

The manuscript has been substantially improved through the successive rounds of revision. In the current version, the modelling chain, the SURFEX–AquiFR coupling, the assumptions related to anthropogenic factors, and the link between climate forcing, groundwater recharge, and groundwater-level response are described more clearly. The additional water-balance results provided in the Supplementary Material also help strengthen the process-based interpretation of the projected groundwater responses. Overall, the main research framework and the presentation of the results are now reasonably complete. However, after carefully reading the latest version, I believe that several issues still require further clarification or verification. Some concern the interpretation of uncertainty sources and the implementation of the statistical analyses, which may affect the robustness of the corresponding conclusions. Others relate to the definition of the modelling framework, the SPLI analysis, and the consistency of quantitative information between the main text and the Supplementary Material. I therefore recommend that the authors carefully address the following issues before the manuscript is finalized.

1. The manuscript acknowledges that a proper evaluation of the contribution from the hydrogeological component would require a multi-model approach, which is not available for most of the AquiFR domain. However, it subsequently concludes that the contribution of the hydrogeological components to the total uncertainty remains negligible compared with that from the climate components. I may suggest authors stating more explicitly that hydrogeological-model uncertainty is not explicitly quantified in this study and moderating the corresponding conclusion. In addition, the subsequent comparison of the CDFs mainly evaluates the influence of the climate projections on the historical reference distributions used to construct the SPLI, rather than directly evaluating uncertainty arising from the hydrogeological models. The scope of this analysis should therefore be clarified accordingly.
2. In Section 2.3.2, the Mann–Kendall test is currently cited as Mann and Whitney (1947), which might be a wrong citation. Please check it. More importantly, detailed information is needed to verify which statistical test was actually implemented in the analysis and clearly specify the input time series and temporal resolution used for both

tests, whether the tests were applied to the original time series or to the 5-year running median. If necessary, the significance results reported in Appendix C should be reassessed accordingly.

3. The Introduction states that Aquifer “is not based on a multi-model approach” because each hydrogeological modelling software package is generally applied to a specific catchment and very few simulated areas overlap spatially. In contrast, Section 2.2 describes Aquifer as “a hydro-meteorological modelling platform based on a multi-model approach.” It is recommended to clearly distinguish between a platform that integrates several regional hydrogeological model configurations and a hydrogeological multi-model ensemble that can be used to assess structural model uncertainty, and then use the terminology consistently throughout the manuscript.

4. A systematic cross-check of the quantitative information in the main text and Supplementary Material is recommended. For example, the model-area range reported in Section 2.2.3 is inconsistent with the values in Table A2; the spatial resolutions of the GCMs and RCMs are expressed in km² rather than km; and the third future horizon in Tables C1–C3 appears to contain an incorrect time-period label. In addition, the origin of the “6.3%” value reported for the green storyline in Section 3.2 should be checked against Table C1, and the statement that extreme wet events increase “by a factor of 7” in Section 4.3 does not appear to follow directly from the values reported in Table C3. I therefore recommend checking the relevant values, units, time periods, and calculations throughout the manuscript and Supplementary Material.

5. Section 2.3.1 explains that monthly CDFs are fitted using a non-parametric Kernel Density Estimation with the scikit-learn package. However, the main KDE settings, particularly the kernel, bandwidth, and bandwidth-selection procedure, are not specified. These details should be briefly provided to improve reproducibility. In addition, it is recommended to use precise terminology in the Results and Conclusions instead of repeatedly referring simply to “10-year wet events” or “10-year dry events”, as was done in the “Methods” section, where the SPLI values were correctly related to the return period of the equivalent event in the historical reference period.

6. Please clarify whether Figure 4 and Appendix C represent mutually exclusive SPLI categories or cumulative exceedance thresholds. Table 2 defines separate categories such as “between 2.5-year wet and 5-year wet,” “between 5-year wet and 10-year wet,” and “>10-year wet.” In contrast, Tables C1–C3 use expressions such as “ $\geq$ 2.5-year wet,” “ $\geq$ 5-year wet,” and “ $\geq$ 10-year wet,” which appear to represent cumulative exceedance thresholds. The authors should ensure that the definitions and terminology are consistent among Table 2, the Methods, the Results, figure captions, and Appendix C.
7. The Discussion appropriately acknowledges that the limited representation of southern France in the Aquifer platform prevents the existence of a France-wide North-South divide from being established with certainty. However, the Abstract still refers more broadly to groundwater changes “over France,” “South of France,” and a “discernible North-South distinction.” It is recommended to make the Abstract, Results, and Conclusions consistent with the more cautious interpretation in the Discussion, for example by referring to the Aquifer domain or to the French aquifer systems represented by Aquifer.
8. The Introduction currently states that the use of the selected projections as storylines to illustrate climate-change impacts on groundwater is an approach that, “to the best of our knowledge, has not yet been carried out in the literature.” Unless this statement can be supported by a sufficiently comprehensive literature assessment, I suggest avoiding such an absolute claim.