

Dear Dr. Bromwich,

We would like to sincerely thank you for your constructive and insightful comments on the initial version of our manuscript. Below, we provide detailed point-by-point responses to each of your comments. The Referees' remarks are shown in italics, and our responses are provided in plain blue font.

Overview Comments

The authors set out to explain the cold bias in ERA5 temperatures over Antarctica prior to 1979 reported by Bromwich et al. (2024) and Dalaiden et al. (2025). The uncertainty in the SST and sea ice conditions (SSC) over the Southern Ocean during this period is explored to determine the SSC contribution to the cold bias in ERA5. About 30% of the ERA5 cold bias is attributed to Southern Ocean SSC uncertainty. The desire for a reliable SSC prior to 1979 is well justified. This is a valuable study that just needs some more contextual content.

We would like to thank the referee again for their positive assessment of our study.

Specific Comments:

Line 30: It is not strictly accurate to call the period prior to 1979 pre-satellite. There were a few primitive satellite sensors in orbit back to the 1960s, as described by Hersbach et al. (2020). It is true that satellite observations became much more extensive starting in late 1978.

We thank the referee for this valuable comment on our use of the term "pre-satellite period". We acknowledge that satellite observations were already available prior to 1979. However, as noted by the referee, continuous and comprehensive satellite coverage began only later, in 1979. We apologize for this simplification. Accordingly, we have replaced the original sentence:

"\cite{Bromwich2024} identified a pronounced cold bias in ERA5 over Antarctica during the pre-satellite era, particularly 1958--1978, when numerous weather stations became available due to the International Geophysical Year."

with the following:

"\cite{Bromwich2024} identified a pronounced cold bias in ERA5 over Antarctica before the continuous, and comprehensive satellite coverage from 1979, particularly 1958--1978, when numerous weather stations became available due to the International Geophysical Year."

We also have removed the term "pre-satellite" throughout the manuscript.

The reconstruction by Fogt et al. (2022) <https://doi.org/10.1038/s41558-021-01254-9> produced a much smaller sea-ice decrease across 1979 than shown by ANT-REC. The seasonal sea ice extent reconstructions were primarily based on monthly mean pressure and temperature records across the Southern Hemisphere extratropics and midlatitudes

from 1905 to 2020, similar to ANT-REC. ANT-REC does provide a much more comprehensive analysis, Fig. 5 for example.

We thank the referee for pointing out the reconstruction of Fogt et al. (2022). We have added a paragraph discussing the comparison with the reconstruction of Fogt et al. (2022), including their explanations in the differences with ANT-REC, as well as a more detailed quantification of the uncertainties of ANT-REC (requested by the second referee) at line 99:

“Another reconstruction of Antarctic sea-ice extent covering the 20th century and beyond has been proposed by \cite{Fogt2022}, based on statistical relationships between mid-latitude atmospheric variability and Antarctic sea-ice extent derived from atmospheric pressure and temperature observations. Compared with ANT-REC, this reconstruction suggests an even weaker decrease in Antarctic sea-ice extent prior to 1979 ($0.18 \times 10^6 \text{ km}^2$ between 1959–1968 and 1981–2000, compared with $0.49 \times 10^6 \text{ km}^2$ in ANT-REC), highlighting the substantial uncertainties in historical sea-ice variability. These differences likely arise from differences in methodology. More specifically, the reconstruction of \cite{Fogt2022} relies primarily on statistical relationships in the atmosphere, whereas ANT-REC dataset used in this study is based on the assimilation of Antarctic and sub-Antarctic atmospheric pressure and temperature observations into Earth System Model simulations, followed by ocean–sea-ice modeling to represent ocean dynamics, thereby explicitly representing oceanic processes.

Beyond differences between reconstructions, uncertainties are also present within the ANT-REC framework itself. Although only one atmospheric reanalysis (based on CESM1) was used to force the ocean–sea-ice model, ten atmospheric reanalyses are available and provide an indication of the reanalysis uncertainty, as they also provide sea-ice concentration fields. Among these reanalyses, the CESM1-based atmospheric reanalysis shows a sea-ice decline between 1959–1968 and 1981–2000 of $0.40 \times 10^6 \text{ km}^2$, which lies close to the ensemble mean of the ten realizations ($0.41 \pm 0.18 \times 10^6 \text{ km}^2$). Finally, the magnitude of the sea-ice decrease in ANT-REC is broadly consistent with independent paleo-based estimates \cite{Thomas2019,Dalaiden2023}, but the discrepancies between reconstructions emphasize the remaining uncertainty in Southern Ocean sea surface conditions prior to 1979, when continuous satellite coverage is available.”

