Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-127-RC2, 2020

Hydrology and Earth System Sciences
Discussions

Interactive comment on "A framework to regionalize conceptual model parameters for global hydrological modeling" *by* Wenyan Qi et al.

Anonymous Referee #2

Received and published: 26 August 2020

The paper compares regionalization approaches at global scale. I like that the paper tests many different regionalization approaches but, unfortunately, one important approach is missing (more below). In addition, the English should be improved by a native English speaking person.

The "spatial proximity method" may well yield the highest KGE, but it cannot be used at global scale due to the lack of gauges in many regions. As an example, grid cells in the southern Ecuadorian Andes would receive parameters from donor catchments located in the Amazon, which doesn't make any sense. So while this approach may give you the best performance scores, it is not actually a method that should be used at global scale. This should be explicitly mentioned in the abstract and the conclusion.

Even if the "mean distance between the donor catchments and the target catchment is

no more than 1500 km" the actual difference in climate and landscape could be huge.

"The k value of the PCRGLOBWB was determined based on the drainage [...]" This was paraphrased (like many other sentences) from Beck et al. (2016), but is this still the case in PCGLOBWB 2.0 (https://gmd.copernicus.org/articles/11/2429/2018/gmd-11-2429-2018.html)?

Table 1: The aridity index can become extremely high in desert regions (higher than 100) so I'm surprised that this isn't reflected in your mean. Did you apply some sort of mask, or did you cap the values before calculating the mean? This info needs to be in the caption.

Table 1: Which datasets did you use for the different attributes?

Section 2.2: Hydro1K is a very outdated dataset. A newer dataset should be used, maybe HydroSHEDS or MERIT.

Figure 1: Which KG map did you use? These are too many classes to distinguish, and the legend is a mess. Better to condense to the 5 major classes.

Equation 1: This is the old KGE, the new KGE has a slightly better formulation of the variability component. See Kling et al. (2012; https://doi.org/10.1016/j.hydrol.2012.01.011).

Section 2.5: Among the five tested regionalization methods, one important approach is missing, and it is most likely the best approach. It is the approach where the model parameters and the regression equations (relating the catchment attributes to the model parameters) are optimized simultaneously. See for example https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD031485 and https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2008wr007327. This approach should definitely be included.

"The regression-based method assumes that a well-behaved relationship exists in the observable catchment characteristics and model parameters" which in reality is almost

never the case due to parameter equifinality and therefore this approach rarely works. Hence my suggestion to test the other regionalization approach.

"The SP method assumes that nearby catchments should have similar behavior for climate and catchment conditions (features) varying uniformly in space." "Nearby" could be several thousand kilometers away so this approach should never be used at global scale.

Another limitation of the study is that lumped catchment attribute and model parameter values are used, despite the often large heterogeneity within catchments. This limitation is addressed in several studies (e.g., Samaniego 2008 and Beck 2020) and should at least be discussed somewhere in the paper.

"The results show that the distributions of model efficiency of four hydrological models are similar to each other and indicate that the difference between hydrological models was negligible in the model calibration and validation, which is in line with previous studies (Beck et al., 2016; Vetter 245et al., 2015; Demirel et al., 2015)." This is definitely not in line with previous studies, which generally found large differences between models and a considerable difference between calibration and validation scores. You actually also obtained substantial differences between calibration and validation scores (figure 2)!

Figures 3 and 10: This figure is impossible to interpret. https://www.climate-labbook.ac.uk/2016/why-rainbow-colour-scales-can-be-misleading/

Figure 5: The figure is a bit difficult to interpret, maybe use vertical instead of diagonal x-axis labels, and apply some coloring to group similar methods together?

Figure 6: There is no information about the number of catchments representing each bar, so a particular bar could be represented by just 1 catchment. I can't think of a solution right now, but this information should not be hidden.

Figure 8: A non-linear color scale might better.

Table 6: "many data may not be directly comparable because of different continental boundaries and averaging periods." A solution would be to only use estimates representing the same area. Also, why report values from a study >40 years old (Korzun 1978)? Considering adding GSCD estimates (http://www.gloh2o.org/gscd/).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-127, 2020.