Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/hess-2021-518-RC1, 2022

Comment on hess-2021-518

Anonymous Referee #1

Referee comment on "Unfolding the relationship between seasonal forecast skill and value in hydropower production: a global analysis" by Donghoon Lee et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-518-RC1, 2022

Summary

The manuscript by Lee et al. explores the relationship between forecast skill and value in the case of the management of hydropower dams. The authors use dam characteristics and forecast skill to identify categories of dams that (1) show potential for improvement or not over climatology-based operating rules, and (2) show improvement or not based on realistic forecasts. A climate classification is further used to "regionalize" the added value of long-term forecasts for the hydropower sector and identify regions where improvements of currently low quality forecasts would translate into added value for dam management.

The paper is of very high quality, is well referenced, well written and scientifically sound. It will be undeniably valuable for the forecasting community, but also has potential to reach hydropower production managers. Along with the manuscript come supplemental materials that further detail the methodology and the results, as well as a dataset and an R script that allow readers to access the datasets for each dam.

For these reasons, I strongly recommend this manuscript for publication. Hereafter, I list some questions to the authors, some recommendations for improving explanations, and mostly minor points.

General comments

Sections 2.1 and 4.4: You decided to use the Köppen-Geiger climate classification. Since you are working hydropower and inflows which are influenced not only by climate patterns but also hydrological ones, a classification based on hydro-climate characteristics and not only climate characteristics would seem more relevant for the goal you are trying to achieve. Please consider using the hydro-climate classification proposed by Knoben et al. (2018).

Knoben, W. J. M., Woods, R. A., and Freer, J. E.: A Quantitative Hydrological Climate Classification Evaluated With Independent Streamflow Data, Water Resources Research, 54, 5088–5109, https://doi.org/10.1029/2018WR022913, 2018.

Section 3.1, Figure 1: It is not clear to me why the authors allow future inflows (t+1, t+2, ...t+7) to be predicted based on future climate indicators (1-8 months ahead). In a true forecasting setting, the ENSO, PDO, NAO and AMO indices for the 1-8 months ahead would not be available, only forecasts of these indices would. Some clarifications would be needed on this aspect. For instance, the authors could re-use the very clear notation t, t+1, …, t+8 to define which time steps they use in terms of climate teleconnection indices with respect to the forecast month t.

Section 3.3.2: Even though the authors argue that MdAPE has a higher correlation and that it provides a value at each time step, KGE, and in particular its components, may have given insights into the forecast characteristics (correct timing, volumes, variations) that influence value. This information would be extremely valuable to guide further model and forecast developments for the hydropower sector, in the same way your investigation of dam characteristics informs dam managers of potential forecast value. I wonder whether this would be something to explore also to address the limitation you note in the Results section L.341 "For dams with poor IDF and high KGE, two features are noteworthy: first, KGE may not fully capture the relationship between forecast skill and value".

Forecasts with horizons up to 7 months are generally probabilistic to account for uncertainties at such long lead times. The authors should discuss the role of uncertainties in their study design, i.e. how realistic it is to consider the value of deterministic long-range forecasts depending on the current state of hydro-climate long-range forecasts, but also on the capacity for hydropower dam managers (whose actions are hypothesized in this study) to inform their decisions based on probabilistic information.

Specific comments

L.132-135 There is a range of models that fall between statistical prediction models and

physically-based model. For instance, conceptual models do not fit in these two broad categories. I invite the authors to revise this statement.

L.137-140 The arguments for choosing a statistical model rather than a physically-based one seem too general. In fact, the statement "the prediction horizon of most physically-based approaches (a few days to 3-4 months) falls short of our preferred lead times up to seven months" only holds when considering currently openly available global reforecasts, and not reforecasts from physically-based (or rainfall-runoff) models in general. There already exists, for instance, global reforecasts up to 7 months ahead and with hindcast periods for at least 30 years (https://hypeweb.smhi.se/explore-water/forecasts/seasonal-forecasts-global/). As the authors rightfully mention in the section on opportunities, "global-scale forecasting systems are gaining momentum", and therefore this part should be rewritten to highlight the impermanence of the statements.

L.141 "Our long-range inflow prediction **model** uses…"

L.145 "For example, forecasts **issued...**"

L.174 Isn't it the goal of the dam inflow prediction model to feed the reservoir model? If so, isn't Q_t not only retrieved from WaterGAP but also from the proposed statistical dam inflow model?

Section 3.2.1 Is there any need for initialization of this reservoir model, and if so, how do you handle this aspect? e.g. which initial values do you use for instance for the reservoir storage?

L.227 "… may **influence** ..."

L.234 "It is reasonable to hypothesize **that** the value…"

L.251-253 "Note that failure implies that the control rules and perfect forecast-informed operations generate a similar amount of hydropower, meaning that information on storage and previous-month inflow are sufficient for near-optimal release decisions." Wouldn't that correspond to an I_{PF} value of 0 rather than to the mean I_{PF}? If this statement is based on the mean I_{PF} value, the reader does not have this information yet, and this sentence is confusing.

L.320-321 "Considering the superior performance of the MP1 model, the forecast skill of MP1 only is retained to represent the overall forecast skill in the following analyses." Since the optimisation uses all forecast horizons, the speed with which skill decreases with the forecast horizon may play a role in the optimization and could have been considered as well.

L.327-329 "Small negative values of I_{PF} are likely a result of the discretization needed by dynamic programming to optimize the release sequence (eq. (4)), hence allowing control rules to outperform perfect forecast-informed operations." Could you please further explain what you mean to help understand the counter-intuitive negative I_{PF} values?

L355-357 "This is attributed to the weekly operations, suggesting that more frequent release decisions may reduce forecast value, since the benchmark operating rules have more opportunities to adjust release decisions." Isn't it the case for all the dams below the horizontal line? Why should these ones (the failing ones) behave differently?

L.446 "… a number of assumptions that must **be** properly contextualized."

Figure 2 It would be more correct to change the caption to "Percentage of dams **whose inflow is** significantly correlated with…" since a dam in itself is not correlated to anything.

Figures 3 and S2 The titles for the color scales in the second and third lines of this figure are confusing. If my understanding is correct, I would suggest changing titles in the first column to "Change in number of predicted months", and in the second column "Change in KGE", with "(b) MP4-MP1" and "(c) MP7-MP1" on the left-hand side.

Figure 5 "red triangles **represent** *failures*"

Figure 6 "meaning that the performance of realistic forecasts is worse **than** the one attained by control rules"