Dear editor and reviewers,

Thank you for your constructive comments and suggestions about our manuscript. We revised the manuscript taking into account your suggestions and comments. Please find attached a point-by-point response, with our answers in blue. We hope that the revised version of the manuscript properly addresses your concerns.

Sincerely,

Ana Oliveira on behalf of all authors

Referee #2 - Warrick Dawes
egusphere-2023-915 "Direct integration of reservoirs' operations in a hydrological model for streamflow estimation: coupling a CLSTM model with MOHID-Land" AR.Oliveira, TB.Ramos, L.Pinto, R.Neves

The work presented in this article is very thorough. It is a valuable addition to the material from Oliveira et al. (2020) illustrating the use of AI/ML techniques, some of which were outlined in Oliveira et al. (2023) when applied to streamflow only. While there are a few peculiarities of wording through the text, the English expression for the most part is very good and it reads well.

We would like to thank referee #2 for the time spent evaluating our manuscript and the positive comments made to our work.

For corrective suggestions, perhaps only the line figures showing flow need to be cleaned up (Figures 2 and 7-9). With daily instantaneous data and small dots connected by lines, the hydrograph becomes a red sludge with the occasional peak that is visually unsatisfactory. Perhaps weekly or monthly volumetric totals would be more distinct, or single years shown as examples of the best/worst fit for the particular solution. This does not mean losing any of the finer daily detail when reporting statistics or minimum and maximum daily flows, as with the current tables.

The original figures will be replaced by new figures where the monthly values are presented, as suggested by the referee. Thus, Figure 2 will be replaced by:

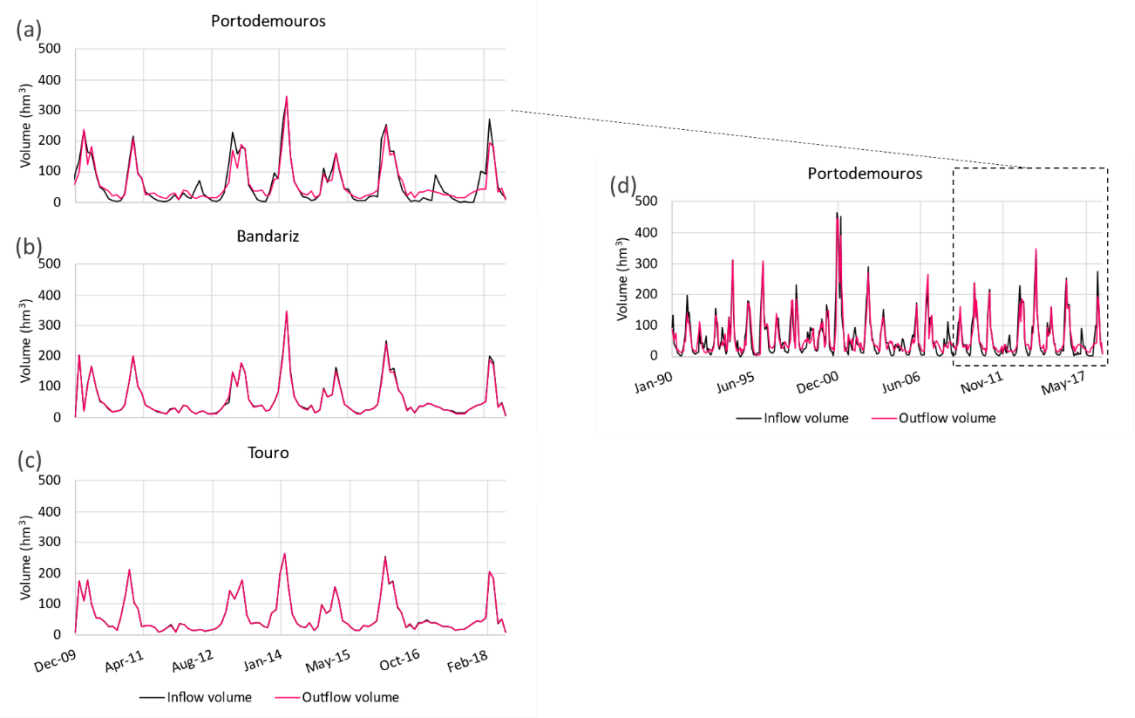

Figure 2 Comparison of inflow and outflow volumes in (a) Portodemouros, (b) Touro, and (c) Bandariz reservoirs for the period 2010-2018, and in (d) Portodemouros reservoir for the period 1990-2018.

And Figure 7-9 will be replaced by:

Figure 7 Comparison of modelled and observed average monthly streamflow in hydrometric stations (a) Ulla-Touro and (b) Ulla-Teo with and without considering the existence of reservoirs. Focus on the daily values for the period between September 2013 and September 2014 in (c) Ulla-Touro and (d) Ulla-Teo hydrometric stations.

