# Report #1

Widespread and increasing violations of environmental flow envelopes

## General comments

Since the last revision, the authors have answered and addressed the most important comments of the reviewers. Especially the inclusion of pre-industrial violation results and the omission of overly strong language has markedly improved the manuscript.

However, there is still need for further clarification and specification in some aspects of the abstract and discussion. By improving these sections this study could be presented more robust. Moreover, there is also an opportunity to improve the writing. Therefore, I recommend minor revisions.

We thank the reviewer for their encouragement, as well as for the further comments and suggestions. We are also grateful for the specific and insightful suggestions on how to improve our writing, which certainly help us to better communicate our results. We hope that our minor revisions and clarifications are adequate in order to improve the robustness and style of the manuscript.

## Specific comments

Note that some of the comments to do with the writing are based on my personal writing experience, and are not “wrong”. However, they may help the authors to improve their writing throughout the document.

Line 28: What uncertainty? Please clarify

This uncertainty is related to using a limited number of both discharge data sets and EF methods, and we have now specified this here. Lines 27–29 now read as follows:

*"Environmental flows (EFs) have emerged as a prominent tool for safeguarding the riverine ecosystems, but at the global scale, the assessment of EFs is associated with high uncertainty related to the hydrological data and EF methods employed."*

Line 29: “The sub-basin specific EFE is” to “Sub-basin specific EFEs are”

We have reworded this; lines 30–31 now read as follows:

*"Sub-basin specific EFEs are determined for approximately 4,400 sub-basins at a monthly time resolution, and their derivation considers the methodological uncertainties related with global-scale EF studies."*

Line 30: “and its derivation considers the methodological uncertainties”, again please clarify

We hope that the change to the previous sentence (comment on line 28) in lines 27–29 will make it clearer that this “methodological uncertainty” relates to the hydrological data and EF methods.

Line 32: Maybe mention Q95 here?

In the abstract, we would like to refrain from explicitly mentioning how the EFE upper bound is determined, as it would unnecessarily complicate the abstract with technical definitions. To keep it less technical, we haven’t provided the technical definition of the EFE lower bound, either.

Line 32: “This” to “This upper bound”

We have reworded this; lines 32–33 now read as follows:

*"This upper bound enables identifying areas where streamflow has substantially increased above natural levels."*

Line 33: “Long-term” is used throughout the manuscript, for different meanings. This is sometimes confusing. See also some comments below. “Long-term” to “annual”?

We agree with the reviewer that the use of “long-term” was sometimes confusing, and we have replaced all instances related to “long-term flow alterations/EFE violations” with “persistent flow alterations/EFE violations”. Here, our intention was to illustrate that commonly, EF studies show violations for one period of time. For example, Jägermeyr et al. (2017) report EF deficits averaged over 1980–2009, whereas we complement this by assessing the trends within 1976–2005.

Line 37: Best to move “global hydrological model outputs from the ISIMIP2b ensemble” to its own sentence (perhaps at the start of the paragraph “Discharge was derived from ...”).

We have reworded this; lines 37–39 now read as follows:

*"Discharge was derived from global hydrological model outputs from the ISIMIP 2b ensemble. We use pre-industrial (1801–1860) quasi-natural discharge together with a suite of hydrological EF methods to estimate the EFEs. We then compare the EFEs to recent historical (1976–2005) discharge to assess the violations of the EFE."*

Line 40: “widespread, occurring” to “widespread and occurring”, as one refers to the spatial extent and the other the temporal extent.

We have reworded this; lines 41–43 now read as follows:

*"The EFE violations are widespread and occurring in half of the sub-basins of the world during more than 5% of the months between 1976 and 2005, which is double compared to the pre-industrial period."*

Line 44: “spatially distributed” to “dispersed” as in the discussion.

We have reworded this; lines 44–45 now read as follows:

*"Indications of increased upper extreme streamflow through EFE upper bound violations are relatively scarce and dispersed."*

Line 46: Ending on a positive note would be stronger, what is the broader application of your study?

