Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-215-RC1, 2020

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "Simultaneously Determining Global Sensitivities of Model Parameters and Model Structure" *by* Juliane Mai et al.

Anonymous Referee #1

Received and published: 3 July 2020

Summary The authors introduce a novel sensitivity analysis method, called Extended Sobol' Sensitivity Analysis (xSSA), that advances upon existing procedures in several ways: (1) it can provide insight into the sensitivity of individual model structure choices; (2) it can clarify the relation between parameter and structure sensitivity; (3) it can account for cases where model parameters are present or absent in different model structures; and (4) it is much faster than alternative methods. The main novelty of xSSA is that it estimates parameter/process sensitivity inside a flexible modelling framework (Raven), which allows the sensitivity estimates to be re-combined through weighting. On any given timestep, thesimualted states and fluxes can thus be based on multiple different parametrizations of the same process, depending on how the weights are set.

Printer-friendly version

Discussion paper

The authors test xSSA against two cases where analytical estimates of sensitivities can be derived (one case where each parameter only occurs once in all possible flux parametrizations, and one case where parameters are shared between multiple flux parametrizations), and in a real-world application of the Raven framework in a single watershed. They find that xSSA converge to the analytical solutions in both test cases, while current methods are only able to converge in the first test case. The real-world test case is used to showcase why process-based SA can be useful. I've read this paper with great interest. Model structure uncertainty is receiving considerable attention and this extension of existing SA methods to take advantage of modern multi-model frameworks is a welcome and timely contribution. Overall, the paper is easy to read but I have outlined various comments that can help the authors clarify their message. In general I think all the required information is there but some polishing would make the manuscript much more accessible for readers who are not so well-versed in Sobol' SA and Raven as the authors are.

General comments

The results section relies heavily on understanding of the Raven functions. It would be very helpful if the authors expand on the model description in section 2.1.2 or appendix C, by including the actual equations or descriptions of each parametrization.

The results section relies on an understanding of each process to interpret model sensitivities. It is not entirely clear to me what each process includes. Can the authors clarify this by briefly explaining what each process in Figure 1C and further figures includes? For example, how does process 8 (potential melt) relate to process 5 (snow balance)? I don't think these explanations need to be very long, but it would be good if they include a bit more detail then the 1-3 words they currently get.

I would encourage the authors to be careful with words such as "appropriate" and "important" in the manuscript. To some of the community, "appropriate process representations" might mean "process representations that are an accurate mathematical

description of the real-world". To the authors (I believe) this instead means "equally sensitive, so equally good choices" (e.g. L538). Similarly, "important" seems synonymous with "high sensitivity" in this manuscript, but I don't think having high sensitivity over an arbitrarily wide parameter range necessarily dictates importance for matching a specific set of observations. Therefore I would strongly recommend the authors to go through the manuscript and either define such words clearly, or avoid ambiguity by being more specific in each case where such words are used (e.g. change "soil and surface processes are of secondary importance for streamflow prediction, . . ." to "simulations are less sensitive to soil and surface processes, . . ."; L556).

Line by line comments

L32. It might be more accurate to refer to "input (forcing) uncertainty" as "data uncertainty" or "observational uncertainty" to acknowledge that uncertainties are also present in model evaluation data such as streamflow observations. See e.g. McMillan et al. (2012, 10.1002/hyp.9384).

L47. It would be helpful to the reader if the authors could summarize the Baroni method in one or two sentences.

L53. "The method introduced . . ." some text is missing here.

L94. This special Raven property is a bit unclear to me. What dimension are the simulated fluxes weighted over? Is this a weighted average across multiple parameter sets, model structures, something else? – If this property is critical to the functioning of xSSA I think it should be explained in more detail here. Perhaps an example can be added.

L103. I appreciate what the authors are going for, but "unconditional parameter sensitivity" is too broad a statement. The answers to questions A-D will be conditional on the catchment(s) being considered. It would be good to acknowledge that somewhere in lines 99-103.

L205. I'm somewhat confused about this statement. One does not need to run 12 fixed model structures but instead needs to run a single flexible structure that contains all the options that are present in the 12 models. How does this reduce the number of computations required? As far as I understand, it's still the same elements that are being tested. If the authors mean that all elements can be tested independently (implying that if and how they are connected to other elements can be ignored), than why would they need to be part of a model structure at all? Why not test each element in isolation and recombine the results through the proposed weighting? This could result in even further computational savings in cases where the same parametrization can be used in multiple processes (quite common in bucket models, possibly also in physics-based models that discretize snow/soil into multiple layers).

