Comment on hess-2021-211
Anonymous Referee #2

Referee comment on "Towards hybrid modeling of the global hydrological cycle" by Basil Kraft et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-211-RC2, 2021

General comments:

This is an interesting paper demonstrating the feasibility of reproducing the simulations of global hydrological models (GHMs) using a hybrid approach (H2M). The latter is based on a toy model consisting of a series of bulk reservoirs, coupled to a statistical model based on machine learning. Results are encouraging. I am not sure that the comparison of H2M with GHMs is completely fair because the precipitation dataset used to force H2M (GPCP) is based on observations, while the one used to force GHMs is derived from the ERA-Interim reanalysis. Another reason why the comparison may be unfair is that spatial resolution of GHMs is degraded from 0.5 degree to 1 degree to be compared to the H2M simulations. Also, it should be emphasized that some GHMs are uncalibrated models. I was not able to do a complete review of this work because some Figures are not readable.

Recommendation: major revisions.

Particular comments:

- L. 5 (Abstract): Is H2M a new model developed in this study? What is the added value of this approach with respect to more traditional modeling approaches? What is the meaning of H2M acronym?
- L. 95 (22 static variables): unclear because 4 lines correspond to static variables in Table 1, not 22.
- L. 113: product?
- L. 166, 174 (softmax, softplus): all readers may not be familiar with these machine learning technical terms. They should be defined.
- L. 199 (model training): more details should be given on the used machine learning approach. Is a local training (one statistical model for each model grid cell) performed or a global training (all model grid cells together represented by the same statistical model)?
- L. 243 (Table 2): CWD and SStor are written here for the first time and were not defined before. A clear definition should be given. The definition of CWD given in the
next paragraph is not clear.

- L. 242 (selection of models): How was model selection made? In Schellekens et al., 10 models are considered.
- L. 284 (Table 3): the period of time for which the comparison was made should be indicated.
- L. 286 (model intercomparison): Could be completed with a water balance Table similar to Tables 7 and 8 in Schellekens et al.
- L. 293: Mean or median scores are little informative in case of non-Gaussian score value statistical distribution. Could you plot score histograms instead?
- L. 340 (Amazon basin): the Amazon area was affected by droughts (2005, 2010, 2015). Are these drought events visible in the simulations performed in this study?