

General comments:

I am glad to see a revised version whose scope is reduced and all my comments are addressed. The paper is well organized but I found many typos. Addition of the conceptual models and their performance CDFs is very informative and more practical as not everyone uses physically-based models. The impressive performance of GR4J, given the few number of parameters, is something I did not expect. I only have minor comments now that are stated below, the authors can incorporate them quickly.

Specific comments:

L239-240: Are the discharges coming from calibrated models? It is not clear. I am assuming they are.

L290-296: Here, the conclusion drawn about differences being more evident in north and south as compared the center are subjective (weak). It would be simpler to just divide them via arbitrary (exact angles need to be determined) lines that divide the area into three portions and then comparing either the means or medians of KGE-NP differences. If this point is the focus of the study, then please be thorough. I don't find it relevant.

L298-305: Yes, but showing relative differences is not a good idea. River width is not shaped by low flows. During low flows, a slight fluctuation in depth can easily translate to above 50% flow increase/decrease. Sedimentation may result in an increase of a few centimeters, depending on the situation. The flow volume that passes through a cross-section is important. Here, another paragraph can be added that translates the relative differences to absolute flow volumes. I think, that would show a different story.

L437-455: The abstract and conclusions deal with the problem stated in the title now. The only problem that remains is getting uncertainty bounds on observed discharges, elsewhere in the world but the results do shed light on what the situation might look like.

Cheers