Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-249-RC3, 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Discussions

Interactive comment on "An integrated multi-fingerprint sensitivity-nested approach for regional model parameter estimation and catchment similarity assessment" *by* Simon Höllering et al.

F. Sarrazin (Referee)

fanny.sarrazin@bristol.ac.uk

Received and published: 18 July 2016

In this manuscript, the authors present a parameter estimation and sensitivity analysis scheme to learn about differences in catchment functioning and to classify catchments. However, many points need clarification, in particular the implications and conclusions of the work. I recommend major revisions of the manuscript.

Parts of the manuscript appear to be quite long (specifically the introduction and results sections) and the reader tends to get lost. A general recommendation to the authors is to select in each section the key elements that contribute to their argumentation and

to clearly connect them. The authors should also try to keep their sentences short to improve readability.

Furthermore, critical information on the implementation of sensitivity analysis is missing (output considered, how the sample size was chosen), which makes it difficult to interpret the sensitivity analysis results.

I provide below more detailed comments for the different sections and some minor comments at the end.

SECTION 1 introduction

1) I think the introduction section is unnecessarily long and not sufficiently linked to the objective section (section 2). Specifically, how do the work of Bardóssy (2006) (p2 L11) Wagener et al. (2007) (p2 L18), Castiglioni et al. (2010) (p4 L14), Yadav et al. (2007) (p4 L18) relate to the work presented in the manuscript? In which way the literature review on catchment classification (section 1.2) is useful to better understand the work presented in the manuscript? I suggest the authors select key elements in the literature review and clearly link them to their study, so to clarify what the contributions of the manuscript are.

2) There is some confusion in Section 1.3 and it needs clarification. A distinction has to be made here between a model and the underlying system it represents. The term equifinality commonly refers to the model and not the system it represents. In fact, Beven (2006) (p21, paper cited by the authors) characterizes the term equifinality as the 'rejection of the assumption that a single correct representation of the system can be found given the normal limitations of characterisation data.' Therefore, equifinality is not an 'inherent' property of a model equations as stated by the author p3 L31, but it is due to the fact that the information content in the data is not sufficient to identify a unique model representation. I am also asking the question: Is the study of physical and hydrological similarities based on observation data or on model simulation outcomes?

3) p4, L20: the expression ‘to raise the information considered in the parameter transfer’ is quite fuzzy. Please clarify.

SECTION 2 Objective and driving research questions

1) As mentioned above, the objectives need to be clearly linked to the literature review, to show the contributions of the manuscript.

2) p4 L30-31: This objective need clarification. First, ‘consistent manner’ is vague and it is required to better explain this expression. Second, the motivation need to be clarified. Is the objective to learn about the model ability to reproduce the data, to learn about the value of observation data etc.? 3) p5 L8-9: the expression ‘within the range of their independent variable’ needs clarification.

SECTION 3 Concept and methodological approach

1) I have reservations regarding the implementation of the sensitivity analysis (section 3.3), specifically:

- p7 L12: How was the sample size chosen? It is necessary to check the convergence of the results, that is to say to assess to what extent the results would change if using a new sample. When the sample size is too small, the sensitivity analysis results can be unreliable (e.g. Sarrazin et al., 2016). A way to assess convergence is to derive confidence intervals on the sensitivity indices using bootstrapping. However, this technique cannot be easily applied when using FAST, given the structure of the sample. Therefore, I would suggest to simply look at the stability of the results when using smaller/larger sample sizes.

- It is required to specify the scalar output used for sensitivity analysis, otherwise the results are meaningless. Sensitivity analysis results can strongly vary when considering different model outputs (e.g. van Werkhoven et al., 2008).

- Why analysing main effects (FAST) only and not total effects (eFAST)? A parameter could have an effect through interactions only and therefore would not be detected

when applying FAST.

2) In Section 3.4.1 several points need clarification:

- p8 L3: ‘sensitivity confined’ is unclear.

- p8 L4: this sentence is vague

- p8 L5: this sentence have to be revised in relation to section 1.3.

3) p9 L10-25: This is quite a long paragraph and I am not sure in which way it connects to the work presented in the manuscript. Again, I suggest to select the relevant information that helps to better understand the present study.

SECTION 4 Study area

1) The authors should present the data used for the simulations and precise the simulation time horizon chosen for the analyses. Also, why presenting some of the data used before presenting the study area? (section 3.3.2, p7 L15-21)

SECTION 5 Results

1) A general comments for this section is that some figures are redundant and could be removed for the sake of brevity (e.g. Figure 6, 9 and 10). Likewise, the text could be more concise.

2) p11 L13: Why is it a ‘qualitative constraint’?

3) p11 L29-30: The behavioural subset is the lower branch of the ‘arrow’ or the points that are within the circles?

