https://hess.copernicus.org/preprints/hess-2022-401/

Review comment posted 8 Feb 2023, with responses from the authors.

Thank you for the opportunity to review Faulkner et al. Modelling non-stationary flood frequency in England and Wales using physical covariates

We are grateful for this in-depth and carefully considered review.

The central problem the manuscript seeks to address (that of extracting decision-useful outputs from non-stationary flood frequency analysis) is compelling problem and I agree with the authors that this issue is often overlooked in the academic literature. The fact that one of the authors is from the Environment Agency of the UK shows that this work is recognized by the appropriate government agency.

That being said, this manuscript is essentially a rewrite of a technical report on the topic of non-stationary flood frequency submitted to the Environment Agency (Faulkner et al., 2020; which is extensively referenced in the manuscript). I did not do an extensive/exhaustive comparison, yet it is clear that some portions of the manuscript are directly copy-paste (or nearly so; e.g., compare lines 81-105 of the manuscript to pgs. 34-35 of the report, lines 119-125 of the manuscript to pgs. 36-37 of the report, Figure 3 of the manuscript to Figure G-9 of the report, and Table 2 of the manuscript to Table 4.7 of the report). I state this (without any judgement) for the purpose of bringing it to the attention of the editor.

It is correct that parts of our manuscript are closely based on the report that we wrote for the Environment Agency, as we state on lines 53-55. The report covered several aspects of non-stationary flood frequency estimation, and in this paper we bring out one, that we believe represents a significant advance and one that is worth wider dissemination. In the paper we add, among others, further development of the concept (e.g. version 5 of the candidate models), a more focused and up-to-date literature review and a wider-ranging discussion of how this sort of analysis can be extended to consider future conditions.

Apart from what is noted in the previous paragraph, the science presented in this paper is incremental. Essentially, this is a case study applying non-stationary frequency analysis to flood events in England and Wales.

The primary advancement, that of an “integrated flow estimate” which allows for representing the non-stationary results in a decision-centric way, is essentially an integration of the non-stationary flood frequency estimate over the domain space of the covariates. I think this advancement, while incremental, is useful and should be made know to the academic/practitioner world. However, it seems it has already been proposed by Eastoe & Tawn (2009). If improvements are needed (lines 127-129), then perhaps this additional paper is justified; but, what were the weaknesses? (They are not mentioned) and does this paper address those weaknesses? Here was a missed opportunity to further justify the significance of this paper.

We acknowledge that on a probabilistic level the advance is incremental. However, from the perspective of statistical application, this is the first application of the concept of the marginal return level to flood frequency analysis. The covariates used by Eastoe and Tawn (2009) primarily exhibited seasonal variation and longer-term trend was not considered. This was one reason given by Faulkner et al. (2020) as to why further investigation is needed before physically meaningful covariates can be applied in flood management practice. Another was the need to consider how to incorporate covariates that have distributions with tails that are not well-represented by empirical estimates, e.g., can have values much larger than the maximum observed values.

An additional innovation is the idea of integrating over some but not all covariates, so that the estimate from the model remains conditional on some covariates, such as the year (the "single-year integrated flow estimate"). We will expand the discussion of this because it can be generalised to any covariate whose behaviour is reasonably predictable such as urban cover.

However, the reason that most convinces us that this advancement should be published in a hydrological journal is our impression, discussed in our introduction, that papers on non-stationary flood frequency typically do not include the step that would allow results to be extracted from non-stationary models with physical covariates.

We will add more justification of the significance of the paper in accordance with these suggestions.

Also, the results associated with this "integrated flow estimate" are fairly minimal (Figures 5 and 6; lines 297 – 322; nothing at all in the discussion and conclusion). Since this is the stated main point of the manuscript, more is needed.

You could walk through a hypothetical example of how a decision-maker or engineer might use Figure 5. You could answer questions such as (and I am sure you can think of additional ones to enrich this portion of the manuscript): What is the implication of a 7x ratio of non-stationary to stationary? Which AEPs are used by UK regulatory standards (and how that does affect the interpretation of these results)? What about the case where the ratio was 0.54 (should we disregard the non-stationary model in that case and use the stationary model since it is more conservative)?

The main point of the paper is to present and discuss the method rather than its results from one particular application, which is included as an illustration, as we explain in lines 51-53. However, we are happy to add more discussion of how the results might be interpreted and applied. One aspect we can emphasise is that there are other factors beyond the model fit as measured by BIC that should be considered when selecting a preferred model for decision-making at a particular site. These include other statistical measures such as AIC and likelihood ratios, inspection of model fit on P-P and Q-Q plots, consistency of model form between locations and hydrological reasoning, confidence limits and comparison between the flow estimates and the recorded flood peak data. We would recommend that practitioners consider all these in combination.

