Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-487-RC3, 2016

Hydrology and Earth System Sciences
Discussions

Interactive comment on "A comparison between parameter regionalization and model calibration with flow duration curves for prediction in ungauged catchments" *by* D. Kim et al.

Anonymous Referee #3

Received and published: 8 December 2016

Review of the manuscript "A comparison between parameter regionalization and model calibration with flow duration curves for prediction in ungauged catchments" by Kim et al.

In this manuscript, Kim et al. compared two regionalization approaches to predict hydrographs and flow-duration curves (FDCs) in ungauged catchments in South Korea. The proximity based parameter regionalization outperformed the model calibration with regionalized FDCs (from top-kriging) for the simulation of hydrographs, whereas their performance was comparable regarding the prediction of FDCs. Given the relative simplicity and the good performance of the parameter regionalization, they recommended to use this approach for the prediction of runoff in ungauged catchments in South Ko-

rea.

I like the idea of the manuscript and read the manuscript with interest. However, I think the current level of the manuscript does not fulfill the criteria for publication in HESS. The manuscript partly introduces a new concept for the prediction in ungauged catchments. It also uses existing methods that should be discussed more critically. Relevant literature is not sufficiently introduced and discussed and references for methods or statements are sometimes missing. The structure of the paper could be improved by better separating methods, results and discussion – as it is now elements of these parts are mixed. To make the manuscript better readable I would also recommend to improve the English which suffers from grammatical errors.

I hope that the comments below will be helpful for the authors to improve their manuscript.

Major comments:

1. In this manuscript two regionalization approaches are compared: the first approach regionalizes parameter sets calibrated with the hydrograph, and the second approach regionalizes normalized FDCs that are used for the calibration of a runoff model. Although I like this second approach it seems to be rather different from the first one. I wonder why the authors did not apply an approach that is closer to the first one such as calibrating the model using the FDC and regionalizing these calibrated parameter sets. Using the suggested approach would make the results more comparable because the uncertainty sources are more similar (e.g. uncertainty due to top-kriging would be eliminated).

2. The manuscript would benefit from a detailed discussion on the sources and the influence of uncertainties related to the different regionalization approaches.

They are crucial for the interpretation of the results.

3. Besides the hydrograph and the FDC also runoff ratio, baseflow index and rising limb density of the ungauged catchments are evaluated. The authors state several times that the calibration of the runoff model against the regionalized FDC and the rising limb density simultaneously would improve the prediction in ungauged catchments. However, I see no strong evidence for this statement based on Fig. 9. I also don't understand why the rising limb density is regarded as being orthogonal to the FDC. I recommend to weaken these statements or to provide good evidence for it. Furthermore, no information is provided on how the rising limb density could be derived for the ungauged catchment. Would you also regionalize it?

4. It could be interesting if you actually tried to constrain the runoff model by the FDC and the rising limb density (or any other suitable runoff signature).

   -kriging is used for the regionalization of the normalized FDCs. Is this approach really a well-established method as you mention? How many studies have used this approach? Is this approach suitable for FDCs and the density of the gauging stations in your study area? Can you give good reasons for not using ordinary kriging?

5. The results from the two regionalization approaches are presented as separate numbers (performance value) or separate boxplots that are next to each other (Fig. 7-9) which is inconvenient for their direct comparison. I would recommend to improve the presentation of the results by using the parameter regionalization approach as a benchmark to which the second approach (calibration with regionalized FDC) is compared. E.g. take the difference between the Nash-Sutcliffe efficiency of approach one and two for each catchment.

6. The snowmelt model used for the calculation of snow accumulation and ablation needs more explanation, especially because snowmelt models based on energy

balance usually are data intensive. From shortly reading the publication from Walter et al. (2005), I don't have the impression that this physics based snowmelt model is simple, as you call it. I agree that the snowmelt model doesn't have parameters that are calibrated, however it has various parameters that have to be estimated (e.g. cloud cover, albedo, windspeed, etc.). It would be worth to discuss whether there is really less uncertainty involved than when using e.g. a degree-day method for the simulation of snow accumulation and ablation.

Moderate comments:

P1 L16-18: I would remove the sentence about the rising limb density and instead add some information about the results of the FDC prediction in the ungauged catchments, because that was one focus of your study.

P2 L1: The study from Seibert and Beven (2009) did not use any regionalization in their analysis. This paper is not the right citation here. Please make sure that you cite properly.

P2 L7: When writing about regionalized flow signatures it would be worth to include the study of Yadav et al. (2007) and Hingray et al. (2010) at this point.

- Yadav, M., Wagener, T., and Gupta, H. V.: Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, Adv. Water Resour., 30, 1756– 1774, doi:10.1016/j.advwatres.2007.01.005, 2007.
- Hingray, B., Schaefli, B., Mezghani, A., and Hamdi, Y.: Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments, Hydrolog. Sci.

P2 L9: You mention that flow signatures have been frequently applied for model calibration. Please give some more examples including runoff ratio, baseflow index and rising limb density.

P2 L16: Please give more examples of studies that regionalize FDCs and explain how they do it. This is important because the regionalization of FDCs is a core method of your study.

P2 L19-26: The information in the first sentence contradicts your subsequent paragraph.

P3 L4-11: Please cite where your information about this paragraph comes from. Is the information of this paragraph for South Korea in general or does it only relate to the study catchments?

