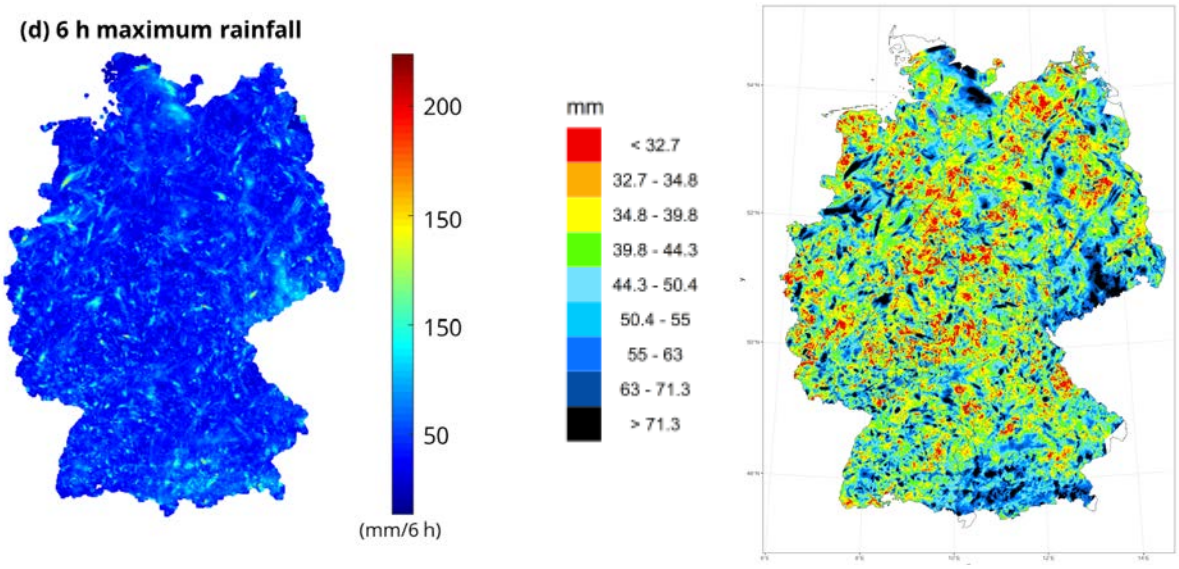
5. If the focus is on the spatial distribution of extreme events, then show it for some durations (the locations of the strongest 10 (red), 100 (yellow), and 1000 (green) events for a duration of x minutes). If the focus is on the low impact of an individual event, then show the frequency distribution of rain amounts for a certain duration, focusing on the most intense 10000 events or so.

Response: We are grateful for the suggestions by the referee, however chose not to change the Figure since it would change our message (spatial distribution of extremes is already shown in Figure 7).

6. Figure 7 is hard to see, especially in a printed version. Figure 7 is in contradiction to section 3.1. The absolute maximum for 6 hours (Fig. 7d) occurs in the area between Ilmenau and Erfurt, roughly. The 1 hour maximum (Fig. 7c) is in the area of southeast Hesse.

Response: We understand the concern of the respected referee. The purpose of Figure 7 is not to identify the location of maxima of maxima, since they are already provided in Figure 4. We

want to show that the longer the duration, the more "organized" the maxima get and big events with long and heavy rainfall are visible on the map as well as seasonal and terrain-related patterns would most-likely become visible. Figure 7 focuses on the overall distribution of maxima. Singular pixels of maxima would be too hard to trace on a small map without zooming in, as the referee pointed out. If the referee would find it more suitable we could change the color scheme as given in the following example (colors represent quantiles):



7. Figure 8: I do not see that Wiesbaden is less fitting to a power law than Stuttgart. The authors do not provide a quantitative or at least objective way to describe the deviation from the power law.

Response: We thank the referee for pointing out this important issue. We placed a reference line in all Figures in order to make the differences more visible for the readers. It is true, that the deviation is a subjective interpretation from the authors of the study. We changed the text from l. 185 a little bit after taking again a look at the Figure with reference line, since we agree that Wiesbaden is not different to the mentioned cities Hannover, Kiel, etc.

Text edition from line 184:

"Figure 8 shows the maximum rainfall–duration relationship of the radar pixels at the major cities of Germany with a blue line as reference to see the differences better. Except for Hamburg and Stuttgart, most cities exhibit slight (Hannover, Kiel, Magdeburg, Potsdam, Schwerin, Wiesbaden) to considerable ....."

