Reviewer 1

The manuscript "Spatio-temporal patterns and trends of streamflow in water-scarce Mediterranean basins" by Laia Estrada et al. proposes an intricate modelling exercise on a set of drainage basins of the Catalan River Basin district, with the main objectives of providing a tool useful for water management and assessing the spatio-temporal patterns and trends of stream flow during the first two decades of the 21st century.

The purposes and the extent of the exercise as well as the contribution of the stakeholders and the innovative design of the model calibration-validation are the main strengths of the manuscript. Nevertheless, there are severe methodological inadequacies that not only put into question model results but also would provide inadequate examples on how this kind of exercises may be correctly done.

Response: Thank you, we found your feedback very helpful! We have done our best to address all your comments, especially the methodological issue arising from excluding the land use changes during the simulated period, to improve this manuscript.

Working hypotheses:

In the Discussion section (lines 397 and subsequent), the authors quote several publications that "report a correlation between afforestation in the headwaters and decreases in streamflow" and indicate that "during the first two decades of the 21st century this trend has continued, with forested area going from 46% of our study area in 2000 to 56% in 2018".

Disregarding these clear alerts, the authors "did not account for land use changes, although we did include the effects of increased evaporation due to warming, which may be more relevant (Buendia et al., 2016)."

Response: Thank you for your comment. We agree that we did not properly justify the exclusion of the effect of land use change on streamflow, so we performed the analysis of temporal trends of the model residuals as recommended. This analysis is presented and discussed in section 1 of the Supplement. We do not find clear evidence that the increase in forest area or density are a major factor influencing flow in our study, as the number of positive and negative trends in the model's residuals are equivalent and cannot be interpreted as a single factor not included in the model being responsible for such trends, despite most of them being significant. Therefore, our assumption that climate variability is the main driver of change in streamflow and that land use change can be omitted from the modelling exercise without compromising results is not incorrect.

But Buendia et al (2016), in the Final Remarks section state: "Overall, results have indicated that increased forest areas are the major driver of reduced streamflows and the magnitude of peak floods."

Response: The sentence cited as to being part of the Final Remarks in Buendia et al. (2016) is not found anywhere in the paper. In fact, although important, increasing forested area was not found to be the major driver (Figure 9). Nonetheless, we have rewritten this section so it is better formulated.

Lines 436-442:

"However, under the assumption that climate variability and not afforestation is the main driver of streamflow reduction (Buendia et al., 2016), we did not account for these land use changes in the model, and we used the forested area in 2018 for the whole simulation period. The analysis of trends in model residuals (see section 1 of the Supplement) does not evidence the presence of a factor other than the ones already included in the model affecting the hydrological response, and thus justifies the exclusion of land use changes in our study. However, it must be noted that despite not accounting for land use changes per se, we do account for the increase in evapotranspiration due to increased temperatures."

The working hypothesis that the increase of forest cover can be omitted as possible cause of temporal trends of streamflow should be stated in the methods section, and the reason claimed by the authors for this omission is untrue.

Response: We agree that we should include our hypothesis in the methods section.

Lines 168-171:

"It should be noted that only a static rather than dynamic land use map is considered in this study, and thus we are omitting the effect that changes in land use during the simulation period may have on streamflow, under the assumption that climate and not land use change is the main driver of the hydrological response. To verify this hypothesis, we performed an analysis of the trends in the model residuals (section 1 of the Supplement)."

Given the results of several previous works, it is likely that land cover change is a more relevant driver of recent hydrological changes in this area than climate warming. Both the overall and spatial flow trends simulated by the model become highly doubtful and are not compared with actual ones.

Response: While we agree that land cover change can have an impact on hydrological changes, we disagree that is a more relevant driver than climate change in this area, as exemplified in the works cited. However, despite the fact that the trends were computed with calibrated streamflow and thus an argument can be made that they are indeed comparable to observed trends, we agree that also comparing them with the observed trends at the gauging stations could be interesting. We added this comparison in section 3 of the Supplement, as well as included the following paragraph in the 'Material and methods' section.

Lines 250-258:

"The advantage of using simulated streamflow rather than observed is working with 999 values for each indicator instead of only 50 (gauging stations), which allows us to better observe spatial patterns. Moreover, some gauging stations present gaps for the period 2001-2022, so simulated streamflow also provides a complete temporal series. Nevertheless, we compared the trends in indicators calculated with observed streamflow with the simulated trends at four gauging stations (see section 3 of the Supplement). We found significant observed trends for 20-29 hydrological indicators (50-72.5% of all indicators), compared to significant observed trends for 11-23 hydrological indicators (27.5-57.5%). Most of significant simulated trends (66.7-73.3%) are also found significant using the observed flow, and the majority of those (82-100%) are in the same direction (i.e. positive or negative trend). Therefore, while we do not capture all the observed trends with the model, the trends that we do capture are comparable to the observed trends."

