

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

### **Anonymous Referee #2**

Received and published: 2 August 2017

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.

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

C1

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.

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.

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?

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

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

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?

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

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)?

