First of all, I apologize for the delay with the decision on your manuscript.
Based on the two reviews of your paper, I suppose that it can be published after some major revisions.
We thank you for the time you dedicated to the review of our paper. In the following of this letter, we detailed the change we brought to the paper, consistently with our responses to reviewers as well as your comments.

In particular, please extend the discussion of different isotope-temperature trends obtained in your work and in the paper by Casado et al. (2023).
We clarified the different isotope-temperature trends obtained in our work and in the paper by Casado et al. (2023), p.6 l.182 to l.185:
« Casado et al. (2023) provide a higher trend from 1950–2005 of 0.11±0.02 ‰ per decade, based on ice core data. Different reasons could explain that mismatch that we are not able to elucidate so far, inter alia: (i) a model discrepancy to resolve processes, (ii) the model resolution, (iii) the geographical distribution of the ice core locations, (iv) the different methods for the SAT – δ18O calibration. »

We also opened the discussion by giving results from published observations : l.220 to 223: *"Non significant relationships were also reported in observations and model outputs. For instance, Goursaud et al. (2018) report no SAT-δ18O relationship at the annual scale over the coast of Dronning Maud Land, the Victoria Land, some of the Indian coast and the Peninsula. An absence of SAT-δ18O relationship derived from firn/ice cores were also published (e.g. Goursaud et al., 2019; Bertler et al., 2011; Vega et al., 2016; Goursaud et al., 2017)."*

If one of the reasons for this discrepancy is "different methods for the SAT – d18O calibration" than is it possible to judge which method is preferable?
We completed the different possible reasons for the discrepancies in the obtained trends, l.201 :
"These disparities could be explained by the different time windows, the different methodologies, the lack of ice core data to make representative regional reconstructions, or a model discrepancy.".

While we cannot check the effect of the grid size as our simulations were all run with a same grid size, we checked the impact of the window length on our simulated regional δ^{18}O trends. The results were integrated in Appendix D.

Please also address the other issues raised by Reviewer 2:
We reported our responses to the second reviewer relative to the four below points you focus on.

1) add a new careful evaluation of your HadCM3 model;
We have added a new careful evaluation of the model in Appendix A, evaluating the simulated Antarctic Surface Air Temperature (SAT), precipitation (P) and precipitation weighted δ^{18}O. The results show as expected a warm bias in the Antarctic interior – this is also observed in other models such as in Polar WRF (Zhang et al., 2022); and a dry bias in coastal regions. Overall HadCM3 performs roughly in line with expectations derived from other similar models, and have a reasonable representation of Antarctic surface climate and δ^{18}O.

In the text, we referred to the appendix l.94 to l.96: *"HadCM3 provides a reasonable representation of Antarctic climate and δ18O (Appendix A, as well as Turner et al., 2006; Tindall et al., 2009; Holloway et al., 2016)."*
Appendix A can be found from page 26:

Appendix A: HadCM3 evaluation of Antarctic surface climate and δ_{18}O

A1 Method

Here, we check that HadCM3 provides a reasonable representation of the Antarctic surface climate and $\delta_{18}O$. Surface Air Temperature (SAT) output data from HadCM3 are evaluated against the AntAWS dataset Wang et al. (2022); a compilation of Antarctic observations from 267 AWS (automatic weather station) operational between some parts of the period from 1980 to 2021. Surface mass balance (SMB) model output, calculated within the model code as precipitation minus evaporation (wind related processes are not accounted for by HadCM3), similarly are evaluated against AnSMB Wang et al. (2021); the most recent quality-controlled published SMB compilation extracted from stakes, snow pits, ice cores, ultrasonic sounders and ground-penetrating radar. Finally, simulated $\delta_{18}O$ values are evaluated using the updated database compiled by GGoursaud et al. (2018); this combines all available firn, ice core, surface snow and precipitation observations of Antarctic $\delta_{18}O$. We show maps and scatter plots (model versus observed values) for SAT, SMB and $\delta_{18}O$. The comparison helps establish if the model underestimates the real spatial heterogeneity across Antarctica. Mean climatological values (20 year averages or more, averaged over the ensemble wherever possible) were calculated at each model grid point, and directly compared to the most equivalent observational climatological value (see paragraph above). The comparison uses output from a closest grid point comparison method.

