Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-249-RC1, 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

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

B. Guse (Referee)

bguse@hydrology.uni-kiel.de

Received and published: 15 July 2016

Review of the manuscript "An integrated multi-fingerprint sensitivity-nested approach for regional model parameter estimation and catchment similarity assessment" by Höllering et al.

In this manuscript, Höllering et al. provide an approach to obtain a better understanding of model parameter behaviour and a more process-based model parameter estimation under consideration of spatial variations. For this, they used at first a fingerprint analysis to investigate how well the model performs for different aspects of the hydrological

Printer-friendly version

Discussion paper

CC BY

system. Second, a temporally resolved parameter sensitivity is used to detect the dominant model parameters along the time series.

Overall, I really like the idea of this approach. However, I think that the core ideas of this study needs to be clarified. The introduction is not appropriate and needs to be reworked (including references). Furthermore, I think that the presentation of the results in the figures can be improved (also in their quality) and condensed, while the interpretation towards the overall benefit for the hydrological community can be enhanced.

Thus, I recommend a major revision of this manuscript. For this, I make several recommendations below.

MAJOR COMMENTS

Introduction:

The introduction needs to be completely reworked and restructured. I see here several reasons for this recommendation.

The introduction is not related to the abstract. It is very surprising that the introduction starts with catchment classification / similarity after reading the abstract. Further, in the methods and results, the performance (fingerprint) analysis and the temporal parameter sensitivity analysis have at least the same priority.

An introduction into the state-of-the-art in parameter sensitivity and more specifically on temporal parameter sensitivity analysis is completely missing. There are several studies in recent years using a temporally resolved sensitivity analysis and trying to extract helpful information for parameter understanding. Please see among others: Guse et al. (2014, 2016), Herman et al. (2013a, b), Massmann and Holzmann (2012), Massmann et al. (2014).

Furthermore, I am not satisfied with the introduction into performance analysis in terms of fingerprint analysis and parameter constraints. Also here, there are several recent

CC BY

advances which need to be considered here to see this study in the context of the state-of-the art in research. Exemplarily please see: Euser et al. (2013, 2015), Gharari et al. (2014), Pfannerstill et al. (2014), Pokhrel et al. (2012), Reusser et al. (2009).

Thus, it is certainly required to extract in a better way the novelty of this approach in the context of temporal parameter sensitivity analysis, performance analysis and constraints for model parameters compared to the state-of-the art.

Objectives:

Following of my comments on the introduction, the objectives are not clearly enough motivated. Please check in the introduction whether all objectives are really motivated in the introduction. According to the current version of the introduction, I do not see a clear reason why it is relevant to constraint parameters in relation to different flow conditions or whether it is required to look at parameter sensitivity to understanding spatial distributed catchment behaviour. Certainly, I agree to both research questions, but I do not agree with their motivation.

Concept and methods:

According to my understanding of the manuscript, the three pillars are not really representing the article. The major point seems to be the third part. The second point is not very specific (streamflow generation). I would recommend to emphasize here more the general idea by shortly explaining how the different parts are related and which benefit is intended to obtain by the different steps. This is currently a bit unclear. Especially when presenting the concept, a clear structure is required. Maybe a flowchart would be more helpful than a list of five steps.

The link between the introduction and the concept is also not clear. Catchment classification dominates the introduction, but is not included in four of the five research steps.

The authors should think about directly considering the relationship between parameter

CC BY

sensitivity and dynamic fingerprints. This aspect is somehow missing at the end and could provide helpful insights into parameter understanding.

Results:

There is a huge amount of figures showing the results. My impression is that due to the content of the figures (and the different subplots), the overall goal of this study is somehow lost. I strongly recommend to focus on figures showing the major outcomes.

In the results (5.3), it is not clear which knowledge is really gained by using the best selected model runs for a certain fingerprint (Fig. 7a). I think that we are more interested in overall best performing model runs combining different indices than in having a model runs which is the best in relation to a single fingerprint.

I had expected a presentation of a joined metric combining all fingerprints such as e.g. shown in Pfannerstill et al. (2014) or Haas et al. (2016), so that a model run could be finally selected which performs well for all fingerprints. In my opinion, this would be a reasonable final result of this part.

