Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-370-AC2, 2017

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Retrospective forecasts of the upcoming winter season snow accumulation in the Inn headwaters (European Alps)" *by* Kristian Förster et al.

Kristian Förster et al.

foerster@iww.uni-hannover.de

Received and published: 16 October 2017

Reply to anonymous reviewer #2

(Reviewer's comments are in italics)

> *This study explores the forecast skill of snow water equivalent (SWE) by using CGCM- driven water balance model simulations over a headwater region. While the topic is quite interesting and some results (e.g., GloSea5-driven forecasts) are potentially promising, the manuscript could be further improved after addressing several comments below.*

We would like to thank Anonymous Reviewer # 2 for his/her detailed review of our manuscript. The comments will help us in the process of improving the discussion paper.

> *Major comments:*
>
> *1. An interesting question that could be answered in this manuscript is whether precipitation or temperature prediction more important for the SWE forecasting over the headwater region. Although precipitation prediction is less skillful than temperature in many cases, the study region shows less skillful temperature prediction, perhaps due to the deficiency in snow or frozen soil processes. To compare their relative roles, precipitation or temperature forecast could be replaced with climatology before driving the SWE model. Such comparison would provide implications in advancing SWE forecasting.*

We appreciate your suggestion! Reviewer #1 also asked us to address the relative roles of temperature and precipitation more explicitly. Indeed, a model experiment that includes both dynamical and climatological forcing data helps to analyse the relevance of temperature and precipitation for SWE forecasts. We will follow your suggestion and we performed such a model experiment. Accordingly, the following runs are performed:

- temperature from models, precipitation from models (this is the configuration we have applied so far)
- temperature from models, precipitation from climatology
- temperature from climatology, precipitation from models

The second and the third model runs is analysed in the same way as already done with respect to the first model experiment. At this stage, we will also keep in mind your

suggestions outlined in the second comment of your review. All performance and skill measures will be summarised in an additional table.

> *2. The study shows that GloSea5-driven SWE forecasting is better than the CFSv2- driven forecasting in terms of pearson correlation for the ensemble mean, but did not tell why the former is better? Some information on precipitation and temperature forecasts could be mentioned in the abstract. Moreover, probabilistic metrics (e.g., RPSS) is needed besides just simply using correlation. Given that this manuscript is not a short communication, I would encourage the authors to have a more comprehensive evaluation for SWE forecasting.*

There are many differences between the CFSv2 and GloSea5 systems that could in principle explain the higher skill of the GloSea5 system and a full answer to this question is beyond the scope of our study, but one likely reason is that the skill of the NAO/AO is higher in Glosea5 (Scaife et al 2014) than in CFSv2 (Riddle et al 2013).

We also added more details on precipitation and temperature forecasts to the abstract:

"Even though predictions for precipitation may not be significantly more skilful than for temperature, the predictive skill achieved for precipitation is retained in subsequent water balance simulations when snow water equivalent (SWE) in February is considered."

We agree that there are many more skill measures which could be addressed in our analyses. We follow your suggestion to add some more metrics. As we are using the ensemble mean instead of using individual ensemble members, we decided to use the deterministic (single value) metrics where appropriate:

- The Continuous Ranked Probability Skill Score (CRPSS) is equivalent to the Mean Absolute Error (MAE) in case of a deterministic (single value) forecast,

  which is why CRPSS is used as measure representing the mean absolute error of forecasts. Here, we compute the MAE skill score (MAESS).
- The Ranked Probability Skill Score (RPSS) is equivalent to the Brier Skill Score (BSS) if a two categories forecast is considered. Thus, we will also compute the BSS values.
- We will also compute RMSE as suggested.

The revised version of the manuscript will include a table that provides these metrics. Similarly, the results of the model experiment suggested in your first comment will also analysed using these metrics. Thank you for these suggestions.

