

Response to Reviewers' Comments

We appreciate the efforts of the reviewers and we thank them for their insightful and constructive comments. We have addressed all concerns in this revised manuscript. Below, we provide detailed responses to each of the reviewers' comments. For convenience, we put the reviewer comments in black font, author responses in blue, and direct quotes from the revised manuscript *in italic*.

Anonymous Referee #1 comments:

The manuscript provides a relevant evaluation of the ERA5 precipitation and temperature data from a hydrological modelling perspective. The dataset is compared with a previous release of the ERA-type of reanalyses. The manuscript is well-written and uses clear language.

Response:

We appreciate the reviewer's comments. Please see the point-to-point responses below.

Specific comments

[page 3, line 72] ERA5 incorporates several improvements over ERA-I. **Please define acronyms before using them.**

Response:

We defined the acronyms [page 2, lines 67-68].

“Amongst all available reanalyses, many studies have shown ERA-Interim (European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis) to be the best or amongst the best performing reanalysis products (e.g. Sun et al., 2018; Beck et al., 2017a; Essou et al., 2017; 2016b), arguably the result of its sophisticated assimilation scheme, 70 and despite a spatial resolution inferior to that of most other modern reanalyses. In March 2019, ECMWF released the fifth generation of its reanalysis (ERA5) over the 1979-2018 period (Hersbach and Dee, 2016). ERA5 incorporates several improvements over ERA-I (see section 3 of this paper).”

[page 3, line 86] Even though the hourly temporal scale brings a lot of potential applications for hydrological studies, a first step in the evaluation of ERA5 precipitation and temperature datasets must be performed at the daily scale. **Avoid exaggeration words like “must”.**

Response:

We have modified this in [page 3, lines 86].

“Even though the hourly temporal scale brings a lot of potential applications for hydrological studies, a first step in the evaluation of ERA5 precipitation and temperature datasets is performed at the daily scale.”

[page 8, line 249-250] ERA5 precipitation is the best reanalysis product, ranking second overall after the observations. It is clear that for hydrological modelling, the ERA5 dataset is a net improvement over the ERA-Interim reanalysis. **This seems redundant with the previous sentence. Consider merging?**

Response:

Thank you. We have merged the sentences [page 8, lines 250-251].

“It is clear that for hydrological modelling, the ERA5 dataset is a net improvement over the ERA-Interim reanalysis ranking second after the observations.”

[page 15, line 485] From all the results in this study, there does not seem to be any reason or indication that ERA-Interim should continue to be used for hydrological modelling applications. **..in north America, at least.**

Response:

Thank you. We have modified the sentence [page 15, lines 491].

“From all the results in this study, there does not seem to be any reason or indication that ERA-Interim should continue to be used for hydrological modelling applications, at least in North America”.

[page 15, line 498] The single streamflow criteria and objective function (KGE), like its Nash-Sutcliffe relative, is weighted towards higher flow events. Other objective functions would return different results, however the fact that ERA5 climate data is generally improved in all areas means that the objective function is unlikely to have a large impact on results. **I believe that KGE is an acceptable objective function, but the justification may need some rephrasing or different argument. I am not convinced that an improvement across regions compensates for an objective function bias toward high flow events. For example if, for every region, a data product is great for high flow events, but terrible otherwise, then an objective function that favours the reproduction of high flow events will deem such a data product to be superior, contrary to an objective function that favours the reproduction of low flow events. The two conclusions would be vastly different. Therefore, observed improvement across multiple regions does not necessarily invalidate a link with the choice of the objective function.**

Response:

We agree that the text was too confident in results that would be obtained with other objective functions. However, we also think that the improvements in ERA5 quality (reduction in biases, etc.) would translate somehow to the quality of simulated streamflow. Therefore, we have rewritten the sentence as follows [page 15, lines 505-507]:

“Other objective functions would return different results, however the fact that ERA5 climate data is generally improved in all areas is an indicator that other metrics could potentially see improved results as well, although no test has been performed to that effect in this study.”