We have reorganised this sentence to end on a positive note and outlined that the globally robust EFEs can inform global research and policies on water resources management. Lines 45–48 now read as follows:

*"Although local fine-tuning is necessary for practical applications, and further research on the coupling between quantitative discharge and riverine ecosystem responses at the global scale is required, the EFEs provide a quick and globally robust way of determining environmental flow allocations at the sub-basin scale to inform global research and policies on water resources management."*

Line 54: “population growth”, not only growth but also agricultural and population development (irrigation expansion and diet changes).

We have added agriculture (especially irrigation water use) as an additional factor contributing to the pressure on freshwater ecosystems, as agriculture is responsible for a major share of all water use, and future unsustainable water consumption is projected to increase (Campbell et al., 2017; Wada and Bierkens, 2014). Lines 56–58 now read as follows:

*"The pressure on freshwater ecosystems is only expected to increase in the future due to population growth, agriculture (especially irrigation water use), and projected climate change (Best, 2019; Campbell et al., 2017; Graham et al., 2020; Thompson et al., 2021, Wada and Bierkens, 2014)."*

Line 63: Probably remove “in addition” as it specifies, and not adds to, the previous sentence?

We have reworded this; lines 66–67 now read as follows:

*"Human actions impact the intra- and interannual variability, which are often considered as parts of the natural flow regime (Richter et al., 2006)."*

Line 124: What are “mechanistic equations”? Probably use “process-based”?

We have reworded this; lines 126–128 now read as follows:

*"Simulating discharge with the GHMs involves modelling the global terrestrial hydrological cycle through process-based equations, as well as forcing the models with observed or modelled climate."*

Figure 1: Could the three month threshold be included here?

We have now included the three-month threshold in Figure 1.

Line 165: What do these levels mean? Maybe omit?

The intention of including the HydroBASINS scale levels in the text was to show that we used a medium detailed sub-basin division in this study, as we were unable to use the highest detail levels, but that we included an adequate level of detail in our analysis. We have now revised the beginning of this paragraph to make this point clearer, and lines 166–171 now read as follows:

*"We used the HydroBASINS sub-basin division, which is a global polygon layer series dividing the world into sub-basins at different scale levels from the lowest detailed level 1 to the highest detailed level 12; we selected the medium detailed level 5 (Lehner and Grill, 2013). Within each level, the geographical areas of sub-basins are relatively equal, and level 5 is the highest level of detail that can be rasterized into a 0.5-degree resolution grid without an excessive loss of sub-basins that are smaller than a grid cell."*

Line 200: Normal distribution cannot be assumed for discharge. I did a quick test for 100 of the world’s largest rivers, and normality could only be assumed for half of them (using the Shapiro-Wilk test on multi-year average measured discharge). Moreover, why is it assumed there are errorous outliers in model simulations that should be removed, how do simulations make errors? This also affects the Q95 upper bound in the study. If so, this should be discussed.

We agree with the reviewer that assuming normal distribution of discharge globally is unsubstantiated, and we have omitted the claim potentially implying so. Here, we did not implicate anything related to the cause of outliers since they can also be extreme natural events (e.g. thousand-year floods). However, we consider that removing outliers, which could largely affect the computation of EFRs and shift the EFE upper bound very high, is justified also considering that some extremely rare natural events may be excluded due to this. If the extreme

outliers were left in the data, the EFE upper bound may rise very high based on extremely rare conditions and potentially mask some of the current EFE upper bound violations. We have revisited the text related to the outlier removal, and lines 204–208 now read as follows:

*"Before computing EFRs, we removed monthly outlier discharge further than three standard deviations away from mean monthly discharge. This procedure only removed extremely deviating discharge values, which could greatly distort the computation of EFRs or shift the EFE upper bound very high if left in the data. Similarly for the resulting EFR distribution, EFRs further than three standard deviations away from mean EFR were removed. This way, we avoided skewing the EFR distribution with extreme outliers in pre-industrial data."*

Furthermore, in Section 4.4 lines 518–520, we've added a sentence outlining that although useful for eliminating potential errors, the outlier removal is also a limitation of the study as it might exclude the most extreme natural events:

*"Moreover, to prevent raising the EFE upper bound extremely high, we excluded outlier discharge prior to determining the EFEs (Sect. 2.2), which may result in excluding not only potential model errors but also extremely rare natural events. "*

Line 328: "figures" to "values" or "results".

