

Reviewer #1: Yetang Wang (comments in red Arial font), 24 Dec 2025

Author responses in black Times New Roman font

By integrating observations, reconstructions, and Earth system model experiments, this study provides a comprehensive insight into changes in Antarctic snow accumulation and its underlying physical mechanisms. The research indicates that increased greenhouse gases are the primary driver of enhanced Antarctic accumulation, but their effect is partially offset by the cooling influence of aerosols. Since the mid-20th century, the high-latitude wind field and meltwater from ice shelves have collectively suppressed Southern Ocean warming, significantly weakening the greenhouse gas-induced increase in accumulation. Both models and reconstructions consistently demonstrate that wind patterns associated with the strengthening of the westerlies and Amundsen Sea Low lead to uneven accumulation distribution across the Antarctica. Furthermore, the authors note that although the historical wind fields in large ensemble simulations align with observations, the corresponding SST and overall Antarctic accumulation responses appear decoupled from these wind field changes. This discrepancy may stem from unaccounted meltwater forcing and/or limitations in model resolution. In summary, the authors have undertaken substantial work. These results are highly valuable for estimating the changes in Antarctic snow accumulation rates and their contribution to global sea-level rise, and will be of great interest to the communities reading Earth System Dynamics. Therefore, I recommend acceptance of this manuscript after minor revisions. Some of my suggestions are as follows:

We thank the Reviewer for their in-depth review, and are encouraged by their positive comments on the value of the work and its relevance to the ESD audience.

Below we detail our responses and discuss the additional work that has been performed for the revised version.

This response highlights the major revisions to the text that address the reviewer's main points. There is a summary of major changes on the last page of this response. Additional minor changes to the text since its original submission are indicated in the tracked changes version that has been uploaded along with this response.

The introduction is logical and concise. However, the summary of recent snow accumulation studies could be enhanced, particularly with regard to the multi-source data employed and the novel conclusions.

The authors should provide a brief introduction to the spatial patterns of Antarctic snow accumulation under different forcings or modes in the results section and compare their differences.

To address both of these requests at once, we add a new summary of previous studies and spatial patterns to the Introduction at lines 89-109:

“Beyond its central role in the AIS mass balance, Antarctic snow accumulation records the history of variations in the Southern Annular Mode (SAM; e.g. Medley and Thomas, 2019; Wang and Xiao, 2023; King and Christoffersen, 2024), teleconnections arising from the El Niño – Southern Oscillation (ENSO; e.g. Wang and Xiao, 2023; King and Christoffersen, 2024), and the strength and position of the Amundsen Sea Low (ASL, e.g. Raphael et al., 2016; Dalaiden et al., 2024), all prominent features of the Antarctic-Southern Ocean climate system that are affected by internal variability and anthropogenic forcing. These

patterns are not necessarily independent – a positive SAM phase (which corresponds with strengthened and poleward shifted westerly winds) is typically associated with a deepened ASL and often occurs during the La Niña phase of ENSO (e.g. Schneider et al., 2012). Since instrumental observations in Antarctica began in the 1950s, the SAM has exhibited a multidecadal trend towards its positive phase (Marshall, 2003) while the ASL has deepened (Raphael et al., 2016), resulting in a snow accumulation trend dipole across the West Antarctic Ice Sheet (WAIS), with more anomalous accumulation in the Peninsula region and less on the Ross Ice Shelf side of the WAIS (Medley and Thomas, 2019; Wang and Xiao, 2023). Integrated Antarctic-wide, there has been a net negative trend in snow accumulation associated with the SAM trend (Medley and Thomas, 2019), damping the expected positive accumulation trend driven by atmospheric warming. Although the SAM trend is commonly attributed to stratospheric ozone depletion (e.g. Gillet and Thompson, 2003; Polvani et al., 2011), greenhouse gas increases also contribute to it (e.g. Kushner et al., 2001; Arblaster et al., 2011), while paleoclimate reconstructions indicate that a significant positive trend occurred in the mid-20th Century prior to the onset of ozone depletion, implying significant modulation by internal variability (Jones and Widmann, 2004; O'Connor et al., 2021)."

Due to length considerations, we do not expand upon every previous study in the Introduction, but instead note the consensus that has been built while pointing out the outstanding puzzles that our study addresses, e.g. the puzzle of the limited change in Antarctic-wide accumulation rates in recent decades and the gap in quantitative evaluations of how models agree with the observed trends under different combinations of external forcings and realizations of internal variability.

In addition, we revise the beginning of Section 3.2 to explain how we evaluate the spatial patterns of snow accumulation, particularly elaborating on our pattern correlation method:

"For spatial analysis, we survey the pattern correlations between the 20th-Century trend in MT19 accumulation and each ensemble member of the Large Ensemble and TPACE, along with the ensemble means of all historical experiments from Table 1 (Figs. B1, B3). Similarly, pattern correlations for the sea level pressure (SLP) trend across 40°-90° S are computed with the same model experiments and using the CESM2-LE PDA as the benchmark (Figs. B2, B4). For illustration, we select the ensemble member whose trend patterns in snow accumulation and atmospheric circulation exhibit the best pattern correlations to the MT19 and CESM2-LE PDA reconstructions, respectively. Specifically, member 40 from CESM2-LEcmip6 has a pattern correlation with the reconstructed SLP trend over 40°-90° S of 0.92 and with the reconstructed accumulation trend on Antarctica of 0.52, giving an average correlation of these two metrics of 0.72. This is the highest average correlation out of 110 individual members from the CESM2-LE, CESM2-LEcmip6, and TPACE ensembles. The following discussion uses member 40 to illustrate spatial patterns but these same relationships are found in several other ensemble members."

Section 3.2 is the longest part of the Results section. It points the reader to Appendix B which illustrates the trend patterns under all forcings and single forcings - (GHG, AAER, EE, GHG + EE, GHG + AAER, BMB) - due to tropical SSTs alone and the canonical La Niña (Figure B5), and across 110 individual ensemble members from the Large Ensemble and TPACE ensembles. We mention in the text of Section 3.2 how the externally forced patterns compare with the pattern in MT19, i.e. while [CESM2-LEcmip6] matches the best of the ensemble means, [GHG] and [EE] are second best, underscoring the roles of greenhouse gases and stratospheric ozone depletion in shaping these trend patterns. Where Figures 10 and 11 are discussed later in the text, the pattern in [AAER] is discussed. We cannot discuss every individual map from Figures B1-B4; instead we highlight member #40 which is the ensemble member that best matches the proxy reconstructed trend patterns.

We have chosen not to use a formal modes of variability framework for this paper. This is partly because modes of variability do not clearly separate anthropogenic signals from internal variability, and that on the 100-year timescale discussed here, the anthropogenic signals are the dominant drivers of the accumulation trends. In addition, a given mode of variability (like the SAM) only explains a fraction of the variance of the system. In this paper, we use wind-nudged experiments to diagnose the contribution of atmospheric

circulation irrespective of modes of variability. Similarly, the zonal wind index used from the proxy reconstruction is of wind speed, not a mode index like the SAM index. It is noted in Appendix A that the proxy reconstruction of wind speed is more skillful than the proxy reconstruction of the SAM index.

That said, to place these results in the context of previous work, in the Introduction and Section 3 we discuss how the patterns we illustrate relate to the SAM and ENSO patterns that have been discussed in work like Medley and Thomas (2019), Wang and Xiao (2023), and King and Christoffersen (2024). In effect, we are considering these patterns as the chosen SLP domain for the pattern correlation metric captures the footprint of the ASL and the 40-90S footprint of the SAM pattern.

In the Discussion, it would be highly beneficial to include a schematic diagram, such as a conceptual model illustrating the physical feedback chain linking "wind–meltwater–SST–snow accumulation". This diagram should include as many atmospheric and oceanic processes mentioned in the study as possible.

The revised Section 3.5 presents a new schematic diagram, Figure 10, with the following discussion of it:

"The foregoing results are summarized in graphical form in Figure 10, which compares accumulation trends and their large-scale contexts across the reconstructions (Fig. 10a), the GHG-only scenario (Fig. 10b), and associated with the proxy reconstructed wind timeseries (Fig. 10c). The maps use a pseudo Spilhaus projection (e.g. Chen et al., 2023) to place Antarctica and the Southern Ocean near the center of the map and emphasize connections across the ocean basins. In Fig. 10a, the 10.5 mm of sea level mitigation occurs in the context of a deepened ASL and strong warming around most of Antarctica except for the Amundsen Sea. The ice-albedo feedback makes these warming/cooling trends larger than anywhere else except for the Arctic. Cold-air advection and the wind-evaporation-SST (WES) feedback propagate the Amundsen Sea cooling signal towards the tropics as the south Pacific subtropical high strengthens. However, this signal is blocked from reaching the eastern equatorial Pacific by the ITCZ and by the competing effects of greenhouse gases (Fig. 10b). Although the eastern tropical Pacific SSTs look El Niño like, the overall SST anomaly pattern results in Rossby waves propagating towards Antarctica that deepen the ASL, reinforcing this pattern (see [CESM2-GOGA] simulated SLP pattern, Fig. 5a), which promotes ice shelf thinning and glacier retreat in the Thwaites and Pine Island region (King and Christoffersen, 2024; O'Connor et al., 2025). Under greenhouse gases only (Fig. 10b), the tropics are decidedly El Niño like and poleward propagating Rossby waves act to weaken rather than reinforce the ASL deepening trend. The south Pacific subtropical high strengthens, promoting the WES feedback, but it does not extend to the coast of South America where it could help stabilize the stratocumulus cloud deck, enhance coastal upwelling, and alter the trajectory of the tropics.

