

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

Simon Schick et al.

simon.schick@giub.unibe.ch

Received and published: 16 August 2017

General comments

The authors thank for the careful review and the clear comments. We think that several comments are related to the terminology used in the article, as we often borrow terms from studies with hydrological simulation models, e.g. the 'ESP' and 'revESP' models, 'initial conditions', 'meteorological forcings', and so on. We do not insist to retain these terms in the article and are open for suggestions.

Specific comments

Is there a reason why the initial hydrological conditions are not included as predictors

C1

(page 3-lines 10-12 and table 2)? Predictors related to storages such as soil moisture content, snow, and reservoir/lake levels all impact future streamflow yet only meteorological predictors are used. I agree that many of these initial storages are affected by the antecedent meteorological conditions, but these connections are not necessarily linear or significant depending on the time frame used. For example, if only predictors for the preceding month are used then there is little connection to snow pack size or reservoir levels and therefore little added value. Thus I miss a description of the time period, and to a lesser extent the domain, for the predictors.

We agree, there exist many other potential predictors. We restricted the set of predictors to precipitation and surface air temperature for practical reasons: These variables are available as gridded products, cover the entire study region (and thus are present in all subcatchments), and are available for a long time period; the assumption of independence is more or less valid; and the regression strategy stays simple. As long as it is reasonable to include precipitation and temperature of the target season in the model, then it does so for the 'initial conditions' too. In fact, using this restricted set of predictors guarantees a fair comparison of the predictor combinations and spatial levels as they all rely on the same source of data.

In case of the 'preceding' predictors (the predictors that act as a proxy to catch the initial conditions via the antecedent meteorological conditions), the time aggregation is allowed to vary between 10 and 720 days. The predictors are defined as catchment area averages. Two example plots showing the regression coefficients and aggregation periods of the refRun model at Lobith and Basel (spatial level 1, predictand=30 day mean streamflow, n=100) are at the end of the document. The interpretation of the resulting pattern throughout the year is tricky, but we can see e.g. an increase in the time aggregation of precipitation in spring (Basel level 1, top right plot), i.e. the model tries to catch the snow volume accumulated in winter.

If the article gets accepted for revision, we will clarify this point; if desired we can also include these two figures (but we suggest in the additional materials rather than in the

C2

results section).

Similarly, I question whether the use of the terms ESP and revESP are technically correct in this paper as it stands. Without any information regarding the initial conditions at the forecast initialisation one can argue that this is not similar to what Wood and Lettenmeier (2008) meant. If it were possible I would suggest the authors include predictors that represented the initial conditions (soil moisture, snow depth, or even streamflow) otherwise they should add a paragraph explaining why the current approach is still an adaptation of the VESPA methodology. I believe that the latter may be difficult to justify especially with respect to revESP.

We agree, the naming of the different predictor combinations is debatable. However, please note that a connection to VESPA does not exist at all (we also do not cite this article). Our intention was to disentangle the predictive power of the 'information prior to the date of prediction' and 'information following the date of prediction'. In our understanding it is exactly this idea that lies at the core of the ESP-revESP framework.

I echo Joost's point where he suggests that the suggestion that the performance may be better for particular months (page 17-lines 15, 16) is unfounded as the article stands now. However, I do expect this to be the case and therefore I disagree with him in that this should be removed. Rather I think it would be of interest to include some results or a section that addresses this variability. This can be done in part in the form of a figure along the lines of the one below (figure 1). Related to this, why are the authors concentrating only on the general performance throughout the year? The usefulness of these forecasts may be much higher, even only, during specific times during the year e.x. during the snow melt period or low flow period.

We wrote the article with the intention to test the MOS and ECMWF's seasonal climate predictions along several spatial scales, and so decided to ignore variations of predictive skill within the calendar year. We strongly expected that skill of the seasonal climate predictions – if present at all – peaks at a particular spatial level. For example,

C3

the spatial averaging of the fields for Lobith at spatial level 1 might remove important information (i.e. where do we have a climate anomaly?). On the other hand, regressing streamflow of small catchments (say 100 km²) against climate predictions that represent spatial averages of about 5000 km² could simply amount to a scale mismatch.

