Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-308-RC1, 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Rapid attribution of the May/June 2016 flood-inducing precipitation in France and Germany to climate change" *by* Geert Jan van Oldenborgh et al.

Anonymous Referee #1

Received and published: 5 August 2016

Manuscript summary

This manuscript proposes a rapid attribution to anthropogenic climate change of a type of precipitation event – location: Seine and Loire catchments, southern Germany; season: April-June; intensity: 3-day maximum/1-day maximum – that actually occurred in May-June 2016. This actual precipitation event notably led (together with wet antecedent conditions) to floodings in France and Germany. The authors make use of a range of GCM/RCM runs as well as observational data to compute the ratio of the probability of occurrence of this type of event (1) between 1960 and 2016 in the actual world, and (2) in 2016 (or 2014/2015) in the actual world or in a counterfactual

world with preindustrial greenhouse gas concentrations and sea surface temperature with the effect of anthropogenic climate change removed. They found an increase in the probability of occurrence in both cases for the Seine and Loire catchments, but not significant in all experiments. Results for Germany appear less robust.

General comments

My first general comment is that the issue of flood event attribution to anthropogenic climate change is scientifically sound and within the scope of HESS. However, the quality of this "rapid attribution *study*" is questionable due to constraints imposed by the short time frame, and the manuscript proposes a "quick-and-dirty" attribution *report* of the study. The reasons for providing this rather strong statements are developed below:

1. First, let me ask the following question: what is the purpose of this "rapid attribution *study*"? (please note that I am not yet referring to the manuscript) The authors claim that it is motivated by demand on such information: "The extreme nature of this event left many asking whether..." (P1L7), "However demand on such information is often in real time, when, for a couple of days, damages and losses raise the attention to the public and media." (P5L3-4). Consequently, as put by the authors, "A challenge is therefore to provide scientifically sound and reliable information in near real time (about a week) about human influence on extreme events" (P5L4-5). The question here is: who expresses this demand? Is it the general public and the media, as suggested by the authors? If so, what is the actual societal use of delivering such information on such a short time? One may argue that it contributes to the public awareness of the local consequences of anthropogenic climate change by resonating with the short-term memory. I personally find it a weak argument, and I believe that it is no sufficient to drive such a "rapid" study. Conversely, I would perfectly understand performing an attri-

bution study – without the "rapid", as I will detail below – as a way for government, local authorities, or regulators to inform climate change adaptation strategies.

2. Coming now to the manuscript. It relates this "rapid attribution study", results of which "were completed and released to the public in one week" (abstract, P1L10). The authors have to be congratulated for this performance. However, do scientists really have to be congratulated for performing a study with such a speed? Would the study have been "quick-and-clean", the answer would be definitely yes. But according to the authors themselves, the short time frame imposed severe constraints:

    - "However, the data required to analyse the event at the sub-daily scale in real time is not yet available to us (although it is publicly available at DWD)." (P3L31-32)
    - "As French precipitation data were not available in real time, the analyses there are based on a relatively sparse subset of stations." (P7L12-13)
    - "These values were taken to represent the observed event in the following as the E-OBS data for 29-31 May 2016 was not yet available." (P717-18)
    - "We have not yet investigated the reason for this." (P8L16)
    - "We did not managed to process the CORDEX simulations for Germany in the near real time window for the study." (P15L13)
    - "The methods to answer this question have not yet been developed enough to answer it in the rapid 10-day time window." (P17L12-13)

    The main issue here is the observational data used. The authors relied on data available in real-time (namely gridded products based on a sparse network of stations), i.e. both sparse and not quality-controlled data. The use of extreme values from such data would at least require checking them against the best available data, which are usually available a month later (for manned rain gauges). The

    fact that no radar data was used in the study was also surprising as these are usually available in real time, even if their quality may be discussed (but at least they offer a detailed spatial view of rainfields contributing to a robust estimate of catchment-average precipitation). A secondary issue here concerns the variable used for this attribution study. In France, it focuses on 3-day precipitation, and not on the actual streamflow values reached during the flooding event. The authors are aware that factors other than this high precipitation intensity came into play:

    - "[...] many other contributing factors are neglected in this rapid attribution study." (P3L15)
    - "Firstly the soil types and saturation levels at the time of this extreme rainfall event have not been captured." (P3L15-16)
    - "This analysis also does not take into account the impacts of the reservoirs [...]" (P3L17-18)
    - "In addition, land cover and associated runoff characteristics have also not been taken into account" (P3L19)
    - "A full attribution of the flood themselves, rather than just the rainfall event, would need to take all of these factors into account." (P3L19-20)

    It appears all the more disappointing from the hydrological point of view that such a streamflow attribution study would have been possible (or at least a subset of the experiments) with the help of catchment hydrologists, had the "imposed" time frame not been so short. As a conclusion to this comment, I would ask this question: is the speed at which a scientific study is performed a positive point for evaluating a corresponding manuscript? Can it compensate other negative consequences of the study resulting from the reduced time frame? I will leave the answer to the editor, but from my point of view, this is clearly no.

