RC1: 'Comment on hess-2022-264', Anonymous Referee #1, 29 Sep 2022

The paper entitled "Producing reliable hydrologic scenarios from raw climate model outputs without resorting to meteorological observations" by Simon Richard and co-authors proposes a new modeling workflow to derive hydrologic scenarios from climate model projections without resort to the usual step of bias correction of climate projections. The topic is well suited for the journal Hydrology and Earth System Sciences and I think the paper raises important and interesting questions at the interface of climate sciences and hydrology.

I found the paper well written and organized, the methods are described in details, and a case study allows readers to have an idea of the performance of the proposed framework for a typical application (in the present case: producing hydrologic scenarios under climate change for a catchment located in Québec - Canada). I enjoyed the comparison with a more conventional approach (i.e. including a step of bias correction of model outputs), and the way the results are displayed and discussed in Section 4: Results.

My only sticking point is the discussion (Sect. 5). In contrast with Section 4, I have the impression that the authors are overselling their approach in the discussion section. Indeed, the general tone of this section tends to make the reader forget that the conventional approach still outperforms the one proposed in the present paper (cf Fig. 6 and 7). I acknowledge the interest of the alternative approach proposed by the authors for both (1) complement the standard approach in areas where meteorological observations are available, and (2) allow hydrologists to perform hydrologic projections where meteorological observations are too sparse to enable reliable climate model bias correction. But I also think that the discussion is too harsh towards climate model bias correction (i.e. the conventional approach), and that the proposed approach does not solve many of the (legitimate) questions raised by the authors in the discussion. More particularly:

We edited Section 5, reflecting the strengths and weaknesses of both approaches in a neutral way. Section 5.1 explicitly presents the proposed approach as a complement to conventional hydroclimatic modeling.

* Integrating meteorological observations in the modeling chain. When meteorological observations are available, I think it is a shame to disregard them. Maybe the conventional approach of bias correction of climate outputs is not ideal (and the authors are right to talk about its limitations), but removing it without replacement is equivalent to do without all the information embedded in meteorological observations. Of course having a method to deal with poorly

gauged areas (in terms of meteorological variables) is a plus, but when data are available I don't see why not using them. So I think that the authors should acknowledge more clearly in the discussion that when meteorological data are sufficient for climate model bias correction, this method is still the one that performs best.

Section 5.1 (lines 449-454) encourages a sound use of reliable meteorological observations when available. It also encourages further exploration of combined approaches (conventional-asynchronous) aiming to maximise available observations at the regional scale or for modelling hydrological processes using more complex physical descriptions.

* Physical consistency of the modeling chain. I agree that bias correction methods disrupt the physical consistency of climate projections. But clearly the approach proposed in this paper does not provide any solution to this problem. The calibration of a hydrological model directly from raw climate outputs will mix climate biases and hydrologic biases (as acknowledged by the authors in Sect 5.3), which results in a completely non-physical hydrological model. So I think that the authors should acknowledge that both approaches are equally breaking the physical consistency of the hydro-meteorological processes involved in the models used to investigate the hydrological response of our environment to climate change, and that where we decide to do it (and unfortunately often to hide it) within the modeling chain is somehow a matter of taste.

The perturbation of the physical consistency between simulated hydrologic processes through parametric compensation is explicitly acknowledged as a limitation in Section 5.3 (lines 501-505).

* Interpretability of climate projections by end-users. The questions raised at the end of Sect 5.2 (l 465 - 472) are interesting and legitimate, but with a very few exceptions can also be addressed with bias-corrected climate model outputs. In a slightly different note, I do not think that bringing expertise in analyzing, selecting and pre-processing climate model outputs is in itself a bad thing. An intense and constructive discussion between climate modelers, statisticians performing bias corrections, hydrologists and stakeholders (I probably forget important participants) is of course essential, but I am a bit skeptical about the idea of a more direct (and therefore possibly less careful) use of raw climate model outputs. From my personal experience, the support from experts in climate models biases and bias correction is essential to avoid misuse of climate projections.

We did not intentionally suggest excluding climate model experts in the analysis of bias. We rephrased line 473, which could be misleading regarding this aspect.

To sum up, I think that the present manuscript is interesting and well written, and I my opinion deserves publication after revisions. My main concern is about the discussion, which I believe can be improved by moderate revisions, mostly by rephrasing this section in a more objective way.

Hereafter are some minor comments/questions I had during my reading of the manuscript:

- L119: how many RCM grid cells cover your study area?

The information is added to the revised manuscript (line 120).

- Fig 1: Maybe add information about topography.

Figure 1 has been edited with topography.

- Fig 2: It would be nice to compare with the conventional approach in the figure (i.e. have 2 workflows in the figure) not only in the caption.

We believe that comparing both workflows would notably impact the conciseness of the manuscript. A detailed description of the conventional modelling approach is provided by Ricard et al. 2020. A reference has been added to caption of Figure 2 (line 159).

- Throughout Sect. 4: also mention relative biases. For instance L 243: "Biases typically range between -1 and +2 mm/day (xx %) depending on ...".

