Persistent Antarctic Sea Ice Biases in CMIP6 Models in spite of the Recent Decadelong Sharp Decline

Lettie A. Roach[a] and Lorenzo M. Polvani[b,c]

[a] Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany
[b] Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York
[c] Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York

(Manuscript received 13 August 2025, in final form 16 December 2025, accepted 20 January 2026)

ABSTRACT: Antarctic sea ice expanded for the first 35 years of the satellite record but in the past 10 years has dropped to record lows. Recent studies argue that the recent decline brings climate models into better agreement with observed Antarctic sea ice trends. Here, we assess this claim by considering Antarctic sea ice area and concentration across a large number of Coupled Model Intercomparison Project phase 6 (CMIP6) simulations, and we reach a different conclusion. Even with many models that sample a large range of internal variability, we find that very few capture the observed multidecadal expansion or the large decline after 2014. Further analysis reveals that models with trends similar to observations typically have poor simulation of climatological sea ice, its variability, and its regional patterns. The continued inability of CMIP-class models to simulate Antarctic sea ice and its observed trends severely limits our ability to understand past changes and to confidently project its evolution in the coming decades.

KEYWORDS: Antarctica; Sea ice; Climate models; Climate variability; Climate change; Model evaluation/performance

1. Introduction

Satellite observations, now spanning 45 complete years from 1979 through 2024, have revealed surprising changes in Antarctic sea ice area. From the beginning of the satellite era until around 2014, Antarctic sea ice steadily expanded at a rate of 2% decade^{-1} (Fig. 1). Such an expansion was rather unexpected, given the concurrent increase in global temperature and the dramatic melting of sea ice in the Arctic. Several papers argued that internal variability alone might explain the multidecadal expansion (Polvani and Smith 2013; Zunz et al. 2013; Singh et al. 2019). Some studies have suggested that the absence of ocean warming (and sea ice decline) in observations may be linked to the background ocean circulation (Marshall et al. 2014; Armour et al. 2016) and possibly eddy-mediated heat transport (Rackow et al. 2022) that have delayed the emergence of forced Southern Ocean warming. Others suggested that sea ice expansion was caused by meltwater from the Antarctic ice sheet (Swart et al. 2023; Schmidt et al. 2023) and possibly by changes in winds (Holland and Kwok 2012), including trends in the Amundsen Sea low (Turner et al. 2016) and Southern Annular Mode largely driven by ozone depletion (Ferreira et al. 2015). However, the relative influence of these processes is still debated (Roach et al. 2023). We note, in particular, that sea ice expansion driven by ozone depletion via a more positive annular mode appears, at this point, quite unlikely (Bitz and Polvani 2012; Landrum et al. 2017; Seviour et al. 2019; Polvani et al. 2021).

We have now observed the start of another surprising phase. After a record maximum in 2014, and another relatively high year in 2015, Antarctic sea ice has subsequently been dropping rapidly with successive record lows in 2016, 2022, and 2023 and with the most recent year, 2024, also remaining low. Hence, an entire decade of anomalously low and rapidly decreasing sea ice area has now been observed (Fig. 1), possibly accompanied by an increase in year-to-year variability (Hobbs et al. 2024). Recent work has suggested that unusual atmospheric conditions (Turner et al. 2017; Wilson et al. 2023; Purich and England 2019) and subsurface ocean warming (Purich and Doddridge 2023) may be responsible for these sea ice lows. However, as for the earlier expansion phase, the relative contributions of internal variability and anthropogenic forcing in the decline phase, as well as the driving processes responsible for sea ice loss, remain unclear.

Coupled climate models are essential tools for understanding coupled interactions, projecting future change, and attribution. A number of previous studies have highlighted model–observation discrepancies in the simulation of Antarctic sea ice (e.g., Turner et al. 2013; Roach et al. 2020; Shu et al. 2020), summarized by Fox-Kemper et al. (2021), spanning several phases of the Coupled Model Intercomparison Project (CMIP), including the most recent phase, CMIP6. The record lows in Antarctic sea ice in the most recent decade have renewed interest in the skill of coupled climate models. Diamond et al. (2024), focusing solely on the exceptional 2023 sea ice anomaly, found such an event to be extremely unlikely within current coupled climate models' simulations. However, taking a broader perspective, two recent studies have reached more optimistic conclusions regarding the abilities of CMIP6 models to represent trends in Antarctic sea ice.