We have reworded this; lines 333–334 now read as follows:

*"These values mean that the typical EFE lower bound violation is caused by discharge falling 19–37% below the EFE lower bound."*

Line 423: "both long-term and recent". "Long-term" here is referring to "more than three consecutive months". However, as long-term is used in combination with "recent", which refers to the last 30 years, which is confusing. As an alternative to using "long-term" I would propose to use "extended" or "persisted" throughout the manuscript?

We have revised this terminology according to the comment on line 33, and here lines 425–427 now read as follows:

*"Given that the change from the pre-industrial period is substantial (Fig. 2d) and all considered violations last three or more months (Sect. 2.3), the EFE violations represent persistent flow alterations during the recent historical period."*

Line 425: If possible, could the absolute values be included in the supplementary?

We will release the R code and data used to compose the results in an open repository. In addition, we will include the results underlying Figure 3 in the supplementary.

Line 434: What is a possible reason for this wider spread? Is it related to the model ensemble?

Methodological differences between our study and Jägermeyr et al. (2017) can be assumed to be a possible reason for the wider spread. Mainly, the differences stem from time periods used in determining the “natural” discharge (1801–1860 vs. 1980–2009), as well as the number (five vs. three) and aggregation (median vs. mean) of the EF methods. In addition, we include a greater number of global hydrological models (GHM) forced with outputs from general circulation models while Jägermeyr et al. (2017) use one GHM (LPJmL) forced with observed climate. In our LPJmL-specific results (Figure S4), the EFE violations are spread similarly, agreeing with other models of our analysis. Hence, the difference in time periods and EF methods could be assumed to be the main cause of differences. In the revised manuscript, we have moved the sentence outlining the methodological differences between the two studies right after introducing the differences in results, and lines 439–442 now read as follows:

*"Our EFE violations are more widespread in large parts of Australia, South America, and Southern Africa (Fig. 2–3) than those reported by Jägermeyr et al. (2017). However, Jägermeyr et al. (2017) determine EFRs based on pristine discharge simulation between 1980 and 2009 and report annual averages, which differs from our seasonal analysis that is based on the pre-industrial period and includes potential climate change impacts."*

Line 437: Can you expand upon these limitation (perhaps with citations)? Or maybe move this to section 4.3, where these limitations are discussed.

We have moved this statement to Section 4.3 in which the limitations are discussed in more depth.

Line 441: “(i.e. not taking potential climate change impacts into account as we do)” would be better at the end of the sentence, were you discuss what your study does. “(...) which is different from our seasonal analysis based on the pre-industrial period and includes potential climate-change impacts”

We have moved the sentence including this phrase upwards and revisited it (see previous comment on line 434).

Line 451: “cautious interpretation of our results” to “cautious interpretation of our results in these regions”

We have reworded this; lines 454–456 now read as follows:

*"Regarding the Pan-Arctic areas in particular, GHMs have recently been shown to perform relatively poorly (Gädeke et al., 2020), which calls for cautious interpretation of our results in these regions."*

Line 465: Would omit this sentence as it is discussed before and is addressed in the results.

We agree with the reviewer and have omitted this sentence.

Line 491: Although the limitations of EFs are described, could you indicate how they are useful? Currently this section seems to undermine the value of your study.

As mentioned in Section 4.4, the benefit of using simplistic EF methods is that they can provide a globally consistent overview of anthropogenic flow alteration. We have revisited this sentence to be more balanced and to highlight the benefits of the simplistic methods, and lines 498–501 now read as follows:

*"For practical use and analyses of sub-basin scale riverine ecosystem integrity, our results should be complemented by local studies using holistic EF methods. However, our globally consistent approach using hydrological EF methods provides a comprehensive global overview on anthropogenic flow alteration."*

Line 509: Important for what? Please elaborate.