3. Coming now to the *contents* of the manuscript. The authors claim that it took them "an additional week to finalise this article" (abstract, P1L10-11). Again, I

believe that the authors should be congratulated provided that their manuscript is clear and sound. However, this is rather not the case:

- The manuscript is definitely not well organized: there is no Data nor Discussion sections. Results and discussions and intertwined. Some discussion-relevant elements also appear in the introduction. Results are commented in the conclusions, etc.
- The central method for deriving risk ratios between now and earlier in the 20th century, and for comparing them with risk ratios derived from factual/counterfactual worlds is not justified nor detailed enough.
- The text is vague in many locations, on methods, on the use of GCM/RCM data, on the interpretation of results, etc. It therefore makes the study not reproducible at all.
- Several figures are not referred to in the text.
- There are several inconsistencies between the text and figures/tables.
- etc.

This manuscript therefore clearly makes the reader feels it has been written in a hurry (like in one week), while it also suggests that the scientific underlying content may be rather valuable (with the restrictions mentioned above). Trying nevertheless to adopt a constructive approach, I took approximately half the time of the authors' writing to identify and list all the points that could/should be improved in the manuscript. This (long) list is given in the Detailed comments section below.

In conclusion, I would recommend the editor to reject the manuscript, and to invite the authors to resubmit a manuscript to HESS. I would recommend this new manuscript to be written without the – presumably – artificial time constraints, and to be based on higher-quality observational data that has been made available since their initial

submission. I would welcome any further discussion with the authors, the editor, and any other scientific contributor, as I believe the issues raised above are of general significance as they allude to possible drifting ethics in science and science publishing.

Detailed comments

Abstract

1. P1L18 “all four climate models”: what are they? Please define clearly the attribution set-up.
2. P1L17 “has increased”: between what and what?
3. P1L19 “The observed trend”: over what period?
4. Figure 1, a and b: It would be great to have the underlying network of precipitation stations. This should be at least available for E-OBS. And please add a scale to each map.

Introduction

5. P2L13 “most severe event”: in terms of what?
6. P2L14 “25% of the flood peak of the Seine”: How is this estimated? Please provide a reference or a method.
7. P2L15 “height above 6.1 m”: At what hydrometric station?
8. P2L16-17 “Some measurement problems...”: What kind of problems? Please clarify this.
9. P2L19 “of about 20 yr”: How is it estimated? Any reference?
10. P2L28 “forced to close”: on what day?
11. P2L28 “without electricity”: where?

12. P3L2-3 “46% above the previous records”: Reference?
13. P3L9 “$-40m^3$.s^{-1}”: for a peak flow of...? As such, this figure is not informative.
14. P3L10 “3-day precipitation”: Well, this France-averaged estimation is not relevant for all basins. It first heavily depends on the catchment size, but also on flood-generation processes which are catchment-specific. The relevant precipitation time scale for catchments located in the Cévennes area (south-eastern fringe of the Massif Central) is much closer to a few hours, whereas it is several (and usually more than three) days for the Seine@Paris due to the buffering effects of large aquifers. Please better justify your choices here, as the whole study depends on it.
15. Figure 2a: What is represented with light blue and orange colours? Please add a legend, and a scale.
16. Figure 2b. Please add a scale and remove the title.
17. Figure 2: The reader should be able to compare the model grid scales and limits (see “which ends at 13^{e}E” in the legend) to the observation network and density. Please add such grid scales in some way to this figure.
18. P3L11 “close to the response time”: how is it estimated? And on what rivers? With what catchment size? Cf. also above comment.
19. P3L31-32 “However, the data required to analyse the event at the sub-daily scale in real time is not yet available to us (although it is publicly available at DWD)”: I don’t understand. If they are publicly available, why did you not use them?
20. P5L14 to P6L8: In my point of view, these paragraphs related to future projections are not relevant in the introduction. Results on observed trends and anthropogenic attribution may however be qualitatively checked against findings for 21st century projections in the discussion.