Some more discussion of the results of Bromwich et al. (2024) is desirable. Their estimate of the cold bias in ERA5 was $\sim 1\text{C}$ for all of Antarctica, compared to 0.7C here based on station observations. Also those authors discussed the widespread Southern Hemisphere issues with assimilation of satellite (atmospheric) observations prior to 1979 and the jump in performance of the Southern Hemisphere forecasts across 1979 (Fig. 8(e) in Soci et al. (2024) <https://doi.org/10.1002/qj.4803>).

We thank the referee for this valuable comment regarding the comparison with the results of Bromwich et al. (2024). As noted by the referee, their estimate of the cold bias ($\sim 1^\circ\text{C}$) is based on Antarctic-wide averages, whereas our estimate ($\sim 0.7^\circ\text{C}$) is derived solely from the locations of available weather stations. In addition, the two studies consider different reference periods: Bromwich et al. (2024) estimate the bias over 1940–1978 relative to the post-1979 period, while our analysis focuses on 1959–1968 relative to 1981–2010. These methodological differences likely contribute to the discrepancy in magnitude. We also would

like to thank the referee for drawing our attention to Soci et al. (2024), in particular their Figure 8, which provides valuable insight into the widespread Southern Hemisphere assimilation issues prior to 1979 related to increased temporal observational constraints and the associated improvement in forecast skill thereafter. In the original manuscript, we noted that the uncertainties in the sea surface conditions alone cannot fully explain the pre-1979 cold bias in ERA5. In the revised version, we have now further detailed the differences in the ERA5 bias obtained in our study and from Bromwich et al. (2024) at line 194:

“Associated with this surface atmospheric pressure bias is a cool bias during the pre-continuous satellite period (Fig. \ref{fig:bias_WS_maps}), also reported by \cite{Bromwich2024}. In their study, the cold bias in ERA5 is estimated to be about 1°C over Antarctica as a whole, whereas our estimate (0.68°C) is derived from the locations of available weather stations and over slightly different period (1940--1978 relative to the post-1979 period), which likely contributes to the difference in magnitude. Our modeling experiments suggest that SSCs account for about 30% of this cold bias within the framework of our experiments, but other factors must also contribute. Among these potential factors, the limited observational constraints before the late 1970s might have contributed to systematic errors in ERA5 for the Southern Hemisphere \citep{Bromwich2024,Soci2024}.”

Dear Dr. King,

We would like to sincerely thank you for your constructive and insightful comments on the initial version of our manuscript. Below, we provide detailed point-by-point responses to each of your comments. The Referees' remarks are shown in italics, and our responses are provided in plain blue font.

General remarks

The meteorological observing network in high southern latitudes is sparse and very few continuous series of observations extend back beyond the International Geophysical Year of 1957/58. For these reasons, studies of interdecadal climate variability and long-term change in the Antarctic regions have relied heavily on atmospheric reanalyses, such as ERA5. Such reanalyses provide a good representation of atmospheric circulation and surface climate from 1979 to the present day. Over this period, satellite sounder data provide a strong constraint on atmospheric circulation, and other satellite observations provide a good reconstruction of sea surface conditions (SSCs, sea surface temperatures and sea ice conditions). It is therefore not surprising that atmospheric reanalyses perform less well over the pre-1979 period when satellite data were not available. In particular, reanalyses show a systematic cold bias in Antarctic surface temperatures (relative to observations) over the pre-1979 period which is smaller or absent in the later period, leading to an artificial warming trend in the reanalysis.

In this paper, the authors carry out a series of experiments with an atmospheric model forced by a new reconstruction of pre-1979 Southern Ocean SSCs to investigate whether errors in the representation of SSCs in ERA5 are responsible for the pre-1979 cold bias in that reanalysis. They conclude that SSC uncertainties in the earlier period could account for around 30% of the cold bias but other factors must also be contributing.

*The paper is well-written and clearly explains the methodology, which exploits a reconstruction of pre-satellite SSCs developed by the lead author. The conclusions are soundly based on the results of the model experiments. I believe that the paper is suitable for publication in *The Cryosphere* after the (mostly minor) comments below have been addressed.*

We would like to thank the referee again for the overall positive feedback on our study.