Greenhouse gases explain the widespread warming and accumulation increase evident in the proxy reconstructions but do not explain the important dynamical details that suppress the accumulation rate increase, give rise to a dipole in snow accumulation across the WAIS, and reshape the warming patterns. The Southern Hemisphere westerly winds are the major source of momentum forcing of the global ocean (e.g. Wunsch 1998). Their potential impact on surface climate is depicted in Fig. 10c. In this depiction, the cooling signal originating in the Amundsen Sea is communicated all the way to the equatorial Pacific along the reverse teleconnection pathway. The subtropical shortwave cloud feedback (Kim et al., 2022) and the ITCZ (Dong et al., 2026) are the key mediators of this connection. Recent studies argue that strengthening winds forced by stratospheric ozone depletion contribute to the La Niña like SST trend in the tropics (Hartmann, 2022; Dong et al., 2025). Similarly, Antarctic meltwater can contribute to a La Niña like SST trend pattern (Dong et al., 2022a; Zhang et al., 2026; Dong et al., 2026). However, none have argued that these high latitude signals are the sole cause of the observed La Niña like trend. It is important to keep in mind that the wind timeseries (Fig. 1b) used in Fig. 10c includes signals arising from internal variability. During the 20th Century, a trend towards the negative phase of the IPO (which favors La Niña) preceded strong forcing from greenhouse gas increases and stratospheric ozone depletion (Dalaiden et al., 2024). As such, this IPO trend is part of the wind-congruent map in Figure 10c.

Nonetheless, the presence of internal signals within this IPO-like pattern does not rule out the possibility of anthropogenic forcing driving a multidecadal shift towards the negative phase of the IPO (Klavens et al., 2025). Regardless of the exact origins of this pattern, as shown it exhibits the mutually reinforcing two-way teleconnections discussed in Dong et al. (2022b). This reinforced circulation pattern enhances meltwater production by the mechanism of O'Connor et al. (2025). To the extent that the meltwater-driven cooling signal reaches the tropics, it is communicated back to the Southern Ocean via the two-way teleconnection.”

This schematic avoids using the meltwater-induced SST patterns from the CESM1-AIS meltwater experiment over concerns that it is not a very realistic experiment for the 20th century. However our discussion of the schematic highlights that the ASL deepening trend is the perfect pattern to drive anomalous freshwater production from the thinning Thwaites and Pine Island glaciers and ice shelves, according to recent studies like O'Connor et al. (2025) and King and Christoffersen (2024). As such, the indicated -4 mm SLE cumulative mass change on the right-column of Figure 10 is a ballpark estimate of what the combined effects of winds and meltwater may have been on the snow accumulation trend.

Since our paper was submitted, Sadai et al. (2025) have published a set of CESM1 experiments that explicitly model the meltwater signal in the climate system as it evolves through the 21st Century, including for Antarctic precipitation. We cite Sadai et al. (2025) in the Discussion and as motivation for using the CESM1 AIS meltwater experiment in Section 2.3. Another paper was recently published – Zhang et al (2026) is a multi-model study using the SOFIA experiments to document the robust Southern Ocean and tropical responses to Antarctic meltwater. They diagnose the feedbacks involved and have a Figure 1c showing the multi-model global precipitation response, indicating a significant precipitation decrease over Antarctica. Both these studies are now mentioned in Section 3.4 to support results from CESM1 AIS-meltwater:

“This result of an accumulation decrease in response to Antarctic meltwater is qualitatively supported by Zhang et al. (2026), who find meltwater-driven precipitation decreases on the Antarctic continent across multiple CMIP6-era models. Additionally, Sadai et al. (2025) find a similar precipitation decrease in a set of CESM1 experiments driven by meltwater fluxes from an offline ice sheet model. Observational uncertainties and model limitations make a highly realistic meltwater simulation very challenging, but there is a multi-model consensus that meltwater leads to Southern Ocean surface cooling and a precipitation decrease over Antarctica (e.g. Zhang et al., 2026).”

Current understanding indicates that Antarctic climate change is strongly influenced by internal variability. Therefore, I suggest the authors enhance the mechanistic explanation regarding the impact of modes of climate variability on Antarctic snow accumulation rates, particularly the teleconnections originating from the tropical Pacific.

We agree that internal variability is important and have already discussed above some places where the revised manuscript mentions internal variability more prominently than before. Here is a summary of where internal variability is highlighted in the revision:

Abstract line 41 – attributing the wind history to both internal variability and anthropogenic forcing.

Introduction line 108 – citing previous work suggesting modulation of the SAM trend by internal variability

Introduction at line 150 – noting that the observed Southern Ocean cooling trend that models cannot replicate might be a product of internal variability and therefore we look at the 100-member Large Ensemble and the prescribed SST/sea ice experiments to probe how internal variability affects the SST and accumulation trends.

Line 165-166: Stating that the large internal variability of precipitation is a motivation for using the cumulative mass metric. (Although not explicitly stated, the 100+ year long record we evaluate makes the internal signals less pronounced).

Line 185 – Noting that TPACE syncs model to observed sequence of internal variability in the tropical Pacific.

Line 221 – Noting that Antarctic wind nudging strongly constrains internal variability in the model

A new Section 2.5 (lines 328-354) summarizing our overall strategy for accounting for internal variability as we attribute snow accumulation trends:

“The broad suite of experiments combined with results from PDA enable a multi-faceted strategy to characterize signals arising from external climate forcing and separate them from the effects of internal variability. For the forced response, the ensemble mean of a large ensemble smooths out the signals of internal variability arising from the initial conditions, revealing the model’s response to the external radiative forcings (greenhouse gasses, aerosols, etc.) imposed in the experiment. In this text, the forced response is denoted by square brackets, for example “[GHG]” for the ensemble-mean of the greenhouse gas single forcing ensemble. The internal component of a given ensemble member is found by subtracting the forced response from that member.

Using large ensembles is only the first step towards a comprehensive accounting of internal variability. This step may not be sufficient because a) models like CESM2 may simulate an unrealistic magnitude of internal variability (e.g. Casado et al., 2023); and b) internal variability in free-running coupled experiments like the Large Ensemble is not synched with real-world internal variability. The second step is the use of nudged coupled experiments, TPACE and CESM2-WNUDGE, which constrain the coupled model by introducing observed anomalies. Third, we use a long segment of piControl to characterize the evolution of accumulation in the absence of anthropogenic forcing. Fourth, the PDA reconstructions of the major surface climate variables that we use are constrained both by well-calibrated proxy records and by the physical dynamical frameworks of CESM1 and CESM2. Finally, we evaluate accumulation trends in prescribed SST / sea ice experiments using again the same dynamical framework of the atmosphere and land component models of CESM. Importantly, results from all of these approaches, presented below, converge upon a single set of conclusions: The accumulation rate increase during the 20th Century that mitigated sea level rise by ~ 11 mm can only be explained by external forcing, namely greenhouse gases. Real-world internal variability most likely has dampened the forced response. The follow-up question we discuss but ultimately leave open for future work is the extent to which internal variability is sufficient to account for observation-model trend discrepancies, or if there is an important role for poorly resolved processes such as the response to anomalous ice sheet and ice shelf meltwater and the sensitivity of SST anomalies to wind stress.”

We believe that this strategy adequately addresses internal variability versus forced responses without formally adopting a modes of variability framing. Modes of variability like the SAM and ENSO are themselves affected by internal variability and anthropogenic forcing. For the purposes of this study, it does not help us to reach our goal of attributing snow accumulation trends by partitioning the variance of snow accumulation into a) Driven by SAM and/or ENSO vs. b) not driven by SAM and/or ENSO. Rather, we can directly use the model framework to evaluate internal variability vs forced responses as discussed in Section 2.5.

(continuing the list)

Section 3.1, lines 416-419: Discussing the ensemble spread of cumulative mass gain and how internal variability can bring the model results in agreement with the trend in MT19, but not explain the mass gain itself.

