Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-124-RC1, 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Global evaluation of runoff from ten state-of-the-art hydrological models" *by* H. Beck et al.

L. Gudmundsson (Referee)

lukas.gudmundsson@env.ethz.ch

Received and published: 5 July 2016

OVERALL RATING:

The paper presented by Beck et al is concerned with the tedious but important task of model evaluation. Overall the paper is interesting in scope, well written and the results are clearly presented. Consequently, I do definitely support the publication of the presented work.

Nevertheless, I do have several comments/suggestions which the authors may wish to consider prior to the publication of the manuscript in a final form. For the sake of clarity, I do list "specific comments" and "small comments" below.

SPECIFIC COMMENT:

Printer-friendly version

Discussion paper

** Specific Comment 1: ** While the paper is generally clearly written and most of the conclusions are supported by quantitative evidence there is a tendency for value statements (e.g. p. 2, l. 26: “NSE . . . is [a] . . . flawed metric”), claims (e.g. p. 1, l. 11: “. . . more effort should be devoted to calibration. . .”) or speculations (e.g. p. 9, l. 32: “. . .performed well. . . due to the lack of baseflow. . .”), which are not clearly highlighted as the authors interpretations or opinions. Although I do value if researchers defend their views on specific topics, I do also belief that it is important to clearly separate “hard facts” (either theoretical or quantitative) from soft interpretation and opinions in a scientific text. Therefore, I would like to encourage the author team to carefully revise the text of the manuscript, aiming at separating opinions from facts that are supported by either theory or quantitative analysis.

** Specific Comment 2: ** I do highly value the analysis presented in Figure 3, as this is a compelling way to investigate the physical consistency of the considered models with respect to the coupled water and energy balance. Unfortunately, the authors did only conduct this analysis for four models that output potential evapotranspiration, Ep, which is used to compute the aridity index.

An alternative approach, which was actually used by Budyko (1974), is to compute the aridity index as the ratio λRn/P where P is precipitation, Rn net-radiation and λ the latent heat of vaporization. This has the advantage that the results are strictly interpretable in the context of the coupled energy and water balance. In addition, this would allow to evaluate the output of all considered models.

In addition, I would like to question the authors conclusion that the models are most realistic if they scatter around the Budyko-curve (ref. p. 12, l. 20). In fact, Budyko (1974) developed the curve on the basis of a limited number of catchment observations and it is a-priory not clear whether models need to be close to this empirical rule. I do, however, fully support the authors conclusions that data-points that are outside the energy and water supply limits are a strong indication for issues in the model physics.

Finally, the similarity Figure 3 with the work of Greve et al (2014, doi: 10.1038/ngeo2247), who used the budyko framework to evaluate the credibility of reconstructions of E, P and Ep and Greve et al (2015, doi: 10.1002/2015GL063449, and sorry for citing myself), who introduced a formal way to account for scatter in the Budyko-space caught my attention. Although I am not sure if this is beneficial in the context of the presented paper, I could imagine that the tools provided in both mentioned studies might be helpful do develop additional quantitative insights to model performance.

SMALL COMMENTS:

Page 1, line 8-9: Repeated use of “(uncalibrated)”. I assume one should be calibrated

Table 1: The way the authors formulate the description of Table 1 reads as this would be a comprehensive review of model validation studies. I am, however, aware of at least two further studies – again, sorry for citing myself - (Gudmundsson et al 2012, doi: 10.1029/2011WR010911, Gudmundsson & Seneviratne 2015, doi: 10.5194/hess-19-2859-2015), that conduct similar assessments. Therefore, I would encourage the authors to either emphasise that the list of studies mentioned in Table 1 is not comprehensive, or to provide a more systematic summary of previous assessments.

Page 4, line 10: “. . . the combination of Penman-Monthie equations. . .” reads strange.

Page 6, line 10: To me it is not clear why the square-root transform is necessary, please explain.

Page 6, line 10: Is Q1 (Q99) a very high or a very low value. In hydrology both definitions are used. Please specify

Equation (1) and associated text: Why do you not use the observations to determine σo? To me this would be much more intuitive and would help to avoid the usage of another dataset which is prone to estimation uncertainty.

Table 3, line 1: Which P dataset was used? Is it the same that was used to drive the models or another one?

Tables 4 & 5: I do like the detailed information, but it would be much more accessible if it could be presented in figures (e.g. bar-plots)

Figure 3: Colour scale for the density is missing.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-124, 2016.