A discussion of the results in the context of the state-of-the-art and of how these results are related to knowledge obtained by former studies in this topic is missing. Concerning this, please see among others the list of references at the end.

I really like the expression Sensitivity duration curve (SDC). It is especially good to see that in this case differences between the catchments were detected. This is contrast to Guse et al. (2014) where a similar presentation (Fig. 6, even when it was not named Sensitivity duration curve) did not show relevant differences (due to a lack of spatial heterogeneity).

Maybe the authors could think about a final figure summarizing the results qualitatively. An example for this could be found in Fig. 9 in Herman et al. (2013a).

Discussion and conclusion:

CC BY

Concerning the stated research questions, I think that the first research question is not really solved. I agree that fingerprints can help in constraining parameter ranges. However, this was also expected, since each hydrologic metric can somehow constraint parameter range. Here, I miss either a method to select overall behavioural parameter sets based on all (or all not correlated ones) fingerprints or a hydrological explanation that a certain fingerprint is able to constrain a parameter range since it represents the associated parameter accurately or something similar.

The sentence on P. 18, L. 26-27 "We further found..." really makes a strong difficulty apparent. The results that the parameter values (or constraints) are largely varying between the different behavioural parameter sets are problematic since it is then difficult to estimate the "best" parameter values. I encourage the author to discuss this point more in detail by suggesting possible way towards an overall behavioural parameter set.

In this context, I think that a more profound discussion of the relationship between the performance of fingerprints and model parameters would be helpful. At the end of the answer to research question 2 (P. 19, L. 6-16), the consistency between sensitivity duration curves and hydrologic fingerprints can be discussed. Do the spatial patterns of both are consistent?

Figures:

All figures: Overall, please check which are the most important figures and which are of less importance and could be removed/reduced.

In relation to this, several figures are not clear enough in terms of their intention and the visibility of the results. In particular, the description of the results in the figures (e.g. Figs. 5 and 7) is sometimes difficult to grasp. Concerning this remark, please see also the following comments to the figures.

Fig. 5: This figure needs to be improved. It is impossible to extract the information

CC BY

of the relationship between same coloured points and the circle. One idea could be a reduction of the selected stations (in this plot) and/or a quadratic plot (increase of the plot height). Another idea could be to add a table (maybe as acknowledgement) stating how many points are within the circles. To summarize this comment: It should be possible to extract the information of how many points are in a circle somehow.

Fig. 6: I do not understand why and how the sample space is defined by the observed quantities in a fingerprint. Here, a clear approach is missing. I can agree that in the best case all simulations should be in the range as defined by the observed values of all stations. However, how is this related to the parameter space? This requires at least a detection of the parameter values leading to the fingerprint values as well as a clear relationship between parameter value and fingerprint. Maybe I understand something wrong here, but for me it seems to be that an information is missing here. It is really crucial for this study to understand this point.

Fig. 7a: This sub-figure needs to be completely reworked and improved. I cannot extract the relevant information. There is too much information: Seven sites as colours, six fingerprints as symbols and two metrics as well as the number of the best performing model run. One idea could be a plot in similar way as Figs. 1-6 in Bastidas et al. (2006). In this case, a separate plot could be shown for each performance metric.

MINOR COMMENTS:

Abstract, first sentence: I think that the first sentence should be more general. It is not clearly apparent why parameters need to be identified in relation to parameter sensitivity and catchment classification.

Abstract, second sentence: This sentence is certainly too long. Please subdivide this sentence into two (or even three) to avoid losing the reader directly at the beginning.

Abstract: The numbering (1), (2) and (3) is not explained and thus not understandable.

The referencing should be consistent. To give an example: On Page 5, Line 25, the

CC BY

three references are neither ordered by occurrence nor alphabetic. Please check this in the whole text.

At several parts of the manuscript, the transitions between the different subchapters are not clear. I recommend to check the beginning and the end of the different chapter and if required add a sentence to relate both chapters. One example for this is the beginning of chapter 5.4.

