Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-65-RC3, 2020

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Adaptive clustering: A method to analyze dynamical similarity and to reduce redundancies in distributed (hydrological) modeling" *by* Uwe Ehret et al.

Shervan Gharari (Referee)

shervangharari@yahoo.com

Received and published: 13 April 2020

I thank the authors for their detailed reply to my comments. For the spirit of open discussion, I would clarify/comment on some of comments/responses:

Comment 2: 2- The Introduction seems to be superficial. I would say the paper is about representation of the system in a model vs computational time/resource. In land [surface] modeling community there is significant body of literature devoted to the effect of grid size (computational burden vs spatial representation) and example of them can be Melsen et al., 2016 (and many more). This is the case in hydrological rainfall/runoff models as well (Liu et al., 2016 and many more). In its current form the Introduction

starts with general reflection on sophisticated processed-based models; then moves to concept of co-evolution (which is not directly relevant to the message this study wants to convey) and then comes down to clustering. I would suggest to re-organize the Introduction to reflect on pervious works on computational burden vs spatial representation, clustering and its application in hydrological similarities and finally make it clear what the reader should expect from this paper.

Reply 2: We thank the referee for pointing at useful literature related to the effect of grid size in land surface modeling. We will integrate it in a revised version of the manuscript. However we do not agree that the introduction is superficial. We discuss conceptual modeling and its main shortcomings to show the merits of distributed modeling. We then discuss the computational challenges of distributed modeling and how they are mitigated by adaptive time stepping adaptive gridding. We discuss the shortcomings of available adaptive gridding methods, namely that they require spatial adjacency, and argue that similarity of sub systems of natural systems is not necessarily limited to neighboring elements. Taken together, this shows that adaptive clustering i) can be useful, ii) it has novel aspects compared to existing methods of adaptive gridding, and iii) similarity is not only an artificial effect caused by representing natural systems in models, but it is also property of the real-world system (although the inevitable simplifications associated with representing real-world system in models can increase similarities). We therefore prefer to keep the structure of the introduction as it is.

Re-Reply2: What the authors mentioned here are clearer than the introduction at its current format. I would say keeping the interlocution as it is, is a disservice to the manuscript. There is only one paragraph about clustering in the introduction. For a reader, the concept of co-evolution is somehow presented as the main topic, while I think the concept of similarities of the sub-system behavior (grouped response units, GRU, hydrological response units, HRU;) should be more elaborated. The examples the authors mentioned, "north facing…", actually fall very well in the concept of GRU and HRU and pave the way for better presentation in the manuscript rather than co-evolution. I agree that co-evolution is the process of creating similarities/rules but in my point of view it should not be the main point of discussion here. I leave this this to the authors and editor to decide.

Comment 6: The are kr and kb identifiable/related? it seems to a redundancy in the two processes/parameters.

Reply 6: Both parameters are indeed retention constants of linear reservoirs, however they strongly differ in magnitude. Kr is the retention constant of a single river element. As all river elements in the model are of about the same length (1 km), we used the same kr for all river elements. Its value of 1.1 hours was found by maximizing the agreement between observed and simulated streamflow in river stretches where up- and downstream gauges were available (this way we could see the effect of translation and retention in the river stretch). Kb is the retention constant of the base flow reservoir. We determined kb values, separately for each geology (see p.9, lines 4-28), by maximizing agreement of simulated and observed streamflow during times of summer low flow. Kb values are in the range of 500 hours (for Schist) and 20000 hours (for Sandstone), i.e. at least two orders of magnitude higher than kr. We are therefore confident that there is little redundancy between these processes/parameters.

Re-Reply 6: sorry I mean Ki, for the fast reservoir, instead of Kb. Indeed, the slow reservoir does not even need to be routed through river network due to its long reaction time. Ki can be said to be at the scale of hillslope temporal response which should be more or less in scale of hours. Interested to know more about Ki and how it is interpreted from the field data and separated from Kr.

