Drought decreases streamflow response to precipitation especially in arid regions

Reviewer 2

The authors study the changing streamflow dynamics of catchments in response to changing precipitation. Disentangling the influence of climatic/weather and catchment properties on the streamflow response is rather complicated. This is regularly found in the efforts to explain the variability of streamflow elasticity or flow duration curve properties between catchments. Both typically end up with a mixture of explanatory variables. The current manuscript is well written and technical solid – as far as I can tell. However, like reviewer 1, I have some questions regarding definitions and the robustness of the results, given those definitions. I also have some questions regarding correlation versus causation.

- We sincerely thank the reviewer for their comments and are very pleased that they found the manuscript well written and appreciate our efforts in developing a solid method. The reviewer provides constructive feedback and suggestions, which we will address in the revised manuscript. Below, we summarize the changes we will make in response to these comments. Our responses are shown in blue, the revised text is shown in *italics*, and line numbers mentioned in this response refer to the current version of the manuscript and they are indicated within brackets [xx].

[1] As the other reviewer also states one relevant question is the definition of "streamflow sensitivity to precipitation". Is the streamflow sensitivity to precipitation well defined when we use the ratio between annual streamflow and precipitation? The mathematical definition of sensitivity is the change in the output due to variations in the input of a system or model. Plots of Q versus P do not quite capture this definition because they only look at the response of the system to the input, without consideration of what state the system was in. The latter is capture by the idea of streamflow elasticity where one quantifies the change in streamflow due to the change in precipitation (or something else) from year to year. Elasticity is notoriously difficult to explain (or regionalize) while the Q-P relationship is often rather stable. Given that this type of analysis was part of Anderson et al. (2024, HESS, "Elasticity curves describe streamflow sensitivity to precipitation across the entire flow distribution"), I think it would be very good if the authors were to make the connection and discuss how their definition in this manuscript differs from previous work (incl. some of the authors) and what consequences the changing definition has.

- The reviewer is correct in pointing out that the ratio between streamflow and precipitation does not provide information about the sensitivity of streamflow to precipitation but rather information on the yearly streamflow response to precipitation (annual catchment response to precipitation). Therefore the use of the terminology: "streamflow sensitivity to precipitation" is misleading and we will change it to: "streamflow response to precipitation". Further, we agree about the need to better define what the yearly ratio between streamflow and precipitation represent and to compare it to other metrics used in the literature (e.g. elasticity). The added text will be for instance:

> [74] Here, we analysed the temporal dynamics of the *annual* streamflow *response* to precipitation (computed as the ratio between annual streamflow and precipitation) in approximately 5000 catchments across the world. *The annual Q-P ratio indicates the fraction of precipitation that is converted into streamflow, providing insights into the catchment's water balance.*
>
> [109] We then computed yearly streamflow-to-precipitation (Q-P) ratio timeseries for each catchment. *This measure represents the annual runoff ratio and is dynamically influenced by climatic and hydrological conditions. By considering an annual timescale, the ratio inherently accounts for evapotranspiration and storage processes within the catchment. However, it is important to note that, first, since the ratio is a lumped representation of these processes, it does not separate individual contributions. Second, in some catchments, storage processes extend beyond a single year, which may influence the annual runoff ratio. This metric differs from other metrics such as elasticity (Anderson et al., 2023; Sankarasubramanian et al., 2001; Zhang et al., 2022). While the annual runoff ratio provides an average measure of how much precipitation contributes to streamflow in a given year, elasticity tells us how streamflow reacts to changes in precipitation (Schaake, 1990).*

[2] My second larger point is about the difference between correlation and causation. The authors work here, and many of the papers cited, use correlation to infer causation. While I fully agree with the type of analysis, I think that it would be good to at least discuss somewhat whether correlation can be used here to infer causation. This also includes the discussion use of some of the references. One example is in lines 487ff. where the authors state: "This decrease could be explained by reduced connectivity among bare patches (Urgeghe et al. 2010)". The Urgeghe et al. study runs a model for a design storm and varies vegetation patches to show their role for runoff behaviour during the design storm. I find it quite a stretch to use this reference in support of long-term catchment water balance behaviour. The authors need to at least explain why they think this connection is valid. The second part of the sentence in lines 487ff. is "and increased soil evaporation due to an increase in solar radiation reaching the ground (Guardiola-Claramonte et al. 2011)." Isn't the latter than coinciding with reduced transpiration? Is the reduction in transpiration not larger than the added soil evaporation (given the deeper capture of moisture through roots)?

This is just an example where I think where the authors could expand their discussion and argument. I just given an example, reflective of the wider discussion section. It would in general be good if the authors were a bit more explicit why the references cited are transferrable to their situation.