New Figure 8:

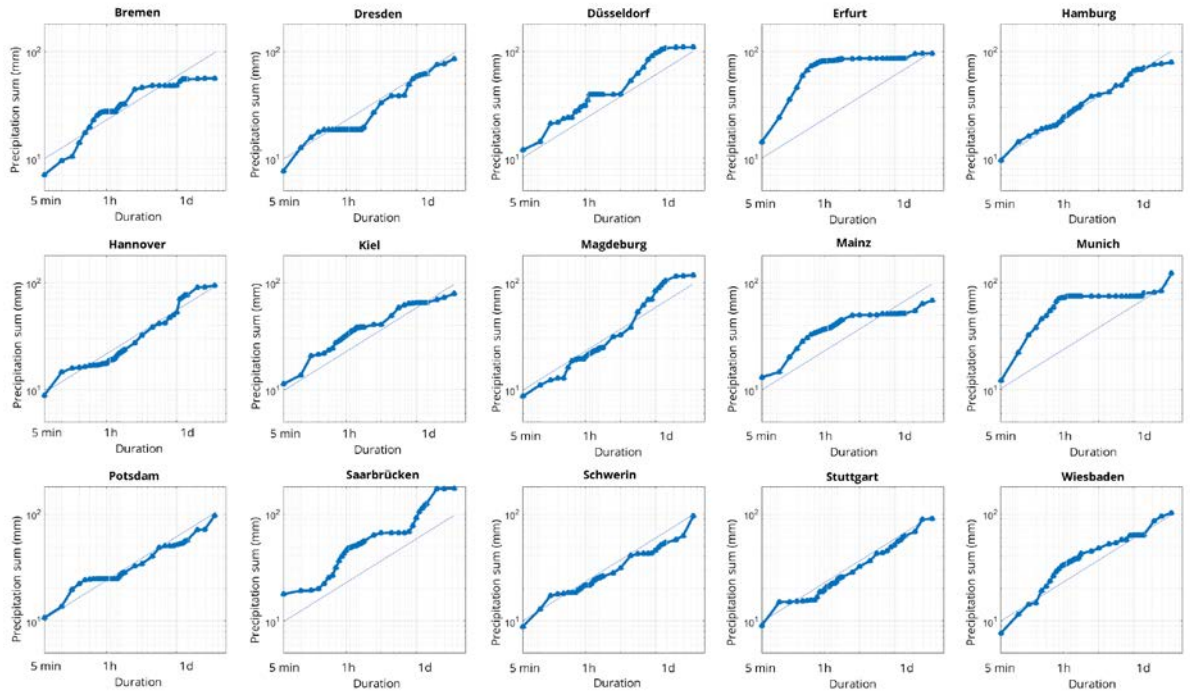


Figure 8. Maximum depth–duration relationships at rain gauge locations of major cities of Germany

8. I'm not convinced by the result of the clustering algorithm as shown in Figure 9. I cannot identify clearly distinct properties between the 6 categories. The authors have the same problem and propose to combine the categories into three new categories. The verbal description of the categories remains vague. Probably it was more helpful to state the properties and group the relations/pixels according to predefined criteria. (Cat 1: All relations with more then 40 mm @ 1 hour. Cat 2: less than 40 mm @ 1 hour but more than 100 mm @ 1 day …) Without a clear description of the categories, Figure 9 lacks a message.

Response: The purpose of the clustering was to automatically group pixels of similar depth-duration relationship characteristics. Thus, we cannot predefine properties before, but evaluate the clustered groups. As the referee points out, we propose to combine the 6 clusters into 3 categories, since we find it more meaningful.

The authors think that Figure 9 proves that the maximum depth-duration relationship for all pixels is mainly driven by singular events, whereas the lower the maxima we look at, the more smooth the curve will get. We put the Category 5 regression line into the sub-figures of all other Categories of Figure 9 in order to make the difference more visible.