Which brings me to the issue that the discussion and conclusion seem disconnected from rest of paper. They read more like a literature review which is found at the beginning of a work, rather than a reflection on what was done. I do not disagree with what is noted in the discussion and conclusion, but it is not in the right place in the manuscript and is not connected with what was shown in the results. Consequently, the reader misses out on a discussion on the actual results that were presented (e.g., answers to the various questions posed in the previous paragraph).

The majority of the discussion section (lines 334 to 390) is about extending non-stationary models into the future. The reason why we included this material in the discussion rather than the introduction is that it is pointing ahead to potential future work, supported by discussion of the current state of scientific practice, rather than introducing the background to the analysis presented in the paper. However, we acknowledge it is unconventional to include so much literature citation in the discussion so we will look to find a more appropriate home for this material in the paper.

Some minor comments

Line 79: To my knowledge, I do not think that Francois et al. 2019 use/reference the East Atlantic pattern (East Atlantic is not found when I do a search). Please double check and correct the reference. As a recent review paper on this topic, Francois et al. 2019 is relevant and should probably be referenced in this manuscript, but referenced correctly.

Correct, the reference about the East Atlantic pattern should have been to Steirou et al. (2019), as for the NAO. We agree Francois et al. is an important paper and will refer to it.

Line 82: AMAX is not defined.

We will spell this out.

Lines 94 – 105: I am not sure that this discussion of Reason 2 for choosing physically-based covariates correctly captures the intent of many of the authors I have read, and the approach as best implemented. The intent of a physically-based covariate to is represent mathematically some physical driving mechanism of floods, whether oceanic, atmospheric, or land based. The confusion of correlation for causation will only happen when an analyst somewhat blindly applies this method: that is, coming up with a suite of possible covariates and trying as many possible and choosing the one with the best correlation. As written, I think this section is somewhat misleading (as if it is the fault of the method, when really this is an error in application of the method).

We have encountered some confusion between correlation and causation in situations other than that suggested. Even covariates that represent a plausible physical driving mechanism may not necessarily be the sole or main driving mechanism for a particular catchment, but if the covariate is increasing along with the floods, there may be a correlation. We will add reference to the three ingredients of trend attribution suggested by Merz et al. (2012) : evidence of consistency, evidence of inconsistency, and provision of confidence level. The same paper makes some specific points about correlation between precipitation and flood magnitude being an insufficient way of identifying a driving mechanism.

We could also cite the work of Montanari and Koutsoyiannis (2014) and Serinaldi, Kilsby, and Lombardo (2018) who argue that a non-stationary model can only be justified where one has deterministic information on the process of change.

We agree that such confusion is not the fault of the method, and will make this clearer. In our experience it is a very common misconception in the method's application. The point that we make that "in principle it would be possible to include any covariate with a trend, whether or not it had any physical connection with the processes that cause floods." is deliberately absurd and so perhaps too easily dismissed. We will strengthen the argument by discussing a more realistic example, which is already outlined in lines 349-356 in the discussion section.

Line 119: The flood frequency estimate is time-dependent regardless of the covariates – either directly dependent on time or implicitly dependent on time via a physical covariate. Wording could be improved here.

Agreed, we will make this change.

Line 198: Why were the models just limited to two covariates? In particular, why weren't the two physically based covariates used in one model – are they too strongly correlated? Need to justify this choice. Based on skimming the report, I suspect your answer might be that it was too many models... perhaps; but the model with the two physical covariates seems important enough to test regardless of the issue of computation.

Both suggested reasons are relevant: cross-correlation of the physical covariates and proliferation of model types. We regarded 88 candidate models as a large enough number to consider. For other applications in which covariates describing catchment land cover are combined with climatic covariates, it may be reasonable to allow combinations of physical covariates. We will add this point to the paper.

Line 219: Why include the GLO if the results are not analyzed? My suggestion is to remove it from the paper, or justify its inclusion (and include results).

We are happy to remove reference to the GLO.

Figure 6: They y-axis scale is confusing. Is it a log-scale? If so please note in the caption. More labels would also be helpful.

We will improve this figure.

Appendix B: Why is this included? It is just a suggestion, without any testing or analysis or use in the manuscript. I suggest to remove it or to more fully incorporated in the manuscript.

We included it in the hope that others could use in future work. But we are willing to remove it.

References

Merz, B, Vorogushyn, S, Uhlemann-Elmer, S., Delgado, J and Hundecha, Ya. (2012) More Efforts and Scientific Rigour Are Needed to Attribute Trends in Flood Time Series. *Hydrol. and Earth System Sci.,* 16, 1379-1387.

Montanari, A. and Koutsoyiannis, D. (2014) Modeling and mitigating natural hazards: Stationarity is immortal!, Water Resour. Res., 50, 9748–9756.

Serinaldi, F., Kilsby, C.G. and Lombardo, F. (2018) Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. Advances in Water Resour., 111, 132-155.