Interactive comment on "Transferring model uncertainty estimates from gauged to ungauged catchments" *by* F. Bourgin et al.

Anonymous Referee #3

Received and published: 8 August 2014

Overview:

The aim of the paper is very clear: estimate global uncertainty of the model output in ungauged catchments. Overall the paper is well-structured. It is also concise, which in general in a good thing. However, at certain points throughout the text further explanation would be helpful to aid interpretation.

Main Points:

1) The Authors aim to estimate total uncertainty. However, in the text (including the title of the paper) they often refer to total/global uncertainty as 'model uncertainty'. The Reviewer thinks this can be misleading, as it sounds like the Authors are trying to assess the uncertainty introduced by the choice of the rainfall-runoff model.

C2981

2) The Authors suggest a way to estimate total uncertainty in an ungauged catchment based on neighbouring gauged catchments. Although the Reviewer does not have a problem with this, the way the Authors implemented this methodology may be faulty. Using the catchments shown in Figure 1 as an example, the errors estimated for the green catchment resulting from transferring information from the yellow catchments (figure 1 B) are probably not representative of the errors expected from the transference of the information from the red catchments to the grey catchment (Figure 1 A). The errors calculated for the green catchment based on the yellow catchments are likely to be smaller as the catchments seem to be nested. On the contrary, the prediction of the runoff hydrograph of the grey catchment uses four catchments from different river branches and therefore the Reviewer expects that the error in this case is higher. Therefore, the Reviewer believes that the way the catchments were selected to estimate the uncertainty is not adequate.

3) The paper lacks a critical evaluation of the methodology suggested.

Minor points:

1) American English and British English are used interchangeably. Some examples (among many others) include: on page 8040, line 21, 'modelling'; on page 8041, line 16, 'behavioural'; on page 8044, line 10, 'optimization'; on page 8051, line 9, 'characterize'.

2) Page 8040, lines 24-25: What do the Authors mean by 'prediction approaches'?

3) Page 8041, line 10: What are the parameter sets constrained on?

4) Page 8041, line 14: hydrographs or hydrograph?

5) Page 8041, line 16: How does the second step relate to the first step?

6) Page 8041, lines 10-19: In a Bayesian approach, like Bulygina et al. (2012) used, there is no distinction between 'acceptable'/'behavioural' and 'non-behavioural' parameter sets. All parameters are acceptable, though some are more likely than others.

C2982

Therefore, the Reviewer suggests the Authors to rewrite this sentence.

7) Page 8042, lines 9-12: This is an example of where the Authors were too concise resulting in an explanation that is not satisfactory. Before reading the rest of the paper, and solely based on this paragraph, it seems that the Authors are suggesting that neighbouring gauged locations are calibrated and the residuals between model prediction and the observed data at these catchments are used/transposed to the ungauged catchment for uncertainty estimation at this location. The Reviewer does not agree with this, as in the ungauged problem there are additional sources of uncertainty when compared to the gauged problem. For instance, additional sources of uncertainty introduced by the transference of information should be taken into account when the final goal is to estimate the global uncertainty of the model output in the ungauged catchment. This is, in fact, highlighted later on by the Authors (Figure 7 and Section 5.2, page 8050, lines 1-3). This needs to be more clearly explained in the early stages (e.g. Introduction) of the paper.

8) Page 8042, line 21: are instead of is.

9) Page 8044, line 2: Why did the Authors select 4 and 7 catchments? What is the justification for using these particular number of catchments?

10) Page 8047, lines 9-21: In general, the definition of sharpness is confusing and should be clarified. The Reviewer interpreted AWI as being [1-average width uncertain bounds/(Q95-Q5)], but this should be better explained. In particular, it is not clear which 'two values' the Authors are referring to on line 13. It is also not clear what the Authors mean by 'compared to the climatology', in line 15. In line 16, what is the percentage reduction of the average width in relation to? Line 17, reduced in relation to what?

11) Pages 8047-8048, Equation 1: It may be worth explaining what range of values would be expected for S, which values correspond to a poor prediction and which values correspond to a better prediction.

C2983

12) Page 8048, line 1: It may be worth clarifying what 'l' and 'u' are.

13) Page 8048, line 5: What does 'unconditional climatology' mean? Please clarify.

14) Page 8048, Equation 2: The Authors have used ISS on the left and on the right hand side. The Reviewer assumes that on the right hand side it should be IS instead of ISS. Please correct this, if that is the case.

15) Page 8048, line 11: Do the Authors mean skill score (IS) or interval skill score (ISS) here?

16) Page 8048, lines 21-23: The Authors say that the median values for reliability for GR4J and TOPMO are 89% and 90% respectively (also shown in Figure 6). Roughly half of the catchments are above the expected 90% value for the 90% prediction bounds, and the other half is below. Therefore, the Reviewer is of the opinion that the Authors should not say that "the prediction bounds are, in most of the cases, able to reflect the magnitude of the errors", when those cases represent only 50% of the cases. The Reviewer suggests that 'in most cases' should be changed.

17) Page 8048, line 24, and page 8049, lines 1-3: This comment links with comment 16. Why do the Authors use CR=0.7 as a benchmark, when they say beforehand that 0.9 should be expected for reliability? Using CR=0.7 as a benchmark is misleading as it makes the results seem better than they actually are. If the aim here is to estimate total uncertainty and a value of 90% is expected for 90%prediction bounds, the Authors should focus on CR=0.9. As said before, approximately half of the catchments present a CR<=0.9, indicating that for 50% of the cases the uncertainty bounds might be too narrow or biased.