Hydrology and
Earth System
Sciences

Discussions

Open Access

Interactive comment on "Identifying the connective strength between model parameters and performance criteria" *by* Björn Guse et al.

R. Arsenault (Referee)

richard.arsenault.1@ens.etsmtl.ca

Received and published: 1 April 2017

This study uses regression trees in a bi-directional framework to estimate the importance of a model parameter in a set of objective functions, as well as the relative importance of each parameter to a given performance measure. The authors conclude that this method permits identifying model parameters with respect to certain performance measures, and I agree with their assessment. I found this paper to be interesting and generally well written. I have a few important comments which I would like to see addressed before considering publication in HESS. I recommend a major revision.

First, the author's RT based method should ideally be compared to other proven techniques to estimate parameter importance, such as Sobol' sensitivity analysis (or any global sensitivity analysis). If not integrated into the work directly, differences in ex-

pected outcomes should be addressed in the literature review. I recognize that the bi-directional aspect of this work is novel but any advantages of this method should be compared to a proven baseline.

Page 4, lines 4-6: This sentence is very confusing, please rewrite differently.

Page 4 lines 7-8: This means that the model parameter has no relevant impact on other performance measures. Perhaps give a clear example of how this can be achieved in the case of a hydrological model with highly interacting parameter sets. My previous work in parameter identifiability suggests that a large part of the relative importance of a parameter on a performance measure comes from its interactions with other parameters.

Page 7, lines 4-5: By using 2000 samples with hypercube sampling, are the authors not effectively working in spaces where parameter combinations might not make physical sense? Usually the model parameters, during calibration, will self-regulate to attain sensible parameter values. With a LHS approach, perhaps some combinations are tested here which are out of the bounds that the model can work with appropriately. More information regarding this aspect would be interesting.

Furthermore, the parameters do not seem to be normalized in their ranges, therefore allowing some parameters more leverage over the performance measures. If I interpreted this correctly, then some of the results would be trivial since the larger boundaries will naturally have more effect on the performance measure and thus the parameter will be more "important". The use of a LHS methodology in an uneven search space will bias the results (as an extreme example, if ESCO bounds were set between 0.995 and 1.005, then the parameter would definitely not be considered important). The choice of boundaries, then, induces a methodological bias in the results. I am not sure how to solve this problem, perhaps by performing multiple calibrations and taking the envelope of the parameter sets, but this also has its drawbacks.

Also, the parameters seem to be evaluated on the entire time series. In a snowmelt-

dominated catchment, the parameters are highly time-variant. How could this affect the method's robustness?

I think Figure 1 can be omitted completely without any loss of information in the paper. It is fairly well described in the text.

Page 12, lines ~20-25: I have the feeling that some of these strong connections are trivial. If I had had to guess in advance, I would have guessed that Evapotranspiration (ESCO) is probably strongly linked to bias (KGE_beta), and that mid flows and lower were also affected by baseflow recessions and to some extent evaporation due to the relative scale of a fixed evaporation rate on total available volumes. Once interactions are important, then the method seems to "get lost" in a sense, as there is no clear path to identifiability (as demonstrated in the discussion). I think sensitivity analyses would provide the same information while also informing on the different order sensitivities.