First, Holmes et al. (2024) evaluated Antarctic sea ice area trends between 1979 and 2023 in CMIP6 historical simulations. They found that while models "exhibit different skills for

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-25-0477.s1.

Corresponding author: Lettie A. Roach, lettie.roach@awi.de

DOI: 10.1175/JCLI-D-25-0477.1

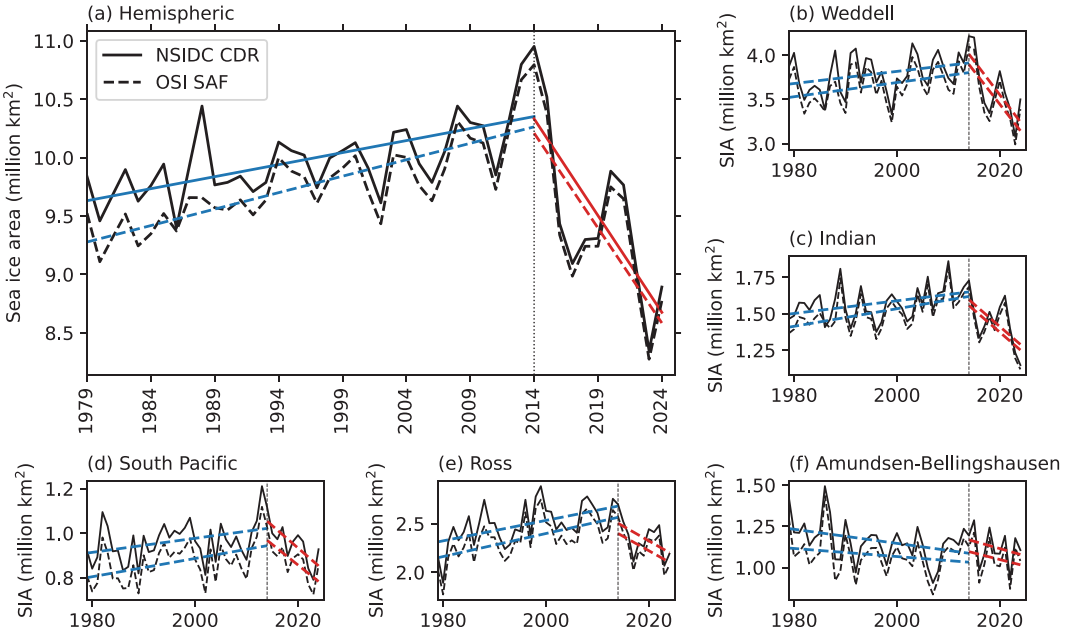

FIG. 1. Annual-mean SIA in the NSIDC CDR (solid black) and OSI SAF (dashed black) observational products. Colored lines show linear fits for different time periods: the expansion period (1979–2014, blue) and the decline period (2014–24, red). (a) Total SIA; (b)–(f) regional SIA for the five NASA regions.

different time scales and periods," the rapid decline in the most recent decade has changed the picture of model–observation agreement. In fact, they conclude that the recent rapid declines "bring observed sea ice area trends back into line with the models."

Second, Liu (2025) identified a small subset of coupled climate models that simulate Antarctic sea ice expansion during 1979–2014 and suggested that natural variability in sea ice enables those models to simulate expansion under historical anthropogenic forcings. Unfortunately, that study specifically excluded the rapid decline period after 2014. By examining average climate conditions in the subset of models that simulate sea ice expansion, that paper concluded that natural variability "reconciles the models with observations."