A2 Results

A2.1 SAT

Turner et al. (2006)'s evaluation of HadCM3 Antarctic climate, including especially near-surface air temperatures, mean sea level pressures and geopotential heights, shows a large warm bias in the Antarctic interior associated with a low-biased modeled orographic height (the heighest model gridpoint elevations do not reach 4000 m asl). This finding remains fully consistent with the newer Wang et al. (2021) observation datasets (Figure A1). The minimum climatological Antarctic plateau SAT value is -37.2 ◦C (Figure A1A), considerably warmer than the AntAWS minima of -64.6 ◦C (Figure A1C). In regions where the observational temperature is above -30 ◦C the model values of SAT match the observations better, although there remains a slightly underestimating (warm bias) in West Antarctica (Figure A1B and top right of Figure A1C). Altogether, although the warm bias in the Antarctic interior contributes to weaken the linear regression between the HadCM3 simulations and the observations (correlation coefficient of 0.76), Antarctic-mean simulated SAT is surprisingly good: Antarctic-mean climatological SAT is -25.1±14.1 and -25.0±9.1 in the observations and the HadCM3 model, respectively.

A3 SMB

Consistent with previous studies, SMB is slightly too low in the Antarctic interior in HadCM3 Turner et al. (2006), suggesting that the warm bias in these regions do not affect the modelled SMB. The largest model SMB errors (dry and wet biases) occur near the coasts (Figure A2B). The dry biases may be due to the

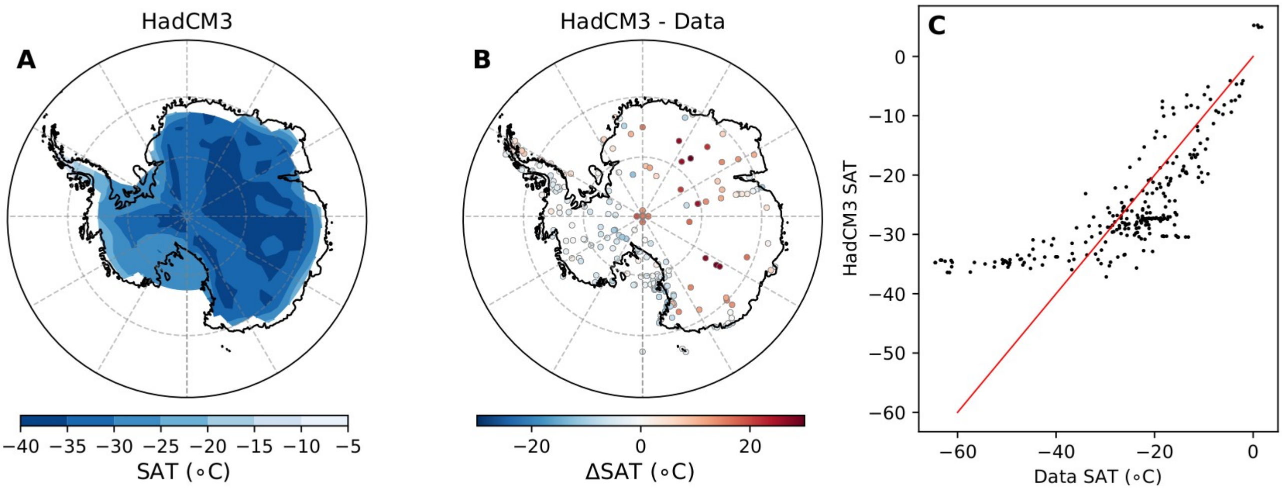

Figure A1. Surface Air Temperature evaluation (SAT): (A) map of the time-averaged HadCM3 SAT distribution over Antarctic resulting from the ensemble mean for the Historical period (in ◦C); (B) SAT difference between the time-averaged HadCM3 outputs from the ensemble mean for the Historical period, coarse HadCM3 grid, altering a realistic orography andand the corresponding SAT observations (in ◦C);

and (C) linear regression between the time-averaged HadCM3 outputs from the ensemble mean for the Historical period, and the corresponding SAT observations (black points). The red line is a 1:1 data-model slope.

representation of the ascending air masses that provide precipitation to these coastal regions. The coarse model grid biases can be seen on Figure A2C as a step representation of the black points compared to an expected linear regression. (Turner et al., 2006) also attribute the wet coastal biases to an overly intense mean sea level pressure field gradient: stronger than observed air flows produce excess precipitation on the west side of the Antarctic Peninsula. These aspects reduce the linear regression correlation (correlation coefficient of 0.70). The Antarctic-mean climatological SMB difference between the observations and HadCM3 is -29.7 mm/month.

