Specific comments: "Retrospective forecasts of the upcoming winter season snow accumulation in the Inn headwaters (European Alps)" by Kristian Förster et al.

Page 1, line 13: I agree that seasonal forecasting is a topic which became a focal point in hydrological forecasting in recent years. But the term "new" isn't adequate in my opinion, as the earliest references e.g. for ESP-based long-term forecasts have been published in the 1970s / 1980s (e.g. Day, G. (1985). "Extended Streamflow Forecasting Using NWSRFS." *J. Wat. Res. Plan. Mgmt.*, 10.1061/(ASCE)0733-9496(1985)111:2(157), 157-170 or Twedt, T. M., J. C. Schaake, Jr., and E. L. Peck (1977). National Weather Service extended streamflow prediction. *Proc., Western Snow Conference*, 52 – 57.). Furthermore on page 2, line 8 you state that "seasonal outlooks […] have been prepared for decades." – This is a contradiction to the term "new", too.

Page 2, line 12/13: This sentence is a bit confusing / misleading, because meteorological data is used for the ESP-approach, too. Furthermore I recommend extending your description of ESP / revESP in order to explain more clearly which components contribute to forecast skill in each approach.

Page 2, line 18: I suggest adding "on the one hand", because otherwise the sentence might be misleading as seasonal forecast aren't solely an initial state problem (as you mention below).

Page 3, line 9: Please be more precise by adding "…based on seasonal predictions in the Alps" (or something similar).

Page 3, line 15: I suggest adding that the Inn basin belongs to the catchment of the Danube and that the Inn is the main tributary of the Upper Danube.

Page 4, line 5: Which "multi-year period" did you chose (the whole HISTALP period)?

Page 4, line 6: Please add a reference and explain what you mean by "randomly selecting valid values".

Page 4, line 13-26: Please add the temporal and spatial resolution of the output from both GCMs you're using in your model.

Page 4, line 24: In line 28 you state, that only re-forecasts starting in November are considered, but here also "25 Oct." is listed as initial start date. Please explain.

Page 4, line 31: What is the grid-size of the AWARE model used in this study?

Page 5, 7-9: I suggest splitting this sentence in order to make it easier to read.

Page 5, line 27: Do you recognize mismatches in summer / autumn, when the reservoirs are filled-up, too? Please comment on this.

Page 6, line 17-20: I suggest adding the number of ensemble-members for each AWARE-run. This will help the reader remembering the set-up you described in section 2.2.3

Page 7, line 16: I don't see the information / added value of Fig. 3 (c) for the reader, because there's no comparison to measured SWE. Please explain why you decided to include this figure.

Figure 3 (d): You should add the errors bars to the legend.

Page 8, line 8: As far as I understand, the hit rate for precipitation is equal (GloSea5) or lower (CFSv2) compared to temperature and higher as you state.

Page 8, line 8: As your finding that “The skill in precipitation predictions is higher [than temperature]” isn’t something I would have expected before, you should extent discussing this result more detailed and refer to similar and contradicting results from other studies.

Page 8, line 11-13: You state that predicting of hydrological storages (SWE in your case) is more robust / skillful than predicting fluxes; You should prove this statement. As your hydrological model also generates flows, I suggest relating this statement to your model results.

Page 8, line 26-28: Without doubt the representativeness of measured SWE and its interpolated on basin scale is problematic. But in my opinion you should mention, that the water balance model is never perfect and that it introduces uncertainties into hydrological forecasts, too. I guess it is your intention to exclude (at this state) the hydrological model-related errors by using a reference run?

Page 8, line 28-29: I suggest to use “GCM-forecast skill” (or something similar) instead of “model skill when using CM-based forecasts”.

Page 8, line 31: The cumulative snow melt is very difficult to recognize in figure 5 as its value is very small. So why did you plot this parameter (you don’t use this information in the text anymore)? I think it can be skipped.

Page 9, line 2: Please add “… and CFSv2-AWARE”.

Page 9, line 10: Please mention, which aspects of your method are really “new” (e.g. predictand, …).

Page 10, line 3: Interesting statement. Could you please comment on the definition of the “target accuracy”. Who defined this Was is defined by users / stakeholders?