In this paper, we offer a different perspective. Given that the trends have reversed in a major way (Fig. 1), suggesting a potential change in the underlying drivers, we here examine the sea ice expansion and decline phases separately. We ask how well CMIP6 models capture the observed expansion and decline phases. As we will demonstrate, very few simulations capture the observed sea ice expansion phase and, for those, large biases in mean state and/or variability are evident, both in basinwide and regional perspectives. Similarly, few models simulate the observed sea ice decline phase and, again, those are affected by similarly poor representation of the sea ice mean state, its variability, and its regional trends. These shortcomings suggest to us that, unfortunately, models are still far from reliably simulating the observed Antarctic sea ice evolution of the last 45 years.

2. Methods

We here focus on sea ice area (SIA), calculated by multiplying sea ice concentration with individual gridcell areas and then summing over the Southern Hemisphere. SIA provides a more consistent picture across different grid resolutions as compared with sea ice extent, the area of all grid cells with more than 15% sea ice concentration (Notz 2014). To account for uncertainty in satellite retrieval algorithms (Ivanova et al. 2014), we employ two independent observational records: 1) the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice (OSI) Satellite Application Facility (SAF) version 2.2 SIA product (EUMETSAT 2023) and 2) SIA computed from the National Snow and Ice Data Center (NSIDC) Climate Data Record (CDR) version 4 sea ice concentration product (Meier et al. 2021).

For model output, we analyze the CMIP6 monthly SIA available from the University of Hamburg (UHH) CMIP6 SIA repository. Although this does not completely reflect all the available simulations from CMIP6 on ESGF, it is the same repository considered by Holmes et al. (2024), allowing for direct comparison of the studies. It is freely available and already processed to a user-friendly format, allowing our analysis to be easily reproduced. As shown below, our results are consistent even when taking subsets of this repository. We consider the CMIP6 historical experiments, which span 1850–2014, and the CMIP6 shared socioeconomic pathway (SSP) 2-4.5 scenario experiments, which span 2015–2100. A full list of models and ensemble members included here is provided in Table S1 in the online supplemental material.

3. Results

Let us start by examining the time evolution of annual-mean SIA in observations (Fig. 1). Observational values from

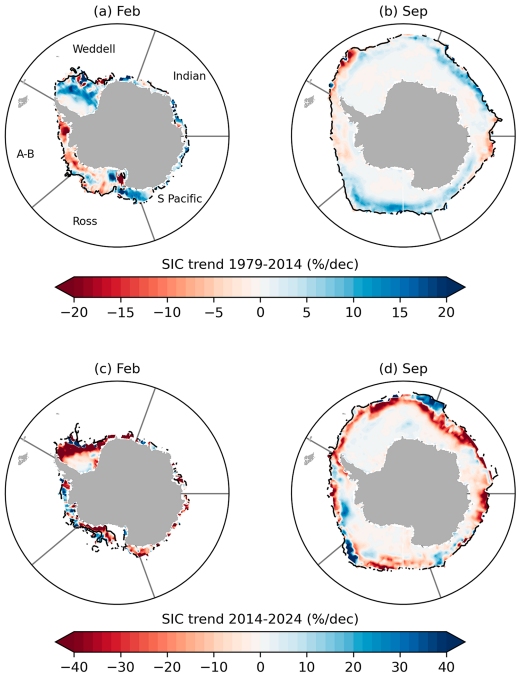

FIG. 2. NSIDC CDR sea ice concentration trends for (a) February 1979–2014, (b) September 1979–2014, (c) February 2014–24, and (d) September 2014–24. The black line shows the 15% sea ice concentration contour for each period's climatology. The sectors used in Fig. 1 are marked with gray lines. Note the doubled color scale between the top and bottom panels.

the NSIDC CDR and OSI SAF satellite products are shown as solid and dashed black lines, respectively. Clearly, the entire observational record from 1979 to 2024 cannot be fitted with a straight line (a simple linear regression yields $R^2 = 0.1$). In fact, no period beginning in 1979 and extending beyond 2021 is captured by a linear regression. Consequently, we avoid using a single long-term linear fit over the whole observational period to quantify agreement between models and observations.

