

HESS Review of:

Virkki et al. 2021 Environmental flow envelopes: quantifying global, ecosystem–threatening streamflow alterations

General:

Title implies ecosystem-threatening alterations are quantified - is this really the case given the lack of ecosystem assessment within the analysis?

English grammar will need a full check and improvement, as some of the points made are hard to follow as written. Shorter, sharper sentences will help.

e.g. Abstract li 39 - The use of less and few.

Introduction li 50 - “rivers upkeep two major...”; li 84 - data are plural

The use of language is strong in some cases, perhaps intentionally? - e.g. threatening, violations - this does influence the tone of the MS.

Abstract:

li 27 - “We present a novel method to determine EFs by Environmental Flow Envelopes (EFE), which is an envelope of variability bounded by discharge limits within which riverine ecosystems are not seriously compromised.”

The authors need to take care not to overstate this point. Presenting the approach as novel is highly problematic in that a sustainability boundaries/envelope approach to eflows was proposed over a decade ago, e.g. by Richter (2009) and then Richter et al. (2012) based on concepts presented in Postel and Richter’s book *Rivers for Life* (2003). The application of the approach globally, at finer basin scale, and the attention to frequency, severity, and trends combined, are arguably where there is novelty in this paper.

In addition, the claim that ecosystems would not be seriously compromised will not be supported by evidence and should be tempered.

Some recent/key references need consideration, incl:

Richter, B. D., M. M. Davis, C. Apse, and C. Konrad. 2012. A presumptive standard for environmental flow protection. *River Research and Applications* 28:1312-1321.

Poff, N. L., C. M. Brown, T. E. Grantham, J. H. Matthews, M. A. Palmer, C. M. Spence, R. L. Wilby, M. Haasnoot, G. F. Mendoza, K. C. Dominique, and A. Baeza. 2016. Sustainable water management under future uncertainty with eco-engineering decision scaling. *Nature Climate Change* 6:25-34.

li 30 “considering also the methodological uncertainties related with global EF studies”. The logical connection between this phrase and the rest of the sentence is not clear.

li 31. “EFE introduces an upper bound of discharge, identifying areas where streamflow has increased substantially.” Such sentences are poorly expressed and need rewriting for technical substance, as well as sharper understanding; increased above what? Presumably natural levels?

li 41 - Because of the way in which you define a violation, you need to compare the result of 5% with the pre-alteration state. Also, the “discharge” cannot “violate” - you need to revise grammar for correctness in these sorts of sentences, perhaps by saying EFE violations.

li 44 - The potential attribution to climate change (cf. other human stressors impacting upper bounds of the flow regime) - supposition, or do you show this analytically using data?

Introduction:

This section needs at least moderate revision.

li 55 - are 5 references needed for this point? The authors need to strike a better balance between under- and over-referencing in this section.

li 65 - I suggest you also emphasise other aspects of regime variability, including intra-annual and interannual variability.

li 66 - the importance of flow extends well beyond physical habitat. See e.g. Bunn and Arthington (2002). In fact, beyond the very general references, this paper lacks attention to ecological dimensions, despite the title suggesting otherwise.

Bunn, S. E., and A. H. Arthington. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management* 30:492-507.

li 76 - You present the 2018 definition which focuses on a regime that meets a set of objectives, and then regress to the now out of date concept of a minimum discharge in the following “refer to the minimum discharge required to sustain healthy and functional riverine ecosystems (Pastor et al., 2014).” Try to draw out the regime-based argument rather.

li 77 - “Hence, the EFR corresponds to a boundary not to be transgressed.” This is overly simplistic in the way it is stated.

li 77 - “Beyond simple EFRs, more nuanced quantification of anthropogenic impacts on discharge based on a multitude of different metrics include e.g. the Indicators of Hydrological Alteration (IHA; Richter et al., 1997, 1996).” This is a poor explanation.

li 80 - The EF assessments are not implemented in legislation as such. Legislation that provides for EF is one of the factors that enables EF to be implemented.

li 84 - Thereby, global studies accommodating EFs...rather use... considering it as a...well-being” This is not a well structured sentence for clarity or grammar. There are many such examples in the text.

li 109-110. There is potential in the approach to undertake such analyses of trends. However, it is overstating the case that EFes are new as a concept.

li 111 “envelope of safe discharge variability that...” That it is safe or even precautionary is really an assumption and that needs to be clear. You are using a composite of hydrological methods that are all limited in their ecological relevance and representation. You do not undertake any actual ecological assessment either, so you need to take real care not to overstate the case.

li 113 - I suggest you reference a paper that explains hydrology methods, past and present, ecologically relevant flow metrics etc. and is more up to date than the IHA refs you provide e.g. Poff et al. (2017)

Poff, N.L., Tharme, R.E. and Arthington, A.H. 2017. Evolution of environmental flows assessment science, principles, and methodologies. Chapter 11. Pp: 203-236. In: Horne, A.C., Webb, J.A., Stewardson, M.J., Richter, B. and M. Acreman, (Eds). *Water for the environment: from policy and science to implementation and management*. Academic Press.