New Figure 9:

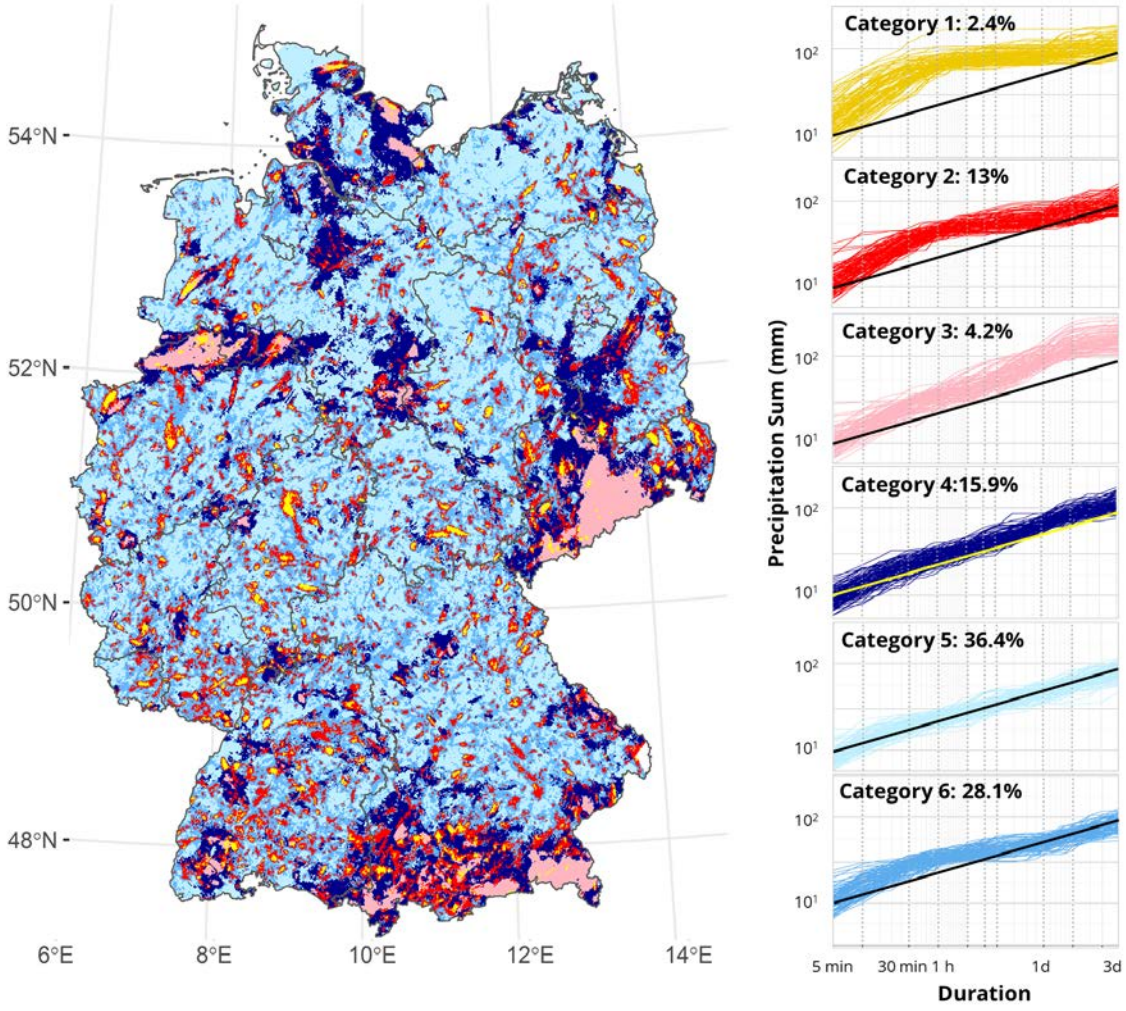


Figure 9. Display of the six categories of maximum rainfall depth–duration relationships a) spatially distributed over the whole of Germany, and b) their corresponding curve shapes, at 100 randomly selected radar pixels belonging to the corresponding of the six categories.

The following was edited (Line 195):

"Figure 9 shows a categorical map of Germany representing each category with a certain color. Additionally, depth-duration relationships at 100 randomly chosen grid elements from each category are shown with the regression line from Category 5 as reference."

Minor

9. Please be more precise in your wording. A "sample" can be a subset of radar pixels, i.e. only locally restricted. It might also be a temporal subset.

Response: The authors generally think that the meaning of "sample" as subset of radar pixels or temporal subset is clear in the corresponding context. However, we have revised all occurrence of sample and changed it at few places.