We attached a figure following the template of Kean Foster for the S4PT model and the MAE skill score with respect to climatology (however, we did not test for statistical significance). As expected by Kean Foster, skill varies within the calendar year. Maybe we can add a short subsection and discuss this variation of e.g. the S4PT model in the results section. To test for significant deviations from zero we might need to resort to a bootstrap in case the distributional assumptions of the test applied to the complete monthly series are not valid.

With regards to H-TESSSEL, Table 4 shows that it has some skill, at least at the spatial level 1. Have the authors tested using these data as predictors in the MOS approach at levels 2 and 3?

Yes, we did. The MAE remains virtually the same: 419, 417, 417 m³/s (Lobith levels 1-3) and 191, 186, 184 m³/s (Basel levels 1-3). Please note that runoff from H-TESSSEL is also included in the S4Q model (but here in combination with the antecedent meteorological conditions and the time aggregation screening). If desired, we can add these values to Table 4.

I am unclear as to whether the S4 data is bias corrected. It is now almost common practice for some sort pre-processing or bias correction of the S4* forecast data before use in hydrological forecasting studies and work. The authors note that the quality of seasonal climate predictions for the study area are low (page 3-lines 20,21) but it is not clear to me whether any attempt to bias correct the data, and if I did miss it by what method.*

We did not apply any bias correction, since we think it is not useful in case of statistical methods (at least we do not know any study that uses bias corrected predictors for

C4

a regression model). The present formulation of the MOS approaches catches any systematic linear error via the regression coefficients. Obviously, this does not hold for nonlinear systematic errors, but we question that e.g. quantile mapping improves the prediction accuracy, since we work with mean values corresponding to at least 5 days.

The statement about the quality of seasonal climate predictions for the study region should be interpreted from a physical point of view: In the midlatitudes, climate is less dominated by the ENSO, resulting in less skillful seasonal climate predictions when compared to the tropics.

Lastly, the authors mention how the uncertainties in forecasts can be reduced when the quantity of interest is controlled by teleconnection phenomena (page 1-line 17-19). I don't contest that this is true but rather question how it is relevant to the paper because there does not seem to be any more references to such modulation activity or its importance in the rest of the paper.

We agree, this statement is not strictly necessary for the article. Rather we tried to sketch the basis for environmental seasonal forecasting in order to start somewhere with the article. Please note that the cited 'slowly-varying and predictable phenomena' are not restricted to the thermal coupling of the oceans and the atmosphere (and potential subsequent teleconnections) – a strong cycle of snow accumulation and subsequent melting or persistence in soil moisture are other examples.

Technical comments

On page 9-line 12 the authors give a secondary citation where I feel that the original citation, or at least inclusion of the original would be strongly advised. Taylor's original article is: Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research: Atmospheres, 106(D7), 7183-7192. The authors are encouraged to check their other sources.

We agree, the original citation is more appropriate.

C5

Lastly, there are some minor grammatical errors in the paper; however these do not detract from the readability or arguments made therein. All the same I do suggest that the authors spend a little time to minimise them if time allows.

We are not native English speakers, but will try our best.