Instead, as suggested by the observations themselves, we choose to divide the sea ice time series into two distinct phases. While formal methods such as changepoint analysis (Purich and Doddridge 2023) or higher-order statistical models could be applied, visual inspection alone reveals a clear transition. First, from 1979 to 2014, we see a clear and steady sea ice expansion: the 1979–2014 SIA trend is 0.21 ± 0.04 million km^2 decade^{-1} in NSIDC CDR and 0.28 ± 0.04 million km^2 decade^{-1} in OSI SAF, where the uncertainty estimate is the standard error in the linear trend over that period. Second, from 2014 through 2024, the trend reverses sharply: the 2014–24 SIA trend is −1.70 ± 0.5 in NSIDC CDR and −1.60 ± 0.48 in OSI SAF observations.

Next, let us consider the regional time series (Figs. 1b–f) for the five standard NASA Antarctic regions: the Weddell, Indian, South Pacific, Ross, and Amundsen–Bellinghausen (A-B) sectors (marked in Fig. 2 and subsequent figures). In the regional perspective, interannual variability is more pronounced. Despite this, regional sea ice trends over 1979–2014 remain reasonably well described by linear regression, with sea ice expansion in four of the five sectors, only the Amundsen–Bellingshausen Sea declining. For 2014–24, all regions show a negative SIA trend as well as substantial interannual variability, similar to total SIA. The largest declines are in the Weddell Sea and South Pacific.

In this study, we will largely focus on integrated Southern Hemisphere sea ice and annual-mean trends, as Antarctic SIA trends are rather similar in all months in observations and in most models (Fig. S1). However, it is also valuable to examine spatial patterns in specific seasons, and here, we highlight September and February, which represent the seasonal maximum and minimum in SIA, respectively. These individual months do not explain the entire seasonal cycle but offer a quick look at spatial patterns and how they vary by season. Maps of sea ice concentration trends for NSIDC CDR in the highlighted months only, for NSIDC in all months, and for OSI SAF in all months are shown in Fig. 2, Fig. S2, and Fig. S3, respectively. As previously explained, we separate the expansion (1979–2014) and decline (2014–24) phases, as a single straight line is not appropriate over the entire 1979–2024 period. It is reassuring that both observational products exhibit very similar spatial patterns. For the 1979–2014 period, sea ice concentration trends in February are mostly positive, strongly so in the Weddell sector and between the Ross and South Pacific sectors, and sea ice loss is limited to the Amundsen–Bellinghausen Sea and parts of the Ross Sea. In September, however, sea ice expansion is seen in nearly all sectors and is mostly found at the sea ice edge, with only small areas of negative trends. For the more recent 2014–24 sea ice loss period, the concentration trends are much larger in magnitude (note the doubled color scale), and the spatial patterns are similar to the expansion period but with opposite sign.

a. Sea ice expansion phase

With the observed expansion from 1979 to 2014 clearly established across multiple datasets and regions, we now assess how well CMIP6 models reproduce this 35-yr-long trend. In Fig. 3, we show linear trends over 1979–2014 for all ensemble members in the UHH repository (markers) and from observations (solid and dashed black lines). Note that many models share components and, therefore, cannot be considered wholly independent (Merrifield et al. 2023). As greenhouse gas, ozone, and aerosol forcing vary in time, and their interplay is understood to be particularly important for Southern Hemisphere circulation (Banerjee et al. 2020), we compare models and observations over the same time period. The sea ice expansion phase 1979–2014 falls within the historical experiment, for which we analyze 462 ensemble members in total from 56 models. This comprises a much larger set of models and simulations than those analyzed in the early report on CMIP6 Antarctic sea ice evaluation by Roach et al. (2020). Unlike Holmes et al. (2024), we retain all ensemble members available for each model to evaluate the range of internal variability, and we examine the sea ice expansion and decline phases separately to more systematically assess model performance.

