

## **Review on: The development and evaluation of a hydrological seasonal forecast system prototype for predicting spring flood volumes in Swedish rivers**

### **Summary of the manuscript**

The manuscript shows, how a hydrological seasonal forecast system prototype is adjusted from the previous version and evaluated on its ability to predict spring flood volumes in Swedish rivers. The aim is to improve water resource management for hydropower decision makers. The study area consists of 84 subbasins in northern Sweden, which have a runoff regime that is strongly influenced by spring snow melt. The skill of the multi-model prototype is compared in cross-validated hindcasts to the historical ensemble streamflow prediction based on measurements between 1981 and 2015. This historical ensemble represents the setup currently used for hydropower reservoir management. The multi-model prototype represents combinations of the historical ensemble, an analogue ensemble (subset of the historical ensemble based on similarities in parameters of interannual climate variability), a dynamic modelling ensemble (bias-corrected seasonal forecast) and a statistical modelling ensemble (downscaled seasonal forecast). Several complementary, statistical measures were used for the evaluation of the new prototype. The prototypes, that combine 3 different ensembles show at best a significant improvement compared to the currently used historical ensemble and at worst a comparable skill.

### **Main assessment**

Based on our assessment, the reviewed manuscript reaches a substantial conclusion based on sufficient results which generally were outlined clearly and used valid assumptions. Overall, we like the clear structure of the paper and the scientific notation. In the abstract and the introduction, it is nicely explained why there is a public benefit behind this research.

When we first read the introduction, we had some problems to understand the differences between the two approaches (dynamical and statistical). When we came to the points where it is explained better it is not a problem anymore. Maybe a quick hint to the section 2.14 and 2.15 would help the readers - or provide some of the clarifications already in the introduction.

It would have been useful to add the research questions to the introduction. The paper describes what has been done, how it was tested and how good the new prototype is for catchments in Sweden but not explicitly what science question or hypothesis is answered. The paper also does not discuss how applicable this method/prototype is for other snow dominated areas.

Thanks to the SMHI, the datasets are very good, but we don't understand why the subbasins with 61, 63, and 73 percent of missing data are also included in these 84 basins that are investigated. Has it at least been checked whether and how the results change when those three subbasins are excluded?

The evaluation part from this study is very elaborate and we like this. However, it could have been more clearly explained what a 6% improvement means, for example in volume of water or economic value for hydropower generation. Is this 6% really significant?

We would have appreciated it if the paper provided a little more information about the seasonal meteorological forecasts that are used. For example something about the uncertainty of these models or why the ECMWF IFS system is used (is there no other meteorological forecast system for six months or is it the best seasonal forecast system)?

Finally, it would have been interesting if the differences for the different catchments would have been discussed. Were the improvements mainly seen for large/small catchments, for high elevation/flat catchments? Some maps would have been nice as well.

### **List of major and minor points**

Page one:

Lines 8-10: In our opinion a very catchy and smart opening.

Line 15 to 18: This sentence is maybe too long and little too complicated for the abstract. (Full stop before 'however'?). Twice 'considered'

Line 23: Unclear reference, presumably 'Statistiska centralbyrån'? For clarity, the abbreviation can be included in the reference (Statistiska centralbyrån (SCB): ...)

Line 27 to 28: The idea or point behind the sentence comes across. But if you first read this sentence, it could be puzzling. We also think that with all these brackets the text looks not as nice as it could. Why not just add 'and vice versa' at the end of the sentence?

Page two:

Line 3: Grammatical: The strategy is to have reservoirs which are then managed.

Line 4: comma: To achieve this, operators ...

Line 6: The meaning of the expression 'sources of predictability' is unclear to us in this situation. Can you explain briefly?

Line 7: Decide on stores within or in the catchment. We suggest using within.

Line 11: Probably : instead of ; after 'forecasts at the seasonal scale'.

Line 15: There is no need for a comma after the closed bracket. Both times.

Lines 19, 20: The second time 'force it' (end of line 19) is not necessary.

Lines 31, 32: Do the references explain how and to what extent historical observations of precipitation and temperature are possible representations of future meteorological conditions in the context of the ongoing climate change? Do the historical data show any significant trend?

Page three:

Line 13: Maybe already mention here the number of catchments and data period. That way we already know something about how these modeling steps are applied. Otherwise the 35 in Line 35 is not so clear.

Line 14 to 15: Twice the word "brief"

Line 19: first improved by Foster et al. (2010) ~~and~~, later improved ~~upon~~ and first tested by Olsson et al. (2016).

Lines 23, 24: Here the manuscript includes already some results but is in the Materials and Methods section. Finish sentence after '... of these four were tested.'

Line 28: Replace 'relevant' by 'respective'.

Page four:

Line 7: Is there a reference on what the seasonal forecasting practice at SMHI is?

Line 9: It is not clear to us how the DBS method is different from the previous method.

Line 26: We are not familiar with the teleconnection approach. A brief explanation would have been useful.

Page five:

Line 2: What is the meaning of  $\alpha$  and the justification for a distance of 0.2?

Line 29 and following: Is there a reference (needed) for the physical support of the asymmetric weighting?

Page six:

Line 24: Typo: overfitted

Line 28: So  $n = 35$  in this case?

Page seven:

Line 27: Why is the relative mean absolute error used (error divided by  $SFV^{\alpha}$ )? Would the absolute error (without the division) not be a stronger focus on the flood peaks and therefore more beneficial for the assessment of the skill of forecasting the spring flood? Please explain.

Line 28: The superscripts suggest that it is SFV to the power of  $\alpha$ . A different notation is clearer.

Page ten:

Line 6: Subbasin.

Line 6: There should be a reference to figure 3 included in the sentence.

Line 10: How is total runoff divided between the 3 subbasins? Would it be interesting to use the new setup also for the two other subbasins?

Line 17: SFV already introduced on page 3.

Line 27: We don't understand why the subbasins with 61, 63, and 73 percent of missing data are also included in the 84 basins that are investigated. Has it been checked whether and how the results change when excluding those three subbasins?

Line 31: What is the PTHBV dataset from SMHI? is this the ptq file for HBV (but without the q)?

Page eleven:

Line 4: Do you refer to the performance measures for the HBV rainfall-runoff model of the historical data?

Line 7: The 'than' near the end of the line is missing one or two adjectives describing latitude and elevation. Decide on either and or or.

Line 22-24: We don't really understand where the 15 and 51 come from. We probably missed that somewhere, or is it not explained anywhere?

Page thirteen:

Line 21: Are these 5% of the catchments located in a certain area or do they have similar characteristics? Large/small, steep/flat, northern/southern?

Page fourteen:

Line 29: Why was exactly this site choose for the analysis of the forecast ensemble sharpness? Just one sentence, why exactly this basin is relevant out of the whole set of the 84 sites.

Page sixteen:

Line 5-18: This is more discussion than real conclusion. Move?

Figures:

Figure 2: In the figure description it is mentioned that the spring flood is in the period between the onset and the last day of July. On the x-axis the days since the 1st January are written. We think that for the reader it would be easier if on the x-axis the first of the months were labelled and the date of the 80 days threshold was indicated (20<sup>th</sup> or 21<sup>st</sup> of March).