

Interactive comment on “Filling the white space on maps of European runoff trends: estimates from a multi-model ensemble” by K. Stahl et al.

Anonymous Referee #1

Received and published: 9 April 2012

General comments

This article is an excellent contribution to the hydrological sciences, specifically the quantification of the ability of models to replicate observed flows from catchments minimally affected by direct human disturbance. The quantification of model errors is done more thoroughly than typical and uses the best streamflow gauges to quantify the interaction of climatic inputs and hydrological outputs. The article is also very useful in describing hydrological trends in areas without any or many streamflow gauges in Europe while fully recognizing the limitations of doing this. The article is well referenced and very well written with clear, complete, and concise text in most places. It should be understandable to non-modelers and non-hydrologists.

The thing that I think is lacking from this article doesn't reflect on the work completed

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper



and described, but is a lost opportunity: describing which large scale hydrological models performed better than others. There may be reasons for not doing this such as limited space or not wanting to be seen as criticizing some models; however, it would be quite beneficial to the modeling community. By comparing which models did better or worse in different regions or seasons, improvements could possibly be made in all models. Also, a new ensemble of models could potentially use only the best models, presumably improving the ensemble results.

Specific comments

p. 2009, line 27: Why was WaterGAP calibrated for this analysis, and would this lead to biases in the results, as compared to other models?

p. 2011, line 12: I agree completely with reporting trend magnitudes and not trend significance, even though a large majority of researchers report significance (and report magnitude only if there is significance). Few people appreciate that hydrological data is very likely not independent from year to year, or even from decade to decade (the issue of long term persistence). Also, few people appreciate the fact that insignificant trends, especially when large, can be quite meaningful to people and aquatic ecosystems.

p. 2012, lines 17-20: I don't fully understand this sentence, please elaborate or explain differently.

Figure 1: It's very difficult to distinguish the ensemble line, and the observation line could be more pronounced. Adding color to this figure would help substantially. It would also be helpful to label the horizontal axis of the upper two panels.

p. 2013, line 1: It's true that the majority of values are within 30%, but this would also be true for other values, since "majority" is a fairly vague term that to me means >50%. I think this should be a more precise statement such as the following statements, or this opening statement could be deleted.

p. 2013, line 10: I'm not sure what is meant by "somewhat more positive trends", both

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

variables have about 60% positive trends.

p. 2013, line 22: I believe the statement “The KS test detected significant similar distributions” is incorrect, since the null hypothesis is that the distributions are similar. The failure to disprove the null hypothesis doesn’t prove it, though it is consistent with it. This may seem picky, but it’s a very important distinction. One way to rephrase this statement would be to discuss the proven (at a certain significance level) failures of the null hypothesis in the summer and then state how the winter months were consistent with the null hypothesis.

p. 2014, line 8: This is an example of my earlier point about the importance of reporting on how well specific models did. It would be very informative to know if any of the models consistently performed better than others.

p. 2014, line 11: One model in February and one in October performed better than the ensemble, thus disagreeing with this statement.

p. 2015, line 1: I assume this refers to the ensemble mean for locations with observations, but this should be stated. Also, in line 2, I see the gradient as southeast to northwest, not southwest to northeast.

Figures 4 and 5: the square and “x” symbols are virtually impossible to see, until I magnified them to 200%. Please make them readable at the scale presented in the article. Being able to see them makes the text discussion of the results (relevant to these figures) much easier to comprehend.

p. 2021, line 8: I believe it would be more correct to change “predominantly positive trends in winter and negative trends in summer” to “more positive trends in winter and more negative trends in summer”. The gradient from positive in the north to negative in the south is still present in both seasons though there are more positive trends overall in the winter (and more negative trends in the summer).

p. 2012, line 6: It would be helpful in the conclusion to specifically mention snowpack

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper



and groundwater as storage processes.

Technical corrections

Figure 1: It would be more clear to refer to the “middle” figure as the “upper right” figure and the “upper” figure as “upper left”.

p. 2020, line 2: would be good for clarity to add “used in this study” after “The large scale hydrological models”.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 2005, 2012.

HESD

9, C737–C740, 2012

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

