

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
<https://doi.org/10.5194/hess-2022-244-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on hess-2022-244

Anonymous Referee #1

Referee comment on "Spatial variability in Alpine reservoir regulation: deriving reservoir operations from streamflow using generalized additive models" by Manuela Irene Brunner and Philippe Naveau, Hydrol. Earth Syst. Sci. Discuss.,
<https://doi.org/10.5194/hess-2022-244-RC1>, 2022

This paper uses generalized additive models (GAMs) to compare pre- and post-dam construction flows as a function of rainfall, temperature and glacier mass changes in 74 high mountain catchments. The approach enables to reconstruct reservoir operations, as well as its hydro-climatic drivers, in situation where inflows and outflows are not readily available (which is often the case). The paper is really interesting, the method is a smart way to tackle a challenging puzzle, and there is a feeling that it is likely novel... but this is incumbent on authors to demonstrate, both by explaining the novelty in more detail, and by validating the method! My main remarks are as follows:

- 1) the novelty needs to be explicitly and precisely stated on the introduction.
- 2) the absence of a formal validation is a real issue that should (and can!) be addressed.
- 3) the methodology is creative, but the explanations are sometimes vague, and needs to be better justified and documented in several places.
- 4) results need to be presented step by step. They are really interesting, but authors need to be more rigorous in presenting how they got to them. I

It might take a substantial revision for authors to address the comments above. Revisions are important because this work has the potential to be a nice paper. I now develop the four points above.

1) Several points need to be clarified in the introduction:

=> What is the exact scope of the paper? Is the type of data authors base themselves on widely available in a range of cases (not just the Alps)?

=> Why are GAM the appropriate methodology here? Why not use machine learning methods for instance?

=> A simple search of “generalized additive model reservoir” on Scopus returns 120 results. Can authors ascertain no similar attempt has been tried in the past? What have been GAM uses in the literature around hydrological and reservoir modelling?

2) Because there is no inflow and outflow data for the catchment authors study, they do not validate the approach using basins for which the post-dam construction reservoir inflows and outflows are known. But there are many mountain reservoirs in other regions (e.g., the USA) for which inflows and outflows, as well as the other data authors use in the Alps. This would provide formal validation. Note they can then choose to split this paper into two papers (i) one that presents and validates the method, (ii) one that applies it to Alpine catchments to derive new knowledge (which is what they have here). I note here that the confidential dataset shared by authors does not address this concern about validation.

3) In order of appearance, my remarks are as follows:

Performance of the pre-dam model: with 74 catchments, it would only be normal that the GAMS model would work more or less well across the sample. The methodology section (lines 126-130 in particular) does not explain what performance measures are used. Related question: do authors investigate the influence of the length of the pre-development record on model quality? Please include a full list (maybe in supplementary material) of the catchments, including relevant characteristics (e.g., flow record length, year reservoir went online, catchment surface, ratio of reservoir storage over natural annual flow).

Authors give no indication whether the possibility of an extended filling period for the reservoirs was considered. Or did they assume this would be negligible? If so, please justify explicitly.

Section 2.2: the method enables one to derive time series of inflows / storage / release. Authors choose a final signal in terms of difference outflows-inflows that is creative and innovative, and could be a key contribution of the paper. Yet it is not justified, especially

in relation to the literature on the role of reservoir in hydrological models that authors claim to contribute to. This leads to several important questions:

- i) How do the chosen signals relate literature trying to specify reservoir outflows as a function of inflows, storage, and reservoir purpose?
 - ii) Why not derive relationships between inflows and outflows once inflows are derived? That is more standard.
 - iii) How do we know smoothing actually is needed? And how do we know how to parameterise the smoothing technique?
- 4) The results section is very interesting in highlighting a category of high-elevation hydropower reservoirs (but authors should comment in the discussion how new this is). But it only reports on what happens once the flow signal is derived, whereas it would be great to see results for the fitting of the pre-dam GAM model, as this is key to getting quality results. Authors should report in detail on goodness of fit, relationship between number of years of record and quality of the results, etc.

Another remark on results at lines 161-163: this is a key point that seems difficult to make without taking a long look at reservoirs' storage capacities. Reservoir with small regulation capacity (ratio of live storage to average annual flow) will naturally have weaker signals.

One last remark concerning results is that it would be great to highlight one or two known water supply reservoirs (for irrigation or any other uses), and show how their method enables quantifying downstream needs. That's not strictly necessary but it would show that the proposed method is versatile (which it is), instead of being a one-trick pony that only identified high-elevation single use hydropower reservoir.

Finally, below are also some minor remarks:

Lines 47-51: You talk about a two step setup and then lose the reader a bit by not making both steps explicit. I'd advise doing that.

Line 56: by convention, please introduce an equation with ":" Several times in the paper please replace "see, e.g." with "see e.g.,".

Line 113: please replace "(GAM) (Hastie" by "(GAM; Hastie" Figure 3.c: the x-axis says this is the early 1950s whereas the legend says early 1910s. Which is it? Also, please fit the limits on the x-axis to the beginning and ending of the record you are plotting (also valid in Figure 4).

Figure 4: please enlarge the figure.

"Glaciermelt": consider writing "glacier melt" instead.

Line 203-207: yes, but it would be great to try to nail the scope in the introduction instead of just speculating in the discussion. How applicable is this data to other places? That can be in principle deduced from the methodological setup alone.