21. P6L9-14: Please summarise the attribution set-up in a few words. The reader may not understand what kind of models are referred to P6L13.

Methods

22. Section 2: The whole "Methods" section is quite unclear. Below are (some of the many) points that need clarification, additional references, etc.
    (a) P6L19 "4-yr smoothed global mean temperature": Please justify the use of the global temperature as an indicator for anthropogenic climate change. For example, why wouldn't you alternatively use the local/regional temperature like in Vautard et al. (2015)? Or maybe the CO2-equivalent greenhouse gas concentrations? This would more in phase with the GCM counterfactual set-up. Less importantly, please also justify the use of a 4-yr smoothing.
    (b) Equation (1): Please define T'. Is it the global mean temperature?
    (c) P6L25 "values larger than about 0.4 are penalised as unphysical": Please give a reference for documenting and justifying this penalised approach.
    (d) P6L26-27 "but take correlations [...] with a moving block technique": Please detail and clarify.
    (e) P6L30 "We evaluate these for the year 2016": I presume this means that you evaluate the cumulative probability density with the temperature of 2016. And so what is this temperature, given that annual temperature for year 2016 is not available? Spring temperature? Please clarify. Same for P7L2-3.
    (f) P7L1: This is actually not a trend detection. This is only a ratio of probability in two specific years, without any formal statistical test, and therefore there is no trend, and no detection. Please rephrase.
    (g) P7L2: Please justify the use of year 1960. Results might have been very different with year 1940 when global temperature first peaked. See also P7L8-9 "the change from 1960 to now...".

23. Section 2: Reading through this section highlights the fact that a "Data" section is definitely missing before it. Indeed, one cannot grasp the meaning of "neighbouring stations or ensemble members" (P6L27), "(analysis and reanalysis)" (P6L29), "Models" (P6L29), "observed record and reliable models" (P7L1), "two models that also have experiments..." (P7L6-7)

24. P7L8-9 "Often we can neglect the effect of natural forcings on these extremes": What does this mean? Please clarify.

Observational analysis

25. P7L12-13: Again, there should be a map of stations used in the gridded products considered here.

26. P7L14-15 "We checked [...] for the past": I don't understand. What is it about decorrelation scales? Why mentioning ERA-Interim here? Is is used at all?

27. P7L15 "Satellite-derived ...": I presume this sentence attempts to justifies the fact that such products are not used here? Please make it explicit and provide some references for their possible low quality of satellite-derived products. And what about radar data? Such products are available in real-time, aren't they? Please provide at least a comment on that.

28. P7L16: How is the CPC gridded product derived from gauges? Please provide some references, and the grid definition. And what about the temporal homogeneity of the underlying network of stations? I mean, is the list of stations used the same in 1960 than in 2016? This is a critical part of the analysis.

29. P7L17, Fig. 3: Please use bar charts for plotting precipitation amounts. Plus, what is the smoother line? Climatology? Over what period? Please add a legend. And remove unnecessary text from the figure titles.

30. P7L21 “The trend analysis”: There is no trend analysis in Eq. 1. Please clarify.

31. P7L21-22 “larger than can be determined with the longest series”: First, what does this mean? Second, what is the longest series?

32. P7L23 “The best fit for the trend is positive”: Please rephrase.

33. P7L23 “However, the uncertainties are large and easily encompass zero”: zero what? Are these uncertainties related to the red vertical bars that appear in (but are not commented nor even mentioned in the legend of) Fig. 4a?.

34. P7L25-26 “We can improve the estimate”: What do you mean exactly by “improving”? Please clarify.

35. P7L25-27 and Fig. 4b: This is really confusing. Precipitation values mentioned above (P7L16-17) are a sum over 3 days, the plot title of Fig. 4c) suggests a daily average, the plot presumably shows the daily average, and the legend says “3-day rainfall”. I would strongly suggest that all plots are made with 3-day sums, not to confuse the reader with the alternative possibility of studying one-day extremes.

36. P7L25-27 and Fig. 4b: Another highly confusing point: the text mentions that the fit is eventually made with a Gumbel distribution. However, the legend of Fig. 4 says that what is plotted is a GEV fit. Please clarify. What adds again to the confusion is the unsuited use of the term “Gumbel plot” for GEV fits...