The number of ensemble members provided to CMIP6 varies considerably between models, from 1 to 73. The small ensemble size for many models makes it challenging to assess consistency

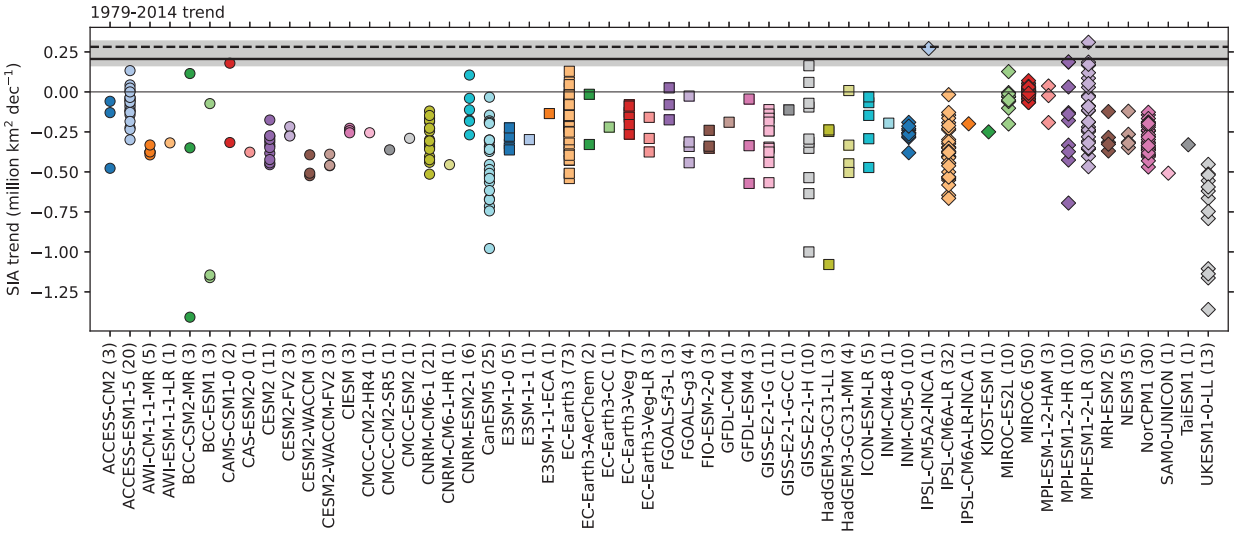

FIG. 3. Annual-mean Antarctic SIA trends over 1979–2014 in historical simulations and observations. Observed trends are shown as horizontal black lines (solid for NSIDC CDR and dashed for OSI SAF), with gray shading representing the widest range from the two observational products' standard error. Colored markers denote individual simulations from CMIP6, grouped by model along the x axis, with the number of ensemble members per model indicated in parentheses. We use colors and three markers (circle, square, diamond) to demarcate individual models.

with observations, and so we compute internal variability in two ways. First, simply considering the range across all available ensemble members for each model, we see that natural internal variability causes each individual model to simulate a range of trends among its ensemble members, even though each member is identically forced. For Antarctic SIA trends, models simulate highly varying amounts of internal variability. For example, contrast the large spread across the GISS-E2-1-H ensemble (10 members) and the smaller spread across the IPSL-CM6A-LR ensemble (32 members).

To test how ensemble size impacts differences in internal variability, we perform a bootstrap-based convergence analysis. For each model with at least five members, we first compute the standard deviation of the sea ice trend using all available ensemble members. We then draw repeated random bootstrap samples of increasing size (from two up to the full ensemble) and calculate the standard deviation for each such synthetic sample. Then, for each sample size, we compute the relative error compared to the original standard deviation with no replacements. We find that an ensemble size of seven to eight ensemble members is sufficient to estimate the spread to within 10% accuracy. Beyond this threshold, increasing ensemble size does not significantly reduce the intermodel spread in internal variability, which instead reflects differences in model physics. Comparing models using synthetic eight-member bootstrapping ensembles (Fig. S4), we find that UKESM1-0-LL and GISS-ES-1-H have the largest internal variability at around 0.3 million km^2 decade^{-1}, while on the low end, MIROC6 and INM-CM5-0 have less than 0.05 million km^2 decade^{-1}. This large intermodel spread, which stems from different model physics, complicates model evaluation and interpretation of trends.

