Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-576-RC3, 2020

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Nonstationary weather and water extremes: a review of methods for their detection, attribution, and management" *by* Louise J. Slater et al.

Anonymous Referee #3

Received and published: 23 December 2020

The authors present a review of methods and metrics for the detection, attribution and management of extreme events. Given the large body of literature on these topics, the authors have identified an area ripe for a solid review that will benefit the community. The manuscript is a very large piece of work that contains a lot of useful and interesting material.

In order for the reader to extract the most from this material, particularly new readers to this field, I would like the authors to enhance the educative components of this review. To this end I recommend improving the graphics (or adding tables) to distill the complexity for the reader into clear, concise overviews of the topics and discussions.

Figures 1, 2, 5b, 6, 7, 8 are good, but Figures 3, 4, 5a should be revised into conceptual diagrams to help the reader. Figure 7 is a good conceptual diagram, but it is not tightly linked to the text, so could be replaced. For example, Figure 2 gives an overview of the workflow in this type of analysis, which is excellent, and the structure of the manuscript follows this workflow. Within each main section, or large sub-section, it would be good to have a conceptual diagram that helps the reader understand the relevant concepts within that section. The text could then be more tightly linked to these figures and then the reader would see the concepts and understand more the logic flow within each section.

Another aspect that I would encourage further work on is linking more to existing reviews of detection and or attribution of extreme events to climate change. I was expecting reference to works like Hulme (2014), Zhai et al. (2018), Ummenhofer & Meehl (2017) and Easterling et al. (2016).

Overall, I am very happy with the content of the review – it is a major piece of work that draws together a lot of diverse material. I think this review could be made even more useful by improving the Figures to help the reader get the most out of this material.

Specific comments Page 2, line 12-13: Is deciding whether a time series should be treated as stationary or not for the purposes of managing extremes "one of the greatest challenges" facing scientists and practitioners today? I agree it is a challenge, but as you point out in this paragraph, the impact of picking the wrong model is a difference in uncertainty. Given the numerous uncertainties in this topic, I am not convinced this is one of the greatest challenges we face.

Page 2, line 14: apparent trends in, and or correlations between, stationary long-memory / auto-correlated / smoothed time series has a long history, which could be highlighted by adding references to Slutzky (1937), Wunsch (1999), Yule (1926) and the Slutzky-Yule effect.

Page 3, line 9: another example of interdecadal to multidecadal hazard-rich and hazard-poor fluctuations is the flood- and drought-dominated regimes of rivers in eastern NSW, Australia (Warner, 1987).

Page 3, line 12: the sentence “However, there is yet no comprehensive, introductory overview of these methods across hydroclimatic extremes, or overarching discussion of key challenges that can arise.” is a challenge to read. How about “However, a comprehensive, introductory overview of these methods across hydroclimatic extremes, including an overarching discussion of the key challenges that can arise, has not been published to date.”

Page 3, last paragraph: in the paper roadmap presented in this paragraph, no mention is made of the “discussion of key challenges” highlighted in the justification for the paper. Is the discussion scattered throughout the sections or presented in a particular location. Please signpost that discussion in the paper roadmap.

Section 2.1: There are a lot of different metrics for each variable extreme presented here. A table with one row per metric that summaries the variables, metrics, inputs, etc would be useful here to distill the complexity for the reader. An alternative to a table would be to revise Figure 3 (see next comment).

Figure 3: many of the terms marked in 3a are not actually used in the text – loss curve, excess rainfall, centroid of rainfall excess. If the Table suggestion above is too difficult, then an improved figure to tightly link the ideas discussed in Section 2.1 with the image shown would help the reader. The idea of example conceptual time series for each variable with features of interest highlighted is good, but the features highlighted need to align with the text better. An improved version of this diagram could enhance the readers understanding of the material in this section.

Section 2.2 & Figure 4: Again there is a lot to digest in this section and I don’t think Figure 4 is helping the reader to understand the concepts. Rather than showing example results, I think a conceptual diagram of some of the concepts discussed would be more useful here. Then the text could be tightly linked to the diagram to help the

reader follow the narrative and understand the differences in metrics.

Page 10, line 7: change “more strongly that the individual” to “more strongly than the individual”.

Page 12, line 11-12: you could mention the CAMELs initiative/papers here, which seek to publish large integrated hydrologic datasets for regions of the world.

Page 12, line 18: the paper by Thyer et al (2006) on how long a record needs to be for stochastic model identification (random, AR1 or HMM) could also be mentioned here as it suggests we need very long records to adequately identify our stationary models (let alone our non-stationary ones).

Page 13, line 2: change “with and perhaps a global "change point" in climate” to “with perhaps a global "change point" in climate”.

Page 13, line 6: Variability of the time series is taken into account in the calculation of significance. It is better to say that for a high variability time series it takes a longer record to statistically identify a change of a given magnitude than the same change in a lower variability time series. See Chiew & McMahon (1993) for a discussion of this issue.