Section 3.2/ Appendix B: Extensive illustration of how internal variability manifests in each ensemble member, and further linking the observed pattern of internal variability to the negative phase of IPO (lines 570-592).

Lines 770-775: In discussing the Figure 10 schematic, it is emphasized that the wind-congruent map contains both anthropogenic and internally driven signals, even though forcing from stratospheric ozone depletion and meltwater could tip the scales towards a La Nina like / negative IPO state.

Lines 830-850: The text makes clear that the ASL deepening trend has been partly internally driven, while the ongoing retreat of glaciers and thinning of ice shelves in the Amundsen Sea region may have been triggered by an internally driven event.

Table 3 top row clarifies that the SST trends driven by winds and meltwater are both anthropogenic and internal.

Line 966: The Conclusion directly states that winds and meltwater histories likely have both internal and anthropogenic drivers.

Regarding teleconnections originating from the tropical Pacific, we assume that the Reviewer is referring to the wave train that manifests in the deepened ASL. The ASL is a circulation feature with large internal variability, but as diagnosed by Dalaiden et al. (2024), it also exhibits a persistent deepening trend driven by greenhouse gas increases and stratospheric ozone depletion.

Changes in tropical Pacific SSTs over the past 40 years have been the subject of much debate in the broader climate science community (e.g. Wills et al., 2022); it is not clear that any tropically driven teleconnection pattern should be assumed to be a product of internal variability alone. Put another way, it would be a mistake to suggest that “modes of climate variability” or “tropical teleconnections” must be purely internally driven.

As for the mechanistic explanation of tropical teleconnections, Section 3.2 lines 553-555 reminds readers that the deepened ASL shown in Figures 4 and 5 is dynamically consistent with poleward propagating Rossby waves and cooling in the central and eastern tropical Pacific owing to the shifted location of tropical deep convection that these SST anomalies imply:

“Considering the dynamics of poleward propagating Rossby waves, a weakened ASL is typically associated with an El Niño – like SST pattern in the tropics, while a deepened ASL is typically associated with a La Niña – like SST pattern (e.g. Raphael et al., 2016).”

The mechanism by which the deepened ASL results in the dipole pattern in snow accumulation trends across the WAIS is mentioned at line 485.

The authors should add a significance testing when conducting trend analyses, such as Figures 4 and 7.

We have added significance testing to the accumulation trend and wind-accumulation regression maps in Figures 4, 6, 7, 8, and 11 . The updated figures use hatching to indicate where the trends are not

significant at the $p < 0.05$ level or better. To keep the figures easy to read and the focus of the paper on accumulation, we have chosen not to show the testing of trends in sea level pressure or surface temperature. We note that Antarctic circulation and temperature trends have been well studied in other work with well-established attributions (e.g. SLP and wind trends driven by stratospheric ozone depletion and greenhouse gases); here they are just used to aid in interpretation of the accumulation trends.

I recommended to add citations or comparisons with results from other CMIP6 models in the discussion.

Although we agree with the value of multi-model comparisons, we are not aware of a comparable study using the broader set of CMIP6 models to attribute Antarctic snow accumulation trends to specific forcings, or to distinguish anthropogenic trends from internal variability. In this work we take advantage of a comprehensive suite of CESM2 experiments that have been conducted over the last ~ 5 years. These CESM2 experiments are not part of the common set of CMIP6 experiments whose results are directly comparable across multiple models. To the best of our knowledge, most of these experiments were performed after NCAR/CESM submitted their experiments to the CMIP6 project.

The revised Discussion mentions this limitation at lines 918-924:

“This work has several limitations, the most prominent being that it features a single low-resolution model, CESM2. The advantage of relying on CESM2 is that it enables a robust assessment of the roles of anthropogenic forcing and internal variability across a hierarchy of experiments from free-running fully coupled to nudged coupled to prescribed SST/ sea ice experiments. We are not aware of comparable studies that have used CMIP6 or other models to make similar assessments of Antarctic accumulation trends across a hierarchy of experiments. As such, we encourage future coordinated efforts like CMIP7 and beyond to facilitate this type of study across multiple models.”

The Discussion (at lines 820-845) cites some older attribution work with CMIP5 models, such as Previdi and Polvani (2016), who evaluate the presence of anthropogenic signals vs internal variability in Antarctic precipitation. The Discussion also mentions more recent studies like Dalaiden et al (2024) attributing circulation trends with multiple models. Inasmuch as circulation trends have driven accumulation trends, we can infer that the attribution of circulation trends has some relevance to understanding accumulation trends.

It is beyond the scope of this work to analyze additional CMIP6 models, however, the revision cites new studies like Zhang et al. (2026) and Sadai et al. (2025) that have evaluated meltwater signals across multiple models. We also added citations for Arblaster et al. (2011) and Purich et al. (2025) with respect to the future of the SAM trend (line 894-985). We added a citation for Kang et al. (2026) in our discussion of new results from high resolution models, along with a suggestion that Antarctic evaluation metrics be added to the evaluation of trends simulated by high-resolution models (lines 944-945).

In a possible final round of revision, we would be happy to cite any relevant CMIP6 evaluation studies that we have overlooked if the Reviewer can suggest specific citations. This would likely be cited in section 4.1 and the text in Section 4.3 would not need to state that there is a lack of CMIP6 studies to cite.

Regarding readability of the figure, it is suggested that the authors increase the font size of text in certain figures to make them more noticeable to readers, as illustrated in Figure 3. Moreover, the representation of statistical significance (stippling) in Figure 5 is not sufficiently clear.

We have checked the readability of every Figure and have made appropriate adjustments to the font sizes. In addition, the figures are not shrunk down as much when inserted into the manuscript document. For

final publication, separate, high-resolution files of each figure will be provided. We also made a more substantial change to Figure 2c.

Regarding Figure 2c, we made a mistaken assumption in the calculation of cumulative mass gain congruent with the wind trend during 1901-2000. The “wind-congruent” and “wind-adjusted” timeseries have been removed from Figure 2c, and the corresponding discussions in the text have been deleted. This correction makes the overall narrative consistent throughout the manuscript, in that both wind and meltwater signals are needed to explain the data-model discrepancies, not just the wind trend. Our revised results concur with Medley and Thomas (2019) in that the positive SAM trend (or stronger winds) does explain some mass loss since the mid twentieth century owing to lower snowfall associated with positive SAM. However, considering the 1901-2000 period as a whole, winds anomalies were relatively weak from about 1901-1940, so they contributed to a relative mass gain during the early 20th century, largely canceling out their contribution to mass loss in the second half of the 20th century. This correction does not change the Abstract’s emphasis on the role of wind and meltwater on dampening GHG-driven trends since the mid-20th Century.

Regarding Figure 5, instead of using white stippling to indicate significance the revised Figure 5 masks out the areas where the wind-SST regression coefficients are not significant. This change highlights the clear relationship between Antarctic wind and tropical/ eastern Pacific SSTs in the observations and reconstructions which is much less evident in the model. Figure A4 shows this same result in the two different instances of PDA and two different members of the Large Ensemble.

REVIEWER #2 – anonymous 29 March 2026

This article analyses ensembles of model simulations to show that the increase in snow accumulation over Antarctica induced by greenhouse gas emissions has been dampen by aerosol emissions and by changes in winds and ice sheet meltwater release. There is a huge amount of work behind this article, with the comparison of several multi-member ensembles of climate simulations. It is an important piece of work that bridges a gap towards the attribution of the total mass change of the Antarctic ice sheet, and that helps understand limitations of the CMIP models. However, several sections of this manuscript are difficult to follow. In my opinion, major revisions would therefore be needed to make this article suitable for publication.

We thank Reviewer #2 for their thoughtful, constructive review. We appreciate the recognition of the importance of the work as well as noticing the substantial effort it took our team to arrive at these results and conclusions. We acknowledge that the clarity of the text in the original manuscript had room for improvement. We are grateful for the reviewer's helpful suggestions on how to make improvements. We think the reviewer and readers will find the revised text to be substantially clearer.

Below we detail our responses and discuss the additional work that has been performed for the revised version.

This response highlights the major revisions to the text that address the reviewer's main points. There is a summary of major changes on the last page of this response. Additional minor changes to the text since its original submission are indicated in the tracked changes version of the manuscript that has been submitted along with this response.

Major comments:

I acknowledge that this article is not easy to write due to the large variety of analysed simulations. However, there is room for improvement, and I find the manuscript difficult to read in many sections. First, the information on the experiments is difficult to find, with some in section 2, some in Appendix, some in the results section (e.g., CESM1-AIS-meltwater), and some in other articles. My recommendation would be to put everything in the Method section. Second, the description of the results is sometimes long and difficult to follow, with some numbers difficult to relate the associated figure or experiment, and a lack of introductory sentences in each section to explain the aim of the following result description (while reading, the readers do not know where these long descriptions are leading them). A short paragraph describing the overall approach at the beginning of section 3 would also help.