L212. Caption of Figure 1. "The three processes are connected through A.B+C (C.D+E) . . . " Text in the brackets should read (D.E+F).

L212. Caption of Figure 1."Processes A (D) and C (E) . . . " (E) should be (F).

L213. Which numerical scheme does Raven use to solve its model equations?

L262. "forcings" > "forcing"?

L269. Why were only 20 years of data used if 56 are available? Wouldn't more data give a more complete assessment because a wider range of conditions is (likely) covered?

L308. It took me a while to figure out that these numbers are: # of models x (# of parameters +2 x K), mainly because the order of operations is reversed compared to L307 (which gives # of parameters first and # of models second) and because the operation K x (N+2) from L303 has already been completed. I'd suggest to clarify this.

L350. The authors use analytically derived Sobol' scores for their shared-parameter model setup. Can these derivations be made part of the appendices or can the authors provide a reference to a paper that provides these?

L358. Single-sentence paragraphs look strange. Suggest to merge with preceding paragraph.

L364, L366. I was under the impression that the shared-parameter models was being tested. Why do these sentences refer to parametrizations A, B and C instead of D, E and F?

L429-441. I find this section difficult to follow, in part because it was not clear to me that the Baroni method uses a regular Sobol' approach. The only mentions of Sobol' so far (I believe) have been in relation to xSSA and the mention of Sobol' analysis on L434 threw me off. I'll repeat my earlier comment that a brief description of the Baroni method would be very helpful in understanding these results.

L435. "This contradiction cannot be resolved." Is it part of the Baroni method to include a single parameter twice? In my (admittedly limited) experience with the regular Sobol' method, one would include any parameter only once, regardless of how many times it occurs in the model processes being considered. This would mean that processes cannot be assessed individually if they share a parameter (which the authors already mention) but getting into this situation in the first place requires that one is looking to investigate processes, not parameters. I think the authors can make their reasoning stronger by repeating here that investigating process sensitivity requires a different approach then parameter sensitivity in cases where parameters are shared between processes.

L451. It might be good to add a reference to sensitivities of non-additive models not summing to 1. I seem to recall this is discussed in Saltelli et al. (2008) for example.

L461. "hence" > "this"?

L499. Suggest to delete "and hence most sensitive"

L509. It might be instructive to adapt the x-axis in Figure 5B, so that it shows which parameters (x-axis in 5A) are included in each process option in 5B. This could clarify

whether process sensitivities can be traced back to particular parameters.

L515. “Same” > “The same”

Figure 5. It might be worthwhile to change the orientation of these plots so that the Sobol’ scores are on the x-axis and the parameters/parametrizations/processes are on the y-axis, so that these are easier to read. I currently need to tilt my head back and forth to read the results in 3.3.2 and compare them to the axes in Figure 5.

L525. “The latter serves as a consistency check of the implementation.” Can the authors clarify what they mean here? – upon reading further, it might make sense to swap this sentence with the one immediately after it.

L527. I admit I’m a bit confused that model outputs of process representations do not change in different model runs. Because this process representation is connected to the rest of the model, and there are changes in the contributions of other processes as a result of different parameter values, wouldn’t it be expected that the model states change as well, and as a consequence, that the contribution of this particular process to overall simulations changes too? Without knowing with SNOBAL_SIMPLE_MELT actually does, I assume that even if it has a constant melt rate, it is still constrained by snow availability and thus cannot produce a time-invariant flux. I would expect such a case (no parameters in a given process, but influenced by other parameters by virtue of being part of a bigger model) as showing in a 0 Sobol’ main effect, but a non-zero Sobol’ total effect. Can the authors clarify this?

L538. “The three infiltration options are equally sensitive and hence equally appropriate.” Logically, only one or none of these infiltration options is appropriate (in the sense of accurately representing the real world). I also doubt that high sensitivity automatically indicates high appropriateness. I suggest to rephrase this sentence.

L538. “quickflow” > should this be “infiltration”?

L545. Can it be said that rain-snow partitioning is a forcing correction function? It does

not change the water balance, only the phase and thus by extent, the timing of liquid water availability.

L556. “This demonstrates that soil and surface processes are of secondary importance for streamflow prediction, . . .” Is this true? As far as I understand, the SA only shows that impact of parameter changes on the variability of the simulations. I don’t think relatively low sensitivity automatically indicates low importance for accurate streamflow simulation, because (1) no simulations have been compared to observations; (2) parameters ranges might be wider during this SA than their “real” range of values and thus much of this variability might occur in regions of the model output space that are far away from the observations. I would recommend slightly more careful phrasing, like used in L559.

L677. These are not author contributions.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-215, 2020.

Printer-friendly version

Discussion paper