Review of HESS paper submission

Ensemble reconstruction of spatio-temporal extreme low-flow events in France since 1871 by Caillouet et al.

In this paper, the authors set out to reconstruct low-flow events in a reference network of catchments across France from 1871 to present. Due to the lack of observed streamflow records prior to the 1950s, downscaled climatological data (20CR) is used to reconstruct streamflow using hydrological models. Using knowledge of more recent low flow events over the past 50 years, the authors validate their novel approaches of low-flow event identification, spatial matching, and hydrological modelling, which succeed in identifying the well-known events of 1976 and 1989-1990. The approach highlights two additional extreme low-flow events in the 19th century: 1878 and 1893 and the authors conclude that many severe, long and widespread drought events occur prior to 1950.

This paper provides a valuable contribution to science in both methodology and results. The methods are many and varied and though multiple subjective decisions are included, they are well thought out and discussed. Whilst the many steps of the spatial matching procedure are quite challenging to decipher, the results are well presented. The combined threshold level is a particularly interesting concept. This is a very long paper, which might daunt the average reader, but the content is valuable and in my opinion the many maps and graphs all convey interesting information. The maps in Figures 15 and 16 are of particular merit, though Figure 15 would benefit from some slight adjustments (see comment below).

If the editor deems the paper too long (which would be understandable), there are several ways in which the main body of this manuscript could be cut down:

1. The introduction could be reduced.
2. The description of the derivation of the SCOPE climate data from 20CR-SANDHY-SUB could be stated much more succinctly in the main text, and a detailed description provided in the appendix.
3. Similarly the description of the spatial matching procedure is lengthy, and could be summarised with a longer description in an appendix. Those who wish to reproduce your methods would be willing to read the appendices, whereas others, who are mainly interested in the results would not need to understand that level of detail.
4. The day of the year analysis in figure 12 could be cut out, however it does show the spread of the start dates.
5. Furthermore, Figure 13 could be removed, as the majority of the information is given in figures 10 and 11. Again however, it does show the linear relationship between severity and duration of events.

HESS REVIEW CHECKLIST

1. Does the paper address relevant scientific questions within the scope of HESS?
   Yes
2. Does the paper present novel concepts, ideas, tools, or data?
   Yes
3. Are substantial conclusions reached?
   Yes
4. Are the scientific methods and assumptions valid and clearly outlined?

SCOPE Climate production and spatial matching require more clarity, see further comments below

5. Are the results sufficient to support the interpretations and conclusions?
   Yes
6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)?
   Not quite – See comment 4
7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution?
   Yes
8. Does the title clearly reflect the contents of the paper?
   Yes
9. Does the abstract provide a concise and complete summary?
   Yes
10. Is the overall presentation well-structured and clear?
    Mostly, see further comments below
11. Is the language fluent and precise?
    Mostly, see further comments below
12. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?
    Except one error in KGE, see comments below
13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?
    Yes, see comments below
14. Are the number and quality of references appropriate?
    Yes
15. Is the amount and quality of supplementary material appropriate?
    Appendices clarifying spatial matching and SCOPE Climate production needed

I recommend this paper to be published, subject to minor corrections, as below:

Comments that <u>must</u> be addressed:

1. The derivation of the SCOPE climate data is not well explained. Page 6 needs a lot of attention. It appears to be a significant amount of work that is not published elsewhere. This procedure needs to be clearly described, and moved to the appendix. For example, it is not clear what variables from the 20CR are downscaled (500hPa geopotential height?)
2. Please indicate the ideal value of KGE (1 I assume?) on both the description of the metric on page 7 and the map given as Figure 2. Furthermore, please check the equation provided on page 7, as this indicates that a low value should be a good score. Gupta (2009) provide the equation you present as the Euclidian Distance (*ED*), whilst KGE is 1-*ED*.
3. Your description of the spatial matching procedure needs some work to improve its clarity. I suggest it is lengthened and sent to an appendix.
4. Page 25 Figure 17 – add a scale/legend to this figure.
5. Page 31 lines 3 and 4 – You state "Result also highlight that the worst events in terms of severity and duration or spatial extent often belong to the pre-1950 period". This statement is not backed up by your graphics. Figure 14 indicates the majority of the most widespread events were post 1940, Figure 15 indicates that 1990 and 1976 were the longest, and similarly in figure 13 only 1945 is picked out among the most severe in the Corrèze catchment. Furthermore, figures 10 and 11 do indicate some long and severe droughts prior to 1950, but the latter period on these graphs seem to show significantly more events and

more serious events than the earlier period. I think what you mean to say is that many severe, long and widespread events occur prior to the 1950s, but I don't think it's true that the majority of the most severe events are in the early part of your reconstructions. Please amend this sentence accordingly.

