Reply to Anonymous Referee #1

Terzago et al. present some encouraging results on seasonal forecasting of snow at levels that could be of commercial use. There are overlaps with the aims of the PROSNOW project (http://prosnow.org/); I haven't seen a published demonstration of seasonal forecasting from that project, but Koberl et al. (2021) on the market for seasonal forecast services for ski resorts is relevant.

Reply: We thank the reviewer for this comment. Indeed there are common aims between this work and the PROSNOW project. The PROSNOW project developed a demonstrator of a meteorological-to-climate prediction and snow management system from a short term forecast covering the first 4 days to several months ahead, specifically tailored to the needs of the ski industry. Stimulated by the comment of the reviewer we have performed an in-depth literature review on the papers related to this project. Ebner et al. (2021) evaluate the accuracy of the piste snow management module implemented within several snowpack models to simulate snow management for selected ski resorts. However no paper has been published on the seasonal prediction of snow resources. We agree that Koberl et al. is a relevant reference that illustrates the demand of climate services from the perspective of the ski resort managers. We have now added a reference to this paper in the Introduction.

L 25: The conclusion that bias correction of precipitation forecasts has little influence on snow depth skill scores might be surprising, or it might be misleading as no observations are used in the bias correction of precipitation (which is not mentioned in the abstract).

Reply: Thank you for this comment which allows us to better explain this point. The bias correction of precipitation forecasts is indeed essential to adjust biases eventually present in the seasonal forecast system outputs. For example, in our study we find that MFS6 largely overestimates total precipitation and this is reflected in disproportionately large snow depth values (Figure 5b of the manuscript): in such cases the application of bias correction methods is essential to reproduce realistic snow depth values and climatology. In particular, the bias correction which we applied to precipitation forecasts, despite being based on ERA5 rather than on observations, successfully adjusts the huge bias of the forecast snow depth climatology and makes it very close to the observed snow depth climatology (Figure 5b). Alternatively, if precipitation forecasts are already realistic, such as in the case of ECMWFS5 (Figure 5a), the application of a bias-correction has little impact on the snow depth climatology. So, in general, the application of bias correction methods to the forcing is necessary to take care of possible biases.

A different point is the influence of the bias correction of the skill of the forecast, and hence on the skill scores. Some skill scores are defined in such a way that they are not sensitive to the bias in the forecasts. The ROC curve, for example, is not sensitive to the bias in the forecasts, and a biased forecast may still have good resolution and produce a good ROC curve: in this case biased forecasts can be improved with bias correction/calibration and the metric gives an indication on the potential usefulness of the forecast. On the contrary, other skill scores such as the BSS and the CRPSS are sensitive to biases in the forecast distribution, so it is important to take care of and adjust biases. In conclusion, the application of bias adjustment methods to the main meteorological drivers of snow processes is needed to reconstruct a plausible snow depth climatology and a distribution as close as possible to the observed one, but it does not necessarily lead to improvements to skill scores. We have added the missing information on the bias correction with respect to ERA5 in the abstract and we have better clarified these points in the discussion of the manuscript.

L70: The GEM15 forecasts used by Bellaire et al. (2011) were only out to 48 hours and were found to overestimate precipitation, so differences in conclusions from this study might be expected. Forster et al. (2018) is a much more direct precursor of this study.

Reply: Thank you very much for the suggestion, indeed, since it is not totally relevant, we decided to remove this reference.

L77: "this study" would more accurately be described as "that study", i.e. Forster et al. (2018), not Terzago et al. (2022).

Reply: Modified as suggested by the reviewer, thank you.

L109: As stakeholders were involved in designing the prototype system, it is curious that none are included in the author list or acknowledgements.

Reply: Stakeholders are now mentioned in the acknowledgements, thank you for this comment.

Table 1: 6 decimal places in latitude and longitude locate the stations to within 10 cm, which seems unnecessary.

Reply: We agree that the information is very detailed, however we prefer to provide the exact location of the station as indicated in the official database and website of the data providers (ARPA Piemonte). Thank you.

Table 1: "Total radiation" here is, I think, net radiation.

Reply: "Total radiation" is the "total incoming shortwave radiation", we have specified it in the main text, thank you.

Table 2 contains only a small amount of information that could easily be incorporated in the text.

Reply: Information in Table 2 has been included in the main text (Section 2.3) and Table 2 has been removed.

Section 2.3.1: I find the bias correction hard to understand. What is the elevation of the ECMWFS5 temperature forecast in Figure 2a? Is it surprising that there is such a large *cold* bias compared with a station at 2410 m elevation? If the green line was produced by quantile mapping the raw data onto the ERA5 CDF, why is it further from ERA5 than the raw data? If the discontinuity in the green line is due to the monthly quantile mapping, why does it only appear in mid-February?