General comments

1. The study is aimed at identifying whether the cold Antarctic bias in ERA5 during the pre-satellite era is a result of errors in specified SSCs. Ideally, one would investigate this by rerunning part of the ERA5 reanalysis with an alternative SSC dataset, but this would be prohibitively expensive so an alternative approach is taken where an atmospheric model is run with both ERA5 and alternative SSCs. This is a fairly standard approach to investigating model sensitivity and, as such, is valuable in itself. However, it is important to recognise that there are differences between CAM and the IFS model which was used to produce ERA5, both in the dynamical core and in the physical parametrisations used. I am not familiar with CAM so I don't know how large these differences could be, but I think that there should be

some discussion of this in the paper. Secondly, while CAM is only constrained by the imposed SSCs, ERA5 is additionally constrained by atmospheric observations. This point should also be noted.

We thank the referee for this relevant comment regarding the differences between CAM and the IFS model. We agree that re-running the ERA5 reanalysis with alternative SSC datasets would be the ideal approach to quantify the impact of sea surface conditions on Antarctic climate in ERA5, but this is currently not feasible (in terms of computation resources and practical access). As stated by the referee, our study adopts a standard AMIP-style sensitivity framework to isolate the atmospheric response to SSC uncertainties using two alternative SSC datasets. We would like to emphasize that our primary objective is to assess the contribution of SSC uncertainty on Antarctic surface climate. According to the referee's comment, we have added a paragraph (at line 202) discussing the differences between CAM and IFS, and we have also specified that our developed framework must be seen as a way to quantify the uncertainties related to the SSC products and not as a way to quantify the bias of ERA5:

“It should also be noted that the sensitivity experiments performed in this study are based on simulations with CAM, whereas the ERA5 reanalysis is produced using the Integrated Forecasting System (IFS). These models differ in their dynamical cores and physical parameterizations, which influence the atmospheric response to prescribed SSCs. In addition, ERA5 is constrained by atmospheric observations through data assimilation – albeit limited in the Southern Ocean before 1979 (Soci et al., 2024) –, while the simulations conducted here are only constrained by prescribed SSCs. Therefore, the results presented here should be interpreted as the response of CAM to different SSC products, and not as a direct quantification of the ERA5 bias.”

We also have adopted more cautious language when discussing the contribution of SSCs to the cold bias in ERA5, by explicitly mentioning, at several places when required, that these results are within our current framework, in particular our model simulations.

2. The authors conclude from their model experiments that around 30% of the Antarctic cold bias in ERA5 could result from errors in pre-satellite SSCs. This is quite a precise statement and, in order to support it, I would like to see some discussion of the uncertainties in the reconstructed SSC product. Is it possible to quantify the uncertainties? Would it be possible to examine the impact of these uncertainties on simulated Antarctic temperatures through a series of perturbed model runs? While I agree with the authors that “Developing a temporally-consistent SSC dataset for the high-latitude southern hemisphere is therefore a priority” (lines 194-195) the lack of satellite observations before 1979 means that the uncertainties in SSCs during the earlier period will always be higher than post-1979. Quantifying simulation uncertainty resulting from uncertainty in SSC reconstructions is thus very important.

We thank the referee for this important comment regarding the uncertainty associated with our estimate that SSC uncertainties account for about 30% of the Antarctic cold bias in ERA5. We acknowledge that this estimate is dependent on the framework and reflects the response of CAM to the specific SSC datasets considered, as well as uncertainties in the alternative SSC reconstruction (i.e., ANT-REC). We therefore agree that this value should be

interpreted cautiously. We have revised our manuscript accordingly. More specifically, we have added two paragraphs discussing the sea-ice changes from ANT-REC, by comparing to other sea-ice reconstructions in the discussion section at line 99:

“Another reconstruction of Antarctic sea-ice extent covering the 20th century and beyond has been proposed by \cite{Fogt2022}, based on statistical relationships between mid-latitude atmospheric variability and Antarctic sea-ice extent derived from atmospheric pressure and temperature observations. Compared with ANT-REC, this reconstruction suggests an even weaker decrease in Antarctic sea-ice extent prior to 1979 ($0.18 \times 10^6 \text{ km}^2$ between 1959–1968 and 1981–2000, compared with $0.49 \times 10^6 \text{ km}^2$ in ANT-REC), highlighting the substantial uncertainties in historical sea-ice variability. These differences likely arise from differences in methodology. More specifically, the reconstruction of \cite{Fogt2022} relies primarily on statistical relationships in the atmosphere, whereas ANT-REC dataset used in this study is based on the assimilation of Antarctic and sub-Antarctic atmospheric pressure and temperature observations into Earth System Model simulations, followed by ocean–sea-ice modeling to represent ocean dynamics, thereby explicitly representing oceanic processes.