37. P7L32-33 “The number [...] recently”: This is confusing. What is the number of stations in 1951 for example? Please clarify and rephrase.

38. P8L2-5: This is clearly not enough supported by details on the procedures or relevant references. Please provide some more details/references.

39. P8L6: What is the “present climate”? Please clarify.

40. P8L7-8 “The probability [...] this area”: This statement is not supported by any figure, or am I wrong? Please clarify.

41. P8L9-10 “... so the spatial heterogeneity”. This is unclear, please rephrase.

42. P8L15 “The re-analysis”: What reanalysis? ERA-Interim? I thought you were discussing about E-OBS? Please clarify.

Sections 4 to 8

43. General comment for these sections: model runs should be presented (period, grid, forcings) beforehand in the data section. And Sections 3 to 8 reorganized in a Results section.

44. P9L4: Please specify what are the forcings.

45. P9L7: How do you deal with the ensemble of realizations in the fitting procedure? I presume you put all realizations together in order to reduce the sampling uncertainty. Please clarify.

46. P9L7 “for the same forcings”: Please rephrase.

47. P9L7-8 “The data... use.”: Wouldn’t it be better suited in the acknowledgement section with a contact or website?

48. P9L1-8: What is the spatial resolution of HadGEM3-A? Again, please make a map of the model grid. This is critical to justify the comparison with observed (well, gridded) data.

49. P9L9-10 “The model ... (table 2)”: Once again, this is really confusing. Is this statement valid for both forcings? Is it relative to the location parameter μ? If yes, I can’t see how you may obtain the figures mentioned in the text from Table 2.

50. P10L3: Table 2 only shows a difference in the probability of occurrence of 55 mm/3dy between 1960 and 2016, not an "increase in extreme precipitation". Please rephrase.

51. P10L5 "at $p < 0.025$": What is the statistical test used? Again P10L13.

52. P10L8-9: Again there is no trend here. Only a difference between estimates for two different years.

53. P10L10-12 "2.0 [...] (0.6 to 7.2)": Please detail how these figures are obtained.

54. P10L13-15: This is unclear. Please rephrase.

55. Figure 6: This figure is not referred to in the text. In the legend the model acronym is not consistent with the text and tables.

56. P11L2 "CMIP5 protocol": Please detail this protocol. I presume this is the *historical* runs until 2005. What about afterwards?

57. P11L6-11: This overall negative assessment is really interesting and useful to the community. Same comment for P13L19-21.

58. P12L3: Please make it clear what is the difference between the first experiment (Climatology) and the other two.

59. P12L10-11 "The 2016 data": I don't understand. Please clarify and detail all data types (variables, etc.) used and their specific purposes for the study.

60. P12L12 "The availability ... attribution." Could you give some examples?

61. P12L13-14 "how it depends ... Eq. 1)": This highlights the lack of comments on that point (noted in an above comment) when introducing Eq. 1.

62. P12L17: Please stick to RR once you defined it. Again P13L4, P13L5.

63. Figure 7: The legend mentions "dots", but plots show also lines. What do they represent? I presume it is an envelop of the individual members, but why not pooling them (as for HadGEM, if I am right)? Please clarify.

64. Figure 7: Please display the observational value for 2016 in the plots.

65. P13L6-7 "Note that ... other analyses.": I don't understand. Please clarify.

66. P13L8: What is GloSea5?

67. P13L8-18: This should belong to a discussion part.

68. P13L9 "We found ... over Europe.": Could you explain?

69. P13L10-11 "However ... event": Again, this is not clearly enough explained.

70. P13L12 "and Seine run-off.": There are presumably missing words in the sentence. If this aims at suggesting that Seine runoff is higher than normal in post-Niño years, please provide a reference.

71. P13L25 "CORDEX": First use of the term. Please define.

72. P13L26 "internal variability": Please specify that this is EC-Earth (and not RACMO2) internal variability.

73. P13L31 to P14L1: Please make sure that bias values are consistent with results shown in Table 2 and Table 3.

74. P14L1-2: Please refer to Table 4.

75. P14L3-5: Please refer to Fig. 8. This Figure is not referred to in the manuscript.

76. P14L4-5: Again, what is the statistical test? What does "very significantly" mean?

77. P14L6-7: Please refer to Fig. 9. This Figure is not referred to in the manuscript.

78. P14L12: Please justify the use of runs forced by RCP8.5. I know this may have very little influence, but I'd like it to be commented.