The challenge in writing this article was not necessarily in the volume of analyzed experiments, but rather in developing an interpretation that fits all of the available evidence. We appreciate the chance to better explain this interpretation and make the overall paper easier to follow.

Following the Reviewer's suggestion above, we moved most of Appendix A to Section 2 (Methods). Table A1 has become Table 2, and Table 2 has become Table 3. This makes the description and purpose of each experiment easier to find.

Regarding the overall approach, the rationale and roadmap for the paper are now stated at:

a) Introduction at line 111... *"The primary aim of the present study..."*

- b) Methods lines 138-152...Summarizing previous work on snow accumulation by Dunmire et al (2022) and related work on climate sensitivity makes clear that we must consider SST patterns and their affects on snow accumulation trends in order to properly evaluate the model and attribute the trends.
- c) The new Section 2.5, quoted above in response to Reviewer 1, lays out the rationale for using this diverse set of experiments for separating forced responses from internal variability and for investigating the role of observed SSTs.
- d) The beginning of Section 3 has an overview of the section to orient the readers to what comes next.
- e) The Abstract and Conclusion are self contained and consistent with each other. Casual readers who only read these sections and browse the Figures should walk away with an understanding of the main messages.

My second major concern is that at several places, the reasoning is difficult to follow because of a lack of description of the experimental design, and/or because some numbers are not clearly sourced. These issues are listed in the detailed comments below.

Please see our responses to the detailed comments below. The revision has expanded and more accessible descriptions of the experimental design. The numbers are clearly sourced following the reviewer's suggestions in the detailed comments.

My third major comment is that the authors seem to consider wind and meltwater changes at the same level as changes in greenhouse gases and aerosols. For example, they conclude that "winds have been working together with ice sheet and ice shelf meltwater fluxes to dampen Southern Ocean surface warming and suppress the GHG-driven snow accumulation increase". Unlike GHG and aerosols, winds and SSTs are not external forcings, they are variables of the climate system. They are affected by GHG and aerosols among others and are also part of the internal variability. Technically, prescribed meltwater may be considered as external forcing, although it is only external due to the absence of interactive ice sheet in the model, and should in theory respond to GHG, aerosols, etc. I think that better introducing all these concepts in the manuscript would make the analysis easier to follow and the results more robust.

It was not our intention to frame winds or SSTs as forcings, and are concerned that the original manuscript may leave this impression on some readers. As emphasized in the Abstract, *"rising concentrations of greenhouse gases (GHGs) have been the underlying driver of increased accumulation."*

In the revision, the star role of GHGs is stated in the first sentence of the Conclusion at line 959: *"Greenhouse gases from anthropogenic activities emerge as the primary contributor to both the long-term Antarctic accumulation rate increase and the larger-scale warming patterns. Other factors have dampened this accumulation rate increase and warming trend."*

This conclusion is foreshadowed at line 350, at line 431 in Results, emphasized in the discussion of the new Figure 10 schematic around lines 759-760, and again in the beginning of the Discussion around line 819.

In this clearer framing, everything besides GHGs falls under "other factors." Due to uncertainties in the observational record and modeling limitations, we *do not conclude with 100% certainty that meltwater has been a major driver of the accumulation and SST trends discussed*. That said, the available evidence does support a hypothesis that meltwater has been involved. The role of the winds is even clearer, as the wind history is relatively well constrained and wind is easy to model in comparison to meltwater. In the list of other factors at lines 962-963, meltwater is placed last after aerosols and winds.

Essentially, *wind and meltwater have to work together* because the anomalous wind patterns are the key driver of the anomalous freshwater fluxes from ice shelf basal melting. The interaction of the winds with meltwater was the point of the Holland et al (2022) attribution study; now there are even clearer wind-meltwater linkages with newer studies like O'Connor et al. (2025), King and Chirstoffersen (2024), and Clark et al (2024).

We agree that winds and SSTs are not external forcings and have carefully gone through the text to make sure that they are not referred to as external forcings. Some authors like Zhai et al (2023) refer to tropical Pacific SST “forcing” but we avoid equating SST anomalies with external forcing. Similarly, when Clark et al. (2024) discuss “external forcing” they are referring to large-scale climate dynamics outside of the Amundsen Sea Embayment. As climate modelers we avoid this usage.

At the same time, we feel it is critically important to discuss the roles of winds and SSTs, as models systematically fail to accurately simulate observed regional trends in winds and SSTs, which in turn negatively impacts the models’ ability to simulate trends in accumulation. We agree that it is appropriate to think about winds and meltwater as knock-on effects to GHGs, with some role also for internal variability and external forcings like aerosols and stratospheric ozone.

Meltwater has been put forward as one of the leading candidates to explain SST trend discrepancies between observations and models (e.g. Dong et al., 2022a). Technically, prescribed meltwater is a forcing and we follow the lead of Schmidt et al (2023) in describing it as such. We expect that this classification will evolve as interactive ice sheet models become more widely available; then meltwater may be treated as and best described as a feedback. We readily acknowledge the absence of an interactive ice sheet in CESM2 and the limitations of the CESM1 prescribed meltwater experiment that we analyze.

The revised Discussion in Section 4.3 mentions the possibility that high-resolution modeling alone might be enough to accurately simulate observed trends without needing to treat meltwater as a forcing, following the new work of Kang et al (2026). However, there is no evaluation of Antarctic accumulation trends or energy balance or effective climate sensitivity in Kang et al (2026) so we suggest in Section 4.3 that more evaluation needs to be done before we can conclude that high resolution alone is enough without accounting for meltwater. Kang et al (2026) present just one ensemble member for a limited period of 1980-2013 and not all high resolution models can accurately simulate the observed SST trend patterns (Dhame et al., 2025). *The revised Abstract at line 44 retains its claim that observed-model trend discrepancy “likely arises from omitted meltwater forcing and/or resolution limitations.”*

Detailed comments:

Abstract, L. 28-29: I found it awkward to describe the role of GHG and aerosols, then to mention that there are other processes and describe the role of winds and freshwater. Changes in winds and freshwater do both emerge from emissions of GHG and aerosols (and internal variability, ozone, ...), so I would not present them as “other processes at work”.

The revised language reads, *“Aerosol-driven cooling partially compensates this overprediction, but the reconstructions provide evidence that poorly modeled processes can explain observation-model trend discrepancies.”*

In the interest of length, the Abstract largely avoids partitioning trends in the winds and meltwater to anthropogenic forcing vs internal variability. These attributions are discussed in Sections 3.5 and 4.1. However, at line 41 the revised Abstract briefly mentions that winds have been driven by external forcing and internal variability.

L. 65: The units seem wrong, Gt yr⁻¹ is equivalent to mm yr⁻¹, not to mm. This number is also misleading because the net SMB is not per se equivalent to sea-level rise, it is only the anomalies that lead to sea level rise.

Corrected mm to mm yr⁻¹. “Sea level equivalence (SLE)” is an accepted term (used in Medley and Thomas (2019) for example) that does not mean the same as “sea level rise.” That is why this line uses the term “Sea level equivalence (SLE)” and not “sea level rise.” If dynamic mass loss exceeds mass input from snow accumulation, the ice sheet can contribute to sea level rise in the same year that a flux of water of 6 mm yr⁻¹ SLE moves from the ocean to the ice sheet.

The revised Methods Section 2.3 mentions that we analyze anomalies in accumulation relative to the pre-industrial control climate. Additional snow accumulation above the pi control baseline contributes to sea level mitigation, while reduced snow accumulation below the baseline contributes to sea level rise.

L. 113-114: This is unclear. Rainfall is also an accumulation term (it fills firn porosity), i.e., a positive term in the surface mass balance, not only snowfall. Is rainfall included in the accumulation term or not?

Rainfall is included in the accumulation term, as was stated in Appendix A at line 845 in the original manuscript. In revision, this is stated at line 280.

L. 170-172: I would like some more explanations about the TPACE and "everything else" (EE) ensembles, and why the role of stratospheric ozone depletion and recovery cannot be isolated with these experiments (if it is important for this article, if not, it could be removed). Why isn't stratospheric ozone investigated in a similar way as GHG and aerosols? Stratospheric ozone is part of the usual Detection–Attribution protocol.

The CESM2 Single Forcing Large Ensemble (Simpson et al 2023) did not include a stand-alone ozone ensemble likely due to limited resources and the emphasis of this ensemble on understanding CESM2's response to aerosols. In the context of this paper, the role of stratospheric ozone depletion and recovery cannot be specifically isolated because there is no CESM2 single-forcing ozone ensemble available to isolate it with. We recognize that stratospheric ozone depletion is a relevant forcing, and that the lack of a stand-alone ozone ensemble is a limitation of this work. Nonetheless, we think this is a worthwhile tradeoff because previous work on Antarctic precipitation has not isolated the roles of GHGs, anthropogenic aerosols, and BMB aerosols.