Beyond differences between reconstructions, uncertainties are also present within the ANT-REC framework itself. Although only one atmospheric reanalysis (based on CESM1) was used to force the ocean–sea-ice model, ten atmospheric reanalyses are available and provide an indication of the reanalysis uncertainty, as they also provide sea-ice concentration fields. Among these reanalyses, the CESM1-based atmospheric reanalysis shows a sea-ice decline between 1959–1968 and 1981–2000 of $0.40 \times 10^6 \text{ km}^2$, which lies close to the ensemble mean of the ten realizations ($0.41 \pm 0.18 \times 10^6 \text{ km}^2$). Finally, the magnitude of the sea-ice decrease in ANT-REC is broadly consistent with independent paleo-based estimates \citep{Thomas2019,Dalaiden2023}, but the discrepancies between reconstructions emphasize the remaining uncertainty in Southern Ocean sea surface conditions prior to 1979, when continuous satellite coverage is available.”.

Regarding the perturbed runs, as suggested by the referee, in the original version of the manuscript, we generated ten simulations for each ensemble (i.e., for each forcing dataset) in order to quantify the influence of internal atmospheric variability. As shown in the manuscript, the ensemble members exhibit very similar responses, indicating that the reported differences primarily reflect SSC forcing rather than atmospheric initialization. We have specified further in the revised manuscript at line 143:

“The small spread among ensemble members indicates that the differences between the ERA5-forced and ANT-REC-forced simulations primarily reflect the imposed SSC boundary conditions rather than differences in atmospheric initialization.”.

Finally, we agree with the referee that SSC uncertainties prior to 1979 will inevitably remain larger than in the period after 1979. Nevertheless, we believe that combining multiple information sources, including early satellite data, weather stations, ship observations, and paleoclimate records, offers a promising pathway toward improved historical SSC reconstructions. In addition, quantifying the impact of both the number and the quality of available observations, and how these uncertainties propagate into climate simulations, is

equally important. We have rewritten the discussion in the final paragraph accordingly (at line 223):

“Given the critical role of Antarctica in the global climate system, particularly through sea-level rise, longer and more homogeneous SSC products are essential for robust assessments of long-term climate change. This is especially important given the large internal variability of the Antarctic climate \citep[e.g.,]{Jones2016}. Developing a temporally consistent SSC dataset for the high-latitude Southern Hemisphere is therefore a priority. However, uncertainties in SSC reconstructions prior to 1979 will inevitably remain larger than after the onset of continuous satellite data coverage. Progress is foreseen through the combination of diverse observational sources, including satellite products, weather stations, historical ship logbooks, as well as indirect climate records (i.e., paleoclimate records), into unified datasets. In parallel, quantifying the impact of the number, quality, and time-varying constraints of available observations in SSC reconstructions is equally important. In this context, future ensemble reanalyses should aim to explicitly sample uncertainties in SSCs, allowing a more robust quantification of their impact on the simulated atmospheric state.”

Specific comments

1. L20 (and elsewhere): Replace “transport of humidity” with “transport of moisture”.

We have applied the suggestion.

2. L46-49: The description of the ERA5 SSC dataset is a bit confusing. Maybe something like “The first dataset is taken directly from ERA5, which uses the HadISST2 dataset before 1979 and the OSTIA product thereafter” would be clearer?

We have applied the suggestion from the referee. We thank the referee for his comment.

3. L49-50: It is not clear what period you are referring to here.

We are sorry for the missing information here. HadISST presents large uncertainties before 1979. It is now specified in the updated version of the manuscript.

4. L120-122: It is not clear how the bootstrap subsamples were extracted.

We thank the referee for pointing out that the bootstrap procedure was not sufficiently described. In the revised manuscript we have clarified it at line 139:

“More specifically, for each station, 1,000 resamples were generated by randomly selecting eight out of the ten ensemble members (without replacement), on which we further compute the mean difference.”

5. L125: “positive” rather than “negative”?

We thank the referee for spotting this error. Indeed, we meant negative and not positive. It has been corrected in the revised version of the manuscript.

6. *Figure 2: It is difficult to see the location of Faraday on the map. Also, note that Faraday was renamed Vernadsky in 1996 following its transfer to the Ukrainian Antarctic Programme.*

We thank the referee for this comment. We were not aware that the station name had been changed. The correction has now been applied. In addition, we slightly adjusted the numbering at that location to improve its visibility and increased the marker size for all station locations in the figure.

7. L137 *“with respect to...”*

As suggested, we have replaced “in” with “with”.