Page 13, line 22-27: I totally agree about the importance of using long records for trend detection and the danger of short record lengths. Another area being explored to extend data back in time for insights into current conditions is palaeo-hydroclimatic reconstructions. For example, Freund et al (2017) reconstructed warm and cool season rainfall in Australia to then investigate recent observed trend magnitude in the context of palaeoclimatic variability. Hydroclimatic reconstructions of the last $\sim$500 years have great potential to place recent observations into a long-term context that is not achievable from short observation based record lengths alone.

Page 14, line 15-23: I think people use the Mann-Kendall test because it does not assume a linear trend. The stated reason here is skewness – however, skewness can be resolved by transforming the data (Box-Cox). Linear regression only identifies linear trends, whereas Mann-Kendall identifies any monotonic trend, which is a much more useful/general feature of the Mann-Kendall test. Also the Thiel-Sen slope is not a test, it is just a way to estimate the magnitude of the trend. Also a reference to Hamed (2009a, b) for applying the Mann-Kendall test to auto-correlated data would be good to add.

Page 15, line 9: it is worth mentioning that the AIC and BIC assess the trade-off between goodness of fit and model complexity. More complex models are penalised, so they have to improve the goodness of fit enough to overcome the complexity penalty. At the moment the sentence is all about better fit, when it should be about better fit even when the increase in model complexity is taken into account.

Figure 6: While I agree the GA nonstationary model has the lowest BIC value and has a nice flat worm, it is interesting to note that the BIC values of the other three models are fairly similar to the GA nonstationary value. If you were to plot 6b for the other three models, would you also achieve an acceptable fit? Are any of the four models not an acceptable fit? It may well be that all four models are “acceptable”, but the GA nonstationary is slightly more acceptable than the rest. If you present this Figure as an example of what can be done with GAMLSS models, then it would be good to more fully explore the results to justify the argument that the GA nonstationary model is the best model and the others are not.

Page 20, line 13+: in this discussion of changes in flooding in a warmer world, it would be good to highlight the finding from Wasko & Nathan (2019, Figure 7) that lower annual recurrence interval floods are more likely to be reduced due to drier antecedent soil moisture conditions, whereas higher annual recurrence interval floods are more likely to increase due to increases in extreme rainfall.

Page 23, line 23: remove the repetition of “to continue”.

Section 7.2: I was expecting to see mention of climateprediction.net and

or weather@home (https://www.climateprediction.net/models/weatherathome/), which are ensemble runs of global or regional climate models that can be used to quantify internal ensemble variance and compare it to observed events.

References Chiew FHS, McMahon TA 1993. Detection of trend or change in annual flow of Australian rivers. International Journal of Climatology, 13(6): 643-653. Easterling DR, Kunkel KE, Wehner MF, Sun L 2016. Detection and attribution of climate extremes in the observed record. Weather and Climate Extremes, 11: 17-27. Freund M, Henley BJ, Karoly DJ, Allen KJ, Baker PJ 2017. Multi-century cool-and warm-season rainfall reconstructions for Australia's major climatic regions. Climate of the Past, 13: 1751-1770. Hamed KH 2009a. Exact distribution of the Mann–Kendall trend test statistic for persistent data. Journal of Hydrology, 365(1-2): 86-94. Hamed KH 2009b. Enhancing the effectiveness of prewhitening in trend analysis of hydrologic data. Journal of Hydrology, 368(1-4): 143-155. Hulme M 2014. Attributing weather extremes to 'climate change' A review. Progress in Physical Geography, 38(4): 499-511. Slutzky E (1937). The summation of random causes as the source of cyclic processes. Econometrica, 5(2): 105–146. Thyer M, Frost AJ, Kuczera G 2006. Parameter estimation and model identification for stochastic models of annual hydrological data: Is the observed record long enough?. Journal of Hydrology, 330(1-2): 313-328. Ummenhofer CC, Meehl GA 2017. Extreme weather and climate events with ecological relevance: a review. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1723): 20160135. Wasko C, Nathan R 2019. Influence of changes in rainfall and soil moisture on trends in flooding. Journal of Hydrology, 575: 432-441. Warner RF 1987. The impacts of alternating flood-and drought-dominated regimes on channel morphology at Penrith, New South Wales, Australia. IAHS-AISH publication, (168): 327-338. Wunsch C (1999). The interpretation of short climate records, with comments on the North Atlantic and southern oscillations. Bulletin of the American Meteorological Society, 80(2): 245–255. Yule GU (1926). Why do we sometimes get nonsense-correlations between time-series? âĂŤ A study in sampling and the nature of time-series. Journal of the Royal Statistical Society, 89(1): 1–63. Zhai P, Zhou B, Chen Y 2018. A review of climate change attribution studies. Journal of Meteorological Research, 32(5): 671-692.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-576, 2020.