Within the SFLE, prescribed ozone forcing is lumped into the EE ensemble, while some ozone depleting substances are greenhouse gases that are included in the GHG ensemble. The strong accumulation signal in EE appears after 1980 (Figure 2c), consistent with the timing of stratospheric ozone depletion, and as mentioned in Table 3 we interpret the cumulative mass gain in EE as “largely due to late-twentieth century stratospheric ozone depletion but with other minor influences.” “Other minor influences” include volcanic and solar forcing as well as tropospheric ozone. It is beyond our scope to evaluate how each of the individual forcings prescribed within EE contribute to accumulation trends. Nonetheless, it is likely that the accumulation signal appearing in EE after ~1980 is largely due to stratospheric ozone depletion.

The revised Methods section states at lines 195-203:

“Stratospheric ozone depletion and recovery is represented in the Large Ensemble, TPACE and “everything else” (EE) ensembles, but its role cannot be isolated with these experiments, as was done in experiments with CESM1 (Lenaerts et al., 2018; Chemke et al., 2020; Schneider et al., 2020). There is no

CESM2 single-forcing ensemble with stratospheric ozone depletion as the sole evolving forcing. Stratospheric ozone is however the dominant time-evolving forcing in EE during the latter part of the 20th Century, during which a strong accumulation signal appears. While the lack of an ozone-only ensemble is a limitation of this work, the greenhouse gas only (GHG), industrial aerosol-only (AAER), and biomass burning aerosols only (BMB) ensembles enable the novel analyses of accumulation responses to these forcings, which to our knowledge have not been previously presented.”

The revision mentions the likely ozone depletion signal in accumulation at lines 435-437 within the results section discussing Figure 2:

“Although the response in [EE] does not strictly isolate the role of stratospheric ozone depletion, the mass gain in [EE] occurs after 1980 (Fig. 2c), consistent with the timing of ozone depletion and qualitatively consistent with ozone-driven accumulation increases found in previous work (Lenaerts et al., 2018; Schneider et al., 2020).”

Given this additional description, readers should not be surprised when stratospheric ozone depletion is mentioned in the Discussion in Table 3 as a driver of the accumulation trend in [EE]. By extension, it is also a component of the all-forcings forced responses in [CESM2-LE], [CESM2-LEcmip6] and [TPACE]. Ozone depletion is frequently mentioned in the text as a driver of the wind trend.

The TPACE experiment is available in the public domain, and is described in the reference of Rosenbloom et al. (2025) which is cited in the main text and the data table in the “Code and data availability” section. This table also points to the webpage <https://www.cesm.ucar.edu/working-groups/climate/simulations/cesm2-pacific-pacemaker> where the nudging mask is displayed (briefly described in the text at line 186-189) and further details on the experimental design are available, with links to the exact codebase that was used. The TPACE experiment generally follows the nudging protocol discussed in the work of Kosaka and Xie (2013), which is now cited in the main text at line 189. TPACE was recently used by O’Connor et al. (2025) as the prior for another PDA reconstruction.

Appendix A / section 3.4: The only information that are provided about CESM1-AIS-meltwater are in section 3.4 (a bit late compared to the method description), and only consists of “The meltwater experiment was designed to represent the spatial pattern of freshwater fluxes from ice shelf basal melt, with the total magnitude of these fluxes set to 2000 Gt yr⁻¹”. I think that more information should be provided about this:

- Is the freshwater injected at the ocean surface? Along the front of unresolved ice shelf cavities?
- Why only mentioning ice shelf basal melt rates given that more than half of the current Antarctic mass discharge ends up as iceberg melting?
- Please compare the 2000 Gt yr⁻¹ to recent estimates, e.g., those in Coulon et al. (2024, their Tab. 1).
- Is there some kind of correction in CESM ensuring that any SMB anomaly is reinjected into the ocean to conserve the ocean mass (e.g., Schmidt et al., 2025, their Fig. 10)? If so, it is worth mentioning it because any increase in SMB will lead to an increase in freshwater release in addition to the additional 2000 Gt yr⁻¹.

The CESM-AIS meltwater experiment is now described at lines 231-250, addressing the concerns raised above:

Anomalous freshwater fluxes from thinning ice shelves are not included in standard CMIP6 historical experiments nor the CESM2 experiments in Table 1. Recent work argues that meltwater is a climate forcing that should be included in such experiments to enhance their accuracy in simulating historical trends and predicting future change (e.g. Rye et al., 2020; Roach et al., 2023; Schmidt et al., 2023; Sadai et al., 2025). Pauling et al. (2016) and Dong et al. (2022a) describe a meltwater hosing ensemble with CESM1 that spans 1980-2013. We utilize this ensemble (using the ensemble mean of 7 members where total anomalous freshwater flux is set to 2000 Gt yr⁻¹, prescribed at the front of ice shelves at depths of 100 - 200 m) for insight into the possible influence of meltwater on snow accumulation. A caveat is that the magnitude of the freshwater fluxes imposed in this ensemble is much larger than observational estimates (Pauling et al., 2016). Although revised freshwater flux estimates have been subsequently published (e.g. Swart et al., 2023; Coulon et al., 2024), the newer values still do not approach the 2000 Gt yr⁻¹ used in this idealized CESM1 experiment. Advantages of the CESM1 basal meltwater experiment are that it captures the spatial pattern of anomalous meltwater forcing including its depth, and that it simulates significant responses in SSTs (Dong et al., 2022a), sea ice extent (Pauling et al., 2016) and transient climate sensitivity (Dong et al., 2022a). We do not consider the surface freshwater (iceberg) experiments from Pauling et al. (2016) because the available ensemble size is small and as mentioned in the Introduction, ice shelf basal melting is the dominant process driving AIS mass loss in recent decades. All CESM experiments conserve ocean mass such that excess snow accumulating on the AIS is returned to the ocean, analogous to river runoff. The freshwater forcing in CESM1-AIS meltwater is implemented as an additional, negative salinity flux (Pauling et al., 2016).

Additional commentary regarding freshwater flux estimates of Coulon et al. (2024):

The total observed flux for 1990-2020 reported in Table 1 of Coulon et al. (2024) is around 2600 Gt yr⁻¹, and in their calibrated model it is 3140 Gt yr⁻¹. Most of this is balanced by the SMB, so the anomalous freshwater flux from ice sheet mass loss is still much lower (maybe 4x lower) than 2000 Gt yr⁻¹. It seems that Coulon et al. (2024) don't update the anomalous freshwater flux estimate that is used as the prescribed forcing in Pauling et al (2016). Pauling et al. (2016) is only cited once in Coulon et al. (2024) at the beginning of their Discussion. Coulon et al. (2024) mention the strengths of the Pauling et al. (2016) experiments in terms of the vertical and horizontal distribution of anomalous freshwater fluxes in these experiments, which they note is quite important for capturing how these fluxes interact with climate dynamics and feedback processes.

Since it appears that Coulon et al. (2024) don't update the total anomalous freshwater flux estimate that is used as the prescribed forcing in Pauling et al (2016) we are unable to provide an updated estimate of what this total flux "should" be. The 2600 Gt yr⁻¹ in Coulon et al. (2024) might imply a revision upward from Pauling et al (2016)'s observed estimate of 119 to 544 Gt yr⁻¹ if the SMB is around 2000 Gt yr⁻¹. Coulon et al (2024) do not really constrain the SMB, making it hard to constrain the anomalous freshwater flux.

In a review paper, Swart et al. (2023) quoting Slater et al. (2021) write that the present-day total anomalous freshwater flux (considering ice shelf basal melting and iceberg calving combined) is 509 +/- 186 Gt yr⁻¹. This is consistent with the statement in the paragraph excerpted above, "*Although revised freshwater flux estimates have been subsequently published (e.g. Swart et al., 2023; Coulon et al., 2024), the newer values still do not approach the 2000 Gt yr⁻¹ used in this idealized CESM1 experiment.*"

In a possible final round of revision, if the reviewer has a specific number in mind from a recent peer-reviewed study that we should cite, we would be happy to try and include it.

Slater, T., Lawrence, I. R., Ootosaka, I. N., Shepherd, A., Gourmelen, N., Jakob, L., Tepes, P., Gilbert, L., and Nienow, P.: Review article: Earth's ice imbalance, *The Cryosphere*, 15, 233–246, <https://doi.org/10.5194/tc-15-233-2021>, 2021

L. 201: “to form the prior” may not be meaningful for readers who have not read O’Connor et al. (2021), please expend a little.