10. A "cell" and a "pixel" refer, as far as I got it, to the same thing: An area of 1 kmˆ2 for which the RADOLAN product provides one rain intensity every 5 minutes. If this is right,

please omit one of the two terms. Otherwise define a "cell". You are not talking of storm cells.

Response: We agree that we should use consistent wording and chose to use “pixel”. We changed it in the whole manuscript.

11. **There are several issues with the figures**: - The y-axis is never precipitation intensity but precipitation sum or precip depth and the unit is mm, not mm/Duration. -

Response: There exist representations of the depth-duration relationships where the precipitation sum is provided as intensity. However, we agree that providing the precipitation sum with unit mm is more meaningful and we changed it in Figures 1, 4, 5, 8, 10.

12. Figure 1 does not show the fit for the Spanish measurements and not the individual measurements for the Eastern German measurements. The caption denotes "regional extremes for Germany", the legend "Eastern Germany". Shall this be the same?

Response: We changed “Eastern Germany” to “Reg. Extremes Germany” to be more consistent. For consistency reasons we removed the individual values and only showed the fit of the measurements, since no individual values are available for the German regional curve.

13. Scaling of the x-axes is difficult. In Figure 1 the structure seems to be clear (minutes, hours, days...) but minor tags are missing (they occur as small gaps in the horizontal grid). In Figure 4 the minor tags are too bright to be seen on a printout. In Figure 5 some minor tags have vanished. Could this be unified in a clear visible way?

Response: We added minor tags to Figure 1 and added “minutes”, “hours” and “days” to Figures 4 and 5 and made the minor tags for Figures 4 and 5 more visible.

14. Line 60f: Lower values of maximum rainfall values on a coarse grid of 400 km^2 gird cells is no underestimation but a known impact of averaging, as Brena-naranjo et al. already mentioned. Whereas "underestimation" indicates a deficit of the measurement or procedure the reduction is physically reasonable.

Response: The authors think that the word underestimation can also be a logical consequence of averaging and the explanation is given in the following sentence in our opinion. However, we respect the referee’s concern and changed the two sentences (from line 60f):

“They showed that the maximum of the areal rainfall averaged over the ~20 km x ~20 km data grid has the scaling exponent of ~0.43 which is similar to that of Jennings (1950). However, the coarse spatial resolution of the satellite data easily misses the small scale rainfall variability that is closely associated with extreme values, thus the found extremes in the satellite data are lower than expected (Cristiano et al., 2017; Fabry, 1996; Gires et al.,2014; Kim et al., 2019; Peleg et al., 2013, 2018).”

15. Line 104: How do you calculate the "imputation bridge"? Radar data are missing so how do you get a rain intensity for these periods?

Response: As it is written in the text, the imputation bridge is the “maximum rainfall difference between right BEFORE and AFTER a data gap”. We thus do not need rain intensity for these periods, but calculate how much would need to be imputed for the highest differences in intensities. The argument would be, that when values before and after a data gap are quite “similar”, most likely no big changes happen in between and if there is a big difference (=high imputation bridge), it is higher likely that we do not know what happens between the two values.

However, as we have further pointed out, the imputation of NA values for rainfall events imposes too much uncertainties regarding the analysis of extremes.

16. The description of the methodology is inconsistent and unnecessarily hard to read. E.g. tau is a duration (see line 115), so it might be given in minutes. np is a number, counting the observations at each location. np thus is unitless. What is np-tau+1 (line 124)? What is min? What is h? (line 117). How did you get results for 3-day-extremes when your analyses is limited to 3 hours (line 116)? Line 130 and following do not indicate how you determine B and b. Equation 7 does not fit to equation 6 (somehow tau is lost). Equation 6 might be meant as log(M)=B+b*log(tau), as the figures show log-log axes. Equation 7 then is M=10ˆB*tauˆb in compliance with Eqn. 1. Besides the mathematical errors there should be more text. E.g.: Mˆ(tau)_max,cell (Eqn. 4) are the individual duration-depth relationships for each pixel. This needs to be mentioned.

Response: The authors agree that the methodology part might be unnecessarily detailed creating more confusion potentially. Thus, we shortened sections 2.2.1 and 2.2.2 into one paragraph 2.2. "Depth-Duration relationships" and removed most of the equations for a better reading.