Here, we aimed to state that temporary rivers are common and important for the local ecosystems supported by them. We have reworded this statement, and lines 523–525 now read as follows:

*"This is a notable limitation particularly in the case of temporary rivers, which have recently been shown to comprise a large part of global rivers, and which are highly important for local ecosystems (Messager et al., 2021)."*

Line 513: Why this specific number? Maybe just omit?

Here, the intention was to clarify that the spatial resolution of our gridded discharge is relatively coarse, and relying on a small number of cells to determine and assess EFEs is highly uncertain. However, providing explicit numbers is ambiguous, and we have therefore reworded this; lines 528–529 now read as follows:

*"Applications at the scale of small catchments consisting of few 0.5-degree grid cells should rather resort to high-detail observed data instead of global data simulated in a coarse grid."*

Line 507: “On the one hand (...) but on the other hand”. These sentences do not contradict, just use “and”.

We have revisited this sentence along with the previous and the next comment.

Line 508: Is there also an advantage to the spatial aggregation (as there is for the temporal aggregation)? If so it would be good to mention here.

Using coarse grid scale could potentially provide unstable results due to high GHM variability in headwater and low-discharge streams. When we simplify the sub-basins to be represented by the outlet cell only, we mask the greatest upstream variability and increase the robustness of our sub-basin scale results. This draws attention to the sub-basin scale instead of showing potentially very uncertain and highly deviating cell-wise results. We have now made it more explicit that this is an advantage, and lines 520–523 now read as follows:

*"Spatially, we consider the sub-basin outlet location as representative for the whole upstream area, which simplifies the sub-basin into one hydrological unit (Sect 2.1). Using the coarse grid scale could potentially provide unstable results due to high GHM variability in headwater and low-discharge streams, which is countered by this aggregation. However, the aggregation also masks local EFE violations that may vary within the sub-basin itself."*

Line 521: Could you elaborate on why the separation of natural and anthropogenic flow alterations would give additional information on the “seriousness” of major violations?

The main reason why this separation would be useful lies in the measures to alleviate major violations. For example, if anthropogenic flow alterations (e.g. water withdrawals) were the main cause of EFE violations in a given region, the measures to alleviate the violations would be very different from those regions in which the main cause is climatic. We have added a sentence in lines 536–538 to elaborate on this claim:

*"Separating natural and anthropogenic flow alterations could prove useful in estimating how major violations – and in which regions – should be deemed as the most serious. This way, the actions to alleviate EFE violations could be best prioritised and targeted to the most affected regions."*

Line 530: “On one hand (...) On the other hand”. These sentences do not seem to contradict. Maybe use “nevertheless” instead of “on the other hand” (and omit “on the one hand”).

We have reworded this; lines 547–550 now read as follows:

*"Operationalising our results at the basin scale requires more detailed data, assimilation of cross-scale information, and interdisciplinary knowledge to more fully portray the ecological and hydrological conditions of each unique river. Nevertheless, our results highlight the need to consider environmental flows in both global research and policies on water resources management as major anthropogenic flow alterations prevail across wide areas."*

# Report #2

## General comments

This reviewer was party to a submission of considerable detail in the first round of reviews. To their credit the authors have clearly attended to most if not all of the comments previously given, which has considerably upgraded the paper. It is thus recommended for publication with minor edits as detailed below.
A major finding of the first review was that the absence of ecological consideration in this hydrological approach was not acknowledged. This has been partially rectified in a clumsy way and deficiencies remain, some of which are indicated below.

We thank the reviewer for their comments on the revised manuscript and agree that addressing the considerably detailed comments improved the initial manuscript substantially. We hope that our further clarifications and revisions will prove sufficient.

## Specific comments

Title –"Widespread and increasing violations of environmental flow envelopes" – this heading does not convey the depth and substance of the paper particularly as the concept of flow envelopes is not immediately intuitive. At least the word "global" should be included.

We agree with the reviewer that especially omitting the word “global” from the title may undermine the value of the study. We have discussed among all authors about the title of the article, and agreed to change it to *“Globally widespread and increasing violations of environmental flow envelopes”.* As discussed in the major revision, we’d be hesitant to include anything related to ecology in the title since we haven’t done any explicit validation for the relationship between EFEs and ecosystem responses.