A4 δ18O

The distribution of the simulated δ18O over Antarctica is similar to observations (Antartic-means of -36.2±9.7 ‰ and -37.4±10.3 ‰ in the observations and the HadCM3 simulations respectively; minimum values of -61.3 ‰ and -57.9 ‰ in the observations and the HadCM3 simulations respectively; maximum values of -3.2 ‰ and -7.7 ‰ in the observations and the HadCM3 simulations respectively). Excessively depleted values occur in the Antarctic interior (Figure A3). These are associated with the warm bias. Overly enriched values are observed over the Peninsula and the Weddell Sea coast, consistently with the wet bias in these region. Nevertheless, the HadCM3 historical simulations do capture the δ18O observations relatively well, as shown by the strong relationship between the outputs and the observations (correlation coefficient of 0.84 and slope of 0.90±0.02 ‰.‰$^{-1}$.

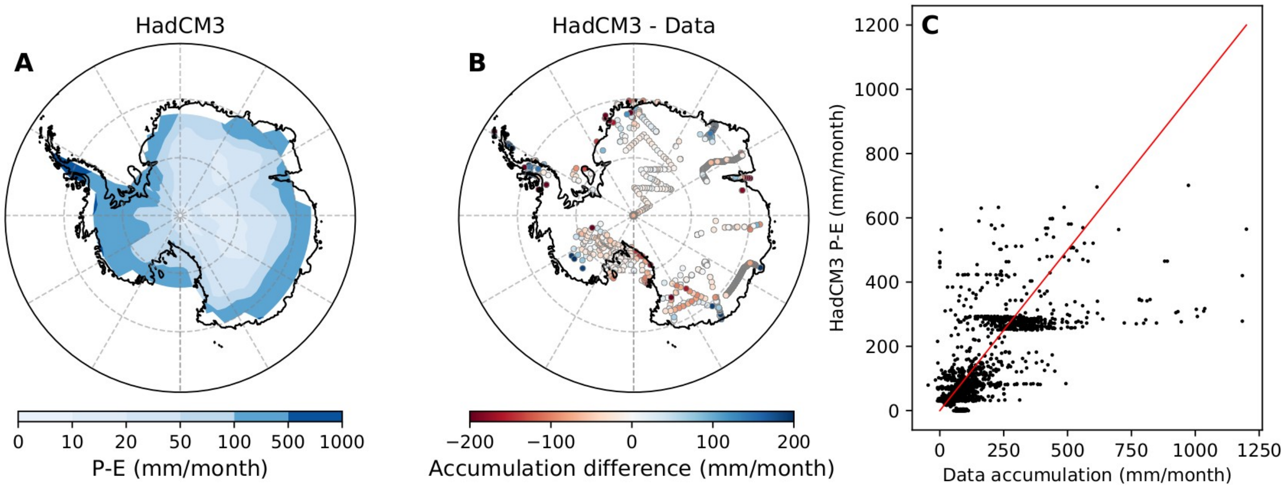

Figure A2. Surface Mass Balance evaluation (SMB): (A) map of the time-averaged HadCM3 Precipitation minus Evaporation (P-E) distribution over Antarctic resulting from the ensemble mean for the Historical period (in mm/month); (B) SMB difference between the time-averaged HadCM3 outputs from the ensemble mean for the Historical period, and the corresponding observations (in mm/month); and ©linear regression between the time-averaged HadCM3 outputs from the ensemble mean for the Historical period, and the corresponding SMB observations (black points). The red line is a 1:1 data-model slope.