Reply: Thank you for this comment, we recognize that this section on the temperature bias correction was not clear and we have modified it in the revised version of the manuscript. Indeed analysis of the ECMWFS5 orography shows that the closest grid point is always at an elevation which is lower than that of the station by several hundreds of meters. Despite the lower elevation, ECMWFS5 shows a large and cold bias compared to observations at Bocchetta delle Pisse (see Figure 1a below) and smaller biases at the other stations, i.e. Rifugio Gastaldi and Lago Agnel (Figure 1b,c). This analysis confirms a systematic cold bias by the model which cannot be attributed simply to differences in elevation.

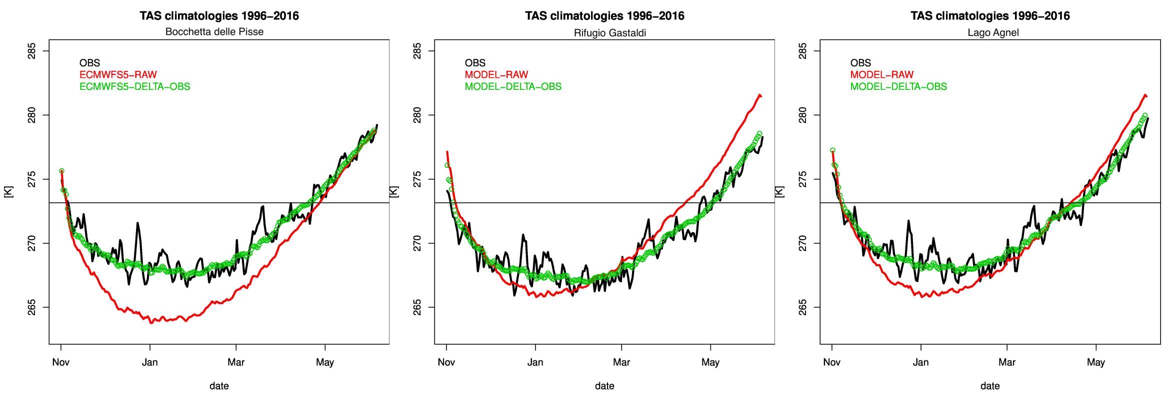

Figure 1 Near-surface air temperature climatology (1996-2016) for the ECMWFS5 seasonal forecast system, before (red) and after (green) the bias-adjustment, compared to surface station data, for the three stations considered, i.e. a) Bocchetta delle Pisse, b) Rifugio Gastaldi and c) Lago Agnel.

Regarding the temperature bias correction, in the submitted version of the manuscript we described two different methods: a more complex one based on the quantile mapping of the temperature forecasts to ERA5, and a simpler one based on the adjustment of the average bias with respect to observations. The first method, based on ERA5, is indeed not optimal mainly because ERA5 is as well biased compared to observations (Figure 2a of the manuscript) and thus not suitable as a reference dataset for the bias correction of temperature. This issue is evident also before applying the quantile mapping (Figure 2a), so the discussion on the application of this method using ERA5 as a reference is indeed not useful to the reader; moreover, Figure 2a contains an error in the quantile mapped data, as promptly noted by the reviewer. In fact, the curves corresponding to ECMWFS5 after the quantile mapping (green) and after the downscaling to the station (light blue) are not consistent with ERA5 temperatures, as both lines are further from ERA5 than the original raw data. Given this issue, and given that this method was discarded and not used further in the analysis, we prefer to remove it in the revised version of the manuscript and to keep only the description of the bias correction method based on the adjustment of 6-hourly average bias with respect to temperature observations, which is effective in reconstructing the observed climatology at all the sites considered (Figure 1) and has been used in the manuscript.

Table 3: Use superscripts for W/m2

Reply: Done, thank you.

Figure 2: The x-axis is labelled in months, not days. "Downscaled data" means different things for temperature and precipitation that is not apparent from the figure or caption. Why does cumulated precipitation appear to decrease in mid-February?

Reply: We have modified Figure 2 according to the reviewer's suggestion. In the new version of Figure 2 the label of the x-axis has been removed. The caption has been modified to include an explanation of the differences among the temperature and the precipitation downscaling. Regarding the discontinuity in the precipitation climatology, it occurs in correspondence of February 29th, when the multi-year average is calculated over the leap years only instead of the full period. In the new plot we removed the data corresponding to the 29th February for clarity.

L264: "(" missing before Matheson.

Reply: We have corrected this, thank you.

L276: subscript perf

Reply: We have corrected this, thank you.

Figure 3: The blue and dark blue lines are hard to distinguish when printed.

Reply: We have changed dark blue lines into black in the final figure.