> *Minor comments:*
>
> *3. Does the AWARE water balance model distinguish the input of liquid or solid precipitation? If so, how to obtain the solid precipitation from global climate forecast model like CFSv2?*

Yes, it does. This information was still missing. We have added the method how the phase partitioning is performed in the model:

"For each grid cell the relative contributions of rainfall and snowfall are computed taking into account two threshold temperature values. If the air temperature falls below the lower threshold temperature, the monthly precipitation depth is assumed to be snowfall only. In contrast, air temperatures exceeding the upper threshold indicate rainfall only. In order to enable the occurrence of both snow and rain, a transition range between both thresholds is defined. Based on air temperature, the fraction of rain and snow is linearly interpolated between these both thresholds."

> *4. What is spatial resolution for the AWARE model over the study catchment?*

We are sorry that is important information was also missing. Reviewer #1 also asked us to specify the spatial resolution of AWARE. It is 1000 m for the Inn headwaters.

5. What is the definition for the benchmark Nash-Sutcliffe efficiency?

In the revised version of the manuscript, we will provide the definition of the benchmark Nash Sutcliffe model efficiency in a new Appendix section along with the other metrics as suggested in comment #2.

6. For the benchmark NSE during the validation period, why does it drop to 0.25? Is it because there is trend or non-stationarity in the time series?

The benchmark Nash-Sutcliffe model efficiency is more sensitive to differences between two time series than the standard Nash-Sutcliffe model efficiency. We will explain the differences in the revised manuscript. The lower performance of the validation period might also be related to reservoir operations. We already addressed the need for a better reservoir representation in the outlook. Moreover, changes in glacier characteristics are not yet fully addressed by the water balance model. Since the calibration period was subjected to negative mass balances, positive mass balances have been observed in the 1980s. This refer to your suggestion to consider possible non-stationaries in the time series. We added the following lines to the manuscript:

Model description section:

"A possible reason for the lower *Eb* value might be the fact that the validation period has seen an advancing of glaciers due to positive glacier mass balances. In contrast, the calibration period is characterised by a shrinkage of glaciers volumes. Both processes are not incorporated in the model so far."

Outlook:

"Moreover, a better representation of changes in glaciated area is currently being investigated through coupling AWARE with a glacier evolution model developed by Marzeion et al. (2012)."

7. Figure 4. Besides correlation, how about the RMSE for the prediction?

Yes, we will add RMSE as well. Please refer to comment #2.

8. Figure 4. Is the model-simulated SWE or observed SWE used for verification? If the former, how to demonstrate the usefulness of the SWE forecasting given the limited skill in SWE simulation with AWARE (where NSE=0.25 in the validation period)?

HISTALP-AWARE computations of SWE were used to assess the skill. Reviewer #1 also commented on model uncertainty involved in hydrological modelling. We used the reference run (HISTALP-AWARE) for verification because further uncertainties are involved in running water balance model. The revised version of the manuscript will address model uncertainty more explicitly:

"However, the comparison between HISTALP-AWARE and the CM-based seasonal forecasts highlights GCM-forecast skill and acknowledges the fact that the water balance model is never perfect since it introduces uncertainties into hydrological forecasts, too."

Please also refer to our reply to your comment #6. We added some remarks regarding possible reasons that might explain the lower model performance.

References

Marzeion, B., Jarosch, A. H., and Hofer, M.: Past and future sea-level change from the surface mass balance of glaciers, Cryosphere, 6, 1295–1322, doi:10.5194/tc-6-1295-2012, 2012.

Riddle, E.E., Butler, A.H., Furtado, J.C., Cohen, J.L., Kumar, A.: CFSv2 ensemble prediction of the wintertime Arctic Oscillation. Climate Dynamics 41: 1099–1116, 2013.

Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., Eade, R., Fereday, D., Folland, C. K., Gordon, M., Hermanson, L., Knight, J. R., Lea, D. J., MacLachlan, C., Maidens, A., Martin, M., Peterson, A. K., Smith, D., Vellinga, M., Wallace, E., Waters, J., and Williams, A.: Skillful long-range prediction of European and North American winters, Geophys. Res. Lett., 41, 2514–2519, doi:10.1002/2014GL059637, 2014.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-370, 2017.