4) p12 L2-3: What are the implications of the statements ‘which portends [. . .] specifically.’ and p12 L10-11 ‘Circles [. . .] varying strength’?

5) How does section 5.3 relates to section 5.1 which is also about constraint on parameter space?

6) In section 5.3, I think it may be more appropriate to identify for each fingerprint and pair of fingerprints not a unique most behavioural parameter set but an ensemble of best performing parameter sets. In fact, observation data are affected by uncertainties and it may not be relevant to make distinctions among the top performing parameter sets.

7) Why keeping correlated fingerprints in the analysis of section 5.3? For instance, CV or HPC could be remove from the analysis since the two fingerprints have the same information content.

8) p14 L12: Please clarify how the five most behavioural parameter sets were determined.

9) p15 L24-26: This statement is not correct. When referring to Sobol' variance decomposition (Sobol', 1990), the sum of main effects is equal to 1 when main effects only contribute to the total variance (no interactions). A sum of main effects for the six parameters analysed equal to 0.43 means that a significant fraction of the total variance (1-0.43=0.57) is due to interactions between parameters. If more parameters were added to the analysis, this would not necessarily result in an increase in the sum of the main effects, since the output variance woul also change. The same applies for the statement p21 L11-12.

10) Section 5.4.1: I have reservations regarding the interpretation of the sensitivity analysis results:

- p16 L4-10: I do not see any clear summer/winter pattern for the sensitivity of Aspect-corrPET and Recharge Coeff for WEN but more 'ups and downs' all year round.

- p16 L26-35 - p17 L1-11: I am not sure about the significance of Figure 11. The authors consider the most influential parameter only. However, the difference in sensitivity among the most sensitive parameters may be small and even not statistically significant given the approximation error in the sensitivity index values. This is all the

more concerning, since all the sensitivity indices take quite small values from Figure 9 (below 0.12).

11) p17 L10-11: The sentence 'Nevertheless [. . .] on streamflow generation.' is very fuzzy. It is required to clarify. I am also still wondering why the authors chose to study the main effects only and not the total effects.

SECTION 6 Discussions and Conclusions

1) I suggest splitting this section in two, with a discussion section and a concise conclusion section. This would help to highlight the contributions and implications of the work.

2) p18 L33: Either an indicator is normalised or it is not normalised but it cannot be 'less normalised'.

3) p19 L10-14 What are the implications of the statements 'The composition [. . .] in the eastern headwaters.'?

4) p19 L18-19 'by model [. . .] of meteorological forcing data' is quite fuzzy. Please clarify.

5) p20 L22: The authors do not actually present any result on catchment classification

MINOR COMMENTS

- title of section 1.3: replace 'inconconsistency' by 'inconsistency'

- there is an error in the reference to Beven (1990). Are the authors referring to

Beven, K: Changing ideas in hydrology - The case of physically based models, Journal of Hydrology, 105, 157-172, doi: 10.1016/0022-1694(89)90101-7, 1989.

Or

Loague, K.: Changing ideas in hydrology - The case of physically based models - Comment, Journal of Hydrology, 120, 405–407, doi:10.1016/0022-1694(90)90161-P, http://linkinghub.elsevier.com/retrieve/pii/002216949090161P, 1990.

- p5 L8: replace ‘of the selected fingerprints’ by ‘to the selected fingerprints’

-p5 L12 and L23: what do the authors mean by ‘dependent’?

- p5 L13: replace ‘of simulated streamflow and of the related fingerprints’ by ‘to simulated streamflow and to the related fingerprints’.

- p6 L21-23: Why introducing eFAST here if it is not used? I think the sentence can be removed to keep the manuscript concise, unless eFAST is actually used.

-p8 L19: aren’t there five dynamic fingerprints?

-p10 L14-15: it is required to reformulate this sentence.

- p14 L31: replace ‘to small’ by ‘too small’

- p16 L25: I don’t think the term ‘interacting inversely’ is correct since only the main effects (and not parameter interactions) are analysed here. Please reformulate.

REFERENCES

Sarrazin, F., F. Pianosi, and T. Wagener. 2016. “Global Sensitivity Analysis of Environmental Models: Convergence and Validation.” Environmental Modelling & Software 79: 135–52. doi:10.1016/j.envsoft.2016.02.005.

Sobol’, I.M. 1990. “Sensitivity Estimates for Nonlinear Mathematical Models.” Matematicheskoe Modelirovanie 2, 112-118 (in Russian), Translated in English (1993). In: Mathematical Modelling and Computational Experiments 1: 407–14.

Van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang. 2008. “Characterization of Watershed Model Behavior across a Hydroclimatic Gradient.” Water Resources Research 44 (1): 1–16. doi:10.1029/2007WR006271.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-249, 2016.