Relative biases computed on daily values are affected with a high variability, especially for precipitations. We converted the constant +0.5 mm/day bias (observed from January to August and from mid-October to December) into relative terms (+27%, see line 243) using the median of daily values over the corresponding period. We believe however that the min-max range (from -17% to +168%) might mislead the reader. We remind that a 5-day moving window is applied to all time series in Figures 3 and 4.

- L 264 and after: Validation of the asynchronous -> I would prefer "assessment" instead of "validation" (maybe personal taste).

As suggested, "assessment" was replaced "validation" in the revised manuscript (line 266).

- L 298: "which can be considered as comparably performant relative to the conventional approach" -> one of the few places in sect 4 where you are not very fair with your results. Consider rephrasing.

The sentence has been rephrased to: " [...] which is comparably performant relative to the conventional approach." (line 300).

- L 369: typo in percentile definition: 0.5 -> 0.05

The typo in percentile definition has been changed to 0.005 and 0.995, respectively (line 371).

- L 386: maybe remind for which period and RCP scenarios the hydrologic scenarios are made.

The description has been rephrased (lines 388-390), indicating the corresponding period and the RCP.

- L422: we validated -> we assessed the performance

" validated " has been changed to " assessed " (lines 426 and 535).

- L 519-521: the way this paragraph is written gives the impression that you consider low and high flows as extreme events. Maybe consider rephrasing?

"extreme events" has been changed to "high and low flow events".

Citation: https://doi.org/10.5194/hess-2022-264-RC1

'Comment on hess-2022-264', Anonymous Referee #2, 15 Nov 2022

Review for "Producing reliable hydrologic scenarios from raw climate model outputs without resorting to meteorological observations". This manuscript seeks to provide a new framework that can use regional climate model projections (CORDEX) to provide reliable hydrological projections. This framework aims to avoid using meteorological forcing data. Although it is an important topic, I feel most of the claimed goals are not well supported.

We acknowledge that the presentation of a novel methodological framework can raise doubts and suspicions. Our intention with this paper is to provide as much arguments as possible supporting the idea that the framework could be useful to hydrologist assessing the impact of climate change on water resource in a situation where meteorological observations are rare. We are fully aware that the scope of the paper only provides a proof of concept and a partial validation. We are confident however that further work could be conducted to provide a more in-depth comparison with conventional hydroclimatic modelling.

1. Although the meteorological data is not used, it still requires streamflow observations. I agree that it still uses less data than "conventional" approaches. However, regions with poor meteorological data are less likely to have reliable streamflow observations as well. Therefore, the benefit of this approach is questionable.

We do not fully agree with comment raised by RC2. An example, in Northern Québec, streamflow data are available at the outlet of some large catchments while almost no meteorological stations are located on the watershed, providing large uncertainty related to precipitation and 2m temperature. We believe, in such cases, that the implementation of an asynchronous modelling framework could provides notable benefits in comparison to conventional modelling.

2. Following my comment above, I think the missing part is: under what meteorological forcing data uncertainty levels the proposed framework is more advantageous? For instance, if we have only one precipitation but good streamflow gauges (not sure if that is realistic), the proposed framework outperforms the conventional approaches.

An assessment scheme comparing the performance of both frameworks using intentionally scarcer meteorological observations is suggested as further work in Section 5.1.

The description of the method requires more details. For instance, in line 170: which parameters are calibrated to minimize nCRPS?

As stated in section 3.3, the calibration is performed using the same objective function, calibration period, and configuration of the optimization algorithm. We specified in the revised version of the manuscript that the same model parameters are calibrated in both frameworks (line 224). However, we think that presenting every parameter for each model and their role would make this section wordy, without bringing key information to the study. A reference has been added to Table 3 for additional information on the model parameters.

3. When we are using regional climate model projections/simulations, we tend to be more interested in the long-term statistics, e.g., trends, standard deviations etc. However, only long-term climatology is discussed.

We limited the analysis to long term climatology for conciseness of the paper.

4. In the title, “scenarios” and “climate models” make me automatically think about climate change and long-term trends. However, these perspectives are not discussed and/or validated. I would suggest the authors to modify the title and the manuscript to avoid any confusions.

We believe that the title fits the scope of the paper. We would like if the reviewer could highlight better the confusing elements. Since future climate and hydrologic conditions cannot be directly validated, a diversity of sound methods validated over the reference historic conditions appears to us as the best compromise to attribute confidence in hydrologic scenarios.

5. Finally, I think it should present the optimized parameters to see if they are physically reasonable.

We did not include a detailed description of the parameter values because the physical interpretation of the global conceptual hydrological model parameters is difficult. Only few parameters could be related to measurable physical quantities. Additionally, the calibration algorithm (the Shuffle Complex Evolution) requires the specification of an upper and lower bound. The parameter values will necessarily lie within the bounds that are forced. Therefore, parameters cannot get a value that is of an order of magnitude different from the ones obtained with a “regular calibration”.

Finally, one should expect the parameter values obtained from an objective function that does not consider the temporal correlation to differ from the ones found with a more traditional score like RMSE or KGE. This makes the comparison of the two sets of parameters complex as the same hydrologic model is expected to behave in a different way in an asynchronous fashion.

Citation: https://doi.org/10.5194/hess-2022-264-RC2