

Reply to the authors' response (AC3) – Referee #3

I thank the authors for their detailed and carefully argued response.

In short, the authors agree with most of my comments and, as is customary, follow almost every agreement with a well-argued "however" that leaves the manuscript largely as it was. The arguments are sound, and several of my comments were questions or requests for clarification rather than identified errors.

I appreciate that the authors have reordered Figs. 1 and 2, so that the surrogates are validated before being used for sensitivity analysis; stated explicitly that the high R^2 values result from in-sample evaluation; made the cost reduction, from $n(p+2)$ to n model evaluations, explicit in Section 2.4 rather than only in the conclusions; added bootstrap confidence intervals to Fig. 4 and Fig. 6a; discussed their results in light of new references, in particular the role of surrogate fidelity for more complex, higher-dimensional models; added a sentence on the generality of the approach beyond hydrology, and on independent validation and transferability of surrogates as future work; clarified the Fig. 2 caption.

The correction of the metamodel type (a GBDT rather than a Random Forest) is also welcome and shows the care taken in the revision.

I also accept the authors' arguments on the remaining points:

- In-sample evaluation: The PVI computation uses perturbed inputs on which the surrogate was not trained, and the resulting estimates agree with the reference T_i . Together, these act as a de facto validation for the intended use. This is consistent with the GSA–ML literature cited.
- Computational cost: The explanation that matrix A serves double duty, while the surrogate replaces the model evaluations on B and AB_i , answers my concern. My comment mainly reflected a lack of explicitness in the original text, which the planned addition to Section 2.4 will resolve. If space allows, briefly reporting the model run times in the paper itself, and not only in the repository description, would make this point even more tangible. I leave this to the authors' judgement.
- Uniform distributions, independence assumption, KGE as output: The authors' justification is convincing: Eq. 8 holds for independent inputs whatever their marginals, and the study is a methodological test of that relationship rather than a full sensitivity analysis of the models.
- SHAP values versus Shapley effects: The distinction the authors draw is clear and useful. It would be worth keeping a sentence on it in the revised text, since readers from the GSA community may have the same initial confusion.

References. The respective roles of Gregorutti et al. (2017) and Wei et al. (2015) are now clear. The notation T_i follows an established convention, and the justification for citing Puy et al. (2022) is acceptable.

- Length and the SHAP section: The precedents given for recent HESS technical notes are convincing. I leave the question of manuscript type to the editor. Given the improvements announced in response to the other reviewers, I understand the choice to keep and strengthen the section on SHAP values.

I have no further requests (except expliciting the model run times, let this to the authors' judgement). I recommend acceptance once the revisions announced in the response are implemented.