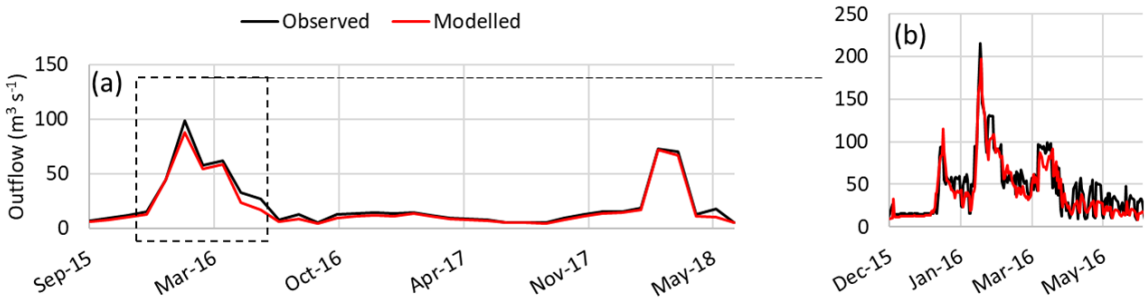

Figure 8 Comparison between modelled and observed Portodemouros outflow considering the CLSTM model: (a) monthly average and (b) daily values between December 2015 and June 2016.

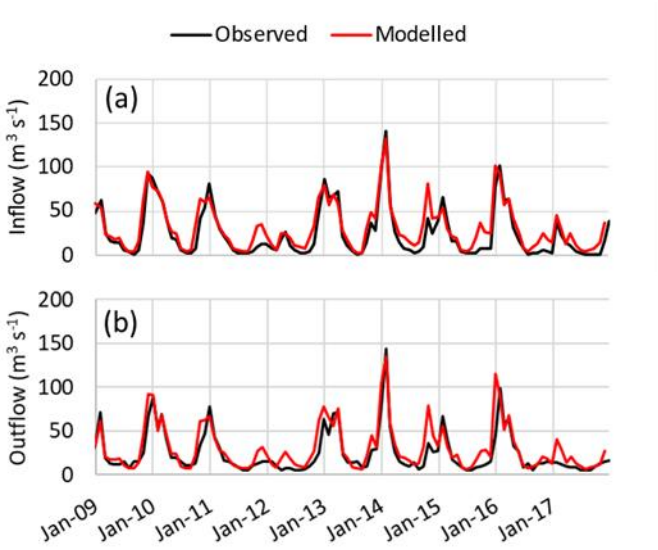

Figure 9 Comparison between the modelled and observed (a) inflow and (b) outflow in Portodemouros reservoir.

Figure 3 is also far too busy with three nearly overlapping lines in each panel. You may need a log scale on the y-axis, omitting all zeroes and the observed data (as the fit is very good).

We imagine that the referee is referring to Figure 7, since Figure 3 just presents the soil maps. Thus, that figure will be replaced as mentioned above.

Most of the questions in my mind are already listed in the Conclusions. Why does the MOHID-Land model not apply evaporation to the reservoirs? This is clearly a significant flux for these structures and part of their water balance. However, the ANN was not trained with reservoir level but only used as an input->output "black box" and it did not have a significant effect on performance. Given the "best" set of weights selected for validation testing, it would be interesting to see the same calculations with the set that best reflected the reservoir level.

This behaviour may be a case of getting a right answer for the wrong reasons (qv. Kirchner, 2006). Clearly it would be preferred that the open-source MOHID-Land model treated the reservoir as a water balance unit with all its attendant fluxes, and that calibration be a combination of flow at a downstream gauging station together with the changes in the significant upstream storages. In the case of this study, the outputs from Portodemouros (gauged?) are the boundary conditions for the river routing through the largely

irrelevant Bandariz and Touro, to the calibration river gauging sites downstream. Thus the ANN only has to ensure that the outputs result in the correct upstream input to do their job. If the modelled flow inputs to Portodemouros are incidentally correlated to the outputs, given the internal storage is not modelled, then the ANN can include them but they may not even be relevant.

As for the addition of irrigation releases, for example, that becomes a very interesting problem with a mix of known storage release rules and the ANN version of other informal flows through the reservoir. It is almost a data integration exercise, where the storage is "managed" with the ANN model but updated periodically with the irrigation rules.

Kirchner, J. W. (2006) Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. Water Resources Research, 42, W03S04, doi:10.1029/2005WR004362

We appreciate the referee's comment and the proposed reference, which was very interesting to read. To take into account the idea presented in the cited work we added the following paragraph in L473:

"This is intimately related with the discussion presented by Kirchner (2006) about obtaining the right results for the right reasons, and where the author explores the limitations of the operational practice of hydrology. In that sense, the coupled system presented here, namely the CSLTM model, seems to be obtaining the right answer but for the wrong reasons. With the behavior of CLSTM model being classified as a "black box", without any physical constraints implemented, its results can be good enough while the model is exposed to conditions similar to those used for its optimization. However, when the forcing conditions go far beyond those used in the optimization, the results of these type of models become unreliable because of their lack of physical realism."