[page 22, figure 1 title] Watershed locations and their mean elevations over Canada and the United-States.
Are these points the location of the streamflow gauges rather?

Response:

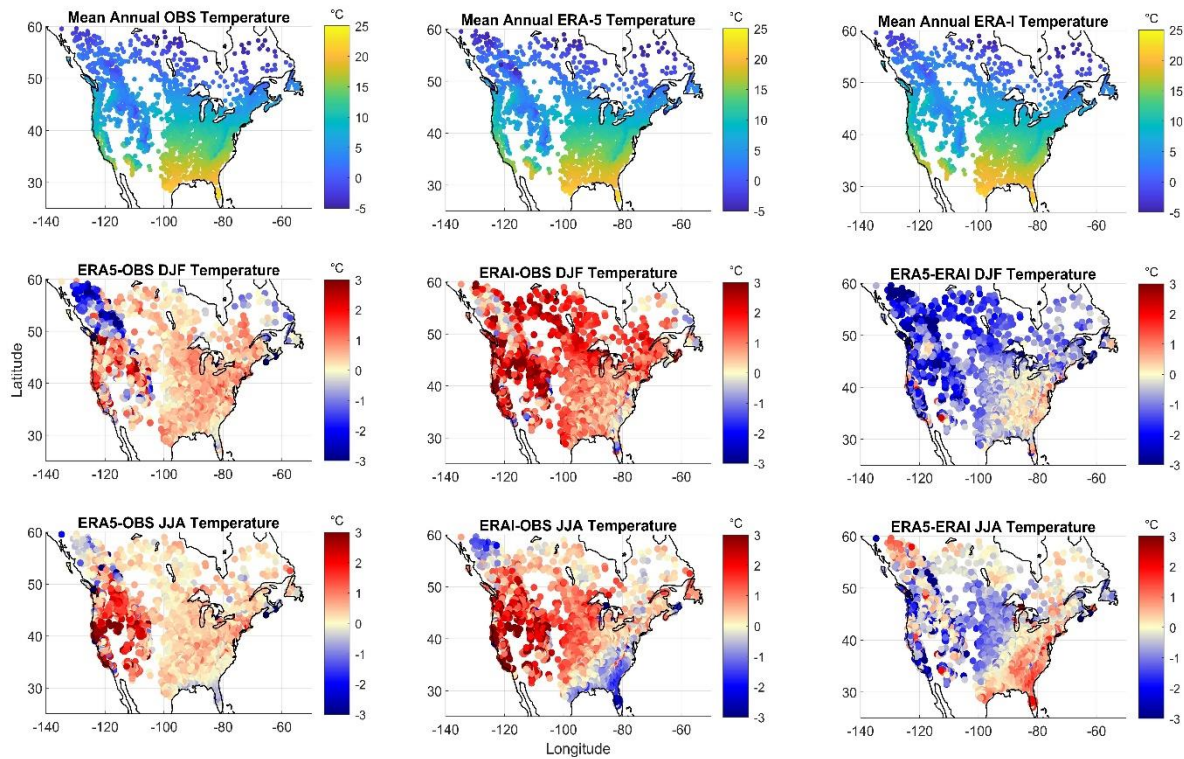
Each dot represents the centroid of the watershed and we have clarified that in the figure description.

“Figure 1: Watershed locations and their mean elevations over Canada and the United-States (each dot represents the watershed centroid).”

[page 23, figure 2] **Formatting:** There is some overlap between the title and the units.

Response:

Thank you. We have formatted the subplot titles.



[page 32, figure 11] **The legend should be included either in the figure or figure description, instead of having the reader hunt for it in the text.**

Response:

We have edited the figure description to clarify what each color refers to.

Figure 11: Results of the Kruskal-Wallis statistical significance test to determine the best dataset for hydrological modelling as observed through the KGE metric, for each climate zone. The green, yellow and red colours respectively indicate the best, second best and worst datasets for each climate zone.