P3 L13-L17: Where is the data of the 29 catchments with high quality from? Please add this reference and also the reference for the inflow data of the multi-purpose dams to the reference list at the end of the manuscript.

P3: Chapter on description of study area and data: where is the evaporation data from that you need as input for the runoff model? Do you need elevation data for the runoff model? Do you have any information about geology, vegetation etc. because you mention this as possible reason for poor top-kriging performance (P9 L15).

P4 L3-8: I recommend to add a short description of the structure of GR4J, information about the required input data and its resolution as well as information about the use of elevation bands.

P6 L24: Why do you evaluate the hydrograph with Nash-Sutcliffe and volume error, but the FDC only with Nash-Sutcliffe?

P6 L28: Give an explanation why you selected runoff ratio, baseflow index and rising limb density as signatures. Why three signatures?

P7 L22: Please explain why you use 5 donor catchments and not 3 or 7.

P8 L9: Again, why do you use 5 parameter sets and not 3 or 7? Does it make sense to give weights to these 5 parameter sets given the uncertainty related to them?

P8 L31: Why do you use NSE(Q) 0.6 as threshold?

P9 L4-10: This paragraph belongs to the methods section and should not be in the results. Do I understand correctly that 5 donors were used for the regionalization of the FDC? If yes, how do you get a total weight of 1? If no, please write this sentence more clearly.

P9 L29: Why a threshold of NSE(FDC) of 0.8 is used? Is it necessary to reduce the number of catchments used in model calibration by the additional constraint of NSE(FDC)? Wouldn't it be better to keep as many catchments as possible?

P9 L31: Figure 5b – how do you explain the scatter in the low flow?

P12 L10: Discussion: The manuscript would strongly benefit from a deeper and more extensive discussion of the results with other studies. Many statements appear in the discussion for which it is not clear where they are taken from - so please cite other studies properly (e.g. first sentence in discussion). There is no chapter in the discussion about the prediction of the hydrograph in ungauged catchments.

P14 L1: Summary and conclusions: I recommend to shorten this chapter. Point 1 (L9-13) is in my opinion no key finding of the study, point 2 (L14-17) is more an assumption than a result and point 5 (L27-30) is also rather a hypothesis than a result and should be formulated as possible further steps.

P15: Please add the information on data availability and author contribution.

P20 L1: If you want to show the parameter ranges used by Perrin et al. (2003), you should also argue why you use different ones. Since there are only 4 parameters the information could also be added in the text.

P24: Figure 4 – please add labels to the FDC-plot. Why did you select catchment 15 which is within the 50

P26: Figure 6 – The authors often mention that the calibration with the regionalized FDC results in hydrographs with poor timing. Such timing issues are not obvious in the plots of Fig. 6. Thus, I would recommend to show time periods or catchments where timing really is a challenge.

P27-29: Figures 7, 8 and 9: These plots all look very similar to me and I recommend to condense or reduce the information of these plots. In my opinion it is not necessary to show the calibration values, I would rather focus on validation efficiency because that's the tougher criteria. I also think it's not ideal to compare the regionalization approaches in this way and I recommend to use the concept of benchmarks for comparison: e.g. make the difference between the benchmark strategy (RFDC_cal) and the PROX_reg strategy. The use of benchmarks results in one single value which can easily be interpreted: e.g. positive values mean that PROX_reg is better than RFDC_cal.

P29: Figure 9: I like the idea of evaluating further signatures and using them as additional constraints in model calibration. However, this plot does not give enough evidence for the conclusions drawn. To show that RLD and RFDC_cal are uncorrelated different methods are needed. I am also not convinced that the additional use of RLD would improve model calibration with the regionalized FDC more than RQP.

Minor comments:

Please use the HESS guidelines for all abbreviations, so that all are done in a similar style as e.g. the abbreviation of the baseflow index. Please also write Figure 9 and not Fig. 9 when you refer to it at the beginning of a sentence.

P1 L1: I would adapt the title: ". . .and model calibration with regionalized flow-duration

curves. . ."

P1 L12: Shouldn't it say "Leave-one-out cross validation"?

P3 L1: Why do you consider the selected signatures as "major signatures"?

P3 L11: Please provide numbers for the percentage of precipitation falling as snow.

P5 L25: Where does this equation for calculating MAP* come from? Why do you need the constant?

P7 L14: I recommend to integrate this whole chapter in chapter 3.1, because it is about regionalization and not about evaluation.

P7 L19: This information is already in the introduction and is not needed in the methods part. Furthermore, you cite different studies here than in the introduction.

P7 L25: I recommend to integrate this whole chapter in chapter 3.2, because it is about regionalization and not about evaluation.

P8 L2: I don't think that the regionalized FDC is used as objective function. It is rather used as reference value against which model simulations are evaluated.

P8 L25: Can you say what the CPI was for these catchments that were poorly modelled?

P10 L8: What about the efficiencies of catchment 13?

P10 L19: Please introduce the abbreviations such as PROX_reg earlier in the manuscript, because e.g. Fig. 5b already uses abbreviations.

P19 L1: Table caption should be adapted because NSE(Q) and NSE(FDC) are not catchment properties.

P21: Figure caption - it's the catchments that are labeled in the center and not the numbers. Also skip ". . .for GR4J model and FDC regionalization"

P23: Figure 3a – I agree that it is important to know that the model is able to simulate runoff in most catchments. However I don't think that a boxplot is needed for that. The median and the range of the model performance in calibration and validation could also just be mentioned in the text.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-487, 2016.