Comment 1:

The article presents an analysis of the sensitivity of root-zone storage capacity to climate variability. The article is generally clear. My main concern on this article is that I did not clearly see the novel insights provided by this study. The authors say a few times that their results corroborates past findings on the limits of the Budyko framework. Actually, I found that they should more clearly emphasize what is new in their results compared to past studies.

Reply:

We highly appreciate the positive overall assessment by Reviewer #2 and thank him/her for the detailed and insightful comments! We will make it clearer what the novelty of this work is. Briefly, the analysis is a direct follow up to the paper by Bouaziz et al. (2022). In their paper, the potential effects of an *assumed, hypothetical future change* in $S_{r,max}$ for stream flow *under a projected future climate* where explored. The novelty here is that we do not assume but instead actually *quantify for the first time past changes* of $S_{r,max}$ based on historical observations and use them to explore the effects thereof on modelled stream flow in the past. In comparison to Bouaziz et al. (2022), whose analysis remained hypothetical, these two new points allow to base the analysis of $S_{r,max}$ changes on *real world observations* and to *evaluate the modelled effects* of changing $S_{r,max}$ on modelled stream flow *against observed stream flow*.

We will clarify that in the revised manuscript.

Comment 2:

L67-69: What about groundwater exchanges? Should not they be considered also?

Reply:

We agree. Ideally, groundwater exchange fluxes should be considered. However and unfortunately, it remains problematic if not at all impossible to meaningfully quantify these fluxes with current observation technology. This is, however, a limitation of the *vast majority* of hydrological studies and not limited to our analysis (e.g. Condon et al., 2020). To keep the uncertainties introduced by potential groundwater exchange fluxes as low as possible, we have therefore excluded all catchments that show clear evidence of a water balance deficit or surplus, which both indicate groundwater exchange (e.g. Bouaziz et al., 2018), from our analysis. This was done by discarding all catchments that plotted (1) above the upper limit of E_A/P ($E_A/P > E_P/P$ and $E_A/P > 1$) and thus outside the physically plausible realm in the Budyko framework (e.g. Fig.3), thereby indicating groundwater export and (2) more than 0.25 below the analytical Budyko solution, indicating potential groundwater import.

Comment 3:

Please explain in simple words what low or high values of omega mean in terms of water balance type. Say omega should be positive.

Reply:

The parameter ω in the parametric Tixeront-Fu equation (Eq.1 in the original manuscript) can range between 1 and ∞ (e.g. Zhang et al., 2001; Greve et al., 2015; Andreassian et al., 2016). The lower the ω of a catchment the lower E_A/P of this catchment (and inversely, the higher Q/P). We will add this in the revised manuscript.

Comment 4:

L141: At this stage of the article, it is unclear why this specific PE formula was used. Maybe this could be justified in a few words (even if some comments are later provided in the discussion section).

Reply:

Agreed. We will add an explanation in the Methods section.

Comment 5:

L151-154: I do not understand why apparently "leaking" catchments where excluded from the analysis. There are probably many catchments in the remaining dataset, which exhibit groundwater exports even if there are within the limits of the graph. The limit IE>IA is arbitrary and does not really corresponds to underlying processes (groundwater exports or not). Would the analysis be very different if all the catchment had been kept to conduct the analysis? Does this catchment selection partly explain the apparently very consistent behaviour of UK and US catchments?

Reply:

Apparently leaking catchments were excluded as the estimates of E_A in these catchments cannot be distinguished from groundwater export with the available data in the water balance, i.e. $P - Q = E_A - Q_{GW,export}$. This may, as correctly remarked by the reviewer, of course also affect catchments within the limits. However, as with the available data there is no meaningful way to quantify groundwater export (or import) otherwise, the best that we can do is to at least exclude those catchments that plot outside the *physically possible* realm and for which we actually *know* that groundwater export plays a role. As the number of catchments that plot outside the $I_E > I_A$ limit remains very low(< 1%), the effect on the overall results of analysis are very minor. As such it can also not explain the consistent behaviour of the UK and US catchments. We will discuss this in some more detail in the revised version of the manuscript.

