Hydrology and Earth System Sciences

Discussions

Open Access

HESSD

Interactive
comment

Interactive comment on "On the choice of calibration metrics for "high flow" estimation using hydrologic models" *by* Naoki Mizukami et al.

Anonymous Referee #2

Received and published: 4 November 2018

1 OVERALL RECOMMENDATION

The manuscript addresses the important topic of the choice of calibration metrics (CM) to be used for rainfall-runoff modeling, and presents results obtained on 492 US catchments. I found the paper interesting, including relevant references. If the presented results are not highly original, the paper is, in his present form, an excellent illustration of the limitation of the use of Nash and Sutcliffe efficiency metric (NSE, 1970) for model calibration. Nevertheless, I do have major comments on the used dataset, the applied methodology and the discussion part. Thus, I recommend to accept the manuscript in HESS with major revisions detailed below.

2. GENERAL COMMENTS

2.1 Description of the studied catchments

Even though the objective of such "large-sample hydrology paper" is not to present results obtained on a limited number of catchments having the same hydro-meteorological characteristics but to have general conclusions on rainfall-runoff modeling, I think the diversity of the studied catchments has to be addressed and quantified. This description is lacking right now in the paper. A presentation of the general characteristics of the studied catchments should be added in the paper, in order to understand the variability of catchments characteristics (catchment area, runoff coefficient, mean annual solid precipitation, etc.), especially in the context of flood modeling: what are different flood processes and dynamics included within this catchments sets (flash floods, snowmelt floods, rain-on-snow floods, groundwater floods, etc.)? Moreover, the timestep considered in the two rainfall-runoff models is not stated in the paper and should be mentioned. Are the models working at daily timestep? Is this timestep consistent with the flood dynamics of every studied catchment?

2.2 Split-sample test

For every catchment, the calibration and validation periods are the same time-periods, 1999-2008 and 1989-1999, respectively. I think that performing a basic split and sample test (Klemeš, 1986) on each catchment would be particularly interesting in this context, especially to address temporal (in)stability of parameter sets obtained with particular CM (topic partially addressed page 6, line 12).

2.3 List of the studied CM

The paragraph listing the studied CM (page 5, lines 5 to 16) is unclear and would be easier to understand if a list (or table) of the five studied CM was added.

2.4 "Application-specific" or "hydrologic signature"?

From page 5 to the end of the paper, APFB is named as an "application-specific" metric, while being introduced as an "hydrologic signature" (see definition of "hydrologic

signature" in the paper introduction, page 2 lines 10 to 24) in the paper objective presentation. What is the difference between an "application specific" and a "hydrologic signature" CM in this context? Finally, is APFB an "application specific" or an "hydrologic signature" CM? Could you address this point?

2.5 Impact of the KGE scaling factors

The limited impact of the different KGE scaling factors used in the paper is very little discussed, while being particularly interesting. This point has to be discussed in the paper. Moreover, why not trying another combination with a larger variability ratio scaling factor, such as (Sr=1, Saplha=20, Sbeta=1), to assess a potential significant improvement of annual peak flow bias? What about another test with (Sr=0.1, Saplha=1, Sbeta=0.1) or even (Sr=0, Saplha=1, Sbeta=0)?

2.6 Figures

In general, the presentation of the figures could be improved for a better understanding:

- The performance metrics plotted have different names in the axis labels and in the figure legends (e.g. it is not explicit that "%biasFHV" is equal to "percentage bias of flow volume above 80 percentile flow duration curve" in the Figure 4) ;

- The name/typography of several performance metrics is changing over figures ("%bias Qpeak" on Figure 2 but "%biasQpeak" on Figure 4) ;

- Why not using Greek letters in Figure 2 x-axis?

- The link between the five CM and the figure legend is never clearly stated, and for example, the reader has to guess that "kge_2alpha" is equal to (Sr=1, Saplha=2, Sbeta=1).

- How boxplots have been constructed? What are the outlier points plotted below and over the boxplots?

3 SPECIFIC COMMENTS

1. Page 4, line 27: please change (Maurel et al., 2002) into Maurel et al. (2002).

2. Page 13, figure 2: please state in the figure legend that results presented in this figure are obtained with the VIC model.

4 REFERENCES

Klemeš, V., 1986. Operational testing of hydrological simulation models. Hydrological Sciences Journal 31, 13. https://doi.org/10.1080/02626668609491024.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I – A discussion of principles. Journal of Hydrology 10, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Printer-friendly version

Discussion paper