Interactive comment on "Toward seamless hydrologic predictions across scales" *by* Luis Samaniego et al.

Luis Samaniego et al.

luis.samaniego@ufz.de

Received and published: 1 May 2017

1. *The authors make a nice case for the value of their multiscale parameter regionalization (MPR) method, analysing several aspects (and advantages) of the method. This is in principle a laudable thing to do. The manuscript itself, however, is quite frustrating to read. One the one hand it remains completely unclear what the novelty is. Large parts of the manuscript essentially repeat what has already been published earlier (as also acknowledged in the references provided).*

We are saddened by the fact that the reviewer consider that this manuscript lacks novelty. We will improve the text to make it clear in the introduction. The novelty of the manuscript is based on the following key elements (see also the response

to Ref.1):

(a) Attempt to describe the progress towards seamless parameterizations in land surface(LSM) or hydrological models(HM). We present a short description of what has been made (the literature on the topic is quite extensive) and provide a simple example to visualize how existing LSMs/HMs are estimating a fundamental parameter such as soil porosity (not found in literature).

(b) We propose, based on our own experience, a way forward that uses MPR and systematize its application by providing a "Protocol for evaluation of model parameterization" (This has not been publish before)

(c) We implement this protocol to PCR-GLOBWB (also new piece of work and unpublished)

(d) Carry out a series of experiments (based on the spirit of the E. Wood's recommendation) to demonstrate how to spot faulty parameterizations (also not publish before).

(e) Compare the effects of the parameterization on three models (mHM, Water-GAP, and PCR-GLOBWB) as part of these experiments (all using the same forcings and underlaying data)

It should be clearly noted that **none of these key elements belong to Mizukami et al. (under review) (hereafter [MCN+2017])**.

2. *On the other hand, the argument remains in places quite imprecise with a lot of quite sweeping (and not necessarily well substantiated) generalizations.*

We will remove generalizations that are not fully substantiated with our experiments. It would be nice to know, however, which parts of our manuscript —according to the reviewer— are not substantiated enough.

3. *In addition, other approaches to parameter selection are quite outrightly dismissed while essentially no critical discussion on potential drawbacks or limitations of MPR are provided.*

We are not dismissing them. We simply do not have space in this study to test all of them. These evaluations (HRUs, Standard regionalization, etc.) have been carried out in independent studies which are cited in our manuscript. Here a short list:

(a) MPR vs. k-NN regionalization:
Samaniego, L., Bardossy, A., & Kumar, R. (2010). Streamflow prediction in ungauged catchments using copula-based dissimilarity measures. Water Resources Research, 46(2), http://doi.org/10.1029/2008WR007695

(b) MPR vs. standard regionalization (no scaling)
Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. Water Resources Research, 46(5), http://doi.org/10.1029/2008WR007327

(c) Lumped HRU, Distributed HRU, vs. MPR:
Kumar, R., Samaniego, L., & Attinger, S. (2010). The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics. Journal of Hydrology, 392(1-2), 54-69. http://doi.org/10.1016/j.jhydrol.2010.07.047

(d) MPR with satellite data (ungauged basin)
Samaniego, L., Kumar, R., & Jackisch, C. (2011). Predictions in a data-

C3

sparse region using a regionalized grid-based hydrologic model driven by remotely sensed data. Hydrology Research, 42(5), 338-355. http://doi.org/10.2166/nh.2011.156

(e) MPR vs. HRU
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. http://doi.org/10.1029/2012WR012195

(f) MPR across scales US basins
Kumar, R., Livneh, B., & Samaniego, L. (2013). Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme. Water Resources Research, 49(9), 5700-5714. http://doi.org/10.1002/wrcr.20431

(g) MPR in Pan-EU (transferability test, evaluation of states and fluxes more than 300 basins)
Rakovec, O., Kumar, R., Mai, J., Cuntz, M., Thober, S., Zink, M., et al. (2016). Multiscale and Multivariate Evaluation of Water Fluxes and States over European River Basins. Journal of Hydrometeorology, 17(1), 287-307. http://doi.org/10.1175/jhm-d-15-0054.1

Limitations and drawbacks of MPR w.r.t. to other methods have been mentioned in all our publications (see above). We will summarize them in the revised manuscript.

4. *In an exaggerated way, the authors present their MPR method, which I think has formidable potential, like in a product promotion folder.*

We politely disagree with the reviewer in this respect, basically for the following reasons. Our Manuscript cannot be consider a "promotion" folder because we put forward a protocol — NOT PUBLISHED before— and then proceed to apply it to

C4

the model PCR-GLOBWB. Then we applied this model to two different scales to test the flux matching condition (all new). In addition to that we remark the deficiencies of current approaches to obtain seamless parameter fields (see fig. 1, also new).

5. *I think the manuscript would strongly benefit from (1) considerably reducing the redundancies with previous work (sections 1-3 can be *substantially* shortened) and (2) taking on a more critical perspective towards MPR. I think that many in the community will agree that it is a great tool. Instead of highlighting this over and over again, it would be more instructive to learn were its limitations are to allow further improvement.*

We will attempt to reduce redundancies. We recapitulate the MPR technique to have a self-consistent manuscript, if we move this section to an appendix, or refer MPR to other manuscripts, perhaps is not the optimal solution for the reader. We will try to summarize as much as possible though. As indicated above, limitations of MPR will be clearly written in the revised manuscript.

