Interactive comment on "Characterizing and reducing equifinality by constraining a distributed catchment model with regional signatures, local observations, and process understanding" *by* Christa Kelleher et al.

M. Hrachowitz (Referee)

m.hrachowitz@tudelft.nl

The manuscript "Characterizing and reducing equifinality by constraining a distributed catchment model with regional signatures, local observations, and process understanding" by Kelleher et al. addresses a topic that is of critical importance for hydrological modelling applications: the respective value of multiple, different data sources and model evaluation metrics to identify meaningful parameters sets within limited uncertainty. The experiment is well-designed and described. I would be glad to see this contribution eventually be published and there are only two, relatively minor concerns I

would like to encourage the authors to address to strengthen the manuscript:

(1) The manuscript would strongly benefit from being proof-read with a bit more care. The present version comes across as a bit sloppy with e.g. wrong figure numberings, wrong or missing references to figures, wrong references to or missing appendices and relatively imprecise figure captions.

(2) Being a highly important and ubiquitous topic in environmental systems models, I was quite surprised that the link to related techniques dealing with ill-posed inverse problems commonly applied in other fields is completely missing. This includes for example regularization which is quite a standard technique to reduce parameter equifinality in e.g. ground water models. It is just that it has not yet found, for whatever reason, its way into mainstream surface hydrology. Linked to that is the complete lack of this manuscript to refer to distributed model frameworks that make use of regularization (although they often refer to it as regionalization, which is the same thing, really), e.g. the work of Luis Samaniego's group on the mhM model (e.g. Samaniego et al., 2010; Kumar et al., 2013)

Detailed comments:

(3) P.1,l.11-12: distributed models do not necessarily require more parameters, if only the input is distributed (e.g. Ajami et al., 2004; Das et al., 2008; Kling and Gupta, 2009; Fenicia et al., 2008; Euser et al., 2015). I would suggest to make this distinction clear somewhere in the introduction section.

(4) P.1,l.18: I find the term "certainty" quite problematic. Given all the different sources of uncertainty, how can we ever think of a parameter as being "certain". What would that mean in that context? "Certain" with respect to what? Reality? Uncertain observations?

(5) P.1,l.27ff: it is not quite clear what the authors define as distributed model in this study. Note that both, physical models but also conceptual models (for example done

with mhM) can be applied in distributed ways. Please provide a more precise definition.

(6) P.2,l.2: maybe rephrase to "Distributed models should represent the...."

(7) P.2,l.10: why would distributed models *require* a single parameter set? This statement goes against much we know about the uncertainties involved in environmental systems models

(8) P.2,l.14-15: not only valid for distributed models but for *any* hydrological model

(9) P.2,l.19-21: not untrue, but the contribution of the developers of the mhM framework, who went to great lengths to exploit the value of regionalization/regularization, should not go unnoticed here (see above)

(10) P.2,l.27: I would suggest to formulate this in a more general way by emphasizing the degree of information that is used on the prior distributions

(11) P.3,l.3: should read as "...to justify the selection..."

(12) P.4,l.18,figure1: where do all the tall trees hide?

(13) P.5,l.2: what does "tandem" mean?

(14) P.5,l.7,table 1: add units in table 1; what is referred to as "ranges"? is it the uninformed prior parameter distributions? Please clarify.

(15) P.5,l.24: please be more specific here – what does shading include? Does it combine shading due to aspect and shading due to topographic features (e.g. mountain ridge) obstructing sunlight to reach certain locations?

(16) P.5,l.26: please be more specific here – what are "allowable" ranges?

(17) P.5,l.30ff: not entirely clear – which period was the model calibrated for? The entire 01/10/2006-01/10/2008 period? If so, I was wondering if it was not more instructive to retain some months for post-calibration evaluation (i.e. "validation") to see the effect of signatures etc. under these conditions as well.

C3

(18) P.8,l.27,table 2: which runoff ratio was used? Over the entire period? Annual runoff ratios? Or seasonal runoff ratios? They carry different information content, and from earlier work we found that in particular the seasonal runoff ratios carry considerable information. This may be worth looking into.

(19) P.8,l.28ff,table 2: PET is constrained with regional data, but how were these regional PETs estimated? Are they more reliable than the penmen-monteith derived estimates here? if so, why? If not, what is the point of using them. I think this point warrants some discussion.