[page 33, figure 12] **What is G-OBS? The streamflow observations? Please define.**

Response:

Exactly, the G-OBS label refers to the streamflow observations. This has been edited in the figure description.

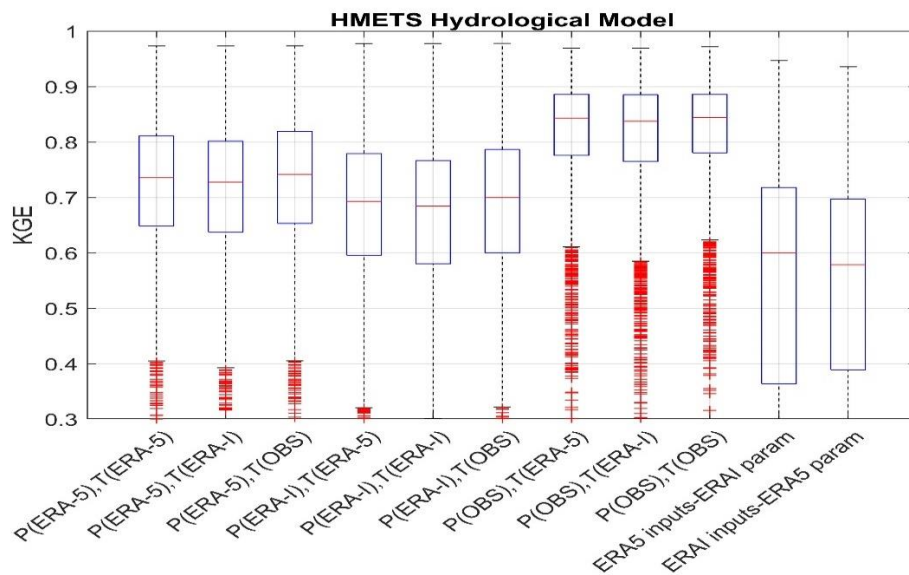
“Figure 12: Difference in hydrological modelling performance, mean monthly precipitation and temperature and mean annual hydrograph using ERA-I, ERA5, observations (OBS) and streamflow observations (G-OBS) on three dissimilar catchments: Ouiska Chitto Creek (top row), Grande Rivière à la Baleine (center row) and Cosumnes River (bottom row).”

General suggestion

While the superiority of the ERA5 dataset over ERA-interim appears clear, the study lacks independent validation measures and opts to use all available hydrological data for calibration purposes. This choice to use more calibration data is justified versus the traditional reservation of an independent period to validate the calibration. However, while the chance of accidental bias of the calibration method toward a particular dataset is likely slim, it is not guaranteed, or rather the method is not set up to minimize its likelihood. My suggestion to circumvent this issue is to interchange model parameters between ERA5 and ERA-I scenarios. For example, if ERA5 inputs with ERA-I parameters still perform better than ERA-I inputs with ERA-I parameters, this would be a very convincing argument against the impact of the calibration method from being responsible for the perceived improvement of ERA5 over ERA-I.

Response:

Thank you for the suggestion. We have interchanged the model parameters between ERA5 and ERA-I. The results show a significant decrease in the overall performance compared to ERAI_data with ERAI_parameters and/or ERA5_data with ERA5_parameters. This is because the model parameters are strongly dependent on the forcing dataset in the calibration process. Calibration parameters are flexible enough to compensate for the potential biases among the forcing meteorological dataset. Therefore, optimal calibration parameters vary across the different study basins and the selected meteorological datasets (Elsner, Gangopadhyay et al. 2014). However, the model calibrated with ERA5 and driven with ERAI data showed slightly worse results than the model calibrated with ERAI data and driven with ERA5 data. However, we believe these results should not be interpreted in any way to comment on the validity or quality of the reanalysis data.



Reference :

Elsner, M. M., et al. (2014). "How does the choice of distributed meteorological data affect hydrologic model calibration and streamflow simulations?" *Journal of Hydrometeorology* **15**(4): 1384-1403.