6. *In general, I think it may be more interesting for a wider audience if the MPR technique was scrutinized and compared to other parameter selection and regionalization approaches *independent* of the model it is used for. In this manuscript it is applied exclusively with mhm if I understand correctly. In my understanding, it is a stand-alone method that should be applicable to any model. Would it not be fairer to be more consistent in the comparisons here, i.e. compare mhm with/without mpr and/or other models with/without mpr?*

We politely disagree with the reviewer in this point for the following reasons. First of all, we can not attempt to repeat all was done in all the publications listed in point 3. Second, in this manuscript we are not only dealing with mHM but with PCR-GLOBWB and WaterGAP. Please see section 5! In this section **we fully**

implement the protocol presented in this manuscript to PCR-GLOBWB. This means with and without MPR in PCR-GLOBWB! Third, comparisons with/and without MPR in mHM have been done, please see Kumar et al 2010, Samaniego et al 2010ab, Kumar et al 2013, etc. and there will be soon a manuscript from Rakovec et al. over 500 US basins showing the effects of MRP/NO-MPR with mHM and VIC. Science is a cumulative enterprise from our point of view. We cannot repeat everything again and again.

7. *The bottom-line is that I have the feeling that two quite independent things are not clearly separated here: the regionalization technique (MPR) and the models (mhm, etc). Here the text needs to become much more precise. Right now it seems to the reader that MPR is compared to e.g. the HBV model. This is not a valid comparison as these are completely different things. In contrast, it would be excellent to make the fact that MPR is a standalone tool clearer, as this may result in more modellers actually picking up the idea for their very own models (which they may not do at the moment due to its perceived exclusive association with mhm).*

We politely disagree with the reviewer with some remarks in this point for the following reasons. First, we are comparing MPR with HBV! We are comparing parameters obtained with whatever method in various models that are related with the water holding capacity and porosity of the top soil and remarking that we have a problem with our LSMs/HMs that we need to be solved if we would like to have scale invariant parameterizations and consistent model simulations. MPR is a possible avenue, a hypothesis that we are scrutinizing over and over again in thousands of river basins across the globe. Second, mHM is an open source code available at www.ufz.de/mhm. Third, the model-agnostic version of MPR is called MPR-FLEX and is presented by [MCN+2017] and has been applied to VIC. Consequently, MPR is NOT exclusive from mHM now. We will improve the text so that this "impression" that "MPR is compared to e.g. the HBV model" vanish.

Specific comments

1. *p.2,l.5: why only over time and not also over space?*

   In this particular case, space is implicit and evolution refers to the development of spatial dependent variables over the time dimension. We will reformulate the sentence in the revised manuscript.

2. *p.2,l.10-12: please avoid subjective terms as "elaborate" or "sophisticated"*

   Done.

3. *p.2,l.28-29: is this actually true? Why would process dynamics that emerge at larger scales and that integrate several processes necessarily reduce "realism"? It is surely possible, but I do not think that it is a physical necessity. In any case, what is the meaning of "realism" in a situation where most of the system is de facto unobservable? How do we know if something is "realistic"?*

   We consider that this sentence is true. We do not want to start a philosophical discussion of what is "reality". We have a pragmatic approach, if a model is able to reproduce surrogate observations in evaluation mode, then we consider that the unobservable states may be plausible. For this reason, we carried out the study reported in Rakovec, et al. 2016 JHM (see above).

4. *p.2,l.33-34: this is a sweeping generalization. What is actually meant by that? Why should an observed quantity, such as for example the stream flow recession constant have no physical meaning? Of course it has, albeit on the scale of the observation.*

   We are referring here to transfer function parameters, for example those constants of the Clapp-Horberger PTF, which are basically found empirically and then used to link soil texture values (observable) with soil properties that may or

   not be observable (e.g., porosity). We are not referring to streamflow recession constant. We will clarify the text to avoid confusions.

5. *p.5,l.15ff and elsewhere: many things are mixed together here and the logic is not convincing. For a meaningful argument they need to be carefully disentangled. Is this about models? About parameter selection/calibration procedures? Parameter region- alization? It reads as if MPR does not rely on calibration, which is not correct. and why should lumped and/or semi-distributed models not be run with MPR-derived pa- rameters? Would this for a, say 100km2 catchment, not be the same as if running a distributed model with a 10x10km2 grid in mhm?*

   Based on this comment, we consider that the text is not correctly interpreted. We will clarify in the reviewed manuscript. Please refer to Samaniego et al. WRR 2010 to see a diagram that represent the steps done to estimate parameter for a given model. A simple conceptual model whose parameters are calibrated fail, in general, to perform well at cross-validation. This is what we are referring to. MPR improves transferability across scales and locations as shown in previous studies. In fact, this is what we demonstrate in Kumar et al. 2010 JoH. MPR could be used to estimate lumped parameters if a single cell covers the whole basin. In Kumar et al., mHM-MPR always performed better than a lumped mHM with no MPR.