Further recommended minor adjustments and comments:

1. Page 2 lines 2 to 6 – references here may be assumed to be based in French catchments due to the outlining of the lack of available data in France, then using the word consequently to start the next sentence. Rephrase to indicate these are not all French studies.
2. Page 3 lines 3 and 4 – "high spatial and temporal scales" and "large-scale atmospheric and oceanic data" high scale means different things to different people. Do you mean high resolution or not?
3. Page 3 line 13 – within which country was Dayon's study based?
4. Page 3 line 16 – downscaling of 20CR – what variables did you downscale and at what resolution?
5. Page 3 line 34 to page 4 line 1 – you mention the use of probabilistic data to account for uncertainty, but very few of your figures display the uncertainty from the ensemble in the results (only really Figures 10 and 11). It would be nice to see some spatial mapping of uncertainty – perhaps a map of one of your events, with density of colour according to how many of the 25 ensemble members identify that event, this map could accompany figure 17.
6. Page 4 line 19 – what does "without direct human influence" mean?!
7. Page 4 line 27 – we are referred to Annex A which mentions Safran Hydro and SCOPE Hydro before the datasets are introduced. I suggest you add "see section 2.2.1 and 2.2.2" to the caption of Figure 18.
8. Page 7 line 20 – I don't think it can be said to give "equal weights", better to say "reduces the bias towards high flows"
9. Page 7 line 24 – how many stations used CemaNeige then? State this. Could you highlight them somehow on Figure 2?
10. Page 7 line 26 – so all the catchments were given the same values for the 2 CemaNeige parameters? Is this realistic? Or were the catchments calibrated several times, and then the median of those given to each individual catchment? This needs to be clearer.
11. Page 8 lines 17-24 – I appreciate that this algorithm comes from reservoir design, but it would be more appropriate to explain it in the context of flow deficit here.
12. Page 8 line 25 - Your *end date* is defined as the time of maximum depletion. Do you think this is representative of the end of an extreme low flow/drought period? What about the time it takes for the stream to return to "normal" flows? See Parry et al (2016) http://ppg.sagepub.com/content/early/2016/06/02/0309133316652801.refs
13. Page 11 Figure 4 – Station 11 shows a single black bar in (a), but a red and a grey bar in (b), is this actually one event or is fig (a) not quite clear enough?
14. Page 14 Figure 7 – why is the late 1990 event not picked up by Safran?
15. Page 14 Figure 7 – this figure takes a very long time to load on my PC – it interrupts scrolling significantly, and almost crashes my browser. Is its file size much larger than Figure 4? If so why? Can it be reduced without compromising its quality?
16. Captions for Figures 8 and 9 – don't think you need to point out the different scales on the y-axes.
17. Page 16 – provide an assessment of the bias in SCOPE Hydro (median) compared with observed flows, for low and high flow seasons. Calculate the percentage of time the obs are within the SCOPE Hydro range for all stations.

