Knoben and Spieler (2021) HESS Review

This work presents a set of (open source - Octave and licensed-Matlab) computational exercises that help teach hydrological model structural uncertainty, particularly model choice as an example of structural uncertainty. As the analysis and coverage of structural uncertainty are limited, this work is a useful contribution. Below are my comments and suggestions.

1. Model adequacy is closely related but different from structural uncertainty (Gupta et al., 2011). Although the authors have indicated the limitation of statistical metrics such as KGE in diagnosing model adequacy (page 4 line 3-18), they used the term adequacy in their core objective plot (Figure 1 – lessons in the three boxes). The work (and Figure 1) is based on the relative performance of two models in two catchments. As such, 'adequate model performance' is not the right phrase to use. I suggest the use of relative terms such as 'better' and/or 'a relatively high' performance.

   Similarly, it is a stretch to use strong words such as 'appropriate' and 'accurate' (on page 4 line 26 and line 31) based on comparative analysis.

2. The manuscript needs to explain why model 'm03' performs better in the two distinctly different catchments while 'm02' performs poorly in one of the catchments. Although the manuscript mentions simulating zero flows and the basis of the models' development, it is important to briefly discuss these points directly. This may support both educators and students to articulate the causes.

3. In this work, 'calibrated' parameters are used to support the comparative analysis. But, it is important to indicate/discuss the non-uniqueness of these parameters and the interplay of parameter and structural uncertainties (Clark et al., 2011; Moges et al., 2021). As separating the two uncertainties is not always straightforward, a brief discussion with references for further reading will be helpful.

Technical comments:
1. It is good not to repetitively use the term "this section describes". If necessary, it is enough to use it once (e.g., the first case on page 3 line 15 – 20). Using this term in other places (e.g., page 3 lines, 24 - 26; page 4 line 2; page 5 line 11) is just a distraction.

2. Page 2 line 21, avoid the use of the term 'For a variety of reasons'. State a few of the reasons or rewrite the sentences.

3. Page 4 lines 19 - 31 referred the catchments and models by their CAMELS and model ID. It is better to first introduce the catchment names, ID and the models' names earlier. Perhaps, on page 3 lines 14 – 15 where the objective of the paper and the experimental designs are indicated.

References:

Clark, M.; Kavetski, D.; Fenicia, F. Pursuing the method of multiple working hypotheses for hydrological modeling. Water Resour. Res. 2011, 47.

Gupta, H.V.; Clark, M.P.; Vrugt, J.A.; Abramowitz, G.; Ye, M. Towards a comprehensive assessment of model structural adequacy. Water Resour. Res. 2012, 48.

Moges E, Demissie Y, Larsen L, Yassin F. Review: Sources of Hydrological Model Uncertainties and Advances in Their Analysis. *Water*. 2021; 13(1):28.