Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC3
https://doi.org/10.5194/hess-2022-345-RC3, 2023

Comment on hess-2022-345

Anonymous Referee #2

Referee comment on "Using simulation-based inference to determine the parameters of an integrated hydrologic model: a case study from the upper Colorado River basin" by Robert Hull et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2022-345-RC3, 2023

This work is an entirely synthetic study to show that what is called the "Simulation-basin Inference" (SBI) can retrieve synthetic parameters, even with some added noise. While there is some value in showing that a parameter inversion procedure works, it should only be a small part of a proof-of-concept paper. This paper has quite some flaws and wasn't correctly marketed.

1. The pure synthetic nature of the study greatly reduced the value of the work. There is a chasm between observations and model space. When going from synthetic dataset to real dataset, you are faced with model mechanism errors and parameter compensations. This may be where machine-learning approaches could help but this paper completely left these challenges untouched. It makes the readers wonder if this approach would ever succeed --- I am not saying it could not, but you should demonstrate you can address the major issues. History has shown that inversion problems can be ill-posed, and a workflow showing perfect performance for a synthetic case can fail fantastically in real-world problems. It is most unsatisfying when you see multiple unaddressed roadblocks in a "proof-of-concept" paper.

2. The abstract and introduction were written that readers may be led to think the model calibrated against observed flows, and the advocated work represents a major breakthrough. However, the lack of true observations undermine many of the statements. Overall, much of the sales language throughout the paper should be significantly revised. To give some examples:
"calibrating them can be difficult" --> it is not calibration if no observations are used.
"confront two recalcitrant issues related to calibrating watershed models" --> both issues remain unaddressed with this paper: the surrogate model isn't perfect and we still don't know how to get *correct* parameters.
"While SBI for parameter determination has shown promise in particle physics","the applications in hydrology have been limited". --> it sounds like SBI is a major solution to our problems. However, many earlier Bayesian methods were proposed in a similar way as SBI. I wonder if it is really necessary to market SBI as a whole or more precisely

emphasize what is novel about it. The main differences from previous methods seem to be (i) previous methods carry distributional assumption while here you have a neural network to generate samples. (however, in reality you still use a Gaussian mixture model in Eq. 4); (ii) you go directly from discharge to these parameter distributions. You can market the NN directly.

3. The authors cited "equifinality". Can the method represent multimodal parameter distributions that can produce the same discharge output?

4. There is again a major gap between the original process-based model and the surrogate model. In this work, even the data for "posterior predictive check" was generated by the surrogate rather than the original model. This leaves everything to the surrogate model. It is well known in surrogate-based model research that surrogate models are never perfect, as the authors actually later showed. However, the author did not do anything to actually address the issue of a deteriorating surrogate model. The authors argued we never know the true conductivity values, but at least they can show how the parameters behave in the original ParFlow model.

Overall, this paper did not comfort me with respect to the potential success of the SBI method going forward. For a proof-of-concept paper, I do not mind a simple case, but at least it needs to be demonstrated that the major roadblocks can be tackled. We do not want to lead the community down the wrong path! I would only consider this paper for publication in HESS when they have a real-world case.