Hydrol. Earth Syst. Sci. Discuss., 11, C2899–C2901, 2014
www.hydrol-earth-syst-sci-discuss.net/11/C2899/2014/

Hydrology and Earth System Sciences
Discussions

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

A. Viglione (Referee)

viglione@hydro.tuwien.ac.at

Received and published: 4 August 2014

In this paper an estimation of the total uncertainty affecting runoff prediction in ungauged locations is performed. The total uncertainty is estimated based on residuals of the estimated runoff at neighbouring gauged catchments treated as ungauged (i.e., in cross-validation mode). I like the pragmatic procedure for the estimation of the total uncertainty. In fact, I was recently involved in editing a book on runoff prediction in ungauged basins (Blöschl et al., 2013, already cited in the paper), where, consistently with this paper, "total uncertainty" was assessed based on the performance of runoff prediction obtained in cross-validation over many locations (see also Parajka et al., 2013, already cited in the paper).

One addition which, in my opinion, would make the results of this paper more interesting for the hydrologic community, would be to stratify the measures of reliability, sharpness and interval score as a function of climatic and catchment characteristics (i.e., aridity index, catchment area, catchment elevation, density of the gauging network, ...). In other words, is the method performing equally well in all France or are there problematic regions? If the latter is true, what could be the reasons? This could also serve to address the concerns of Reviewer #1 (Denis Hughes).

The Authors have chosen to assess the reliability, sharpness and interval score for the 90% prediction intervals. Does the method perform equally well for other prediction intervals? More generally, since the method gives an estimation of the empirical distribution of the error for different flow groups, why haven't the Authors checked the goodness-of-fit of these distributions, for example through an uniformity test of the non-exceedence frequency of the actual error values (see e.g., Laio and Tamea, 2007, pages 1272-1273)?

Overall I think that the paper is well written and sufficiently clear, even though I agree with Reviewer #1 in that some points could be better clarified. Even though I've asked to add some analyses, I think that a minor revision should be sufficient for that. Some specific comments follow.

Page 8044, line 5: I would suggest to shortly discuss here in what are the two models different. I understand that this may be found in the previous papers, but for readability I would summarise the main differences here too.

Page 8045, line 5: I have a concern about the "output averaging option". Since averaging many signals results into a smoother one (also in the case that they are correlated), are the extremes well predicted? If so, are the results in this paper affected by that? This could be checked, for instance, applying the procedure for the 98% interval.

Page 8047, Section 4: here the Authors introduce the concepts of "reliability", "sharpness" and "interval score". Regarding the first two, unless the concepts are new, which

is not the case, I would suggest to add references here to where these concepts are extensively discussed (e.g., statistical books?).

Page 8047, line 13: The sentence about the “two values” is a bit confusing here. The Authors intend the two average widths of the uncertainty bounds and of the historical flow quantiles, while at first I confused the two values to Q5 and Q95. I see that also Reviewer #1 had a problem with this sentence.

Page 8047, line 15: What is the climatology?

Page 8048, line 9: That’s related to the previous point. I do not understand what a climatological interval is.

Page 8048, “interval score”: I have difficulties to understand what S measures. Maybe some more information should be given to help the reader. I’ve seen that also Reviewer #1 has concerns about this.

Page 8049, line 21: Same here. What is the unconditional climatology?

Page 8049, Section 5.2: The results obtained using donor catchments as gauged are not surprising. They descend from the fact that the procedure is wrong, since calibration removes biases. It is interesting, though, to see the results from a wrong procedure. However I would stress in the section that the procedure would be “wrong” since the uncertainty of runoff prediction in ungauged catchments is of interest.

References: LAIO F., TAMEA S. (2007), Verification tools for probabilistic forecasts of continuous hydrological variables, HYDROLOGY AND EARTH SYSTEM SCIENCES, European Geophysical Union, pp. 11, 2007, Vol. 11, 1267–1277, ISSN: 1027-5606, DOI: 10.5194/hess-11-1267-2007

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 11, 8039, 2014.