

## ***Interactive comment on “Seasonal streamflow forecasts for Europe – II. Explanation of the skill” by Wouter Greuell et al.***

**Anonymous Referee #3**

Received and published: 9 February 2017

The manuscript is the 2nd of a companion paper addressing streamflow forecast performance across Europe. The methodology is fairly well described and presented, however substantial clarification and justification is necessary in numerous areas. The authors predominantly limit themselves to evaluation of only a few performance metrics, and report many findings for the whole of Europe. The overall contribution contains meritorious aspects, particularly the performance of this dynamical system, however these need to be highlighted and clarified significantly. Also, distinction and improvement between this and prior studies (e.g. Bierkens and van Beek) is not sufficient.

Specific comments: 1. The title indicates seasonal streamflow prediction, yet the paper focuses on Monthly results for streamflow, temperature, and evaporation. Title not entirely indicative of manuscript focus.

C1

2. Given the spatial heterogeneity of Europe, the authors should provide better justification for reporting predominantly spatially lumped results.

3. Auto-regressive effect (streamflow persistence) not explicitly mentioned or discussed. Is it assumed to be (partially) accounted for in initial soil moisture? For most rivers, particularly large rivers, this is a dominant feature.

4. GRDC has discharge stations downstream of reservoirs, where regulations and management of discharge is often evident. But these have not been corrected (or even noted) in the dataset. Europe is full of situations like this. How have those been accounted for?

5. Unclear (no explanation) of what the ratio of actual/theoretical skill means. Clear that they are closer for larger basins (no surprise) but does the fact that both are far from 1 indicate less “realistic” outcomes? Or does this have little bearing on skill metrics (comparing apples to apples.) Please clarify.

6. How do the three “conditions” relate with the pseudo-obs? For example, report R or RPSS between soil moisture and streamflow by grid? Or snow and streamflow? Could assess for at least a sub-set of locations. This would also give insights as to the value added (or not) by VIC.

7. In addition to reporting the % of cells where R is significant, consider also reporting the mean and standard deviation of R in those cells. The number of significant cells does not necessarily represent the quality of the relationship (e.g. % of cells could increase, but mean decrease. . .) And then discuss.

8. RPSS is mentioned early in study, but results are not presented. Such categorical skill scores are worth exploring.

9. The authors lightly compare their study outputs with others, namely Bierkens and van Beek, indicating lower performance, likely attributable to the latter’s use of semi-statistical forcing. While there are still other meritorious aspects to this current contri-

C2

bution, the authors do not adequately discuss the implications of poorer performance. Are there reasons that the proposed methodology is advantageous as compared with others? Should the GloSea5 approach be used in lieu of the one proposed here? More discussion is needed.

10. In the Conclusion, the authors mention the potential improvement of assimilating soil moisture or SWE to VIC. Why was this not performed and analyzed?

11. Challenge to follow train of thought in some parts. Could benefit from the writing be tightened up overall - and simplified in some places. Word choice also needs to be improved in many places (e.g. "Fig. 8 analyses [sic] a remarkable feature." Figures cannot analyze. Figures can illustrate.)

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-604, 2016.