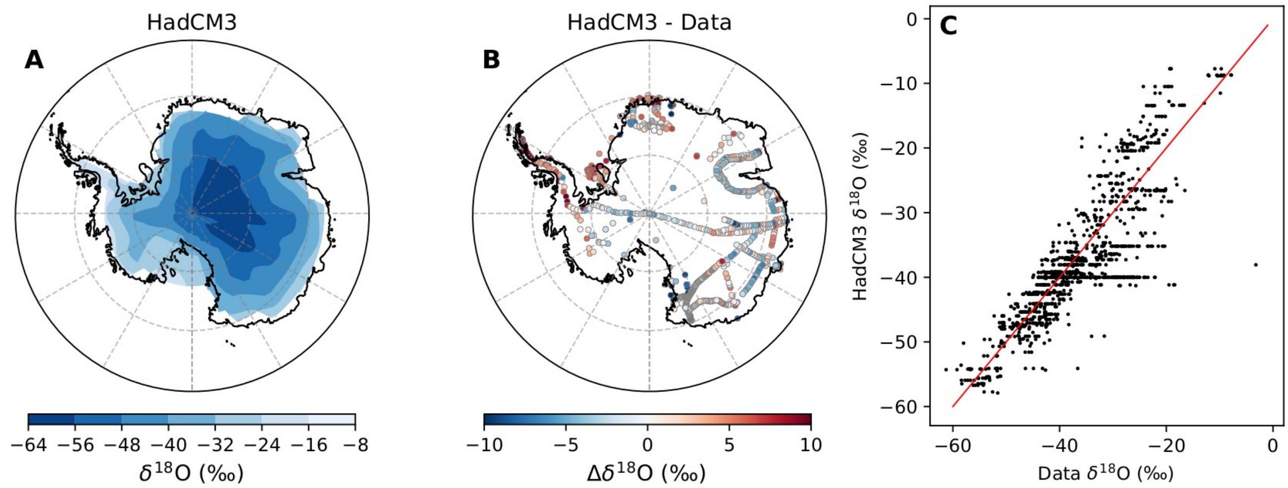

Figure A3. δ18O evaluation: (A) map of the time-averaged HadCM3 δ18O distribution over Antarctic resulting from the ensemble mean for the Historical period (in ‰); (B) SMB difference between the time-averaged HadCM3 outputs from the ensemble mean for the Historical period, and the corresponding observations (in ‰); and (C) linear regression between the time-averaged HadCM3 outputs from the ensemble mean for the Historical period, and the corresponding δ18O observations (black points). The red line is a 1:1 data-model slope.

2) provide comparison with ECHAM6 rather than with ECHAM5;

We replaced the analysis with the latest generation of the AGCM ECHAM equipped with water isotopes: ECHAM6-wiso (Stevens et al., 2013, Cauquoin et al., 2019). As stated by the reviewer, compared to ECHAM5-wiso, the performance of the water isotopes in ECHAM6-wiso is clearly improved. This is attributed to: (i) a modification of the supersaturation parameters ; (ii) that the kinetic fractionation at the evaporation over oceans is now assumed to be independent of the wind speed in order to better represent the d-excess versus deuterium relationship from the Antarctic Snow reported by Masson-Delmotte et al. (2008) ; and finally (iii) that the sublimation processes now accounts for the isotopic content of snow over sea ice. Based on the evaluation of global simulations against ERA-interim and ERA5 reanalyses, Cauquoin and Werner (2021) report that the nudging does not significantly change the simulated isotope values, while increasing the resolution generally improves the performance of the simulations. However, the evaluation of the simulated water stable isotopes in precipitation over Antarctica remains rather qualitative (Figure 1, Cauquoin and Werner, 2021).

Having obtained this new model output data from the newer version of ECHAM, we performed the same analysis as previously applied to ECHAM5 and HadCM3. As implied by the reviewer, using the newer version of the ECHAM indeed entirely resolve the discrepancy between the models – ECHAM6-wiso and HadCM3 (in the newer ECHAM6 version) now have equivalent SAT-δ^{18}O surface air temperature relationships.

We thus made the following changes in the text:

- In section 2 (“Data and methods”), l.106 to 114:

“Our Historical SAT–δ18O linear relationship at the regional scale are compared with the regional slopes and correlation coefficients that we computed from the AGCM ECHAM6-wiso equipped with water stable isotopes (Cauquoin et al., 2019). The water stable module of this last generation of the model ECHAM was updated compared to its predecessor, especially (i) the supersaturation parameters, (ii) the kinetic fractionation at the evaporation over oceans, now assumed to be independent of the wind speed in order to better represent the d-excess versus deuterium relationship from the Antarctic Snow reported by (Masson-Delmotte et al., 2008), and finally (iii) the sublimation processes now accounting for the isotopic content of snow over sea ice. Here, we use a simulation run at a T127L95 resolution (0.9° x 0.9° horizontal resolution and 95 vertical

levels) and nudged towards the ERA5 reanalyses (Hersbach et al., 2020) over the period 1979 – 2022 Cauquoin and Werner (2021).”

