Dear Editor,

We thank you for your time in reviewing our revised manuscript and your further comments on it. We have now implemented all your suggestions to improve the clarity of the manuscript. Furthermore, we have analysed the influence of the correction procedure for precipitation, in order to rely less on previous studies and provide evidence from our data. Please find more details below in our point-by-point replies to your comments (italic) and the changes we did in the manuscript to address them (underlined). We further attach a revised version of the manuscript, with tracked changes as well.

Best regards,

Giulia Bruno and co-authors

Technical Comments and Clarification for Improvement:

1- Clarification on PDJF and Summer Low Flows (Referee #1, Question 4):

In the author's response to Referee #1 regarding Question 4, further clarification is needed on how increases in PDJF (precipitation during December–February) could have indirectly contributed to decreases in summer low flows in the Eastern cluster. Specifically, how can increased precipitation potentially result in lower streamflow during summer? This seems counterintuitive and should be better explained or corrected.

We argued that increases in winter precipitation (P_{DJF}) might have indirectly contributed to decreases in summer low flows in the eastern cluster via feedbacks with evapotranspiration (E). Specifically, increases in P_{DJF} might have supported increases in storage (S) recharge in winter and thus in S levels at the beginning of the growing season. In turn, increases in S levels at the beginning of the growing season potentially contributed to increased vegetation growth and therefore E, by further contributing to reduced summer low flows. While a rigorous attribution of increases in E to their drivers would require additional data (e.g., on vegetation gorwth) and we acknowledge that this negative correlation may also be spurious, we rephrased L355–357 to make our line of reasoning clearer.

L355–357: <u>Increases in winter storage recharge may also promote increases in vegetation growth and E, and thus decreases in summer low flows, as the negative correlation between trends in P_{DJF} and in summer low flows in the eastern cluster may suggest (Fig. 7), even though this relationship may be spurious.</u>

2- Trend Analysis in Abstract and Main Text (L20–21):

The statement: "Summer low flows decreased (increased) significantly in 31% (2%) of the catchments, with a median trend of -3.7% per decade across all catchments." needs further examination. If 67% of the catchments showed no significant trend, what is the rationale for computing the median trend across all catchments, including those without significant trends? Additionally, it would be more informative to report the median trend separately for the 31% of catchments showing decreasing trends and the 2% with increasing trends. This would help to clarify the magnitude and direction of significant changes.

We decided to compute trend statistics on all catchments, and not on catchments with significant trends only, because we believe that trend slopes can be informative of the general tendency to the increase or decrease even if statistically non-significant at a predefined level and for a specific trend test. In the case of summer low flows, trend slopes are negative for most catchments with a median trend of -3.7 % decade^{-1} and an interquartile range of -7.5/-0.6 % decade^{-1}, as reported in the manuscript, which points to a general tendency to the decrease of summer low flows in the study region between 1970 and 2019 (Fig. 4c, white boxplot). We acknowledge that this was not easy to grasp from the Abstract alone, which we rephrased as reported below. Furthermore, we agree that statistics on significant trends only may provide a more comprehensive picture of the observed changes in the region and therefore, we added them in a table in the supplement (Table 1 here, new Table S1 in the supplement), which we now refer to in the main text as reported below.

Table 1: Trends in summer low flows (7d$Q_{min,\ JJA}$) and their predictors (annual evapotranspiration, E, precipitation over summer, P_{JJA}, spring, P_{MAM}, and winter P_{DJF}) over 1970–2019.

Variable	Catchments with positive trends [%]	Catchments with negative trends [%]	Median trend across all catchments [% decade^{-1}]	Catchments with significant positive trends [%]	Median trend across catchments with significant positive trends [% decade^{-1}]	Catchments with significant negative trends [%]	Median trend across catchments with significant negative trends [% decade^{-1}]
7d$Q_{min,\ JJA}$	23	77	-3.7	2	7.4	31	-9.7
E	66	34	1.1	27	4.6	5	-4
P_{JJA}	52	48	0.2	0	-	0	-
P_{MAM}	34	66	-0.6	1	4.8	3	-6.9
P_{DJF}	90	10	3.6	10	6.9	0	-

Abstract, L20–22: Summer low flows generally showed a decreasing tendency (median trend of -3.7 % decade^{-1} and interquartile range of -7.5/-0.6 % decade^{-1} across all catchments), significant negative trends in 31 % of the catchments, and significant positive trends in 2 % of them only.

