We would like to thank the editor for the handling of our manuscript. We would also like to thank the reviewers for their comments and suggestions, which, without a doubt, have helped with the improvement of the manuscript by clarifying important points such as the difference between the altitude of the weather stations and the reanalysis grid.

We respond to your comments in more detail below:

Reviewer 1

Comments:

Lines 41 – 43. Are the datasets independent, so that the observations have not been assimilated into the ERA5 reanalysis? Could this be explicitly mentioned in the text?

R: The ERA5 dataset and the bias corrected reanalysis dataset are not completely independent since the latter is obtained from applying an elevation (and other monthly-based climate biases) corrections. We now specify this in the text.

Lines 41 – 43. How large is the area where these 17 AWSs are located?

R: There was an error upon submission of the manuscript and Figure 1, which shows the map of the location of the stations and the scale of the area, was not included. We apologize for this. The manuscript has now the correct Figure 1.

Lines 47 – 49. Are all observations compared with the same grid cell in ERA5?

R: Each AWS was compared to the closest ERA5 grid node. The wording in the manuscript was wrong and did give the impression that we only used one node. We have fixed the text to clarify this, it now reads:

"The ERA5 grid nodes used in comparisons to each individual AWS were selected by minimizing the haversine distance between each AWS and all the nodes in the reanalysis grid."

Lines 59 – 62. How representative are observation sites and are geographical properties similar in ERA5 as in observation sites? I think you could write a sentence or two about how similar surface properties are around the stations and in ERA5 because surface geographical properties remarkably affect near surface temperature in the polar regions. As surface inversions are common, near surface temperatures are sensitive to even small scale features in surface topography.

R: We hope that the corrected Figure 1 can bring some insight into answering this question. The AWS geographical locations range from a variety of distances from the sea and they also vary in elevation from almost at sea level to about 1.8 km altitude. In this sense, they are geographically representative of the Dry Valleys region. As can be seen in Figure 1, the closest ERA5 grid nodes are quite well correlated in space with the AWS. However, this is not the case with the BCR dataset. This is due to the fact that the ERA5 grid has a resolution of 0.1° X 0.1°, whereas the BCR was downsampled to a grid resolution of 0.5° X 0.5°.

Lines 59 – 62. What is the altitude range in ERA5 grid cells? Is the surface elevation almost the same in reanalysis as in reality in an observation site?

R: This is a very good question. We were previously, erroneously, missing applying a lapse-rate correction for the difference in altitude between the AWS and the closest grid cell. As you can see in the modify version of the manuscript, we now apply this correction and in general it reduces the bias but it does not completely eliminate it. Also, interestingly, there are a few cases where the altitude correction makes the bias even worse. Given all of this, we now present the results of the two datasets (ERA5 and BCR) with and without the altitude corrections.

Discussion. Could you suggest any reason for negative bias in ERA5? If you have any ideas why temperatures are underestimated in ERA5, you could maybe a little bit speculate with them in discussion. Especially, why ERA5 perform well in the Southern Antarctic Peninsula - Ellsworth Land region in the study of Tetzner et al. (2019) but not in the McMurdo area?

R: There are a number of possibilities of why there could be a negative bias in the McMurdo Sound compared to the South Antarctic Peninsula-Ellsworth Land. For example, the Peninsula is further north than the McMurdo Sound; there is a much greater mass of sea ice that accumulates in the McMurdo Sound compared to the Peninsula (which we suspect is an important effect in the bias); the difference could also be due to differences in regional climatic conditions, such as cloud coverage.

However, since these are all speculative explanations, we will refrain from adding them to the text.

Lines 89 – 90. “TheERA5 temperatures show a large overshoot during the summer, with an average difference of 6.7 ± 0.8 °C (e.g., Figure 2).” What do you mean with overshoot? A large negative bias?

R: Yes, here we mean that during the summer months the bias is larger than compared to the rest of the year. We have removed the word “overshoot” and rephrased the paragraph to avoid confusion.

Figure 1. Figure is wrong. According to text and caption the figure should show a map of the region but there is a time series of temperature in the figure. I think the map could give answers to some of my questions. Therefore, I hope that the map includes information about surface conditions (i.e. satellite image in summer at background) topography (i.e. surface elevation contours), sizes of ERA5 grid cell (grid cell boundary lines) and locations of AWSs.