- In section 4 (“Temperature versus δ18O relationships”), l.211:
“To enable a consideration of model dependency, we also compare our Historical ensemble against a nudged ECHAM6-wiso simulation (Table 1).”

- In section 4.2 (“Stability over the Historical period and model dependancy”), l.242 to l.253:
“Interestingly, the ECHAM6-wiso simulation and the last 50 years of our HadCM3 simulation display similar SAT-δ18O relationships. ECHAM6-wiso simulates slightly stronger relationships with a mean correlation coefficient difference of 0.04, while gradients tend to be slightly higher in HadCM3 with a gradient difference of 0.13 ‰/°C. The only notable differences are for Dronning Maud Land and the Indian coast with stronger relationships and higher gradients simulated by HadCM3 (Table 1). Thus, whilst it is unclear whether the nudging of ECHAM6 towards ERA5 reanalysis, the model resolutions or differences in sea ice behaviours, are the main reason for these discrepancies, it is clear that simulated temperature versus δ18O relationships have low but significant uncertainties. These need to be considered, both regionally and for the most relevant climate state, before being undertaking any inferences of past temperatures using isotopes measured in ice cores.”

In section 6 (“conclusions”), l.307 to 308:
“Interestingly, we find similar but slightly weaker SAT-δ18O correlations and slightly higher gradients compared to ERA5 –nudged ECHAM6-wiso simulations at the regional scale.”

Table 1 is updated to reflect the replacement of ECHAM5 with ECHAM6 output.

Table 1. Historical SAT–δ^{18}O relationships at the regional scale. Slope (in ‰/°C) plus or minus the standard error, and the correlation coefficient (into brackets) of the surface-weighed average of surface air temperature against the surface-weighed average of δ^{18}O for the Antarctic regions as defined in the PAGES Antarctica2k project (Stenni et al., 2017): the plateau, the Indian coast, the Weddell coast, the Peninsula, the WAIS, Victoria Land and Dronning Maud Land, simulated by the ECHAM6-wiso model (, over the period 1979-2022, 44 points, 'ECHAM6-wiso') and simulated by HadCM3 over he last 50 years (1955-2004, 50 points, 'last 50 years of HadCM3'), and over the whole historical simulated period (1851-2004, 154 points, 'Historical HadCM3') using the ensemble mean of the six simulations (see methods). All the relationships are significant (p-values<0.05).

	ECHAM6-wiso	last 50 years of HadCM3	Historical HadCM3
Plateau	0.48±0.07 [0.71]	0.61±0.14 [0.52]	0.57±0.07 [0.53]
Indian coast	0.29±0.08 [0.48]	0.55±0.15 [0.46]	0.67±0.07 [0.59]
Weddell coast	0.49±0.11 [0.57]	0.57±0.11 [0.59]	0.57±0.07 [0.57]
Peninsula	0.37±0.05 [0.74]	0.28±0.06 [0.52]	0.31±0.02 [0.71]
WAIS	0.56±0.07 [0.75]	0.60±0.12 [0.58]	0.50±0.05 [0.61]
Victoria Land	0.43±0.13 [0.46]	-	0.30±0.12 [0.19]
Dronning Maud Land	0.43±0.13 [0.46]	0.76±0.12 [0.69]	0.49±0.05 [0.60]
West Antarctica	0.49±0.11 [0.59]	0.50±0.10 [0.57]	0.70±0.07 [0.62]
East Antarctica	0.48±0.08 [0.69]	0.49±0.10 [0.57]	0.56±0.06 [0.58]
All Antarctica	0.45±0.09 [0.59]	0.67±0.13 [0.60]	0.57±0.06 [0.62]