(20) P.9,l.15,table 2: does this refer to the maximum flow of each year? In many cases in particular the extreme events are subject to unproportionally high observation errors. Thus, do you not run into the risk of forcing the model to reproduce a value that is relatively likely to be considerably off from reality (cf. epistemic errors!!)? I also think that this point needs some more reflection.

(21) P.9,l.16,table 2: the equation for the error in peak timing (again: does this apply for the timing in each individual year?) seems to involve some magic – I could not figure out how a ratio of days over days would result in days. In addition, I am not sure if this ratio makes sense, as we talk about the day of the year, at least as I understand it. Would be probably more meaningful to just minimize or constrain the absolute errors, i.e. $abs(T_{q,s}-T_{q,o})$

(22) P.10,l.8ff,figure 4: this is not entirely clear. Do the grey and coloured bars show the parameter sets that are retained or those that are discarded? The text and figure captions seem to somehow contradict each other. Please check.

(23) P.10,l.10-11: what about the effect of observational uncertainty in precipitation (in particular snow accumulation!!) and peak runoff - see comment (20)??

(24) P.11,l.11,figure 6: I am struggling with this figure. Caption seems to describe an entirely different figure. Where is the black dotted line (the one I see merely seems to

C4

separate the two hydrological years)? Where are subplots (a) and (b)?? what is the light blue line (observation?)? please clarify.

(25) P.12, l.3, figure 8(?): figure number seems to be wrong. Is this not figure 7? if not, where is figure 7? Please indicate years on x-axes. What are the grey shaded areas? Are these model ensembles or uncertainty ranges? If the latter, how were they constructed and, in particular, where do the gaps come from??

(26) P.12,l.4: what is meant by "relative timescales"?

(27) P.12,l.17: I cannot find Appendix D (and C not either, for that matter).

(28) P.12,l.23: please reconsider the use of the term "parameter certainty" (see above)

(29) P.12,l.30: what is meant by "original" distribution? Is it the uninformed prior distribution (i.e. table 1)? If so, please call it also like that. In addition, although I see and understand the intention here, I am wondering how much this also depends on the choice of the parameter range in the prior distribution. If a narrow prior range was chosen, the normalized values shown here may be less constrained than if a wider prior range was chosen. in addition, are the individual parameter sets likelihood weighted to give more wight to better solutions? if not, why not? Any thoughts on this?

(30) P.13,l.6: where are these predictions included? Figure b3? Please specify.

(31) P.13, l.13, figure 9: the figure and captions are a bit confusing. What is meant by set 1/5/9? And how is set 1 drier than set 9? Please clarify.

(32) P.18,l.6: I think the wrong reference is provided in the reference list. The paper you are referring to is Euser et al. (2015, HP), but in the reference list Euser et al. (2013, HESS) is provided.

(33) P.18, l.9: some references to mhM would fit in nicely here.

Best regards, Markus Hrachowitz

References:

Ajami, N. K., Gupta, H., Wagener, T., & Sorooshian, S. (2004). Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. Journal of Hydrology, 298(1), 112-135.

Das, T., Bárdossy, A., Zehe, E., & He, Y. (2008). Comparison of conceptual model performance using different representations of spatial variability. Journal of Hydrology, 356(1), 106-118.

Euser, T., Hrachowitz, M., Winsemius, H. C., & Savenije, H. H. (2015). The effect of forcing and landscape distribution on performance and consistency of model structures. Hydrological Processes, 29(17), 3727-3743.

Fenicia, F., Savenije, H. H., Matgen, P., & Pfister, L. (2008). Understanding catchment behavior through stepwise model concept improvement. Water Resources Research, 44(1).

Kling, H., & Gupta, H. (2009). On the development of regionalization relationships for lumped watershed models: The impact of ignoring sub-basin scale variability. Journal of Hydrology, 373(3), 337-351.

Kumar, R., Livneh, B., & Samaniego, L. (2013). Toward computationally efficient largeâĂŘscale hydrologic predictions with a multiscale regionalization scheme. Water Resources Research, 49(9), 5700-5714.

Kumar, R., Samaniego, L., & Attinger, S. (2013). Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. Water Resources Research, 49(1), 360-379.

Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a gridâĂŘbased hydrologic model at the mesoscale. Water Resources Research, 46(5).

C7