- We thank the reviewer for this comment and fully agree on the need for a more careful use of references and caution in inferring causation from correlation. In the discussion section, our intention was to assess whether our results align with findings from other studies and to explore how these findings have been explained in terms of underlying processes. Upon revisiting the references, we acknowledge that the use of Urgeghe et al. (2010) may indeed be an overextension in this context. As well as the use of Garreaud et al., 2017, so we will delete it.

Additionally, upon further review of the literature, we recognize that drought in arid regions does not always lead to a decrease in hydrological connectivity. For example, (Ruddell & Kumar, 2009) highlight cases where connectivity decreases, whereas other studies (Goodwell et al., 2018; Liu et al., 2024) document increases in connectivity under certain conditions. These differences depend on factors such as the type, timing, and duration of drought, as well as vegetation type and other catchment characteristics, making generalizations difficult.
We also agree with the reviewer that negative NDVI anomalies typically reflect reduced transpiration compared to the system's baseline (Johnson et al., 2009). However, variations in transpiration and soil evaporation during negative NDVI differ across systems (Dijke et al., 2019; Lawrence et al., 2007).

We propose to revise the discussion as follows:

[485] Spatial differences can also be found in the influence of negative NDVI anomalies on the Q-P relationship, *though the overall influence remains small (less than 5%)*. While the response of the Q-P relationship generally increases during negative NDVI anomalies, in arid and semi-arid catchments, this response *slightly* decreases (Figure 2b). This decrease could *partially* be explained by reduced *hydrological* connectivity among bare patches *(Jaeger et al., 2014)* and increased soil evaporation (Guardiola-Claramonte et al. 2011). *However, these processes are highly dependent on the type, timing, and duration of drought, as well as catchment-specific characteristics (Goodwell et al., 2018; Liu et al., 2024), making generalizations challenging. Furthermore, we acknowledge that reduced transpiration, typically associated with negative NDVI anomalies, may also take place (Johnson et al., 2009). The interplay between these processes likely drives the observed variability, underscoring the need for caution when interpreting these results.*

[3] Influence of length of about 30 time steps on the robustness of the stationarity test? It seems like a very short time span for such tests. And given the widely discussed limitations of using statistical significance for justification. This is not a criticism of the approach, but a question of how one can assure robustness of the results?

- We agree with the reviewer that the robustness of the ADF test can be influenced by the length of the time series. With a short time series (~30 points), the ADF test is more likely to misclassify a stationary series as non-stationary. The test might indeed not find enough evidence to reject the null hypothesis (H0: "the series is non-stationary").

In our study, the identified non-stationary time series are used to investigate possible shifts in streamflow response to precipitation, transitioning from one steady state to another. To ensure the robustness of our results and confirm that these time series present a significant shift and, hence are truly non-stationary, we performed in our study additional tests (see section 2.4 of the manuscript). Indeed, we examined whether these time series exhibit a linear, curvilinear, or abrupt (characterised by sudden changes) trend. Abrupt changes were specifically tested using a threshold regression approach (see lines 256–267). The best-fitting trend was then identified using the Akaike Information Criterion (AIC). Further, to account for potential uncertainty due to short time series and data noise, we bootstrapped each time series 100 times without replacement and compared the model results across iterations (lines 268–276). Finally, to further increase confidence in detecting step changes, we applied a series of restrictive criteria (see lines 268–299). Through these additional steps, we increased robustness of our results and hence that the time series used in our step-change analysis are non-stationary, as they exhibit a significant shift in their structure.

[4] Temporal connectivity of drought events? Is there relevance to the temporal sequence of drought periods for this analysis. Even though I appreciate that the short time series might make this difficult to study.

- As the reviewer suggests, we could not fully account for the temporal connectivity of drought events due to the relatively short time series used in the analysis. However, we partially explored temporal connectivity by analysing the influence of droughts occurring in the preceding year. These influences were incorporated into our analysis through Equation 1. We wrote a reflection on this in the discussion:

  > [568] *Although drought is a continuum, with temporal connectivity between events* (Van Loon et al., 2024)*, our analysis treats droughts as independent events, summarizing their characteristics at a yearly scale to facilitate comparison with the yearly ratio of Q to P. We only partially accounted for drought connectivity by incorporating drought characteristics from the preceding year into our analysis. However, their influence was minimal (less than 5%), with meteorological drought showing a slightly higher influence compared to other drought types.*

[5] How relevant are conclusions that show differences of 2-3%? This should be quite below the amount of uncertainty one would expect in precipitation and/or streamflow observations even in good circumstances.

- If the reviewer is referring to cases where our results differ by 2–3%, such as the 30% and 27% influence of hydrological and soil moisture drought on the Q-P ratio (lines 344–345), we agree that this difference is minimal and likely irrelevant given the inherent uncertainties in precipitation and streamflow observations. We will explicitly acknowledge this in the limitations of our study.