The revised text briefly summarizes the prior as follows: *“We first use the reconstruction presented in O’Connor et al. (2021) that uses members of the CESM1 Large Ensemble (Kay et al., 2015) to form the prior, which is a physically based assumption about the state of the system without accounting for observational or proxy data.”*

Fig. 3 is not perfectly clear to me. Does each marker represent the mean temperature of a given year in 1901-2000 (X-axis) and the surface accumulation from 1901 to that year (Y-axis in panel a)? In a context of global warming, wouldn’t there be a good correlation between anything that accumulates in time and global temperature? Wouldn’t it be more relevant to plot the global annual mean temperature vs annual mean accumulation (in Gt/yr)?

Fig 3b plots the Antarctic annual mean temperature anomaly vs annual mean accumulation anomaly (relative mass) for the 100 years from 1901-2000. Each marker is a year. This follows the practice of calculating Antarctic accumulation sensitivity to Antarctic temperature in previous work like Monaghan et al. (2008) and Frieler et al. (2015). As such, this is the most relevant plot that readers interested in model simulations of Antarctic accumulation trends would expect to see.

Fig 3a plots the cumulative mass timeseries vs the global mean temperature anomaly timeseries. In general, we agree that two timeseries with a positive trend may exhibit a “good” correlation even if they do not have any real physical relationship. In this case there are sound physical reasons to expect a good correlation. Correlation does not prove causation, but here we present a whole series of physical model experiments to back up this correlation and to connect GHG-driven warming with the accumulation increase.

Throughout the manuscript: numbers could be provided only in sea level equivalent, providing both Gt and mm SLE everywhere does not improve the readability.

To improve readability, for Sections 3 and 4 the revised text only provides SLE. For completeness, Gt and SLE are used in Table 1.

The Introduction mentions the conversion between Gt and SLE (e.g. 2000 Gt is approx. 6 mm SLE).

L. 264: “has a mass gain of 6079 Gt”. Add “from 1900 to 2000”.

Addition made; now at line 408 using 16.8 mm SLE not 6079 Gt.

L. 265-268: I had to read these two sentences several times to try to figure out what was the meaning and the point of these sentences.

Sorry for the confusion. There are two halves of the Large Ensemble that have different biomass burning forcing protocols. The default CMIP6 biomass burning forcing results in spurious warming (Fasullo et al., 2022), which we argue in turn results in a spurious accumulation rate increase and excessive cumulative mass gain.

The revision states, *“The original members of the Large Ensemble with the CMIP6 biomass burning protocol exhibit a slightly higher ensemble-mean value of 17.8 mm SLE, physically consistent with the spurious warming that arises from CMIP6 biomass burning emissions (Fasullo et al. 2022).”*

L. 269-270: It is ok to make such a statement here, but the statistical significance of the difference between two means is not only based on the standard deviations, it is also based on the number of years and/or ensemble members used to calculate the mean (see, e.g., two-sample t-test).

The uncertainty ranges of MT19 and the spread of the CESM2-LE represent fundamentally different quantities. The CESM2-LE spread reflects internal variability sampled by the ensemble, whereas the MT19 uncertainty bounds arise from propagated observational and methodological uncertainties in the reconstruction, rather than from a distribution of independent realizations.

As a result, a formal two-sample statistical test is not directly applicable in this context. Our intention here is therefore limited to a qualitative comparison of the relative magnitudes and overlap between the simulated ensemble spread and the reconstruction uncertainty, and then to provide an interpretation of the differences in their means.

L. 270-272: please help the reader to see where these numbers are from (which panel? which experiment?).

The revision states (line 414), *“Interpreting the [CESM2-LE] value of 16.8 mm SLE as the forced response to all forcings combined, a mass loss of 6.3 mm SLE is required to match the value of 10.5 mm SLE from MT19. This mass loss can be interpreted as driven by internal variability as 10.5 mm lies within the ensemble spread of CESM2-LE, but falls on the extreme end of the ensemble distribution outside of the 5% lower whisker (Fig. 2b). A similar distribution occurs in CESM2-LEcmip6, with 10.5 mm SLE lying on the extreme end (Fig. 2d).”*

L. 327-337: add a sentence at the beginning to explain what you’re doing with a single member. How was member #40 identified as the best member? Was the selection done in an objective way, i.e., based on some metrics?

Yes, the selection was made in an objective way. The expanded first paragraph of Section 3.2 states, *“For spatial analysis, we survey the pattern correlations between the 20th-Century trend in MT19 accumulation and each ensemble member of the Large Ensemble and TPACE, along with the ensemble means of all historical experiments from Table 1 (Figs. B1, B3). Similarly, pattern correlations for the sea level pressure (SLP) trend across 40°-90° S are computed with the same model experiments and using the CESM2-LE PDA as the benchmark (Figs. B2, B4). For illustration, we select the ensemble member whose trend patterns in snow accumulation and atmospheric circulation exhibit the best pattern correlations to the MT19 and CESM2-LE PDA reconstructions, respectively. Specifically, member 40 from CESM2-LEcmip6 has a pattern correlation with the reconstructed SLP trend over 40°-90° S of 0.92 and with the reconstructed accumulation trend on Antarctica of 0.52, giving an average correlation of these two metrics of 0.72. This is the highest average correlation out of 110 individual members from the CESM2-LE, CESM2-LEcmip6, and TPACE ensembles. The following discussion uses member 40 to illustrate spatial patterns but these same relationships are found in several other ensemble members.”*

Section 3.2: Are these results robust if the second-best member is used instead of member #40? Even for a perfect model, it seems impossible to get a perfect match with all observations for a member out of only 50. So how much are the conclusions based on the specific characteristics of member #40?

The results and conclusions do not hinge on the interpretation of a single member.

Results from selecting the second-best member would be qualitatively the same but may exhibit slightly different quantitative values. By our average pattern correlation metric, member #2 is second best. From Figure B3, member #2 also has a snow accumulation trend dipole across the WAIS. From Figure B4, we see that #2 has a deepened ASL, which is basically required in order for the model to simulate the accumulation trend dipole. We see similar relationships in the third-best member, which is member #99 (member 49 of CESM2-LE, shown in Figures B1 and B4). These relationships are pretty robust, even extending to the tropical Pacific SST anomalies that go along with the high-latitude SLP and SMB trend patterns.

We do not claim that member #40 is a “perfect” match, only that it is the “best” match out of the 110 individual members. A minor caveat is that the PDA was performed using the CESM2-LEcmip6 as the prior, so it may not be by chance that the “best” member comes from the CESM2-LEcmip6 half of the ensemble. However, we would have a similar story if CESM2-LE were the prior and member #99 was the “best” match. Figure A4 shows that #040 and #099 exhibit similar wind-SAT relationships, in a similar pattern as the PDA reconstructions but of weaker magnitude.

Alternatively to the pattern correlation method, we could follow the method of Holland et al. (2022): subtract the reconstructed trend patterns from the forced response given by [CESM2-LEcmip6] to obtain the “internal” trend patterns. This would still support the same story that the observation-model trend discrepancies are associated with a deepened ASL and negative IPO-like SST trend pattern across the Pacific basin.

L. 437:439: The sentence “Negative accumulation trends occur in Wilkes Land, where the thinning Totten Glacier is located and in West Antarctica over the thinning Thwaites and Pine Island glaciers” might suggest that the accumulation trend explains the mass loss of these glaciers, but the increase in grounding line discharge is by far the leading driver of their mass loss (e.g., Rignot et al., 2013).

The revised text should avoid any possibility of misleading readers (lines 599-605):

“Significant negative accumulation trends occur in Wilkes Land over the Totten Glacier and in West Antarctica over the Thwaites and Pine Island Glaciers, two regions where independent observational studies show strong evidence for circulation-driven variability in snow accumulation that affects year-to-year variations in the mass balance of these basins (King and Christofferson, 2024). Nonetheless, long-term mass loss in these basins has been predominantly driven by ice-ocean interactions and grounding line discharge, not by surface processes (e.g. Shepherd et al (2019), Lu et al. (2025)).”

Section 3.4: please comment on the realism of the meltwater perturbation in these experiments. If it is over or under estimated, what would a more realistic perturbation induce?

As discussed above, the revised Methods section discusses the caveats of the CESM1-AIS meltwater experiment. We invite further studies with more realistic meltwater experiments, but it is beyond our scope to perform those experiments and/or to analyze their results here. The recent studies of Sadai et al (2025) and Zhang et al (2026) qualitatively support a precipitation decrease over Antarctica in response to meltwater perturbations. This is noted at line 676:

“This result of an accumulation decrease in response to Antarctic meltwater is qualitatively supported by Zhang et al. (2026), who find meltwater-driven precipitation decreases on the Antarctic continent across multiple CMIP6-era models. Additionally, Sadai et al. (2025) find a similar precipitation decrease in a set of CESM1 experiments driven by meltwater fluxes from an offline ice sheet model. Observational uncertainties and model limitations make a highly realistic meltwater simulation very challenging, but there is a multi-model consensus that meltwater leads to Southern Ocean surface cooling and a precipitation decrease over Antarctica (e.g. Zhang et al., 2026).”