6. *p.5,l.19 and elsewhere in the manuscript: much is made of "discontinuities". How- ever, the authors do not provide a clear definition of what they mean. Nature is, in places, discontinuous (e.g. forest vs. grassland, north vs. south aspect, sharp transitions in geology, breaks topography, etc). thus it is not clear why mod- els should not represent these discontinuities. I suppose that the authors want to say that between individually calibrated catchments discontinuities can occur, where there are in reality no discontinuities. But this needs to be made clearer.*

   An example of artificially induced discontinuities by parameter calibration is

shown is Fig.4. We agree that there are natural discontinuities, we expect however, that it is unlikely that everywhere the model parameter and fluxes/state fields follow exactly the boundaries of the drainage area at a given location (see Fig.1 below). We call this negative effect calibration imprint, and we attempt to remove it with MPR. This artificial boundaries is what we call discontinuities. In the revised manuscript we will clarify our definition to avoid confusions. Nevertheless, we provide references to literature in p.5 l.19 to illustrate our definition. Please see the obtained parameter fields in Fig.1 (below) as obtained by Merz and Bloeschl 2004 and by MPR in Rakovec et al. 2016 JHM.

7. *p.5,l.21-23: sure, but is this not also the case for distributed models and dependent on the calibration/parameter selection method?*

   This is the case for any model even if one uses MPR on a single basin. This is the reason for showing the Fig.4a. Parameter estimation implies to have a representative sample. For this reason we attempt always to perform parameter estimation on several basins simultaneously, see Fig.4b. Single basin calibration is disadvantageous for any parameterization method because artifacts of the data can be "over-learned" which in-turn would induce large bias somewhere else.

8. *p.6,l.29 and elsewhere: "CONUS": not necessarily every reader will be exposed to large scale studies employing these terms. Thus please avoid the use of fashionable abbreviations without first defining them.*

   Done

9. *p.8,l.7: a question cannot be postulated. Please rephrase.*

   Thank you for the remark. We mean "put forward". It will be rephrased in the revised manuscript.

10. *p.8,l.10-11: what is meant by "poor". How do you define it?*

C9

   A poor parameterization does not lead to flux-matching, exhibits low model performances (say KGE) in cross-validation experiments across scales and locations, and exhibits artificial "discontinuities", i.e. non-seamless fields. This definition will be clearly mention in the revised manuscript.

11. *p.9,l.3: over-parameterization is only addressed in MPR if simultaneously calibrated to a high number of catchments and/or objective functions. Thus, it depends on how MPR is implemented and applied. Please rephrase.*

    We will rephrase this sentence in the revised manuscript.

12. *p.10,l.17-18: how do you know that the parameters are "realistic"? See also comment above. Does this not also strongly depend on the assumptions in the upscaling relationships? It is always a question of how MPR (or other parameter selection techniques) are implemented and not a defining proprietary feature of MPR.*

    This is a good question. It depends on many assumptions, PTFs, upscaling relationships, parameter estimation methods, etc. Visual impression may be useful but it is subjective. For these reasons, we need a formalized approach such as that described in Sec.3.3: **Protocol for evaluation of model parameterization**, which was put forward in this manuscript, and depicted in Fig.2. The experiments presented in Sec. 4 were introduced to addresses this question.

13. *p.13, section 4: in many parts of the section it is unclear what is meant: the individual models or rather the parameter selection/regionalization techniques in the different model applications? These are different pairs of shoes and need to be carefully separated.*

    We will clarify this section in the revised manuscript.

beta (-)

Period: 1987-97
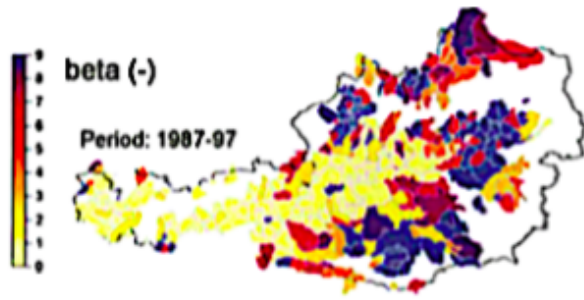
HRU based on: Merz & Bloschl, JoH 2004
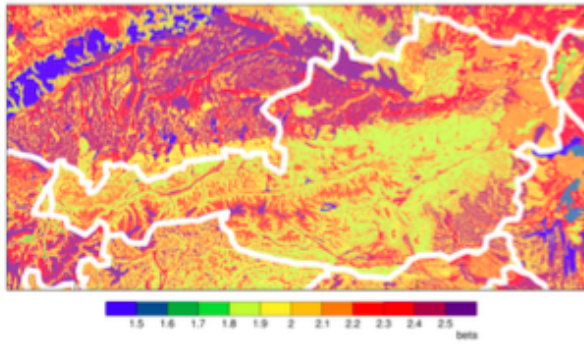
MPR based on: Rakovec et al. JHM 2016

Fig. 1. Fields for the "Beta"parameter estimated for HBV and mHM. We consider that obtained with mHM and MPR a seamless field.

C11