We define as "consistent with observations" any individual simulation that falls within the range of the observed trend plus or minus its standard error, taking the widest range from the two different observational products. As shown in Fig. 3, few models have any simulations consistent with observations during the observed sea ice expansion. Out of 462 total ensemble members, only 8 ensemble members from 4 models are consistent with observations. These are one ensemble member from CAMS-CSM1-0, one ensemble member from IPSL-CM5A2-INCA, two ensemble members from MPI-ESM1-2-HR, and four ensemble members from MPI-ESM1-2-LR. This represents less than 2% of the simulations we consider here.

Of course, it is easy for models to fail this consistency test if only a few ensemble members are available. Small ensemble size is also difficult to account for, given the large intermodel spread in internal variability estimates. To test robustness of our conclusions, we compare with internal variability estimated using the bootstrap approach. We first compute the multimodel mean of the bootstrap-based estimates of internal variability (σ = 0.14 million km^2 decade^{-1}). For each model, we then add this estimate to its ensemble mean (or its single value, in the case of n = 1 ensemble members). Using one standard deviation as the measure of internal variability (for the sake of simplicity), we find that no additional models are consistent with observations. Using two standard deviations suggests that two additional models might be consistent with observations (FGOALS-f3-L, MPI-ESM-1-2-HAM)—a relatively minor increase.

In section 3b, we will consider the SSP2-4.5 experiments to assess the decline phase, for which fewer models and ensemble members are available compared to the historical experiments. To ensure that the comparison between the two phases is fair, we here subset the historical runs to include only models that have SSP2-4.5 simulations in the UHH repository (218 ensemble members from 36 models). In doing

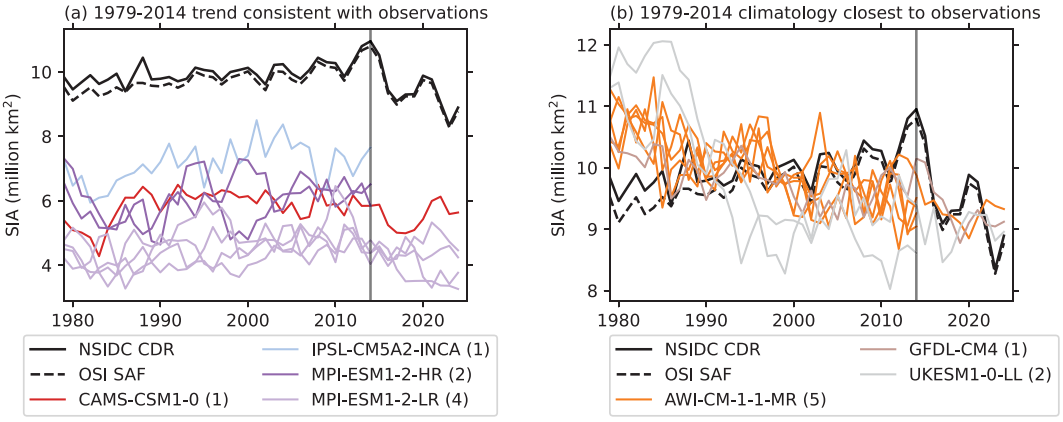

FIG. 4. (a) The eight historical simulations whose 1979–2014 annual-mean SIA trend is consistent with observations. (b) The eight historical simulations whose 1979–2014 time-mean SIA is closest to observations, based on RMSE. Note that we also show the SSP2-4.5 extensions were available to span the full observational period, with a vertical gray line marking 2014. In the legend, the numbers in parentheses indicate the number of ensemble members included here. Both panels include the observed SIA time series from NSIDC CDR and OSI SAF.

so, we find that only five ensemble members from two models fall within the observed trend range over the expansion phase. This represents around 2% of the subsetted simulations, approximately the same fraction as from the full dataset, indicating that the result is not strongly sensitive to model selection. Thus, we can confidently say that, irrespective of model selection, only a very small proportion of simulations capture the observed sea ice expansion under historical forcing.