Analysis of modelling results:

Contrary to its recurring attribution as a 'physically based model', SWAT is an empirical model without a sound physical basis. The core of SWAT is the Curve Number Model that is undoubtedly an empirical model.

This is not just a rhetorical question but is relevant to the interpretation of the modelling results. In a physically based model there might be some hope that the internal model variables (stores and fluxes) are acceptable if the simulated discharge is so (but Anderton et al., 2002). However, when a conceptual or empirical model is calibrated using streamflow data, "Model performances measure the correctness of estimates of hydrological variables generated by the model and not the structural adequacy of the model vis-à-vis the processes being modelled" (Klemes, 1986). In other words, the model not necessarily gives the "good answers for the good reasons" (Grayson et al., 1992; Beven, 2002; Kirchner, 2006), so model fluxes not directly used for calibration are highly suspect of being model artefacts.

Response: Thank you for your insight. However, we disagree that SWAT is "an empirical model without a sound physical basis". While some aspects of the SWAT modelling process might have empirical roots, it is a very complex model which simulates many different physical processes to the best of its ability, and its soundness and usefulness have been proved in many studies (see many referenced works throughout the paper). There are no perfect models, but an imperfect model can be useful, which is the case for the application of SWAT+ in this study.

Furthermore, the uncertainties associated to the model simulations (at least those used for calibration) must be analysed to provide the users with estimates of the risks in decision making (Grayson et al., 1992; Beven and Binley, 1992; Beven 2006; Herrera et al.,2022...).

Response: Thank you for your comment. We have included the quantification of the uncertainty associated to simulated streamflow in section 2 of the Supplement.

We also separated Fig. 3 into two figures, and for Fig. 4 (Annual Flow) we added the distribution of non-standardized Sen's slope and of the standard deviation.

Lines 241-243:

"We also assessed the 95PPU uncertainty bands and their metrics P-factor and R-factor (Abbaspour et al., 2015, 2018) for representative gauging stations of each main basin (see section 2 of the Supplement)."

Lines 272-274:

"The comparisons between observed and simulated daily streamflow for representative gauging stations of the main rivers of the CRBD demonstrate good model performance (Fig. 2, see section 2 of the Supplement to visualize the uncertainty represented by 95PPU bands)."

Lines 297-299:

"We don't observe a specific spatial pattern on the distribution of the Sen's slope standard deviation for total annual flow (Fig. 4c), except for the few significant trends in the Tordera basin, where the standard deviation is generally high, so overall we can conclude that the uncertainty in Sen's slopes for all CRBD is similar."

Figure 4:

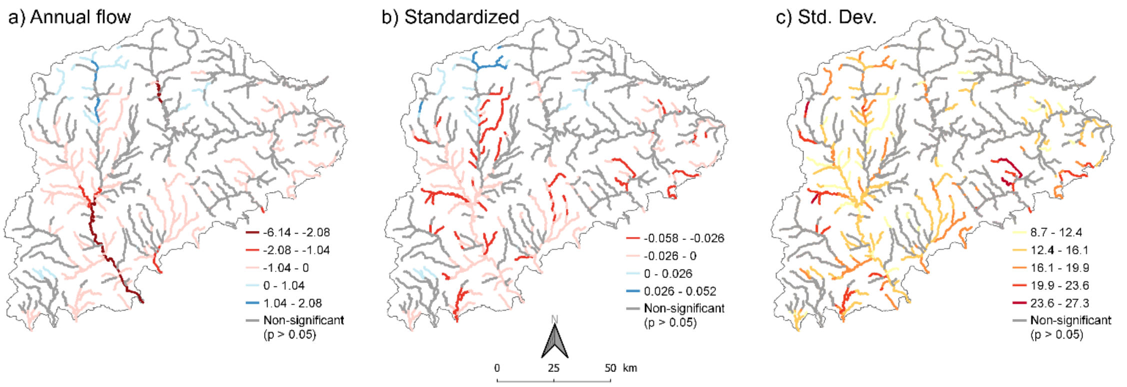

Figure 4: Spatial distribution of Sen's slope (a, units hm3/year), standardized Sen's slope (b) and standard deviation (c) for the hydrological indicator total annual flow.

Finally, In a sub-section section of the 'Materials and Methods' section named 'Data analysis', the authors included the calculation of many hydrological indicators, but contrarily to the title of the sub-section, this analysis was made (if I am not in error) not on the original 'data' but on internal (not calibrated) model results. Therefore there is no assessment on how these indicators represent the ones of the actual hydrological regimes.