Methods and Data:

All assumptions and caveats e.g. violations, need to be well explained.

The technical argument for “violations” is weak, particularly in relation to the comparisons of near-natural (largely unmodified by humans), present-day and future flow regimes. A more rigorous explanation is needed. One of the main challenges with the approach and the terminology is with the violation of Q90.

With an eflow of Q90, that flow will be naturally ‘violated’ 10% of the time. Given the authors’ choice of a low flow eflow that is the median value of hydrological methods ranging from Q90 to the mean monthly value of dry months, the so-called ‘violations’ will occur even more frequently. It is thus unclear as to what proportion of the ‘violations’ described are simply a result of the hydrological approach used. This information needs to be provided, as any violations resulting from human impacts on the flow regime should be in excess of the natural ones.

li 150 - You need to provide referencing for HydroBASINS when it is first mentioned, especially as its use is critical for the sub-basin divisions applied. You only provide the reference at li 150. The introductory paragraphs before section 2.1 could focus more on the general steps. Suggest you reword for clarity, so that there is no confusion about what is meant by “into consistently sized” and the hydrological nested approach can be well understood.

Reasons for the selection of models need to be give, for the hydrological, climate change and the five hydrological eflow methods.

Greater explanation of the advantages and limitations of the input datasets from ISIMIP2b is needed.

Similarly, the reasons for the selection of 1801–1860 as the reference flow period need to be expanded on. What is meant by “quasi-natural”? How was this ascertained, or was it assumed based on the pre-industrial era? Likewise, for the present-day/recent data period. What about flow alteration post 2005? For some regions/basins very recent trends are significant - what are the implications of not including this most recent part of the time series?

There are also several good references addressing criteria for hydrological record selection that are surprisingly not mentioned in this section on methods.

li 153-155. You actually exclude a fairly significant number of basins - did you test what the implications might be? At the least you need to objectively acknowledge this fact and its implications, rather than word it in such a way as to downplay it.

li 163 - The EF methods. The strengths and limitations of these methods are not provided, but need to be clear to the reader. Table 1 just lays out the methods. There are a number of more recent hydrologic methods that are less simplistic and/or include a greater diversity of ecologically meaningful flow metrics. You have not used any of those methods in the composite approach. The

reasons need to be given, especially given the assumptions made in the paper around ecological threat/safety.

The paper would benefit from some discussion of why more sophisticated methods including regional approaches (ELOHA based) were not possible to apply at this time.

Li 180-183. This section of text needs careful attention. “As the EFE lower bound, we selected the median of the EFR distribution...potentially consist of unrealistically low or high EFR estimates, caused by either highly deviant discharge provided by certain GCMs or distinctively different representation of ecosystem water needs in the EFR method”. These two points need to be broken down and carefully argued. They are based on assumptions without any apparent ecological rationale. This needs to be clear, e.g. is the choice of median as the lower bound arbitrary? It designates the low-high flow boundary sometimes hydrologically, but that is not the same as a boundary designated for eflows purposes. Also, the last point on “distinctively different representation of ecosystem water needs” is hard to follow - what are you intending to say here?

Li 207. "we excluded time periods during which the EFE is violated for less than three consecutive months"- by definition you would then exclude floods. Can that be correct?

li 213. “ excluded sub-basins with extremely low flow from our analysis” This is a limitation that warrants more explanation or returning to in the discussion section. You also need to clarify the extent to which temporary rivers might be excluded as a result and to discuss the implications (see new Nature paper by Messenger et al. 2021 on the global importance of such systems) Messenger, M.L., Lehner, B., Cockburn, C. et al. Global prevalence of non-perennial rivers and streams. Nature 594, 391–397 (2021). <https://doi.org/10.1038/s41586-021-03565-5>

li 214. You excluded basins if the mean flow was less than 10 cumecs – that is rather high as so many regions have the bulk of their rivers smaller than this. Interestingly the resolution of your maps does not show these smaller exclusions. By way of illustration – the Mgeni River in South Africa supports 6 million inhabitants in two cities Pietermaritzburg and Durban. The river just above the estuary in its natural condition only exceeded 10cumecs for 38 days a year, so this entire river should be excluded from your map – yet it does not show as excluded. This is concerning and should be discussed.

li 225 - “classified each month of record into low ($Q < 0.4MAF$), intermediate ($0.4MAF \leq Q \leq MAF$), and high ($Q > MAF$) flow classes”. On what basis? What evidence is here from the literature to support this categorisation or was it somewhat arbitrary?

li 234. Statistical procedures require supporting references. e.g. for Kendall rank correlation coefficient.

li 236 - “Finally, we combined the EFE violation frequency and severity throughout the recent past time series with the linear violation trend slopes and performed a fuzzy c-means clustering (Bezdek, 1981) to each flow season separately.” This sentence needs breaking down into clear points. Reasons need to be given for choices made e.g. for fuzzy c-means clustering. This is important. e.g., as Fig 5 relies heavily on the clustering method adopted.