18. Page 17 line 12 – not sure I agree that the observations are "most of the time" included in the SCOPE Hydro range – the obs seem to be on the periphery most of the time – address with comment above.
19. Page 18 Figure 10 and page 20 Figure 12 – the dashed lines are visible on the computer, but not on the print out for me. This could easily be a problem my end, but please double check it prints correctly for you.
20. Page 18 figure 10 – the "Whiskers" (note spelling!) do not extend to 1.5 times the IQR, they extend to the largest (and smallest) observations still within 1.5 times the IQR
21. Page 19 line 17 – "A higher number of extreme events and higher severity values are simulated after 1940 for the Corrèze" – do you have any stats to prove this? Mann-Whitney U test for step change? Is there statistical significance?
22. Page 20 line 9 – "There is no visible trend on the seasonality of start dates" – did you do any statistical tests?
23. Page 22 Figure 14 – You have displayed months and years on this plot whilst all other plots just categorise by year. You could remove the months from this plot to make the plot clearer.
24. Page 23 Figure 15 – this is a great Figure, but it really isn't colour-blind friendly. Remember 1 in 10 men is colour-blind. I know this will be difficult to take into account with a map like this, but I suggest you make sure that the 3 or 4 main events you are picking out are in contrasting colours that colour-blind people can differentiate from the others. At the moment, for those with Protanope colour-blindness (the most common) 1990 is clearly visible, 1893 and 1985 look the same as each other, 1878 and 1978 look the same as each other, 1943 and 1949 look the same as each other, and the rest are all very similar. Try using the mobile app "CVSimulator" to test your images (I've run your image through this app, see graphic at end of document). On this note – Figures and 4 and 7 use orange and green which are indistinguishable to the colour-blind. Reconsider this colouring if possible. Figures 16 and 17 are OK, Figure 20 isn't but I don't think the issue can be avoided here sadly.
25. Page 23 line 4 – I think you have 1878 and 1893 the wrong way round in this sentence
26. Page 23 lines 7 to 9 – comment on the 2003 event – this is a major finding, I suggest you add this to the conclusion!
27. Page 26 lines 4 to 9 – this discussion seemed to suggest that 1990 wasn't a particularly severe event, whereas Figure 16 indicates it really was just as severe as the 1893 and 1976 events, especially in the massif central regions.
28. Page 26 lines 10 and 11 – you state that the 1990 event start dates lie within a common 6 month period for <u>all stations concerned</u> – however I can definitely see some green dots on that lower right map in Figure 16! I suggest rewording "all stations concerned" as "for the vast majority of stations"
29. Page 30 section 6.6 – no mention of parameter uncertainty and model structural uncertainty – e.g. GR4J and GR5J.
30. Page 30 conclusion – no mention of the use of Safran Hydro
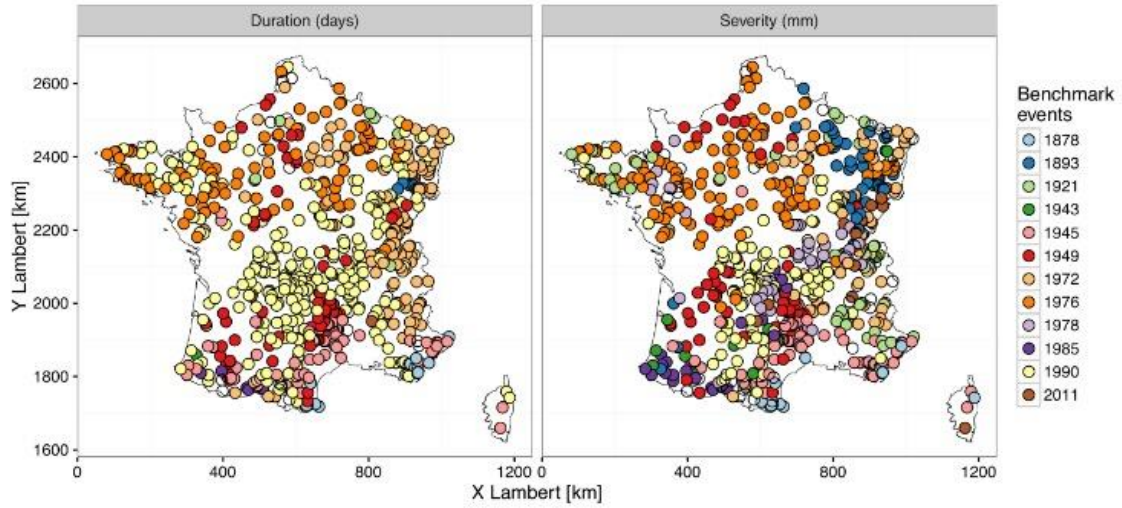
Spelling/Grammatical Errors:

1. Page 1 line 1 – replace "historical depth" with "period" or "length"
2. Page 1 line 7 – change "downscaling of the" to "downscaled from the"
3. Page 1 line 16 – replace first word "on" with "of"
4. Page 1 line 25 – change "a deep knowledge" to "a comprehensive knowledge"
5. Page 2 line 4 – remove the word "only"