The methodology section is now shortened as follows:

"2.2 Depth-Duration relationships

Maximum rainfall values for each duration τ between 2001-2016 were calculated with rolling sums applied over moving windows using the R package RcppRoll (Ushey, 2018). Durations of up to 3 d were chosen for the analysis, with multiple steps for minutes and hours out of our interest for sub–hourly and sub–daily pattern. The records may include non-rainfall data thus do not imply continuous precipitation for the period considered. Values were not aggregated spatially, since this usually reduces the maximum intensity values (Cristiano et al., 2018).
First, the extreme values for each pixel and duration $M_{max}^{\tau,pixel}$ are calculated. Afterwards, the overall maxima for whole Germany for each τ ($M_{max}^{(\tau)}$) is extracted from these calculated extreme values. Based on these results, the depth–duration relationships can be build for each pixel as well as for the whole of Germany."

Section 2.2.3 is thus changed to 2.3.

min and h (line 117) are the journal's abbreviations for minute and hour, but we have written it out as "minutes and hours" now.

"3 h" in line 116 was mistakenly written and was replaced with "3 d" (for days).

17. Figure 4: The publication of WMO, 1994 indicates the value for 3 days should be at 3130 mm (not roughly 4000 mm) and the points at 30 min./200 mm and 3 h/>700 mm are not given there (Table II.5.6). The study of DWD, 2002, cites "Ertel and Schmidt, 1999", as source for their records (without giving a findable reference). DWD 2016 is not accessible. The caption cites a Spanish study but the figure does not show the data.

Response: The authors added NWS 2016 (https://www.nws.noaa.gov/oh/hdsc/record_precip/record_precip_world.html) as a reference for the world records which provides updated values. The 3 days value of 3929 mm corresponds to an event in La Réunion in 2007. The values for 30 mins and 3 h (724 mm, USA 1942) are shown in the graph, which might be better visible after we have changed the x-axis. However, we

mistakenly twisted numbers for the 30 min value and show 208 mm instead of the correct value of 280mm. We updated Figure 4 (see attachments Fig. 3)