L229–232: Trends in 7d$Q_{min,\ JJA}$ were significantly negative in 31 % of the catchments and positive in 2 % of them only (Fig. 4b, Table S1). Across all catchments, trends in summer low flows had a median (interquartile range, IQR) of -3.7 (-7.5/-0.6) % decade^{-1} (Fig. 4c). Focusing on catchments with significant negative trends only, a median decrease of -9.7 % decade^{-1} was observed (Table S1).

L318–320: Summer low flows showed a median trend of -3.7 % decade^{-1} (IQR of -7.5/-0.6 % decade^{-1}) across all catchments, with significant negative trends in 31 % and significant positive trends in 2 % of the catchments (Fig. 4b and c).

3- Ambiguity in Trend Percentages (L225–227):

The sentence: "Trends in 7dQmin, JJA were significantly negative in 31% of the catchments and significantly positive in 2% of them (negative in 77% and positive in 23%, Fig. 4b), with a median (interquartile range, IQR) of -3.7 (-7.5/-0.6)% per decade across all catchments (Fig. 4c)." is confusing. It is unclear how the 31% and 2% (which refer to statistically significant trends) relate to the 77% and 23% figures (which seem to include non-significant trends). Please clarify the distinction between statistically significant trends and overall direction of trends, and how these percentages are used in analysis and interpretation.

Trends in summer low flows were negative in 77 % of the catchments and positive in the remaining 23 %. Moreover, trends were significantly negative in 31 % of the catchments and significantly positive in only 2 % of them. As better explained above, we used both the trend slopes and their significance for the interpretation of our results, since we believe they provide complementary information. While the magnitude of trend slopes across all catchments provides an indication of the general tendency to the decrease/increase, the significance of these slopes provides information regarding the strength of these changes. We see, however, that reporting both percentages in the same sentence may hamper its readability and thus, we decided to move the percentages of positive and negative trends in the new Table S1 (see reply to comment #2).

L229–230: Trends in 7d$Q_{min, JJA}$ were significantly negative in 31 % of the catchments and positive in 2 % of them only (Fig. 4b, Table S1).

4- Over-Reliance on Previous Studies:

While referencing previous literature is important for contextualization, several of the author's responses rely heavily on prior studies without providing sufficient evidence from the current analysis. It is recommended to strengthen the manuscript by presenting clearer links between your results and your arguments, supported by direct evidence from your own data.

We would like to point out that we performed a number of new analyses of our dataset for our first round of responses, rather than mainly referring to previous studies as this comment seems to suggest to us. Specifically, we added (i) the catchment-scale attribution of trends in summer low flows through multiple linear regression for each catchment (Referee #1, Comment #3), (ii) the analysis of trends in annual streamflow and its timing (Referee #2, Comment #1), and (iii) a formal verification of the applicability of the method we used to estimate the dynamic storage of the catchments (Referee #2, Comment #2). For all these analyses, we introduced methodological details and subsequent results in the manuscript or the supplementary material. By carefully going through our previous replies, we furthermore noticed that we referred to previous studies in the replies to:

1. Referee #1, Comment #1, to support our assumption of little influence of the *P* correction procedure on our results;
2. Referee #1, Comment #2 and Referee #2, Comment #5, to support our choice of using *P* in the preceding seasons as proxy for *S* conditions;
3. Referee #2, Comment #3 and Referee #2, Comment #7, to refer to previous works which already investigated what suggested by the Reviewer;
4. Referee #2, Comment #4, to support our choice of using the water-balance approach for *E*.

We do not list here replies where we used previous studies as references for our methods (e.g., Referee #2, Comment #1) or as additional evidence for our findings/argumentations (e.g., Referee #1, Comment #4 and Referee #2, Comment #1), since we believe this is not what you are pointing to here. Among the above-listed instances, we see only the first one as an assumption, based on previous findings rather than new data analyses, whereas we see the others as appropriate references to works using similar methods (points 2 and 4) or investigating partly-related topics (point 3).