[572] *The accuracy of the percentage values representing the influence of a certain drought type on the yearly Q-P ratio is affected by uncertainties in precipitation and streamflow observations. Although these percentage values are not exact due to observational uncertainties, the relative magnitudes provides meaningful information, allowing us to identify which drought types have the strongest influence on the Q-P ratio.*

If, instead, the reviewer is questioning the relevance of drought types with a low influence (~2–3%) on the Q-P ratio, such as NDVI anomalies (line 309), we argue that the broader finding remains valid despite observational uncertainties. A relatively small influence suggests that this specific drought type has minimal impact on catchment response. In contrast to the 20–30% changes observed for other drought types, this lower effect may indicate that these catchments are more resilient to changes associated, for instance, with NDVI anomalies.

References

Anderson, B. J., Brunner, M. I., Slater, L. J., & Dadson, S. J. (2023). *Elasticity curves describe streamflow sensitivity to precipitation across the entire flow distribution*. https://doi.org/10.5194/hess-2022-407

Dijke, A., Mallick, K., Teuling, A., Schlerf, M., Machwitz, M., Hassler, S., Blume, T., & Herold, M. (2019). Does the Normalized Difference Vegetation Index explain spatial and temporal variability in sap velocity in temperate forest ecosystems? *Hydrology and Earth System Sciences*, *23*(4). https://doi.org/10.5194/hess-23-2077-2019

Garreaud, R. D., Alvarez-Garreton, C., Barichivich, J., Pablo Boisier, J., Christie, D., Galleguillos, M., LeQuesne, C., McPhee, J., & Zambrano-Bigiarini, M. (2017). The 2010-2015 megadrought in central Chile: Impacts on regional hydroclimate and vegetation. *Hydrology and Earth System Sciences*, *21*(12). https://doi.org/10.5194/hess-21-6307-2017

Goodwell, A. E., Kumar, P., Fellows, A. W., & Flerchinger, G. N. (2018). Dynamic process connectivity explains ecohydrologic responses to rainfall pulses and drought. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(37). https://doi.org/10.1073/pnas.1800236115

Jaeger, K. L., Olden, J. D., & Pelland, N. A. (2014). Climate change poised to threaten hydrologic connectivity and endemic fishes in dryland streams. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(38). https://doi.org/10.1073/pnas.1320890111

Johnson, G. L., Sinclair, T. R., & Kenworthy, K. (2009). Transpiration and normalized difference vegetation index response of seashore paspalum to soil drying. *HortScience*, *44*(7). https://doi.org/10.21273/hortsci.44.7.2046

Lawrence, D. M., Thornton, P. E., Oleson, K. W., & Bonan, G. B. (2007). The partitioning of evapotranspiration into transpiration, soil evaporation, and canopy evaporation in a GCM: Impacts on land-atmosphere interaction. *Journal of Hydrometeorology*, *8*(4). https://doi.org/10.1175/JHM596.1

Liu, C., Chen, Y., Huang, W., Fang, G., Li, Z., Zhu, C., & Liu, Y. (2024). Climate warming positively affects hydrological connectivity of typical inland river in arid Central Asia. *Npj Climate and Atmospheric Science*, *7*(1). https://doi.org/10.1038/s41612-024-00800-4

Ruddell, B. L., & Kumar, P. (2009). Ecohydrologic process networks: 1. Identification. *Water Resources Research*, *45*(3). https://doi.org/10.1029/2008WR007279

Sankarasubramanian, A., Vogel, R. M., & Limbrunner, J. F. (2001). Climate elasticity of streamflow in the United States. *Water Resources Research*, *37*(6). https://doi.org/10.1029/2000WR900330

Schaake, J. C. (1990). From Climate to Flow. In *Climate Change and U.S. Water Resources*.

Van Loon, A. F., Kchouk, S., Matanó, A., Tootoonchi, F., Alvarez-Garreton, C., Hassaballah, K. E. A., Wu, M., Wens, M. L. K., Shyrokaya, A., Ridolfi, E., Biella, R., Nagavciuc, V., Barendrecht, M. H., Bastos, A., Cavalcante, L., de Vries, F. T., Garcia, M., Mård, J., Streefkerk, I. N., … Werner, M. (2024). *Review article: Drought as a continuum: memory effects in interlinked hydrological, ecological, and social systems*. https://doi.org/10.5194/egusphere-2024-421

Zhang, Y., Viglione, A., & Blöschl, G. (2022). Temporal Scaling of Streamflow Elasticity to Precipitation: A Global Analysis. *Water Resources Research*, *58*(1). https://doi.org/10.1029/2021WR030601