From all the work on the meltwater signal that we have seen, there is consensus that meltwater leads to a Southern Ocean cooling, leading to Antarctic precipitation decrease. The uncertainties lie in the magnitude of the forcing, the sensitivity of the model to the forcing, and the timing of when the anomalous meltwater forcing became relevant. At line 804, we note that a more realistic meltwater experiment covering the pre-satellite era would have relatively small anomalous freshwater fluxes concentrated in the Amundsen Sea Embayment region.

Note that the SOFIA experiments analyzed in Zhang et al (2026) aren’t necessarily “more realistic” than the old CESM1-AIS meltwater experiment. Zhang et al (2026) report an observed estimate of 0.1 Sverdup yr-1 vs 1 Sverdup yr-1 used in the SOFIA experiment protocol. We simply do not know yet what a more realistic perturbation in a more realistic model would induce. We speculate in the limitations paragraph within the Discussion that high-resolution models may be more sensitive to meltwater perturbations than low-resolution models (line 937).

Regardless of the exact quantitative role of meltwater in driving SST trends, it is clear from the results of the CESM2-GOGA experiments that observed SST and sea ice trends do truly matter for Antarctic accumulation trends.

Title of section 3.4: why “hidden”? Is it more hidden than any other of the investigated effects?

“Hidden” was used because the role of meltwater is much more subtle and difficult to tease out of the data than say, the role of greenhouse gases or atmospheric circulation. The section is now titled “Potential roles for meltwater and two-way teleconnections” to emphasize the uncertainty in the role of meltwater.

L. 531-536: This is very unclear. First, it needs to be understandable without having to read Dong et al. (2022a). Then, in which figure, experiment or analysis is the trend multiplied by five? And why five and how is this quantitatively linked to the evolution of Pine Island and Thwaites ice shelves in the mid 20th century? From the caption of Fig. 9b, I understand that a trend over a decade is extended to 50 years, is it just that a trend estimated over 10 years is applied to 50 years or is the slope of the trend multiplied by five? Please expand and reformulate.

Dong et al. (2022a) express SST trends (rates of change) attributable to meltwater in units of deg. C 10 yr-1, using the CESM1-AIS meltwater ensemble in comparison with the CESM1 Large Ensemble. They provided the meltwater-induced SST field to us, as shown in Figure 9b. But instead of showing deg. C 10 yr-1 we scale it to deg C 50 yr-1, e.g. the regression coefficients (slope) at each grid box are multiplied by 5. We do this because the period of 1951-2000 is 50 years long. From the observational evidence in Clark et al. (2024) it is argued that anomalous freshwater flux from the Amundsen Sea Embayment region has

occurred since around 1950. This is a back-of-the-envelope type calculation with several caveats, and we are using cautious language to discuss it. Yet it is a plausible story that ties together all the recent observational evidence and modeling work.

The revised description of Fig. 9b states (starting from line 706), *“The meltwater-induced SST anomaly pattern in Figure 9b was originally estimated by Dong et al. (2022a) as a decadal trend; here we multiply their decadal trend by five to account for 50 years during which Antarctic meltwater has plausibly been a relevant forcing. Ice shelf thinning and grounding line retreat in West Antarctica commenced in the mid-20th Century, several decades prior to the modern satellite era (Smith et al., 2017; Clark et al., 2024). Since the mid-20th Century, an anomalous contribution of freshwater to the ocean has been sustained by ice shelf thinning and glacier retreat driven by northerly wind anomalies in the Amundsen Sea Embayment associated with the ASL deepening trend (O’Connor et al., 2025).”*

L. 601-602: The internal variability in the CESM model may be weaker than the natural variability of the real world, as suggested by Casado et al. (2023) for a bunch of CMIP6 models. How would this affect the conclusions of this article?

The possibility that internal variability may not be accurately represented in CESM2 is indeed a motivation for evaluating this broad suite of experiments, which introduce some observational constraints. This idea is stated in Section 2.5: *“Using large ensembles is only the first step towards a comprehensive accounting of internal variability. This step may not be sufficient because a) models like CESM2 may simulate an unrealistic magnitude of internal variability (e.g. Casado et al., 2023); and b) internal variability in free-running coupled experiments like the Large Ensemble is not synched with real-world internal variability.”*

That said, we cannot find within Casado et al. (2023) a specific quantification of how much CESM2’s internal variability differs from the real world. Casado et al. (2023) look at a 6-member CESM2 historical ensemble from the CMIP6 archive, not the 100-member CESM2 Large Ensemble. Note that the Large Ensemble is both a “macro” ensemble (in which the ocean initial conditions are changed across members) and a “micro” ensemble (in which just the atmospheric initial conditions are changed across members), as such it samples a broader range of possible evolutions of internal variability than does the much smaller CESM2 ensemble submitted to CMIP6.

We have a multi-pronged approach for accounting for possible errors in the representation of the model’s internal variability: The TPACE experiment prescribes actual observed SST anomalies, constraining the model to have the observed magnitude of SST anomalies and sequence of internal variability at least in the tropical Pacific. We have quantified the effect of these SST anomalies as a mass loss of 6.4 mm SLE over 1901-2000. In other words, real world internal variability results in a mass loss, not a mass gain that could explain the observed trend of 10.5 mm SLE. Anthropogenic forcing, mainly greenhouse gases, is required to explain the observed mass gain. Second, the PDA reconstructions also contain real world internal variability within the context of the dynamical framework of CESM2. Our various regressions using the reconstructions vs the free-running model data show that the real world may have a little bit different wind-SST-accumulation relationships than does the model world. This makes sense in the context of Casado et al (2023) who discuss a mismatch between model fidelity in getting global-mean trends broadly correct and model fidelity in getting the polar regions correct.

As diagnosed in O’Connor et al (2021), the PDA reconstructions show that anthropogenic forcing is required to explain mid and high latitude atmospheric circulation trends. All in all, we do not believe errors in representing internal variability could significantly change our conclusion that anthropogenic forcing has played a strong role in recent Antarctic climate trends. This conclusion is broadly consistent with Casado et al. (2023) who suggest that internal variability *and the forced response* are underestimated by models.

Our study hints at a related concern – that the forced atmospheric circulation trends (shown in Figures 4c and 7c) are underestimated by the model. If interested, see Klavens et al. (2025) for a discussion of this signal-to-noise problem in atmospheric circulation trends over the North Pacific:

Klavens, J. M., DiNezio, P. N., Clement, A. C., Deser, C., Shanahan, T. M., and Cane, M. A.: Human emissions drive recent trends in North Pacific climate variations, *Nature*, 644, 684–692, <https://doi.org/10.1038/s41586-025-09368-2>, 2025.

Klavens et al. (2025) is cited at line 777 in the discussion of internal versus external signals in the negative IPO pattern shown in Figure 10.

High-resolution global modeling may be one path towards solving this signal-to-noise problem (e.g. Yeager et al., 2023). We hope to do some future work with high resolution models, as alluded to in Section 4.3.

L. 657-660: I am sorry but I would need further explanations to understand this: “As diagnosed by Simpson et al. (2023), positive ice-albedo feedbacks in [AAER] act on a cold pre-industrial background state, leading to strong cooling. These feedbacks are not as pronounced in [CESM2-LE] in which aerosols are introduced only after the climate has warmed somewhat due to increasing greenhouse gases. The strong cooling in [AAER] is artificial ...”.

It also took us some time to grasp what the difference in results from AAER + GHG vs all-forcings CESM2-LE are telling us. Strong, nonlinear ice-albedo feedbacks in the AAER experiment lead to strong polar and global cooling. Simpson et al. (2023) suspect that these feedbacks are particularly amplified in AAER because the pre-Industrial climate state on which they act is rather cold. In the all-forcings CESM2-LE, aerosols are competing with greenhouse gas increases, whose effects tend to win out over time. We suggest that the strong cooling in AAER is less representative of the real world than the all-forcings CESM2-LE. When used in combination with the GHG ensemble, the strong cooling in AAER could compensate for the cooling hypothesized to be driven by missing meltwater forcing.

For brevity, we have shortened the discussion of cooling in AAER and moved it to the end of section 3.5.

Lines 799-803:

“Both the aerosol-only and meltwater experiments have limited utility as explanations for the 20th-Century history. The strong cooling in [AAER] may be partially spurious because of the experimental design in which aerosol forcing acts alone on a cold pre-industrial background state, triggering ice-albedo feedbacks to amplify the cooling more than it would be when greenhouse gases increase at the same time (Simpson et al., 2023).”

L. 683: correct “20250”.

Corrected.

Section 4.4: what about the effect of stratospheric ozone recovery in the projections to 2050?