R: We apologize, we have added the correct Figure 1. Indeed, figure 1 was always meant to show not only the location of the stations and the ERA5 nodes, but also the satellite image of the McMurdo sound where the topography of the Dry Valleys region can be appreciated.

Figure 2b. Could you use the same temperature scale in both axes or add line T_obs = T_ERA to help comparison between ERA5 and observations.

R: Thank you for the suggestion, we have added a T_obs = T_era line in the correlograms of figure 2 (and in all subsequent supplementary figures). It does help show the biases more clearly.

Reviewer 2

The manuscript presents evaluation of ERA5 reanalysis 2-m temperatures (T2m) in Antarctica. The topic is important, as ERA5 is very much applied in climate research, and is often considered to represent the best available information from regions with no or sparse observations avaialbale. However, the manuscript suffers from major shortcomings, and requires a major revision before eventual acceptance.

Major comments

1. The difference between observations and ERA5 T2m is surprisongly large and the bias is negative (cold), although previous studies from Polar regions have indicated predominantly positive (warm) biases. This makes me to suspect that the difference between the real ice sheet elevation and that in ERA5 has not been taken into account.

This is a fundamental shortcoming, as it does not make sense to compare observed and reanalysis-based T2m values if the model orography does not fit with the true orography. See, e.g., papers by Bromvich group on how to take the elevation difference into account comparison of observations and model products. First, the authors should quantify the elevation difference between each observation site and the nearest reanalysis grid cell. Note that even if the ERA5 biasdoes not depend on the topography (lines 85-86), it may well depend on the elevation difference between ERA5 grid and reality.

R: Thank you very much for your comment. Indeed, we were erroneously missing applying the lapse-rate correction for the differences in altitude from the AWS and the ERA5 grid. We have applied the correction using the standardly used value of 6.5 °C/km. The altitude correction does reduce the average bias observed in the region, but it does not completely eliminate it.

2. It seems that the authors do not fully understand the concept of an atmospheric reanalysis. A reanalysis is not derived by a combination of climate data assimilation and climate simulations (lines 24-25) but from a combination of data assimilation and short-term (mostly 6 hours) simulations applying an operational numerical weather prodiction (NWP) model. Also the text on lines 77-79 may give an impression that ERA5 is applied in seasonal-scale simulations, which is not the case, and the summertime ERA5 T2m values do not have any effect on the annual melt rate of snow, glaciers, and permafrost in Antarctica (i.e., text on lines 91-92 is misleading). The authors are correct that modelling of snow and ice melting using ERA5 temperature as atmospheric forcing is a problem, if ERA5 T2m has a large bias.

R: We have rephrased the definition of the ERA5 reanalysis, the text now reads:

“ERA5 dataset represents the fifth iteration of ECMWF (European Center for Medium-Range Weather Forecasts) global hindcasting based on the Integrated Forecasting System (IFS) Cy41r2 derived by a combination of data assimilation and short-term simulations applying an operational numerical weather prediction (NWP) model (Hersbach et al, 2020). “

Regarding the seasonal variation, it seems that the wording in our text was confusing, since we do not suggest that the ERA5 is applying in seasonal-scale simulations, but rather that when analyzing the decades-long time series and the correlogram, it is clear that the is a

seasonal dependency on the bias. The correlogram shows a hysteresis that is present for all stations, and the larger bias during the summer season is also evident from the time series and the correlograms. We have rephrased our text with the hope that this is now more clear.

3. Figure 1 is missing, and Figure 2 is presented twice.

R: We apologize, there seems to have been a mistake upon submission. We have now added the correct Figure 1.

Minor comments

4. Specify what you mean by "ground temperatures" (line 32)

R: By “ground temperatures” we referred to the 2-m air temperatures, so it was misleading and we have removed it from the text.

5. Better explain the near-surface bias corrected reanalysis dataset (line 45)

R: We added an explanation of the near-surface bias corrected reanalysis dataset and added a number of references where the dataset and the corrections applied are documented in detail.

6. In the rightmost column of Table 1, give the difference reanalysis minus observations so that positive biases are seen as positive values. Consider the observation accuracy of automatic weather stations in Antatrctica, and remove the irrelevant second decimales from the temperature data.

R: We have changed the order of the comparisons to show the reanalaysis minus the AWS data. We removed the second decimals and we also now include the altitude corrected values for both ERA5 and the bias-corrected reanalaysis datasets to show the effect of the correction.