Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-372-RC3, 2017

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Monthly streamflow forecasting at varying spatial scales in the Rhine basin" *by* Simon Schick et al.

J. Beckers (Referee)

joost.beckers@deltares.nl

Received and published: 22 August 2017

Simon Schick et al. on the point of seasonal forecasting: Yes, we agree, the term 'season' needs to be used more carefully, especially when it comes to the actual results of the study (e.g. page 1, line 14). In the introduction, however, we tried to give a general overview of the approaches that one could apply to integrate GCM output for seasonal (or subseasonal) streamflow forecasting: The traditional model chain GCM - hydrological model; runoff from the land surface component of the GCM run; and PP and MOS based forecast strategies. We do not claim any of these approaches to be superior nor that using GCM output is superior to the ESP approach and its variants. Rather it is an attempt to summarise the existing toolbox, without stating that any particular model is skillful at the (sub)seasonal time scale. The fact that skill of the

tested methods in case of the Rhine basin is restricted to one month ahead forecasts does not invalidate this list – these options still exist (from a technical point of view) and might work or not, depending on the study region, skill of the GCM, validity of the model assumptions, and so on. A forecast at the seasonal time scale remains one, be it skillful or not.

Response from Joost Beckers: Yes but using the term seasonal forecasting creates expectations. Since no skill is found beyond the one-month lead time, I would recommend to keep references to seasonal forecasting minimal.

Simon Schick et al. on the suggestion that the performance of MOS may be better for particular calendar months: We agree; it was not our purpose to suggest that the models might better perform in particular months, but that the findings are valid only for forecasting all calendar months. Moreover, we should stress that the model might perform better or more worse in particular calendar months, so the results represent somehow an "average"month.

Response from Joost Beckers: I agree with the remark by K. Foster that it would be even better to demonstrate this variation in performance in an additional figure or table.

Simon Schick et al. on the limited skill of H-TESSEL at the seasonal time scale: Yes, we agree, this conclusion is too loosely formulated (again it boils down to the correct usage of the term 'season'). It also implies several arguments that we would like to reformulate here. The present study tries to test approaches for seasonal streamflow forecasting that have not received much attention in the scientific literature so far, i.e. a MOS approach and runoff simulations from a coupled GCM. The latter simply emerged out of curiosity: While downloading the precipitation and temperature hindcasts from MARS, we realised that runoff is also contained in the list of available parameters. We then decided to request runoff as well and started to use it as a benchmark for the MOS approach. To avoid any misunderstanding: We did not run H-TESSEL at all, we only applied a linear bias correction. Thus, the simulations correspond to the model configuration used within ECMWF's S4 setup. To our surprise it turned out that these runoff simulations are quite hard to beat, for which reason we included them in the present study (please note that the MOS approach works on an acceptable level when fed with the best available input data, independently of the lead time). Clearly, this finding does only hold for the Rhine basin and H-TESSEL, but nevertheless we think that this point is interesting both from an operational as well as scientific perspective. Operational, since the runoff simulations are already present after the GCM finishes his run (so no need to run a model by oneself, ECMWF and other institutes already did the job, and notably in a fully coupled way). Scientific, since the land surface model is not aimed to forecast streamflow, but to provide a lower boundary condition for the simulation of the atmosphere. Thus, if we already beat streamflow climatology with runoff from the land surface model, to which extent can we benefit by using models specifically tailored to forecast streamflow? For example, what are the benefits of including river routing algorithms, groundwater or lake modules, or a higher horizontal resolution? Studies that use GCM output for seasonal streamflow forecasting need to download the necessary parameters anyway, so why not add runoff and use it as a benchmark?

Response from Joost Beckers: Thanks for that elaborate answer. Indeed the H-TESSEL is a excellent benchmark and it was good to include it in the analysis.

Simon Schick et al. on the testing of MOS against ESP- or GCM-driven hydrologic models: Indeed, it would be very interesting to contrast the two tested approaches with the 'traditional' forecast strategy. Since to the best of our knowledge no study exists that we could use for such a comparison, it would be left to us to set up a model and run it with the seasonal climate predictions. This, however, is in our opinion not feasible in the context of the present article, but could lead to a next study.

Response from Joost Beckers I understand that this would go beyond the scope of your project. Note, though, that there are organizations that have such Rhine models running operationally (e.g. BfG, Germany and Rijkswaterstaat, Netherlands).

Simon Schick et al. on the version of H-TESSEL: Yes, we agree. With respect to the Rhine at Lobith and Basel, the runoff simulations from ECMWF's S4 in combination with a linear bias correction score the same skill as the tested MOS approach does. This needs to be more stressed in the summary as well as in the conclusion. However, we think it is not valid to conclude in general that MOS methods are not a viable option for operational streamflow forecasting. Neither can we speak for the whole MOS family nor for other input data nor other parts of the world. Rather, we argue that MOS in general is part of the "seasonal streamflow forecasting"-toolbox – with all the disadvantages and advantages of a black box model.

Response from Joost Beckers: I agree that a general conclusion about the viability of any of the methods would go too far. Your last sentence puts it in the right perspective.

Simon Schick et al. on the point of resolution and ESP-revESP: We would like to take up again the point concerning the resolution of the lead time, since we probably misunderstood the suggestion of J. Beckers in his review. In the following, we use 'lead time' as the time interval between the release of a forecast and the onset of its validity. For example a mean streamflow forecast for April 21 - 30, produced at March 31, has a lead time of 20 days. The MOS approach of our study is based on the assumption of linearity and thus needs a certain time window to average the actual predictand. For monthly streamflow averages this assumption seems to be more or less valid. Increasing the temporal resolution of the predictand (e.g. the prediction of 5, 10, 15, ... day mean streamflow at zero lead time) could be an interesting experiment, but rather to test the assumption of linearity than for a detailed ESP-revESP analysis. This is clearly a disadvantage of using regression instead of a hydrological simulation model. However (and this was eventually already proposed by J. Beckers) shifting the time window in steps of 5,10,15... days (that is using short lead times) could also reveal some insights. It does not help concerning the ESP-revESP experiment, which remains unresolved at a submonthly time scale, but to detect the skillful time range of the seasonal climate predictions. For example, if the monthly streamflow forecasts based on the seasonal predictions arrive at the MAE of the ESP model at 15 days lead time, we could argue that skill of the seasonal climate predictions is restricted to the first 15 days. Do you agree with that line of argumentation?

Response from Joost Beckers: Yes, I agree. Your proposed investigation would reveal how the skill at zero lead time decays to the ESP value at some longer lead time. If it the ESP value is already reached after one week, then the conclusion must be that the skill at zero lead time is entirely due to a good prediction of the flow for the first few days after forecast time. Possibly, you can even reconstruct the MAE for these first few days from the monthly flow MAEs at zero and 1-week lead times. However, this would be a somewhat indirect way of determining the skill at higher temporal resolution. I do not understand why a more direct analysis of the skill for the first week is not possible. Why does the MOS method require a flow averaging over a monthly interval? If you would apply the method to weekly or even daily flows, surely the results will become more noisy, but on average the linear relationships would still hold, or not? If such an analysis at a shorter interval would be possible, this would also enable to investigate the crossover from ESP to revESP in more detail.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-372, 2017.