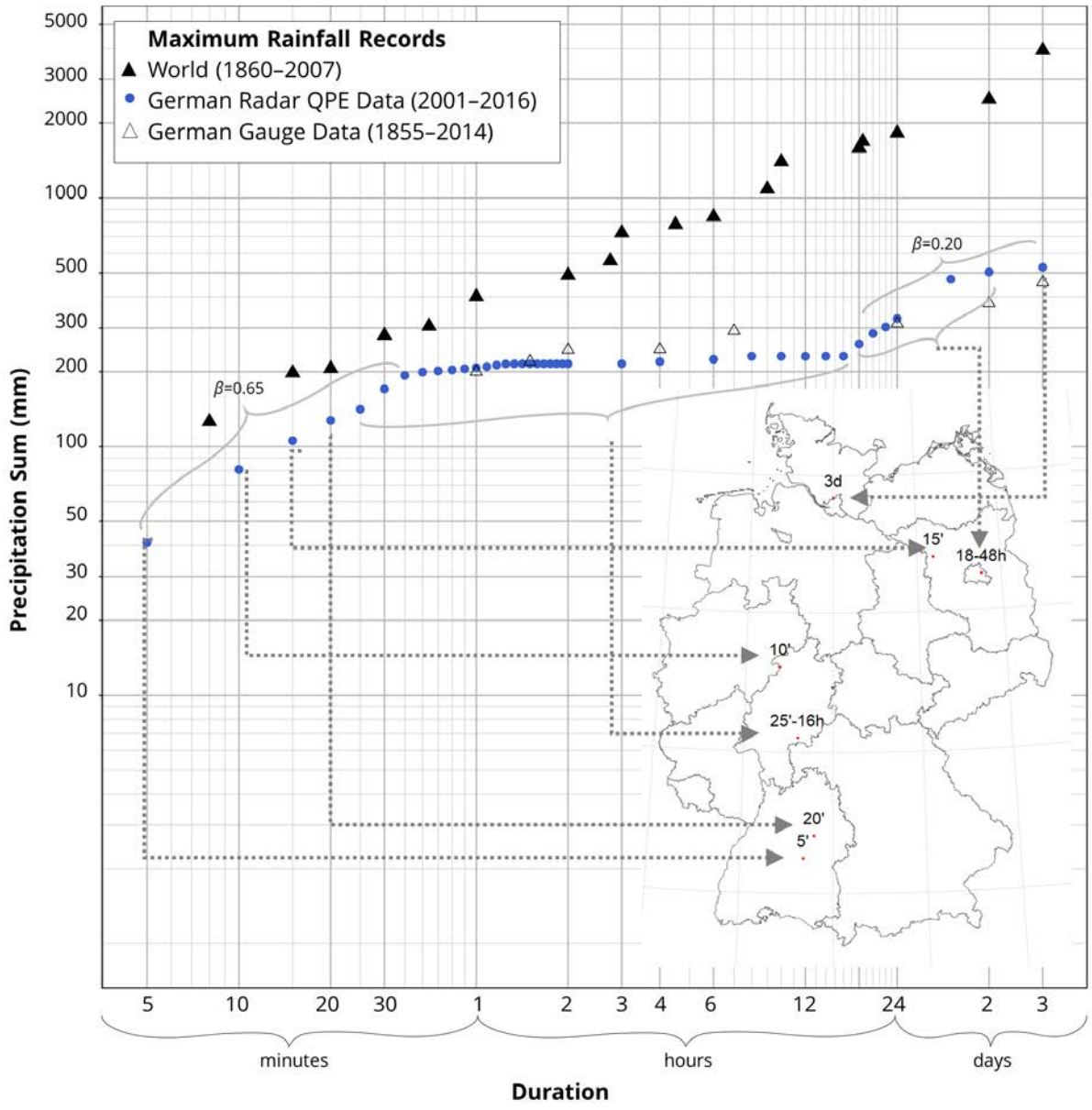


Figure 4. Maximum depth–duration relationships and locations of rainfall maxima. Dots: German radar derived data of this study. Non-filled triangles: German ground network (Rudolf and Rapp, 2003; DWA, 2015; DWD, 2020). Filled Triangles: World records (World Meteorological Organization, 1994; NWS, 2014). Map: Locations of rainfall maxima (based on the German radar data) for the considered duration.

We contacted the DWD about accessing the DWD 2016 source and they informed us, that the same Table from DWD 2016 is also available in "DWD (2020): Nationaler Klimareport. 4. korrigierte Auflage, Deutscher Wetterdienst, Potsdam, Deutschland, 54 Seiten. DWD. Stand Druckversion: 04/20 " (p. 37) accessible via https://www.dwd.de/DE/leistungen/nationalerklimareport/download_report_auflage-4.pdf?__blob=publicationFile&v=11. We changed the reference accordingly.

We agree to replace the study DWD 2002 with the report of B. Rudolf and J. Rapp (2003), "Das Jahrhunderthochwasser der Elbe: Synoptische Wetterentwicklung und klimatologische Aspekte" (values provided in Table 2), accessible via

https://pdfs.semanticscholar.org/fdfc/a0eb2c7ac37d2d80ddd2700b3f710a7fed79.pdf, with no cross-reference of another study, but the source is the DWD.

The Spanish study indeed is not included in the Figure, so we removed it in the caption.

18. Figure 7: There are inconsistent color scales, showing 150 mm twice in (c) and (d).

Response: When comparing to Figure 4, it is clear that the maximum values do not change significantly between 1 hour and 6 hours. Thus, the range of values remains similar for both Figure 7 (c) and (d). However, when looking at Figure 7 (d) we can observe that more pixels have higher values than for Figure 7 (c), which is what we would expect, since more rainfall is happening for longer time periods.

19. The unit should always be mm.

Response: Though we think that it would be clearer to use mm/duration of the Figure, we changed all units to "mm".

20. Figure 8, section 3.4: How did you choose the pixels for the cities? All of them are larger than 1 kmˆ2. Please consider to draw all 15 lines of Figure 8 into one pair of axes (different colors, different line style), making it possible to compare the curves.

Response: We added a reference line in Figure 8 in order to compare them with each other (compare updated Figure in comment 7.). Drawing all 15 lines into one plot will result into a rather chaotic display in our opinion.

Rain gauge data can be downloaded from a DWD repository including a list of all rain gauge stations with x-y coordinates. We took these coordinates and transformed them into the RADKLIM/RADOLAN projection. Since we are reprojecting point data (gauge) and match it with raster data (RADKLIM) there might be very slight geographic shifts between gauge locations and the corresponding 1 km² raster grid pixel. We included a half-sentence in Figure 8 (compare comment 7).