Therefore, we have now formally tested our initial speculation regarding point 1. Specifically, we repeated all our analyses with uncorrected *P* for gauge undercatch. Using uncorrected *P* results in a median trend in *P* equal to 0.4 % decade^{-1} and in *E* equal to 1.2 % decade^{-1} across all catchments (Fig. 1 here), as compared to 0.3 % decade^{-1} and 1.1 % decade^{-1} respectively when using corrected *P*. Importantly, these differences in trend magnitudes do not affect the attribution of trends in summer low flows to their main predictors, with summer *P* (P_{JJA}) still the dominant predictor for the pre-Alpine cluster and *E* as a significant predictor for the eastern and northern clusters, with P_{JJA} in the latter case, according to both the analyses that we performed to this end (Fig. 2 and 3 here). Also the results regarding the changes in the *P*-*Q* relationship during the multi-year drought between 1989 and 1993 still hold when using uncorrected *P*, with the multi-year drought detected for 14 % of catchments and changes in 33 % of these, as compared to the 15 % and 26 % obtained with corrected *P*, and a median trend in *E* for catchments with changes equal to 5.9 % decade^{-1}, as compared to the 6.1 % decade^{-1} for corrected *P* (Fig. 4 here). Thus, our conclusions do not depend on the use of the correction procedure. We added this analysis in the revised version of the manuscript, by introducing it in the Methods, Results, and Discussion as follows, and reporting Fig. 1 and 4 here as new Fig. S6 and S10 in the supplement.