P. 7 L. 3: Why do you use only six parameters which only explain less than the half of the variance? Later in the text (P. 15, L. 24), it is mentioned that even on the day with the highest sum of the partial sensitivities, this value is lower than 0.5. I think that a good reasoning for this is required. Even though that I am aware that it is not useful to do a temporal sensitivity analysis with all model parameters, I am curious whether it is possible to increase the explaining variance by using e.g. 8 or 10 parameters. Or otherwise, could you explain why a higher number of parameters is not beneficial/possible?

P. 7, L.11: I would recommend to write here: "In the case of using six parameters, the FAST method requires altogether 91 model runs..." It should be highlighted that the number of model runs depends on the model parameters and is provided by the FAST methods meaning that the same number of model runs is required for the same number of parameters independently from model, catchment or parameter selection.

P. 7, L.25: Here, 91 model runs are used to identify values for six parameters. This approach is certainly in contrast to typical model calibration algorithms using a significantly higher number of model runs for the same number of parameters. Even when I agree of using this approach for this study, I think that it is required to mention that a lower number of model runs is acceptable here according to capture all goals of this study. Or other way round, it is required to say that this number of model runs has to be certainly higher when only focusing on model optimization.

P. 8, L. 6: Which are the five classes?

CC BY

P. 8, L. 17: There are more than six hydrologic fingerprints in Table 2. Could you explain why you mentioned here "six“ hydrologic fingerprints?

P. 8, L.19: Which one (of the dynamic fingerprint)?

P. 8, L. 26: I would recommend to write: "from each of the simulated...“

P. 8, L. 30: I think that here and maybe also in the introduction a discussion of the PAWN method is missing as proposed by Pianosi and Wagener (2015) since the role of different performing model results within the sensitivity analysis is directly included in PAWN.

P. 9, L.-10-25: I recommend to also refer here to the work from Pfannerstill et al. (2015) and Pokhrel et al. (2012). In both studies FDC and their segments are used to identify (constrain) parameter values.

P. 11, L. 12-13 (Fig. 5): Could you explain why you used circles assuming that the variation are similar for both variables. Is it maybe more useful to use an ellipse (which it would be in the case of a quadratic plot)?

P. 11, L. 12-16: I did not understand how the radius is selected and why it has this size.

P. 11, L. 21: Could you explain why you mentioned both distance measures? Is it maybe more appropriate to select one (the best) of them?

P. 12, L. 21-28: Which result (figure) supports this text passage?

P. 13, L. 20: Please explain why you have selected these seven gauging stations

P. 14, L. 2: Why do you selected four fingerprints and calculate an Euclidean distance between them. Why not using all fingerprints?

P. 14: I have expected a clearer description of the intention of each subplot of Fig. 7 and a presentation of the major outcome of each subplot.

P.16, L. 26: Please add here or later in the discussion that the relationship of parameter

CC BY

sensitivities and FDC (or sorted discharge) was already captured e.g. in Herman et al. (2013b) and Guse et al. (2016) if not already included in the introduction after revision.

P. 16, L. 29: Is there a reason why two observed and only one simulated station are used in Fig. 11?

P. 20, L. 21: I would recommend to structure the discussion in two sub-chapters to avoid a misunderstanding evoked by a double-use of the numbers 1-3 in the discussion. The second sub-chapter in the discussion could start at this line.

P. 21, L. 12: Why not directly increasing the number of parameters in this study?

Fig. 1: I do not see the relationship between the FAST sampling design the parameter values in the calibration. Why do you show both in one plot? Which information can be derive from this relationship? Furthemore, due to the different ranges of the parameters, the interpretation of the parameter values is rather difficult.

Fig. 1, caption: Please changed to "parameter values in the 91 model runs according to the FAST sampling“ or a similar expression.

Fig. 2: Please increase the labels a and b in the figure.

Fig. 2: I strongly recommend to subdivide this figure into two plots showing separately hydrologic fingerprints and the dynamics response.

Fig. 2b: The lines in the dynamic response fingerprints are unclear.

Fig. 3: Do you really need this figure? I do not see the real benefit.

Fig. 3: Yilmaz et al. (2008) made a FDC segmentation at 20% and not at 30% of flow exceedance.

Fig. 3: Are you showing here the 91 model runs as FDC? In this case it would suggest to clarify this by stating this.

Fig. 4: Since the gauges and their abbreviation are used several times in the

CC BY

manuscript, I strongly recommend to increase the labels in size. Maybe a white background (for the labels) would be helpful in addition.