We now briefly place the model–observation trend comparison in the context of other biases in models' simulation of Antarctic sea ice, notably the simulation of sea ice climatology and interannual variability. The observed annual-mean SIA over the expansion phase is approximately 10 million km^2, yet 10 models simulate less than half that value (Fig. S5a). Some models show negative climatological SIA biases for the winter maximum in September, and many models show negative climatological SIA biases for the summer minimum in February (Figs. S5b,c). In February, 33 models have SIA less than half the observed value, and six models have SIA less than 0.1 million km^2. A difference of 5 million km^2 in SIA corresponds to a difference of 10° in the zonal-mean ice edge latitude (Eisenman et al. 2011), so these biases are substantial.

We also assess interannual variability, noting that the observational record may not be long enough to capture multidecadal variability. We compute interannual variability for each model, and for the observations, as the standard deviation of the time series over the expansion period only, after removing the linear trend (regardless of its significance) (Fig. S5d). Some models (e.g., MIROC-E2SL and MIROC6) have lower variability than observed. Others (e.g., BCC-CCSM2-MR, MPI-ESM1-2-HR, and especially GISS-E2-1-H) have higher interannual variability than observed, falling well outside what could be considered a plausible range.

With these clear biases in mind, we now ask to what extent do the eight individual simulations whose trends are consistent with observations over the expansion period resemble observations in other aspects? To answer this, Fig. 4a shows the eight modeled time series of SIA over 1979–2024 in color, together with the observed ones in black. For five of these eight simulations, the experiments are extended in the figure through 2024 for reference, using the SSP2-4.5 scenario to show the full observational period. Although the magnitude of sea ice expansion in these simulations is broadly comparable to the observed trend, the simulated expansions are generally less well described by a linear trend. Moreover, the climatological SIA in these simulations is substantially underestimated, typically around half, or even less, of the observed value. As a counterpoint, Fig. 4b shows the eight simulations with the closest match to the observed 1979–2014 climatological SIA, based on the root-mean-square error (RMSE). (We choose to show eight simulations to match the number shown in Fig. 4a.) These runs, as one can see, fail to reproduce the observed expansion trend and in fact simulate strong declines. Taken together with Fig. 4a, this illustrates that no model captures both the observed sea ice expansion and the observed climatology of Antarctic sea ice.

Next, we examine the spatial patterns of sea ice gain and loss. To give a sense of sea ice behavior during winter and summer, we focus on selected months. Figure 5 shows the September sea ice concentration trends over 1979–2014 in those eight simulations that capture the observed overall sea ice expansion over this period, with observations in the top left. Unlike observations, where both sea ice gain and loss are confined to a narrow band near the ice edge, most models exhibit broad, diffuse regions of sea ice concentration increases, often accompanied by large areas of sea ice loss. The largest trends in the models are in the Ross and Weddell Seas. Figure S6 shows February trends over 1979–2014 in these models. Because most models simulate very limited summer sea ice cover by February, little change is evident, unlike in the observations, which show notable trends during the summer season (Fig. S6a).

b. Sea ice decline phase

Next, we turn to the sea ice decline phase, i.e., 2014–24. Fewer simulations were conducted for this period using the