79. P15L4 "biases": I presume on precipitation?

80. P15L4 "simple scaling to a common mean": This is the second time in the paper that this procedure is refereed to, but it is still unclear what is this common mean. Do you simply divide by the observed mean? In that case, at what spatial resolution? Please clarify.

81. P15L6-7 "The uncertainties take that into account": Well, this is far from being sufficient as an explanation, and far from being reproducible. Please detail.

82. P15L8 "The basin averages over the Seine and Loire": I presume you mean "the distribution of basin-average April-June 3-day maximum precipitation"... Please try and be more accurate.

83. P15L8-12: Please refer to Fig. 10. This Figure is not referred to in the manuscript.

84. P15L10-11 "This is ... other models": Is it a formal statement or a more qualitative one?

Conclusions

85. P16L2 "Floods on the Seine are rare this time of the year" and P16L2 "only two late spring/summer floods have been recorded before in over 500 years": Well, I disagree factually. Out of the 30 remarkable flood events identified for the EU Floods Directive (EU, 2007) in the "Seine-Normandie district", eight occurred during April to June (Lang and Cœur, 2014, p. 386): April-May 1983, 16-17 June

1997, April-May 1998, 7-13 May 2000, March-April 2001, 1 June 2003, 7-8 June 2007, 14 June 2009. Among these only, already two show a very similar pattern of soil saturation followed by intense rain, on areas close or very close to those hit by the 2016 rainfall event: the 10-15 April 1983 flood particularly hit the Essonne subcatchment (Lang and Cœur, 2014, p. 404-405), the April-May 1998 flood hit the Yonne and Loing (mentioned P2L7 for 2016) subcatchments (Lang and Cœur, 2014, p. 415-416). But there is also (for example) the 16-23 March 1978 flood that hit the small tributaries south of Paris, including the Yvette river mentioned P2L7 (Lang and Cœur, 2014, p. 403-404).

86. P16L7 "The observational records are too short to establish a trend over the last 65 years.": I don't understand. Is 65 years too short a period to derive a robust trend? As for the length of observational records, the precipitation series available from Météo-France over the Seine basin allow for a computation of basin-scale daily average as least as reliable of this from E-OBS, and for a much longer period. Please rephrase.

87. P16L1012: This should belong to the Methods section. Plus, this is rather unclear as such.

88. P16L14-15 "We just ... result": Please rephrase and detail.

89. P16L14 to P17L2: Some of this should also belong to the Methods section.

90. Table 1: Why does it appear as Table 1 as it is only commented after the two other ones? Plus, I presume you meant "natural-2016" on row 8 column 4.

91. P17L3-13: This belongs to the Discussion section.

92. P18L5-7: I am not sure that comparing the trends in extreme precipitation values in Germany with that of the Cévennes range (with high orographic effects) and Jakarta (in a tropical setting with monsoon influence) is necessarily relevant...

93. P18L11 “the two analyses”: What are they?

Technical corrections

1. P1L14: “return time” → “return period”, and throughout the whole manuscript. This is the most commonly used terminology in hydrology.
2. P1L16: “once roughly” → “roughly once”
3. P1L17 “Seine a factor“: probably a missing word
4. P2L2: “rainfalls” → “rainfall”
5. P2L22: “less” → “lower”
6. P2L22: I believe you mean “March 2016”
7. P2L23: the official name is “EU Sequana 2016”
8. P2L11: “seizes” → “sizes”
9. P7L3 and P8L12: “ration” → “ratio”
10. P7L16: “55,mm” → “55 mm”
11. P12L13: “assumptions to” → “assumptions on”
12. P16L2: “flood crest” → “flood peak”
13. P16L14 “his” → “its”
14. P18L10 “his” → “this”
15. P23L2: “precipitations” → “precipitation”

References

EU (2007). Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks. *Official Journal of European Communities*, L 288/27.

Lang, M. and Cœur, D., editors (2014). *Les inondations remarquables en France – Inventaire 2011 pour la directive Inondation*. Quae.

Vautard, R., Yiou, P., van Oldenborgh, G.-J., Lenderink, G., Thao, S., Ribes, A., Planton, S., Dubuisson, B., and Soubeyroux, J.-M. (2015). Extreme Fall 2014 precipitation in the Cévennes mountains. *Bulletin of the American Meteorological Society*, 96(12):S56–S60.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-308, 2016.