3) extend the discussion of the SAM impact on the isotope signal;
new analysis of the impact of the SAM is given in Appendix G. This shows that HadCM3 reproduces the impacts of SAM on SAT and P reported in previous studies (Clem et al., 2016; Fogt et al., 2020), *i.e.* colder and drier conditions in a positive SAM. For δ^{18}O, HadCM3 simulates depletion in most areas of the Antarctic continent while the SAM is in a positive phase, but these

results are associated with relatively low correlation coefficients with means of -0.26±0.11 over the Historical period and -0.27±0.12 for the period 1950 – 2004. We thus conclude that our simulations cannot establish a robust link between the SAM and the Antarctic precipitation weighted $\delta^{18}O$. This result is supported by the diversity of $\delta^{18}O$ measurements from precipitation and firn/ice cores on different Antarctic locations (*e.g.* Vega et al., 2016; Kino et al., 2021; Servettaz, 2022; Dreossi et al., 2023). Moreover, it was shown that SAM impacts are different with the ENSO phases (Wilson et al., 2016), and that other modes affect Antarctic climate (*e.g.* Shields et al., 2022). Further analysis on the impact of the atmospheric circulation on Antarctic precipitation weighted $\delta^{18}O$ for the Historical period would need to be the subject of a future study. The new results are references in Section 5 ("Drivers") p10 l.292 to l.295 as:

"The dynamic processes behind the sea ice extent induced δ18O changes are complex and multiple. Although the Southern Annular Mode, leading mode of the atmospheric variability in the Southern Hemisphere, might explain part of these δ18O simulated changes (Appendix G), a more comprehensive study might investigate the impact of the atmospheric circulation changes."

In the conclusion, l.310, we replaced:

"We identify three processes [...]" by *"We suggest [...]"*, meaning that an extended study is necessary to check the atmospheric processes at the origin of our simulated results.

Here is our new Appendix G:

The Southern Annular Mode (SAM) is the leading mode of atmospheric variability in the Southern Hemisphere (Thompson and Wallace, 2000). Especially, it describes the position and the strength of the polar jet position, the southern westerly belt and the associated storm tracks. A positive (negative) phase of the SAM is associated with an intensified (weakened) pole-ward (northward) shift of the polar jet. The SAM is thus the preferred studied mode to investigate the Southern Hemisphere teleconnection with lower latitudes. Here, we used the definition of the SAM index following the approach of Gong and Wang (1999), as the difference between the normalized monthly zonal mean sea level pressure between 40°S and 65°S. Here we used the period 1961–1990 as a reference interval.

$$SAM = \frac{P_{40} - \mu_{40}}{\sigma_{40}} - \frac{P_{65} - \mu_{65}}{\sigma_{65}}$$

where P40 and P65 are the monthly mean sea level pressure at 40°S and 65°S, µ40 and µ65 are the mean of the monthly mean sea level pressure at 40°S and 65°S over the reference interval 1961–1990, and σ40 and σ65 are the standard deviations of the monthly mean sea level pressure at 40°S and 65°S over the reference interval 1961–1990. We computed the linear regressions between the calculated SAM and our climate variables (Figure G1): (i) the surface air temperature (SAT), (ii) the precipitation (P) and finally (iii) the precipitation weighted δ18O (δ18O). These linear regressions were computed over the whole Historical simulated period, as well as for the recent period 1950–2004, at the annual scale.

