Authors' Response to Reviewers of "Guidance on evaluating parametric model uncertainty at decision-relevant scales" by Smith et al.

Reviewer comments are in black

Author responses are in blue

Proposed text edits are in red

RC2

1. L8-10: "We evaluate six sensitivity metrics that align with four decision objectives; two metrics consider model residual error that would be considered in spatial optimizations of engineering designs." This sentence is confusing -- do the six sensitivity metrics add up to the four decision objectives + two model residual error metrics?

   Yes, that is correct. We propose modifying the sentence to:

   We evaluate six sensitivity metrics, four of which align with decision objectives and two of which consider model residual error…

2. L39: extremes or high flows?

   Engineering controls can work to lower high flows and raise low flows. We propose changing to:

   Extreme high and low flows

3. It would be helpful to somewhere define "engineering controls".

   We propose adding examples to the first sentence in the introduction.

   (e.g., green and gray infrastructure)

4. L 116: here are you write flooding, low flow, reservoir water supply objectives, but earlier you wrote flooding, low flow, and all other flows. If these are the same, that should be explicitly stated.

   We propose changing L116 to:

flooding, low flow, and all other flow objectives

5. L320: “The goal of this sensitivity analysis is to inform the selection of parameters to calibrate a RHESSys model that could be used in such a reforestation optimization.” is this the overall goal of this paper? If so, this should be stated in the introduction section much earlier.

   We propose to state this at the end of the first paragraph of the introduction:

   In this paper, we evaluate the influence of decision-relevant and calibration-relevant sensitivity metrics on parameter selection for calibration, and discuss the potential implications on subsequent model calibration and optimization of water management decisions.

6. L 328: are the elementary effects for all the parameters normalized on a percentage basis? Why compare the 95th percentile for the elementary effects to the overall mean of all parameters’ elementary effects, if that is what is being explained in this sentence? What does the 95th percentile estimate for the elementary effects mean?

   As explained in Section 2.5, we completed bootstrapping of the elementary effects to generate a distribution of mean absolute values of elementary effects for each parameter. This recognizes that there is uncertainty in our estimate of a parameter’s mean absolute elementary effect. From that distribution, we obtain the 95th percentile estimate of the mean absolute value EE for each parameter. Normalization in each panel of figure 2 is completed by taking the maximum 95th percentile EE value across all parameters and setting that to 1 (i.e., all EEs are divided by this value). The minimum is 0. So, the figure normalization is not on a percentage basis. It should also be noted that the EE is a normalized metric to begin with, as it is the absolute value of the change in the output metric per change in the input parameter, where the change in the input parameter is 50% of its range in our study.

   Also stated in Section 2.5, we do not compare 95th percentile EEs to the overall mean across all parameters EEs. We sort the mean absolute EE values across parameters from largest to smallest and find the top X%. We then compare each parameter’s 95th percentile estimate of its mean absolute EE to the X%-ile. Any parameter whose 95th percentile mean absolute EE estimate is above that threshold is selected for calibration.

7. Figure 3 seems to be referred to before figure 2 (L335).

This reference to Figure 3 is to state that we will discuss a point in more detail later in the same section. We plan to leave this as-is unless asked to change.

8. L349: does an elementary effect value of exactly 0 mean that this parameter has no effect on the stream flow or hillslope metric? It would be helpful to state this explicitly.

   Yes, it does. We propose adding the following to the end of the sentence:

   (i.e., these parameters do not affect model-predicted streamflow)

9. The text discussing figure 2 is useful, (line 348 in the rest of this paragraph), but without knowing what the specific parameter numbers are in figure 2, I'm not sure what to take from this graphic.

   The supplementary material provides the full list in number order in a spreadsheet. We discuss the parameters with the largest elementary effects within the text that you mention.

10. Line 480: I see now that engineering designs are not explicitly evaluated in this paper. My earlier comment (comment 3), asked about what engineering controls meant. The focus on engineering controls in the introduction section led me to believe that this paper would be about engineering controls. Rather it seems that this paper has implications for where to locate engineering controls but does not directly investigate this placement. If this is accurate, then I would suggest deemphasizing engineering controls from the introduction section.

    This is correct that we use siting of engineering controls as a motivating reason for doing a spatially distributed sensitivity analysis. Calibrated models are used to optimize these and other water management decisions, and parameter screening is used to reduce the dimensionality of the search to make the calibration more tractable. We propose broadening the introduction to say "engineering controls and water management decisions" to be more applicable. We propose changing the first sentence of the introduction to:

    Spatially distributed hydrologic models are commonly employed to inform water management decisions across a watershed, such as the optimization of locations of engineering control measures.

11. From what I understood of this article, the first main finding was that parameters describing watershed characteristics are sometimes important for modeling hillslope

hydrologic response even though they do not affect the streamflow at the model outlet much. The authors state that this might be important in the spatial location of engineering controls. There are many other reasons why getting the hydrology right within the watershed is important (modeling of spatially distributed soil moisture, etc.), but a major limitation is that we don't normally have data to compare to within the watershed, so in practice it would be hard to calibrate these parameters that don't affect streamflow much. The second main finding was that commonly used metrics (e.g., NSE) are not as sensitive to the decision relevant streamflows that we would want them to be. These are both important findings and points to make, but I found the article overall hard to read and understand. The authors may be served by focusing the text on the main findings and reducing discussion of peripheral topics.

These are two key points of the article. Because there are so few papers on decision-relevant sensitivity metrics, we thought it would be useful to provide an extended discussion that describes the importance of having the decision objectives in mind when completing a sensitivity analysis and subsequent calibration and optimization. Spatial sensitivity analysis is also often limited by data in sub-catchments and can lead to calibration challenges. We think it is important to discuss how the resulting parametric uncertainty can be used in robust optimizations.

We propose a better framing of the intent to use SA results to inform calibration of a model that is used to optimize decisions. We propose adding this sentence to the end of the first paragraph (same sentence as mentioned in reply to comment 5):

In this paper, we evaluate the influence of decision-relevant and calibration-relevant sensitivity metrics on parameter selection for calibration, and discuss the potential implications on subsequent model calibration and optimization of water management decisions.