Fig. 7: Please think about the benefit of each subplot.

Fig. 7: The legend to the gauges belongs to Fig. 7a and not 7f.

Fig. 7b: Please discuss in the text why the best performing run in relation to SLFDC is among the worst runs related to NSE.

Fig. 8: Please add in the figure caption that the numbers in brackets in the legend are the numbers of the model runs.

Fig. 10: Please explain in a better way: "highest parameter sensitivity related observed hydrograph".

Fig. 11: It seems to be that the major information from these plots could be extracted in a simpler way. I do not think that the grey lines are required. What about showing only the changes in the dominant parameters as a line (or a row) for each gauge.

Fig. 12: Maybe the legend could be shown only once and outside of the plot at the right side (only a very minor comment).

Technical corrections: P.1, L. 10: sensitivity P.4, L. 16-18: This sentence does not read well. P. 5, L.6: It recommend to use the paper of Reusser et al. (2011) instead of the dissertation work (Reusser 2010).

References:

Bastidas LA, Hogue TS, Sorooshian S, Gupta HV and Shuttleworth WJ (2006): Parameter sensitivity analysis for different complexity land surface models using multicriteria methods, J. Geophys. Res., Vol.11, D20101, doi:10.1029/2005JD006377.

Euser T, Winsemius HC, Hrachowitz M, Fenicia F, Uhlenbrook S and Savenije HHG (2013): A framework to assess the realism of model structures using hydrological sig-

CC BY

natures, Hydrol. Earth Syst. Sci, 17(5), 1893-1912.

Euser T, Hrachowitz M., Winsemius HC and Savenije HHG (2015): The effect of forcing and landscape distribution on performance and consistency of model structures, Hydrol. Process. 29(17), 3727-3743.

Gharari S, Shafiei M, Hrachowitz M, Kumar R, Fenicia F, Gupta HV and Savenije HHG (2014): A constraint-based search algorithm for parameter identification of environmental models, Hydrol. Earth Syst. Sci., 18, 4861-4870.

Guse B, Reusser DE and Fohrer N (2014): How to improve the representation of hydrological processes in SWAT for a lowland catchment – temporal analysis of parameter sensitivity and model performance. Hydrol. Process. 28: 2651–2670.

Guse B, Pfannerstill M, Strauch M, Reusser D, Lüdtke S, Volk M, Gupta H and Fohrer N (2016): On characterizing the temporal dominance patterns of model parameters and processes, Hydrol. Process., 30(13), 2255-2270.

Haas M, Guse B, Pfannerstill M and Fohrer N (2016): A joined multi-metric calibration of river discharge and nitrate loads with different performance measures, J. Hydrol., 536, 534-545.

Herman JD, Kollat JB, Reed PM and Wagener T (2013a): From maps to movies: high resolution time-varying sensitivity analysis for spatially distributed watershed models. Hydrol. Earth Syst. Sci. 17: 5109–5125.

Herman JD, Reed PM and Wagener T (2013b): Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior. Water Resour. Res. 49. DOI:10.1002/wrcr.20124.

Massmann C and Holzmann H (2012): Analysis of the behavior of a rainfall–runoff model using three global sensitivity analysis methods evaluated at different temporal scales. J. Hydrol. 475: 97–110.

CC BY

Massmann C, Wagener T and Holzmann H (2014): A new approach to visualizing time-varying sensitivity indices for environmental model diagnostics across evaluation timescales. Environ. Model. Softw. 51: 190–194.

Pfannerstill M, Guse B and Fohrer N (2014a): Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. J. Hydrol. 510: 447–458.

Pianosi F and Wagener T (2015): A simple and efficient method for global sensitivity analysis based on cumulative distribution functions, Environ Model Softw. 67, 1-11.

Pokhrel P, Yilmaz KK and Gupta HV (2012): Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures. J. Hydrol. 418: 49–60.

Reusser DE, Blume T, Schaefli B and Zehe E (2009): Analysing the temporal dynamics of model performance for hydrological models. Hydrol. Earth Syst.Sci. 13: 999–1018.

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

Printer-friendly version

Discussion paper

CC BY