Note that, as done in the main corpus of the manuscript, we computed these linear regressions using the stack of the ensemble members, resulting in 918 points for the Historical period (1851–2004) and 324 points for the period 1950–2004. Within the frame of the CMIP5 project, the ability of HadCM3 to reproduce the SAM was evaluated (Zheng et al., 2013). As for all the CMIP5 models, HadCM3 overestimates the SAM index variability (Zheng et al., 2013; Zhang et al., 2022). Nevertheless, it reproduces the decadal variability of the SAM index and displays the best correlation coefficient between modeled and observed detrended SAM index (Zheng et al., 2013). Previous studies reported, based on observations, that main of the Antarctic continent is globally colder and drier while the SAM is in a positive phase, as the stronger southern westerly wind belt reduces the exchanges with warmer air masses from midlatitude regions, at the exception of the Peninsula (Clem et al., 2016). These effects are reproduced in our HadCM3 simulations, as shown *by the correlation coefficient values between the SAM and the SAT that are positive over the northern Antarctic peninsula, but negative over the rest of the continent, especially on coastal areas (Figures G1A and G1D). Similarly, it was shown that there is less southward moisture advection towards the Antarctic interior in a positive phase of the SAM, reducing precipitations. In our simulations (Figures G1B and G1E), this effect is enhanced over the Antarctic plateau, Victoria Land and Marie Byrd Land. At the opposite, the Antarctic peninsula receives more precipitation. However, the discripancy in the HadCM3 orography unables the «shadow effect» decreasing precipitation on the Eastern part of the peninsula due to the presence of mountains (Fogt and Marshall, 2020). The link between water stable isotopes and the SAM is less settled. A couple of publications displayed a correlation between the water stable isotope content in ice cores and the SAM index, but no systematic method allowed an established link. For instance, (Servettaz et al., 2022) suggest some impacts of the SAM on the isotopic content of the Aurora Basin North ice core over the last millennium, although not on the whole length of the core. Also, over the Fimbull Ice Sheet, Vega et al. (2016) suggest that the absence of correspondence between water stable isotopes and SAT might be explained by changes in atmospheric circulation, supported by a high correlation between d-excess measured in the KM and BI ices cores and the SAM index. Kino et al. (2021) showed the contribution of SAM over precipitation weighted δ18O at the daily scale simulated by the MIROC5-iso model nudged toward the JRA-25 reanalyses, over the period 1981–2010 at Dome Fuji. However, they warn that it does no prevail on all antarctic locations of the Antarctic plateau. For instance, Dome C is less sensitive to SAM compared to possible other teleconnections modes (Dreossi et al., 2023). In our simulations, the correlation coefficients between the SAM and precipitation weighted δ18O are significant and negative over the whole continent (Figures G1.C and G1.F), but remain week, with a mean of -0.26±0.11 over the Historical period and -0.27±0.12 for the period 1950–2004. From our simulations, we thus cannot neither establish a robust link between the SAM and the Antarctic precipitation weighted δ18O. However, studying the impact of the atmospheric circulation change on Antarctic precipitation weighted δ18O should not be boiled downed to the link with the SAM. For instance, only some El Nino Southern Oscilation (ENSO)/SAM combinations (El Nino/negative SAM and La Nina/positive SAM) contribute to strenghen the Amundsen Sea Low (e.g. Wilson et al., 2016), as observed through the analysis of the the Roosevelt Island Climate Evolution (RICE) δ18O Emanuelsson et al. (2023). SAM-induced processes impacting Antarctic precipitation weighted δ18O are also not trivial: SAM changes SAT, precipitation regimes but also the sea ice in a more complex manner (Fogt and Marshall, 2020). Other modes affect the Antarctic atmospheric circulation and might explain the δ18O changes, as for the Indian Ocean Dipole in phase with El Nino through the production of atmospheric rivers (Shields et al., 2022).*