Response: The analysis was indeed made with model results, albeit they are calibrated, and thus it can be argued that the indicators analysed are comparable to those computed with observations. Moreover, using calibrated streamflow instead of observed allows us to obtain a more comprehensive spatio-temporal analysis than the one the observed record can provide, as it presents gaps in space and time. However, following a previous comment, we have already added the comparison between observed and simulated trends for some gauging stations in section 3 of the Supplement. We have also changed the section title to 'Trend analysis' to avoid confusion.

Overall manuscript assessment.

In spite of the valuable strengths stated above, the modelling exercise is based on the inadequate working hypothesis that warming is the main driver of hydrological trends in this area and manages several principal and internal model outputs as actual data without any assessment of the uncertainty associated with these simulations.

Response: Thank you for your comment. As mentioned in other responses, the hypothesis that climate change is the main driver or hydrological trends in this area is not inadequate, as discussed in section 1 of the Supplement.

Concerning the uncertainty, we have quantified and shown the uncertainty of simulated streamflow, as well as for one of the main indicators Annual Flow (Fig. 4 and Supplement section 2).

Recommendations.

Both the importance of the objectives and the magnitude of the modelling exercise deserve finding some feasible way to improve the soundness of the project.

Response: Thank you for your comment. We hope that with your recommendations we managed to improve this manuscript.

The fact that the encroachment of forest cover in the studied catchments is a likely or very likely driver of the hydrological response involves a difficulty for the modelling exercise but an opportunity for water management. Indeed, if climate were the main driver of the hydrological response, management strategies for adaptation to the climate change would be limited. Conversely, if forest cover is the main driver, it can be managed to reduce the 'green water' consumption and increase the 'blue water' delivery (Falkenmark, 2000) as a climate change adaptation strategy.

Using SWAT for simulating the hydrological response to forest cover change is a cumbersome and risky task, taking into account the poor or very intricate examples available (Haas et al., 2022; Karki et al., 2023).

Response: While we have verified that climate is our main driver of the hydrological response, we do not disregard the effect of land use change and management, evidenced by its inclusion in the discussion of our article. In fact, the SWAT+CRBD model was used in Garcia et al. (2024) to assess the impact of forest management on water resources availability. We have included the following into the discussion:

Lines 361-369:

“Another consideration on model inputs is the fact that land use change during the simulation period is not considered in our study, due to having determined that land use change and in particular afforestation is not a main driver of hydrological response for the scope of this study (section 1 of the Supplement). However, it can still be an important factor at the local scale, and its consideration represents an opportunity for future management practices. Forest cover can be managed to reduce “green water” (i.e., water stored in the soil and vegetation and that is then consumed) and turn it into “blue water” (i.e., runoff), increasing water availability in potential areas suffering from water scarcity. Garcia et al. (2024) used the SWAT+ CRBD model to assess the effect of forest thinning on water yield, and results highlighted the potential of forest management to enhance “blue water” availability.”

But the flow simulations made may be used to test the null hypothesis that the climatic forcing is sufficient to explain the observed flow records, analysing whether there are time increasing model residuals that could be attributed to the role of increasing forest cover extent or density. This exercise may be made in most of the gauging stations used, providing a map of the hydrological changes attributable to the encroachment of forest cover. The statistical significance of trends should be made following the recommendations issued by the IPCC (Mastrandrea et al., 2010).

Response: Thank you for your recommendation. As mentioned in previous comments, we have performed the trend analysis of the model residuals to test whether the assumption that climatic forcing is enough to explain the observed flows. We have assessed that climate variability is the main driver of change in streamflow in our study, and therefore land use change can be omitted from the modelling exercise without compromising results.

Unfortunately, the hydrological indicators analysed in the manuscript may be obtained for the flow records at the gauging stations, but any comparison with the simulated ones is expected to give inconsistent results because it is not possible to determine if the differences are attributable to modelling errors or to the role of the hydrological role of forest encroachment.

Finally, the maps of figures 4 to 7 should be discarded because these results are highly suspect of being modelling artefacts because do not take into account the role of forest cover change and these are internal model outputs not calibrated and of unknown uncertainty.

Response: As we have determined that the increased forested area is not a main driver of hydrological change in our study, the hydrological indicators obtained from calibrated streamflow are adequate for the analysis of trends and patterns. However, a comparison with observed trends has also been included, as well as the quantification of the uncertainty. Thus, we do not believe Figs. 4 to 7 (now Figs. 5 to 8) should be discarded.