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

C6

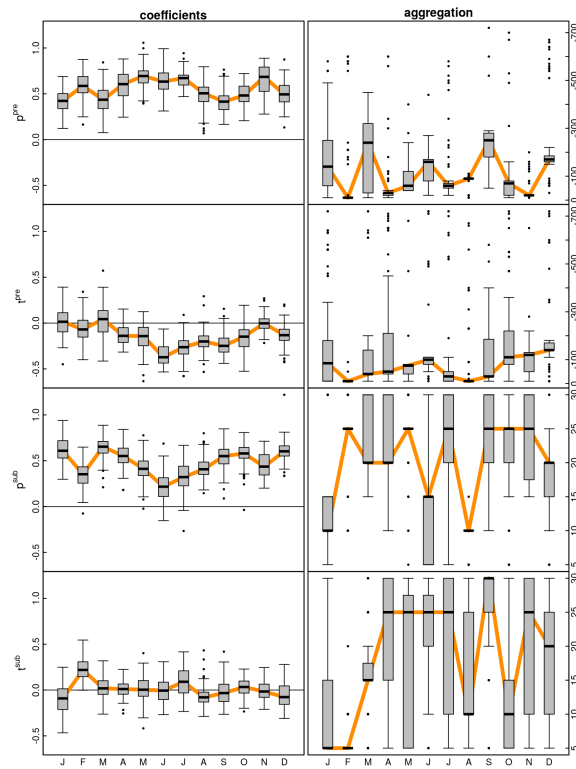


Fig. 1. Regression coefficients and aggregation time periods of the refRun model at Lobith and spatial level 1 in case of 30 day mean streamflow forecasts..

C7

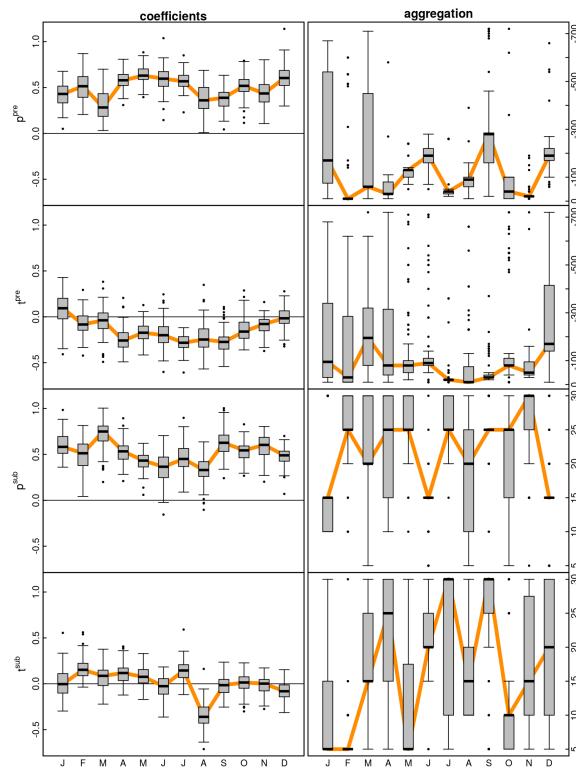


Fig. 2. Regression coefficients and aggregation time periods of the refRun model at Basel and spatial level 1 in case of 30 day mean streamflow forecasts..

C8

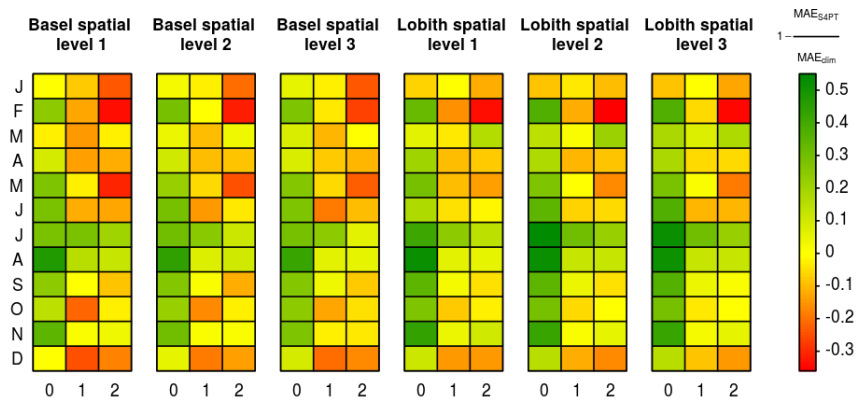


Fig. 3. MAE skill score with respect to climatology of the S4PT model at different lead times and date of predictions.