175 – quasi natural – while dams may not have been prevalent, canals were common enough in a few areas e.g. in England the construction of canals began in the mid-1700s. This would have impacted streamflow but only in selected areas.

We agree that minor anthropogenic modification of rivers may have already existed by 1860, however, e.g. large-scale hydropower dams and extensive irrigation schemes followed only after the industrial revolution. We have revisited this statement to be less absolute, and lines 178–182 now read as follows:

*"We defined the EFEs based on the pre-industrial period (1801–1860). While some flow alteration (e.g. canals) may have already existed by 1860, large-scale human modification of rivers has prevailed mainly after the pre-industrial period. For example, area equipped for irrigation has increased sixfold since 1900 (Siebert et al., 2015), and many of the globally*

*largest dams have been commissioned during the 20th century (Lehner et al., 2011). Therefore, we presumed that this time period is quasi-natural – i.e. near the natural flow regime."*

250 – it would be good to include at least a sentence of summary of the supplementary material.

We have added a sentence explaining the effect of changing the minimum streak length in lines 251–255:

*"In addition to results presented in the following section with a minimum three-month sequence of violations, we repeated the analysis with other minimum lengths of the violation streak. The results of this sensitivity analysis are presented in the supplementary material (Fig. S1–S3); shorter minimum violation streaks extend the violations to wider areas, and increasing the minimum violation streak limits the violations to relatively small regions."*

480 - these purely hydrological methods aim to establish the hydrological conditions that would be acceptable to ecosystems, but do not include any metric whereby this may be tested. Suggest this perspective be included.

We have added this perspective to the sentence, and lines 482–485 now read as follows:

*"Our method – and EFs in general – assumes that violating or respecting the EFE is associated with degrading or preserving riverine ecosystems. However, this might not hold for simplified hydrological EF methods as they lack metrics of assessing the correlation between presumably adequate hydrological conditions and ecosystem responses (Poff and Zimmerman, 2010; Richter, 2010; Richter et al., 2012; Mohan et al., 2021)."*

484 – "This is because the ecosystem response to altered flow regimes varies across spatial and temporal scales, as well as between different species" – this sentence does not cover the issue so suggest mentioning that altered flows affect a range of ecological characteristics from sediments, to stream morphology including riparian banks, to biodiversity and community dynamics of most fauna and flora. Any of the reviews of EF will spell this out.

Regarding the impacts of flow alteration on ecosystems, we have added more detail according to the suggestion, and a reference to Poff and Zimmerman (2010). Lines 487–489 now read as follows:

*"This is because the ecosystem response to altered flow regimes varies across spatiotemporal scales and different species due to the impact of altered flow regimes on e.g. sediment transport, stream and riparian bank morphology, and community dynamics of fauna and flora (Biggs et al., 2005; Poff and Zimmerman, 2010; Poff et al., 1997; Rolls et al., 2018)."*

487 – "Though fish make up only a part of a riverine ecosystem, these studies support incorporating water quality–related factors in a comprehensive EF definition." It is not clear why water quality is given prominence here – when in most EF studies it is the response indicators

(biota) that are most important as they are the best indicators of the success of the EF. It is also complex to include WQ in any comprehensive way in determination of EF because the presence of WQ issues may be and generally is completely non flow-related.

We agree with the reviewer that water quality is not the only factor missing from comprehensive EF determinations, but water quality was the main determinant considered in the referred studies. Our intention was to state that many factors beyond discharge could benefit the determination of EFs, and we have now revisited the paragraph to less prominently focus on water quality only. Lines 493–498 now read as follows:

*"Recently, quantitative water flows have been shown to be less important than water quality and invasive species for assessing rivers' ecological status, determining fish biodiversity, and driving fish habitat loss (Barbarossa et al., 2021; Grizzetti et al., 2017; Su et al., 2021). Though fish make up only a part of a riverine ecosystem, these findings underline that discharge alone cannot provide a comprehensive EF definition but other factors should be considered, as well. Holistic EF methods that include these factors – and also observation of biotic responses – correlate much better with ecosystem states, but require in situ data, ancillary variables, and local expert knowledge (Poff et al., 2017; Tharme, 2003) that are not available at the global scale."*

494 – floodplains - this seems a strange example, because from an ecosystem point of view you would hardly call the necessary flooding of a floodplain a violation of the environmental flows.