21. Line 193: Did you only cluster the 15 depth-duration relations from Figure 8 or from all pixels?

Response: We added a half-sentence that Figure 8 clustering is for all pixels as described in the corresponding section 2.2.3

Edition of first sentence in line 193:

"The maximum depth–duration relationships for all pixels within Germany were clustered since Fig. 8 indicated that they might show similar shapes."

Technical

22. Line 3 and 65: 1 kmˆ2 (it's an area, not a distance)

Response: We changed it.

23. Line 7: A smooth scaling behavior was (not were).

Response: We changed it.

24. Line 20/22: If probable maximum precipitation is abbreviated by PMP you should introduce this abbreviation: Probable maximum precipitation (PMP)...

Response: We changed it according to the suggestion.

25. Line 35ff: Bad deal with units! (1) is a numerical value equation, thus you need to indicate units of all variables.

Response: We changed it.

26. Line 37: In (1) duration is d not D.

Response: We changed it to D in all of the occurrences.

27. Line 37: Are the values 425 and 0.47 found by Jennings (1950)? If so how did that curve change during the last 70 years? Shouldn't it be "alpha and beta had the values", as climate change might have changed these values?

Response: The values found by Jennings remained surprisingly stable over the last 70 years, however, Gonzales et al. (2017) updated the envelope for the world records and we included the update in the paragraph.

Text edition for comments 25. – 27. (from line 33):

"Jennings discovered that this unique scaling behavior holds at the rainfall duration between 1 min through 24 months. Paulhus (1965) showed that the same power law relationship holds after addition of new world rainfall record observed at the island of La Réunion at the duration between 9 h and 8 d. The envelope for this extreme values can be expressed as:

$P = \alpha D^{\beta}$ (1)

where P is the maximum precipitation (in mm) occurring in duration D (in h), the coefficient α (425 in Paulhus (1965)) represents the value at one hour of the depth–duration relationship plotted on the log–log plane, and the exponent β (0.47 in Paulhus (1965)) is the parameter characterizing the scaling behavior of the depth–duration relationship. The Spanish study of Gonzales and Bech (2017) updated the world envelope's slope to 0.51, showing a remarkable stability. Multiple exponents describing the scaling property of ....."

28. Line 77: DWD is running (not has been running) a radar network. They still do.

Response: We changed it.

29. Line 85: There is no need to mention RW product. There are more RADOLAN products but RW and YW.

Response: Yes, there are several RADOLAN products available from the DWD. However, RADKLIM-RW and YW are products from a separate project from the DWD as mentioned in the section. We feel the need to mention the RADKLIM-RW product, since the product that we used (RADKLIM-YW) is more or less directly derived from this hourly product and not an independently calibrated/adjusted product.

30. Line 114: What is a radar cell? If a pixel is meant, call it a pixel. Are you aware of the difference of a pixel (beam volume element) within the basis data (sphere coordinates, growing size with distance) and a RADOLAN grid element?

Response: In this study we are talking about a RADKLIM grid element as defined by the referee before: "An area of 1 kmˆ2 for which the RADOLAN (RADKLIM) product provides one rain intensity every 5 minutes." We follow the referee's suggestion as written as reply to comment 10 and consistently used the word "pixel" to avoid further confusion.

31. Line 135: Please, add "Eqn. (4)" to the first sentence.

Response: We added it.

32. Line 143: There is no blue solid line in Figure 4. There is no red dotted line in Figure 4.

Response: We changed it to the characteristics shown in Figure 4: Triangles (filled/empty) and dots

33. Line 146: As well as for one (not some) sub-daily values. Only for 1 h radar has a higher value than rain gauges.

Response: We changed the wording according to the referee's suggestion.

34. Figure 5: Please consider to add the curves of this figure into Figure 4 and to remove the inlay with the map of Germany into a new figure.

Response: We are thankful for the suggestion, however, prefer to keep the Figures separate since Figure 4 focuses on the maxima of maxima, whereas Figure 5 is more related to Figure 6 and does transport a different message than Figure 4.

35. Line 169: Which curve? There is no reference.

Response: We added the reference. It is the two bottom lines of Figure 5. (compare response to comment 4, final sentence in the text edition")

36. Line 185: Stuttgart, not Stuttugart

Response: We changed it.

37. Line 193: depth-duration (not dept-duration)

Response: We changed it.