Comment 9: is the satellite based evaporation a result of more sophisticated model (such as a land model)?

Reply 9: Yes, the ET estimates are produced by forcing a SVAT model (a simplified version of the ECMWF TESSEL SVAT scheme) by Land-SAF radiation products (DSSF, DSLF and AL) and ECMWF meteorology. A detailed descrip-

tion is given in Trigo et al. (2011), section 3.1, and on the LSA- SAF pages (https://landsaf.ipma.pt/en/products/evapotranspiration-energy-flxs/met/).

Re-Reply 9 – So good that you gave a low weight to its NS value.

Comment 13: Section 2.3.2 is again called adaptive clustering, similar to section 2.2. and again, in Section 2.3.2 the authors are referring to CATFLOW and MIKE SHE, etc.

Reply 13: Respectfully, we do not understand the concerns of the referee here.

Re-Reply 13: I meant both Sections have the same title. Also again in the middle of the manuscript it is referred to CATFLOW and other models (page 18, lines 3-4).

Comment 15: Page 20 line 9, why it is “striking” that the entropies are lower than the uniform? I would always expect so. It is also expected that the entropy is lower for the recession and higher for rising discharge. This is kind of similar to the heteroscedasticity assumption on the error as well (more diffused with higher discharges). If only observation is used with varying error assumption, higher streamflow will have higher entropy and lower streamflow will have lower entropy.

Reply 15: We agree that it is no surprise to see entropies below the entropy of the corresponding uniform distribution. ’Striking’ here refers to the fact that entropies are well below the uniform entropy, and often close to zero. In our opinion this is indeed noteworthy, and it shows the high potential for adaptive clustering. In a revised version of the manuscript, we will make this point clear. We also agree with the referee that hydrologists have known since long that the degree of similarity between sub systems varies with the hydrological situation. The point we want to make here is that i) entropy of normalized, binned distributions of states and fluxes expresses this in a conveniently dimensionless way, and ii) that we make use of this knowledge.

Re-Reply 15: I would say it is not still sticking. In my point of view any model set up (even worst ones) can easily show very good behaving entropy as they are mostly affected by forcing and memory of the forcing rather than parameters.

Comment 19: 6- For me personally, moving from the world of conceptual models to land models, I would like to question the motivation of this study. Although saving time is valuable but having method that needs model re-run or updating for more complex model is terribly cumbersome. This is the reason why the authors have chosen to use SHM rather than CATFLOW for example.

Reply 19: The referee correctly states one main motivation for this study: saving computation time. This is already a more than sufficient reason, as the referee will surely agree, as he also works on concepts to make high-resolution land surface modeling more efficient (Gharari et al, 2020). In addition, the concept of adaptive clustering yields valuable insights in the time-and space patterns of similarity among sub systems, which, we daresay, is a useful contribution to hydrology research. We have chosen SHM for the proof-of-concept as any hydrologist can easily connect to it, and hence we can show the effects of adaptive clustering in an environment familiar to most hydrologists. We agree with the referee that implementing adaptive clustering in more advanced models will be more demanding, but also the gains will be higher (see p. 29, lines 5-9). So it will be well worth the try.

Re-Reply 19: Yes, testing the method for the more sophisticated models is desirable of course. I just wanted to draw the attention of the authors that to the fact that running a more complex model means more technicalities. Given those technicalities, and time/resources to fix them, it is not really clear if the final gain will be higher. The technicalities can be how to efficiently read/write/update this adoptive clustering; how to efficiently do a warm start for a model; how to pass this over various processors if needed; do the mentioned models' capabilities allow such an approach? Etc.

I give an analogy of the sensitivity analysis of land models. Land models may fail (crash) for some given parameter sets therefore may not result in output values (objective functions) which are essential given the struct of parameter sampling method. This may cause issues for sensitivity methods which should be thought through. I would say adding one or two sentences on thse technicalities/obstacles at the end might be useful

for the reader.

With regards,

Shervan Gharari

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-65, 2020.