We agree with the reviewer that necessary flooding of a floodplain should not be considered as an environmental flow violation. However, if the flooding occurs during a period in which the natural conditions are dry (EFE upper bound violation), the ecosystem dependent on distinct dry and wet periods will degrade (Hayes et al., 2018; Junk et al., 1989; Schneider et al., 2017). We have made the mechanism behind this example more explicit, and lines 504–506 now read as follows:

*"The link between EFE upper bound violations and ecosystems exists, since, for example, floodplain ecosystems in monsoon flood pulse systems require distinct dry and wet periods, and disturbing the dry period by increased discharge may degrade the ecosystems (Hayes et al., 2018; Junk et al., 1989; Schneider et al., 2017)."*

500 – " detrimental to riverine ecosystems outside monsoon regions" – it is not clear why there is the limitation of monsoon areas. Does this imply that they are detrimental to ecosystems inside monsoon areas?

As outlined in the previous comment, studies suggest that EFE upper bound violations during the naturally dry period will indeed be detrimental to ecosystems inside monsoon areas. Here, we have complemented the previous addition by making this sentence more explicit; lines 509–511 now read as follows:

*"Hence, EFE upper bound violations are strong signals of increased upper extreme flows, although it cannot be inferred from this study whether these are detrimental to riverine ecosystems beyond regions with distinct dry and wet periods, such as the monsoon areas."*

516 – "In the future, the EFEs should be developed by complementing our global analysis with more advanced EF methods and more detailed hydrological data that better correlate with riverine ecosystem status" – this is a presumptuous way of stating this, presumptuous in that it is your model that should form the basis of future EFEs

We agree with the reviewer that the sentence was formatted poorly, and we have revisited it to be less presumptuous. Lines 531–532 now read as follows:

*"In the future, global analysis with more advanced EF methods and more detailed hydrological data that better correlate with riverine ecosystem status could further develop the EFEs."*

# References


Campbell, B., Beare, D., Bennett, E., Hall-Spencer, J., Ingram, J., Jaramillo, F., Ortiz, R., Ramankutty, N., Sayer, J., Shindell, D., 2017. Agriculture production as a major driver of the Earth system exceeding planetary boundaries. Ecology and Society 22. https://doi.org/10.5751/ES-09595-220408

Hayes, D.S., Brändle, J.M., Seliger, C., Zeiringer, B., Ferreira, T., Schmutz, S., 2018. Advancing towards functional environmental flows for temperate floodplain rivers. Science of The Total Environment 633, 1089–1104. https://doi.org/10.1016/j.scitotenv.2018.03.221

Jägermeyr, J., Pastor, A., Biemans, H., Gerten, D., 2017. Reconciling irrigated food production with environmental flows for Sustainable Development Goals implementation. Nature Communications 8, 15900. https://doi.org/10.1038/ncomms15900

Junk, W., Bayley, P., Sparks, R., 1989. The Flood Pulse Concept in River-Floodplain Systems. Canadian Special Publication of Fisheries and Aquatic Sciences 106, 110–127.

Poff, N.L., Zimmerman, J.K.H., 2010. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. Freshwater Biology 55, 194–205. https://doi.org/10.1111/j.1365-2427.2009.02272.x

Schneider, C., Flörke, M., De Stefano, L., Petersen-Perlman, J.D., 2017. Hydrological threats to riparian wetlands of international importance – a global quantitative and qualitative analysis. Hydrology and Earth System Sciences 21, 2799–2815. https://doi.org/10.5194/hess-21-2799-2017

Wada, Y., Bierkens, M.F.P., 2014. Sustainability of global water use: past reconstruction and future projections. Environ. Res. Lett. 9, 104003. https://doi.org/10.1088/1748-9326/9/10/104003