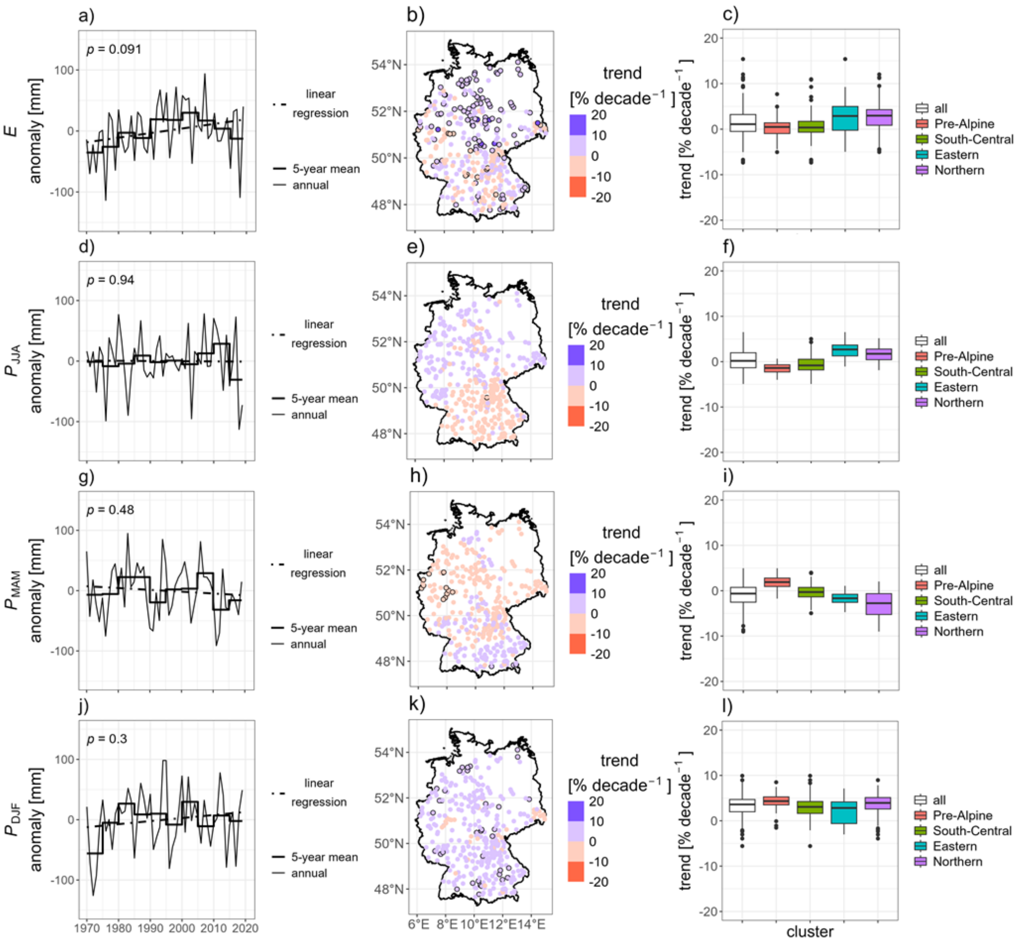

Fig. 1: Long-term variations in predictors of variations in summer low flows over 1970–2019 (annual evapotranspiration, *E*, panels a–c, precipitation over summer, P_{JJA}, panels d–f, spring, P_{MAM}, panels g–i, and winter P_{DJF}, panels j–l) from precipitation not corrected for gauge undercatch (Sect. 2.2). (a, d, g, and j) Average anomalies across the catchments. (b, e, h, and k) Maps of catchment-scale trends (black edges if significant). (c, f, i, and l) Boxplots of trends for all catchments and by cluster.

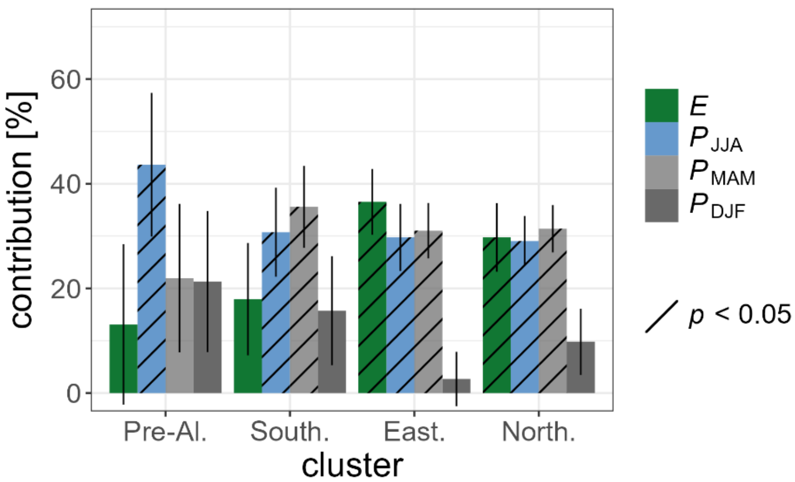

Fig. 2: Attribution of long-term variations in summer low flows to their predictors (contribution to temporal dynamics) from precipitation not corrected for gauge undercatch (Sect. 2.2): relative contribution of annual evapotranspiration (*E*), summer

(P_{JJA}), spring (P_{MAM}), and winter precipitation (P_{DJF}) to the predicted long-term dynamics of summer low flows ($7dQ_{min, JJA}$) from multiple linear regression (Sect. 2.5) for the different clusters. Pre-al. refers to pre-Alpine, South. to south-central, East. to eastern, and North. to northern cluster. Vertical lines indicate the uncertainty of the regression coefficients.

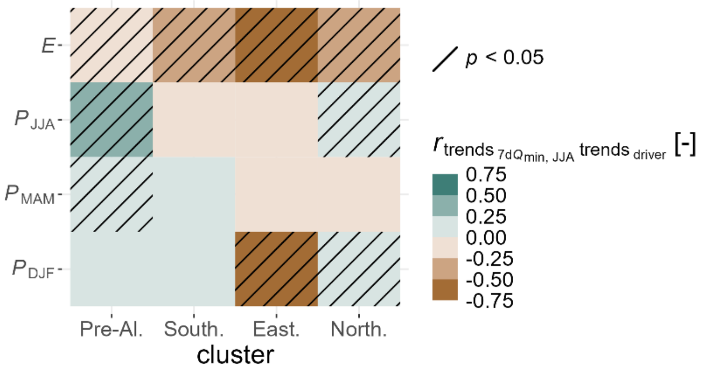

Fig. 3: Attribution of long-term variations in summer low flows to their predictors (strength of spatial coherence) from precipitation not corrected for gauge undercatch (Sect. 2.2): Pearson's correlation coefficients (*r*) between catchment-scale trends in summer low flows ($7dQ_{min, JJA}$) and in potential predictors (annual evapotranspiration, *E*, summer precipitation P_{JJA}, spring precipitation P_{MAM}, and winter precipitation P_{DJF}) over 1970–2019, for the catchments in the different clusters. Pre-al. refers to pre-Alpine, South. to south-central, East. to eastern and North. to northern cluster.