Comment 6:

Section 4: I felt a bit confused in reading the results section. In the first part of the analysis, the Meuse basin seems to be somehow outlier in its behaviour compared to the UK and US datasets, but no clear explanations on this behaviour is found. Therefore the added value of this small catchment sub-set in this part of the analysis is unclear. In the second part where the process-based model is applied, only the Meuse basin is used. Though I understand it is difficult to apply such a model on large datasets, I found it makes the study less clear and the overall conclusions more difficult to draw.

Reply:

We appreciate that the reviewer points out that our choices are not completely clear. Briefly, we have included the catchments of the Meuse basin for two reasons: (1) as described in our reply to *Comment 1* above, this analysis is a direct follow up to the work of Bouaziz et al. (2022), who investigated the effects of future E_A (using $I_E = E_A/P$) on $S_{r,max}$ and the associated stream flow. Their study was based on *projections of future climate* and the *assumption* that catchments largely remain on their parametric Budyko curve defined by a constant parameter ω. They explored future effects and could therefore in their study not test whether these assumptions hold. As the Bouaziz et al. (2022) analysis was executed in the catchments of the Meuse basin, we here decided to zoom in on these same catchments to allow a direct comparison with that previous study. (2) Future estimates of E_A, and thus by extension also future estimates of $S_{r,max}$, depend on the deviation of catchments from their specific parametric Budyko curve, i.e. ΔI_E over time. To our knowledge, there has so far been no systematic analysis to quantify these distributions of deviations ΔI_E over time. In other words, we do not know to which extent the assumption of Bouaziz et al. (2022) that catchments move along a specific curve over time actually holds. We have therefore decided to not only quantify these distributions for the Meuse basin but to provide a wider context of which magnitudes of deviations need to be expected over a larger sample of catchments. The fact that the Meuse catchments showed a rather pronounced change in I_E, as visible by the skewed distribution of ΔI_E, while the samples of catchments elsewhere were characterized by much lower changes, on average, gave us the opportunity to quantify the effects of both, extreme and average changes, on $S_{r,max}$ and thus on predicted stream flow in the Meuse basin. Please note, that here we did not seek to identify the reasons for the pronounced ΔI_E in the Meuse basin. This was previously explored by Fenicia et al. (2009), who found that changes in land use management are the most likely reason for the observed pattern. Instead, our objective was to explicitly quantify the change and to analyse which knock-on effects it has on $S_{r,max}$ and eventually on stream flow predictions. We will make this clearer in the revised version of the manuscript.

Comment 7:

Section 4.1: I found figures 9 to 11 not very useful. Maybe the authors could say what were their results without showing figures (or putting them in SM).

Reply:

We agree. We will condense some of these figures and show them in the Supplementary Material.

Comment 8:

Sections 4.3.2 and 4.3.3: I found these parts of the article tedious to read. A lot of detailed information is given, which makes it difficult to draw the overall picture.

Reply:

Thank you for this observation. We will rework these sections and make them more concise and readable.

Comment 9:

Conclusion: see main comment above

Reply:

Agreed, we will make it clearer in the Conclusion what the novel findings of our analysis are

Comment 10:

14: I found this figure not very useful (overall I found there are too many figures in this article and their number could probably be reduced).

Reply:

We agree. We will move this figure to the Supplementary Material.

Comment 11:

The reference Hulsman et al. (2021) is missing in the list of references.

Reply:

Thank you for pointing this out. We will add the reference.

Comment 12:

UK is used in notations, but GB is used for the CAMELS dataset. Maybe use a single abbreviation to be consistent.

Reply:

We absolutely agree. This will be corrected.

References:

Bouaziz, L. J., Aalbers, E. E., Weerts, A. H., Hegnauer, M., Buiteveld, H., Lammersen, R., ... & Hrachowitz, M. (2022). Ecosystem adaptation to climate change: the sensitivity of hydrological predictions to time-dynamic model parameters. Hydrology and Earth System Sciences, 26(5), 1295-1318.

Fenicia, F., Savenije, H. H. G., & Avdeeva, Y. (2009). Anomaly in the rainfall-runoff behaviour of the Meuse catchment. Climate, land-use, or land-use management?. Hydrology and Earth System Sciences, 13(9), 1727-1737.