6. Page 2 line 6 – replace "or on few stations available on longer periods" to "or on a few stations available for longer periods"
7. Page 2 line 8 – replace "breadth and depth" with "spatial coverage and record lengths"
8. Page 2 line 13 – I would use the word "recovered" rather than "reconstituted" in this context
9. Page 2 line 19 – replace "it has to be noted" with "it should be noted"
10. Page 3 line 9 – replace "reconstructions on the 20th" with "reconstructions of the 20th"
11. Page 3 line 11 – replace "depth" with "extent"
12. Page 3 line 11 – replace "based on a analogue" with "based on an analogue" – check remaining script for correct usage of "a" and "an"
13. Page 3 line 14 – replace "combined to a physically" with "combined with a physically"
14. Page 4 line 21 – add an "s" to influences and remove "a" for "with good quality"
15. Page 5 line 14 – replace "paving" with "covering"
16. Page 6 line 35 – replace "19 diagram" with "Figure 19"
17. Page 8 line 18 – replace "if" with "is"
18. Page 10 line 17 – "would lead to too much aggregated events", do you mean "would lead to too many aggregated events"? If not reword for clarity.
19. Page 10 lines 27 to 29 – this sentence doesn't make sense – it needs re-wording.
20. Page 10 line 30 – replace "earlier" with "earliest"
21. Throughout manuscript – "Cemaneige" written in manuscript, "CemaNeige" is naming convention in IRSTEA documentation.
22. Page 12 lines 19 and 20 – another sentence that doesn't make sense.
23. Page 15 line 4 – replace "goes" with "is"
24. Page 15 line 7 – replace "available on this" with "available in this"
25. Page 15 line 10 – replace "This actually concerns" with "These are the"
26. Page 15 line 16 – replace "an" with "a"
27. Page 15 lines 21 and 14 – replace "centre" with "center"
28. Page 15 line 27 – replace "concerned" with "affected"
29. Page 15 line 29 – replace "estimates" with "estimate"
30. Page 16 line 2 – add "Then for all x > 0," before "the cdf can be computed as follows", and remove the $\forall x > 0$ from the equation
31. Page 16 line 7 – replace "based on a L-" with "based on an L-"
32. Page 17 line 6 – add the word "and" before "split-sample experiments"
33. Page 17 line 13 – replace "good reliability" with "reasonable reliability"
34. Page 18 line 2 – add the work "generally" before "included in the SCOPE Hydro"
35. Page 22 line 14 – reword to "highlights the fact that events covering more than 70% of France have only occurred after 1940".
36. Page 23 line 1 – replace "hit" with "affected"
37. Page 23 line 14 – replace "On the contrary" with "In contrast"
38. Page 23 line 15 – replace "largely" with "widely"
39. Page 24 figure 16 – replace "median of the number of days" with "median start date, as the number of days"
40. Page 26 line 3 – replace "excepted" with "except"
41. Page 26 lines 1, 3 and 18 – replace "hit" with "affected"
42. Page 26 lines 2 and 11 - replace "On the contrary" with "In contrast"
43. Page 26 line 3 – replace "excepted" to "except"
44. Page 26 line 5 – remove "on the contrary"
45. Page 26 line 10 – add apostrophe to 1893 and 1990 events' start dates
46. Page 27 line 21 – replace "on" with "of" in subheading

47. Page 27 line 27 – replace "largely documented" with "extensively documented"
48. Page 27 line 19 – replace "hit" with "affected"
49. Page 27 line 30 – replace "1975 on" with "1975 onwards"
50. Page 28 line 30 – replace "hit" with "affected"
51. Page 29 line 2 – replace "in a spatio-temporal view" with "from a larger scale spatio-temporal perspective"
52. Page 29 line 2 – replace "want to draw a more" with "want to make a more"
53. Page 29 line 4 – replace "matching of event" with "matching of events"
54. Page 29 line 6 – replace "consist in" with "consist of"
55. Page 29 line 7 – replace "this can lead to identify multiple" with "this can lead to identification of multiple"
56. Page 29 line 7 – add comma after "dataset" followed by "but"
57. Page 29 line 18 – replace "the right way forward" with "a promising way forward"
58. Page 30 lines 6 and 10 – you use GR6J-Cemaneige and then GR6J+Cemaneige – do you mean the same thing by these? If so be consistent.
59. Page 30 line 11 – replace the word "evolved" with a more appropriate word "expanded"? "decreased"?
60. Page 30 line 20 – place comma after hydrometeorological studies and remove "-"
61. Page 30 line 21 – remove "-" so comma follows "flows"
62. Page 30 line 21 replace "the limited historical depth of surface" with "the limited historical amount of surface"
63. Page 30 line 28 – replace "across a set of stations" with "space"
64. Page 30 line 32 – replace "qualified" with "quantified", or say "these events are qualified, quantified and compared …"
65. Page 31 line 2 – replace "improve our knowledge on historical events by taking advantage of the higher historical depth of upper-air atmospheric data" with "improving our knowledge of historical events by taking advantage of the more abundant historical upper-air atmospheric data"
66. Page 31 line 9 – replace "deeper" with "more comprehensive"
67. Page 31 line 10 – replace "proxys" with "proxies"
68. Page 31 line 20 – replace "match generally well the observed ones" with "generally match the observed ones well"
69. Page 31 line 27 – remove "in" in "used for in the spatial"

C

P