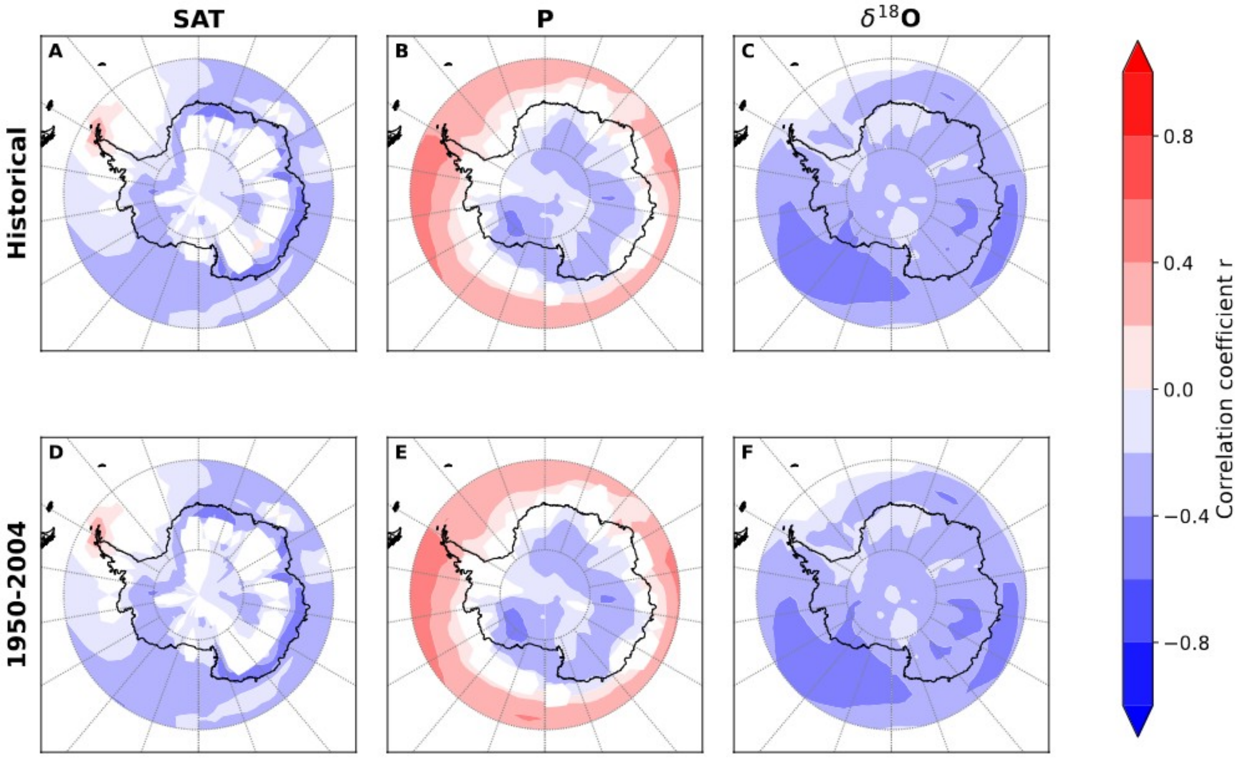

Figure G1. Correlation coefficients between the Southern Annular Mode index and the Surface Air Temperature ("SAT", A and D), the precipitations («P», B and E), and the precipitation weighted δ18O (C and E) simulated by the HadCM3 model at the annual scale for the Historical Period (1851–2004, first row) and the 1950–2004 period (second row). Only significant relationships are shown (p-value<0.05).

4) better describe the model and the simulations setup, and, in particular, was the sea ice simulated or prescribed? (also requested by the first Reviewer);

The paragraph is re-ordered as requested. Parts relating to CMIP6 are removed, to prevent confusion, and instead the details of the protocol are given, as requested, after the model description. This includes, as requested some more information on the HadCM3 coupled Atmosphere-Ocean model, and that sea-ice is not prescribed but calculated. The paragraph dedicated to our model description and simulations is now p.3 l.79 to 90:

"Here, we use the Hadley Center Atmosphere-Ocean general circulation model (HadCM3; AOGCM), to run six transient Historical simulations. HadCM3 is a version of the coupled Atmosphere-Ocean UK Met Office climate model (Pope et al., 2000; Gordon et al., 2000), which means that sea ice is prognostic. The model is equipped with stable water isotopes (Tindall et al., 2009). Its horizontal resolution is 3.75° × 2.5°, and there are 19 vertical levels (Pope et al., 2000; Gordon et al., 2000; Tindall et al., 2009). The setup of the Historical simulations is described in (Schurer et al., 2014), and follows the recommendations of the third Paleoclimate Modelling Intercomparison Project (PMIP3; Schmidt et al., 2011)(PMIP3; Schmidt et al. 2012). Each simulation is forced with time-varying orbital, solar, volcanic, land-use and well-mixed greenhouse gas forcing. As above, sea ice is not prescribed, rather calculated by the model. Changes in orbital parameters were calculated following (Berger, 1978). Volcanic forcing is that described in (Crowley et al., 2008). The solar forcing follows (Shapiro et al., 2011). Changes in CO2, N2O and CH4 were set following the PMIP3 standard (Schmidt et al., 2011). Changes in the abundances of 6 Halocarbons were prescribed following (Tett et al., 2007). Changes in land-cover were prescribed by reclassifying the Global land cover reconstruction developed by (Pongratz et al., 2008). Each of our simulations were only altered by starting each simulation a year apart."

5) address a number of minor comments and questions.
We integrated the changes we reported in our responses to the reviewers'minor comments, as spotted in the track-change version.

Also, as requested by Reviewer 1, please improve the figures, in particular Figure 2.
Finally, we improved the figures increasing label and title sizes. In Figure 2, we tried to simplify the reading by relocating the subplots and adding the names of the regions.