The revision discuss the possible role of stratospheric ozone recovery at line 890:

“Although these maps (and the all-forcings CESM2-LE, not shown) indicate that the westerlies will continue to strengthen through the mid- 21st Century due to continued greenhouse gas increases, this could be a model-dependent result stemming from CESM2’s high climate sensitivity and a scenario-dependent result associated with the high-emissions SSP3.7 scenario. Under a lower emissions scenario

and/or in a model with lower climate sensitivity we might expect stratospheric ozone recovery to lead to a weakening of the westerlies by the mid- 21st Century (Arblaster et al., 2011). However, on the annual-mean basis the positive SAM trend is projected to continue across many models in intermediate to high-emissions scenarios (Purich et al., 2025)."

L. 692-697: Specify that surface melting and runoff only become important beyond 2050 (e.g., Kittel et al., 2021; Jourdain et al., 2025).

The revision cites these studies but emphasizes caution in relying on model results alone to anticipate future trends, also citing Mottram et al. (2025): (around. line 885):

"While the uptick in intense atmospheric river events, recent sea ice loss and warm Southern Ocean conditions could indicate that increased snow accumulation will provide significant sea level mitigation, strong warming will mean more mass lost to surface melting and runoff (Gilbert and Kittel, 2021) especially after 2050 (e.g. Kittel et al., 2021; Jourdain et al., 2025). As warming continues, some precipitation will fall as rain (Vignon et al., 2021). The observed uptick in atmospheric river events that bring intense heat and moisture to the ice sheet (e.g. Wille et al., 2025) is a worrying trend that may mean that rainfall and mass loss processes become important sooner than model simulations would suggest (Mottram et al., 2025)."

L. 759-760: How does "the snow accumulation history" show that there is causality?

As summarized in Table 3 and expanded upon throughout the text, the snow accumulation history is interpreted as the *"response to greenhouse-gas driven warming, offset by aerosol-driven cooling and SST trends (internal and anthropogenic) associated with winds and meltwater."* A pressing question in the climate modeling community has been, "What causes the SST trends that have damped global warming (e.g. Andrews et al., 2022), and by extension, what has damped Antarctic snow accumulation trends?" Snow accumulation provides a proxy archive of the effects of winds, GHGs and meltwater on SSTs. Given the relationships illustrated here, the same drivers that have caused tropical Pacific and Southern Ocean SST trends must also explain Antarctic snow accumulation trends. By interpreting (or "decoding") the snow accumulation history we also provide an interpretation of the broader SST trends.

As far as the influence of the Southern Ocean on the tropics, the revised text removes the word "causality." The middle paragraph of the revised Conclusion section emphasizes that our results support a rapidly growing area of research on Southern Ocean - tropical connections: *"In contrast to previous attributions of Antarctic trends that have emphasized that what happens in Antarctica depends on the evolution of the tropical Pacific (e.g. Holland et al., 2022), this work adds to the growing body of literature emphasizing a significant influence of Antarctica and the Southern Ocean on the tropics (e.g. Hartmann, 2022; Dong et al., 2022b, 2025, 2026; Kim et al., 2022; Kang et al., 2023, 2026; Yeager et al., 2023; Zhang et al., 2026). One key piece of evidence we add is that spatial trends in Antarctic accumulation and atmospheric circulation over the 20th Century cannot be explained by tropically driven teleconnections – the nudged TPACE experiment worsens the replication of these trends compared with the unconstrained Large Ensemble. From the Large Ensemble and single-forcing large ensemble we infer that the high-latitude circulation trends are explained by a combination of greenhouse gas increases, stratospheric ozone depletion, and internal variability. The "internal" component does not fit the response to the observed tropical SST trend. Rather, we have discussed evidence that the reverse is true – the observed tropical SST trend pattern is consistent with the influence of the Southern Ocean on the tropics. Although the nudged TPACE experiment does not help to explain the spatial pattern of accumulation or atmospheric circulation trends, it dampens the Antarctic-wide accumulation rate,*

suggesting that cooling signals from the high latitudes have reached the tropics, consistent with recent results from high-resolution models that have relatively strong shortwave cloud feedbacks and explicit treatment of heat transport by ocean eddies (Kang et al., 2026)."

L. 787-788: "The CESM2's ensemble-mean version of accumulation increase would have been a welcome addition to the ice sheet's mass". What do the authors mean here?

We meant that if there had been an accumulation-related mass gain of 16.8 mm SLE during the 20th Century (as simulated by [CESM2-LE]) instead of the observed 10.5 mm SLE, the ice sheet would be close to in balance if not gaining mass. The usage of the model is largely to test sensitivity (it is not a perfect predictor of reality); the objective of this paper is to bring these worlds a little closer together.

This specific sentence has been deleted for the revision. The reformulated final paragraph gently reminds users of CESM model output that taking the model output at face value without evaluating how it compares to the observed world may lead to overconfident or otherwise erroneous predictions of snow accumulation trends.

Additional references

Casado, M., Hébert, R., Faranda, D. and Landais, A. (2023). The quandary of detecting the signature of climate change in Antarctica. *Nature Climate Change*, 13(10), 1082-1088.

Coulon, V., De Rydt, J., Gregov, T., Qin, Q. and Pattyn, F. (2024). Future freshwater fluxes from the Antarctic ice sheet. *Geophysical Research Letters*, 51(23), e2024GL111250.

Jourdain, N. C., Amory, C., Kittel, C. and Durand, G. (2025). Changes in Antarctic surface conditions and potential for ice shelf hydrofracturing from 1850 to 2200. *The Cryosphere*, 19(4), 1641-1674.

Kittel, C., Amory, C., Agosta, C., Jourdain, N. C., Hofer, S., Delhasse, A. and others (2021). Diverging future surface mass balance between the Antarctic ice shelves and grounded ice sheet. *The Cryosphere*, 15(3), 1215-1236.

Rignot, E., Jacobs, S., Mouginot, J. and Scheuchl, B. (2013). Ice-shelf melting around Antarctica. *Science*, 341(6143), 266-270.

Citation: <https://doi.org/10.5194/egusphere-2025-3666-RC2>

Summary of Major relevant changes:

Abstract has minor wording changes to a) Correct language that Reviewer 2 found to be awkward; b) State that the wind history has been driven by internal variability and anthropogenic forcing.

Introduction includes a new paragraph starting at line 89 to better summarize previous studies and large-scale drivers of spatial patterns in snow accumulation following the request of Reviewer 1.

Last paragraph the Introduction is substantially revised to more clearly state the objectives of the study.

At lines 185-204 the TPACE and EE ensembles are more clearly described following the request of Reviewer 2.

Starting at line 214, the descriptions of CESM2-WNUDGE, CESM1-AIS meltwater, CESM2-GOGA, and CESM2-TOGA have been moved from Appendix A to the Methods Section 2. The description of CESM1-AIS meltwater in particular has been expanded following requests of Reviewer 2.

Section 2.3 describes how snow accumulation is calculated; this has been moved from Appendix A.

Section 2.5 is a new description of our strategy for assessing forced responses vs internal variability, following advice from both reviewers.

The beginning of Section 3 has a roadmap to the section following the request of Reviewer 2 to make the manuscript easier to follow.

Throughout Section 3.1 only SLE is used for the cumulative mass gain results to make it easier to read. The numbers are more clearly sourced.

Figures 4,6,7,8,10 indicate statistical significance of the accumulation trends following the request of Reviewer 1.

In Section 3.4, additional supporting evidence for the meltwater signal in accumulation is added (Sadai et al., 2025; Zhang et al., 2026).

There is a brand new “Narrative and Graphical summary” of the major results in Section 3.5, accompanied by the new schematic diagram, Figure 10, following requests of Reviewer 1. The very recent studies of O’Connor et al. (2025) and Kang et al (2026) are included in this discussion.

The Discussion section has been streamlined into three subsections - 1) a succinct overview and interpretation of the results clarifying the roles of external forcing vs internal variability; 2) an outlook towards the future with additional citations following the advice of reviewers; 3) Limitations and future work.

The Conclusion has been substantially rewritten into 3 paragraphs: 1) A concise summary of results with clear attributions; 2) The context of our results emphasizing Southern Ocean - to - tropics teleconnections that the accumulation trends support but are not well represented in current-generation models; 3) Statement on how projections might be improved and caution against using existing projections at face value.

There are only 3 parts of the Appendix now. Figure A4 added to illustrate that a) Choice of CESM1-LE PDA or CESM2-LE PDA does not substantially affect the spatial trend patterns associated with winds ; b)

Regardless of the choice of ensemble member that most closely matches the PDA reconstructions, the wind-SAT relationships in the free-running model are weaker than in the PDA reconstructions.

Gemma O'Connor and Zaria Cast have been removed from the author list for reasons stated in confidence to the editorial team.