

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC3
<https://doi.org/10.5194/nhess-2022-171-RC3>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on nhess-2022-171

Anonymous Referee #3

Referee comment on "Comparison of machine learning techniques for reservoir outflow forecasting" by Orlando García-Feal et al., Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2022-171-RC3>, 2022

The paper is well written and the figures are of high quality.

The research is well structured and explained: the problem, the methods, the data and the results.

However, the interest of the topic is not clear. The input variables considered exclude essential information on the reservoir outflow, which limits the predictive accuracy for any algorithm. Operation strategies play a key role in the outflow values. During dry periods, only the environmental flow is probably discharged in dams without hydropower units. In other situations, the flow will be obviously influenced by the strategy applied for optimizing power generation. Other restrictions such as freeboards in wet periods also have an influence.

Predicting reservoir outflow without considering this information is difficult, and the use of one technique or the other has a lower effect in accuracy.

The authors actually mention this in the introduction:
"it can be a good approximation in flood scenarios during wet seasons, especially in small reservoirs or when they are nearly full and have little margin to alter the natural flow of the river."

The results confirm this intuition:
"Looking more closely at the data in Table 3 a tendency is detected in reservoirs with a higher capacity to have worse statistics than those of lower capacity."

I suggest modifying the statement in the introduction. I would rather say that the proposed approach can only be applied with reasonable accuracy to small reservoirs during wet seasons in wet climatic conditions. All reservoirs considered are located in Galicia, an area with higher rainfall rates than other Spanish regions. The approach is probably less applicable in dry regions. This is an important piece of information for the community.

The authors may consider performing some kind of variable selection. For instance, adding inputs such as the season, the month or the day of the year, could serve as proxies to the operation rules. Also, the gradient of pool level could be informative (the operation of the reservoirs is probably different for increasing than for decreasing pool level for a given value of the stored volume). This may result to be more useful for increasing accuracy than the use of some specific, complex ML algorithm: the results show that differences among algorithms are mostly negligible and that even a very simple regression (MLR) is comparable to sophisticated techniques.

Nonetheless, operation strategy is the essential element in outflow. Overall, the usefulness of predicting outflows from reservoirs without information on the operation strategy is questionable. Every river basin authority should have such information available. The authors should clearly explain the scope of application of the approach, i. e., under which conditions the operation strategy is not known by the water resources management authority.

The pdf file attached includes additional comments and suggestions.

Please also note the supplement to this comment:

<https://nhess.copernicus.org/preprints/nhess-2022-171/nhess-2022-171-RC3-supplement.pdf>