Results:

li 239 - "EFE violations are widespread around the world, concentrating on lower bound violations in the arid and dry temperate climate zones" The pre-alteration context needs to be given, for comparison and for these results to have meaning. Also, there are limitations of the approach for rivers with very low to no flows, and this needs covering in the results and discussion.

Again, there are issues around grammar and the expression of information that detract from the sense of this sentence (violations that concentrate on lower bound violations). While I understand what is intended, such sentences need to be corrected.

li 245. "Therefore, the EFE is rather violated by insufficient than excessive discharge, and regional patterns are more clearly visible in EFE lower bound 245 violations whereas EFE upper bound violations are more dispersed into individual sub-basins." I think the results need breaking down further to make this argument. How much of this pattern is a result of: using hydrology based approaches that tend to focus more on the low flow regime (Q90) or monthly time step, and/or less on event-based influences such as floods; or a regional effect attributable to naturally more arid conditions rather than a function of the operation of infrastructure?

How sensitive are the results to the somewhat arbitrary setting of the upper and lower bounds, and to natural flow regime characteristics?

li 273 - Niger River - for a named river, the R should be upper case. Tigris-E Basin? Northern China - the N might need to be lower case unless it is a specific region.

li 259-276. The results are interesting and some of them could feature more substantively in the discussion.

li 287 - "Respectively..." Such sentences are poorly structured and compound, and need to be rewritten for clarity.

li 300 - "increasing discharge alleviating EFE lower bound violations may turn out to be amplifying for EFE upper bound violations in some regions and downplay the positive indications of decreasing EFE lower bound violation trends." Such sentences cannot be understood as written.

Discussion:

The discussion needs greater depth in content, including in relation to ecological limitations, and the implications of the results.

Draw in some points on more arid river systems (where the approach seems particularly weak) and on temporary rivers.

li 346 - Suggest you differentiate between ecosystem change and threat - e.g. in Europe many rivers are already ecologically degraded, not simply at threat of degradation.

li 348 - "Since these areas show relatively little EFE violations, it can be inferred that even though the quantitative discharge would be within the EFE, the anthropogenic flow alteration can still be major." Should read "relatively few/minor". This is a key point that needs more elaboration - otherwise, it suggests that the approach taken/methods used are insufficiently ecologically representative or attuned to reality.

li 359 - "can largely be attributed to anthropogenic impact..." But one might argue that you simply see potential congruency rather than having demonstrated an attribution/causal link through the data?

li 362 - What is meant by "remotely teleconnected" here?

Li 366. This is suggesting that drought is the main cause of increasing violations – not increased abstraction? That is strange as demands in these regions are high.

li 384 - "flaws" - suggest reword as "limitations" and ref. Poff et al. 2017 here. If they are flawed, then this puts their use into question.

li 384 - "global fish biodiversity has shown that several other factors, such as water quality and the presence of invasive species, may be more important in maintaining riverine ecosystems than quantitative flow (Su et al., 2021)." Yes, but the context is important to get right here and in general. Moreover, fish are one group of species, not directly equivalent to an ecosystem or its functions. This section on ecology is weak and needs a very careful read-through to address assumptions/oversimplifications.

li 386 - "well-being" is typically used for people, and integrity or health for ecosystems.

li 394-404. This is a mix of natural flow paradigm, novel and hybrid systems, designer flows, shifting baselines etc. and needs to be sorted out into a more cogent argument.

li 396 - "its globally equal absoluteness". This is not English as written.

li 408 - Section 4.3 needs a serious edit. Please use simpler sentences that make clear points, to help with the logic of your argument.

li 409 - You need to provide a reference or better introduce "ensemble thinking" in this context - here or under the methods.

li 427. Terms such as "excessive flows" need to be expressed using more scientifically appropriate wording.

li 447 - "quantification of the riverine ecosystem responses to prolonged and excessive flows through case studies would benefit the development of the EFE upper bound." Suggest include some supporting references here for a couple of e-flow case studies that address this aspect.

Li 455 – it would be useful to comment on the application of this approach at the local scale compared to the global scale, and the limitations associated with that. In line 466 you mention "operationalising our results at the basin scale", but just what this means could be clarified.

li 466. Can you be more specific on ways to make the approach more meaningful operationally i.e. from an implementation standpoint? There are obvious limitations, given the characteristics of the EF methods used. So the appropriateness and the limits of such simple global scale, hydrology-based methods must be kept in mind in these concluding remarks.

Figures and tables:

The paper follows a nice, clear and logical layout and the figures are generally good.

Fig 3. It is difficult with these colours to differentiate the low violations cases from the less than 10% frequency case. Also, is it well explained in the text as to what is meant by the shorthand “no months in season”?

Fig A1 - There is too much small text in the figure.

Recommendation

Moderate revision.