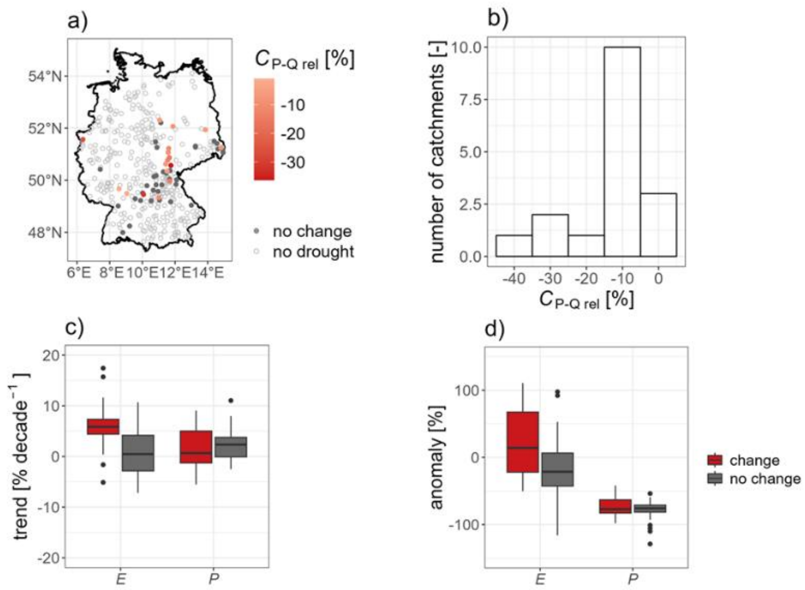

Fig. 4: Changes in the annual relationship between precipitation (*P*) and streamflow (*Q*, *P-Q* relationship) during the multi-year drought between 1989 and 1993, and their potential predictors, from *P* not corrected for gauge undercatch (Sect. 2.2). (a) Map of the magnitude of changes in the *P-Q* relationship ($C_{P\text{-}Q\ rel}$), across the study catchments. (b) Histogram of $C_{P\text{-}Q\ rel}$. (c) Boxplots of trends in annual catchment actual evapotranspiration (*E*) and *P* over 1970–1993, for catchments with change and no change in the *P-Q* relationship. (d) Boxplots of mean anomalies in *E* and *P* over the drought, for catchments with change and no change in the *P-Q* relationship.

Methods, L145–147: We then computed area-weighted catchment average time series from P_{corr} (*P* in the following) and P_{uncorr} fields for each catchment. We repeated all analyses (Fig. 1) for both *P* and P_{uncorr} to verify that the use of the correction procedure does not affect out results.

Results, L255–256: Similar trends, for both *E* and seasonal *P*, were obtained when using P_{uncorr} (Fig. S6).

L295–296: When using alternative low-flow metrics ($30dQ_{min, JJA}$ and $7dQ_{min, M\text{-}O}$) instead of $7dQ_{min, JJA}$ or P_{uncorr} instead of *P*, results were comparable to those presented here for both attribution analyses (not shown).

L308–309: The use of P_{uncorr} led to comparable results (Fig. S10).

Discussion, L434–435: This correction procedure may lead to further uncertainties, but we achieved comparable results for all analyses for both P and P_{uncorr} (Sect. 3.2 and 3.3), meaning that our conclusions do not depend on the use of the correction procedure.

5- Language and Grammar Corrections:

Please revise the manuscript for language accuracy. For example, in Section 4.4: "Due to the generally coarse resolution of satellite-derived E estimates, we computed E from observed P and Q, and estimates of the Sdyn of the catchments from Q data as a first order approximation of S." This sentence could be rewritten for clarity and correctness. A suggestion: "Given the generally coarse resolution of satellite-derived E estimates, we derived E using observed P, Q, and estimates catchment dynamic storage (Sdyn) from Q data, as a first-order approximation of S."

Thank you for this valuable suggestion on how to improve the clarity of our writing. We had another careful round of proofreading for language polishing and we refer to the tracked-changes version of the manuscript for the edits we implemented to this end.