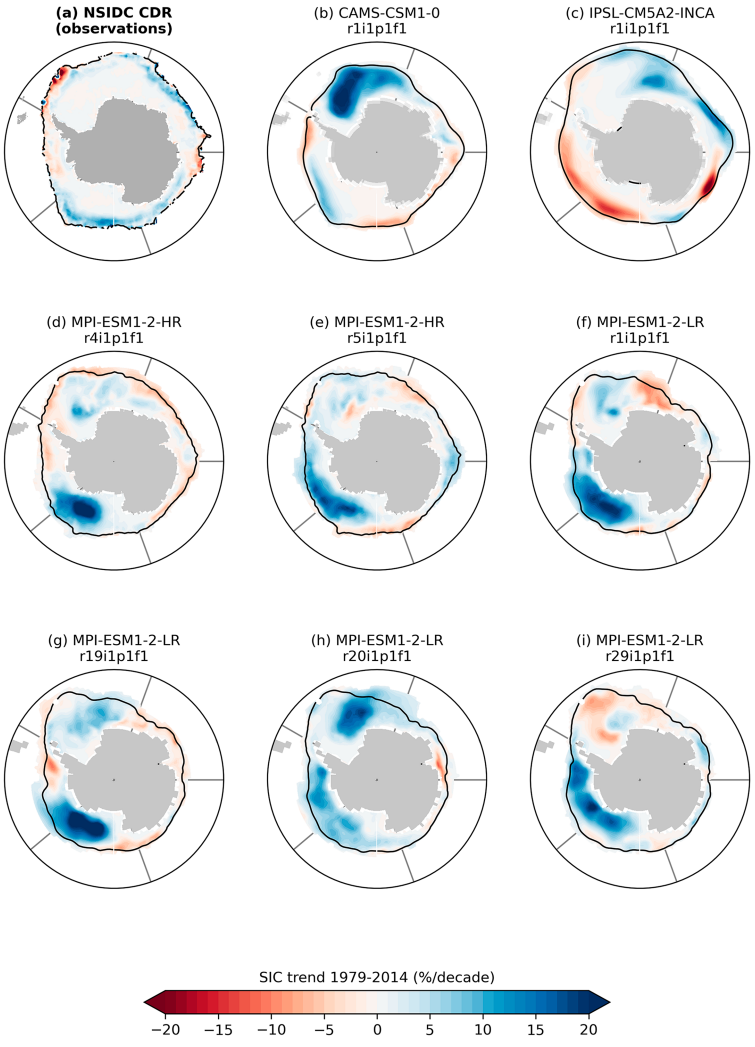

FIG. 5. September 1979–2014 sea ice concentration trends in (a) NSIDC CDR observations and (b)–(i) the eight simulations whose 1979–2014 annual-mean SIA trend is consistent with observations. The black line shows the 15% sea ice concentration contour for the 1979–2014 climatology. Gray lines show the sea ice sectors.

SSP2-4.5 scenario compared with the historical scenario. The UHH CMIP6 archive includes 36 models with a total of 218 ensemble members for SSP2-4.5. Figure 6 shows the SIA trends over this period for both models and observations. Over this shorter time period, the observed trend is more variable than over 1979–2014, which results in a larger standard error (gray shading). We find that 12 simulations from 8 models capture the observed decline over 2014–24, representing 6% of the total number of simulations. If instead we use the SSP5-8.5 scenario, with 185 ensemble members from 36 models, a similar result emerges, with about 5% of simulations falling within the observed trend range (not shown, models listed in Table S1).

Comparing to the previous subsection, it is, therefore, 2–3 times more likely for models to capture the 2014–24 decline than the 1979–2014 expansion. However, the periods are of different lengths; it should be easier for models to capture a shorter trend than a longer one. Comparing all 11-yr periods between 1979 and 2024, we see that the fraction of models capturing the observed trend is in fact nearly at its lowest during the recent decline phase.

As for the expansion period, we assess whether the individual simulations whose trends are consistent with observations over the decline period, 2014–24, resemble observations in other aspects. Figure 7a shows these 12 modeled time series of SIA over 1979–2024 in color, together with the observed ones in black. Notably, some ensemble members from MPI-ESM1-2-LR also capture sea ice expansion over 1979–2014 but have very low SIA values and variability that does not resemble observations. Among the remaining models, most simulations show multidecadal declines, in contrast to the observed behavior. ACCESS-CM2 and CNRM-CM6-1 stand out slightly, remaining fairly steady until the mid-2000s before decreasing. As a counterpoint, Fig. 7b shows the 12 simulations with mean

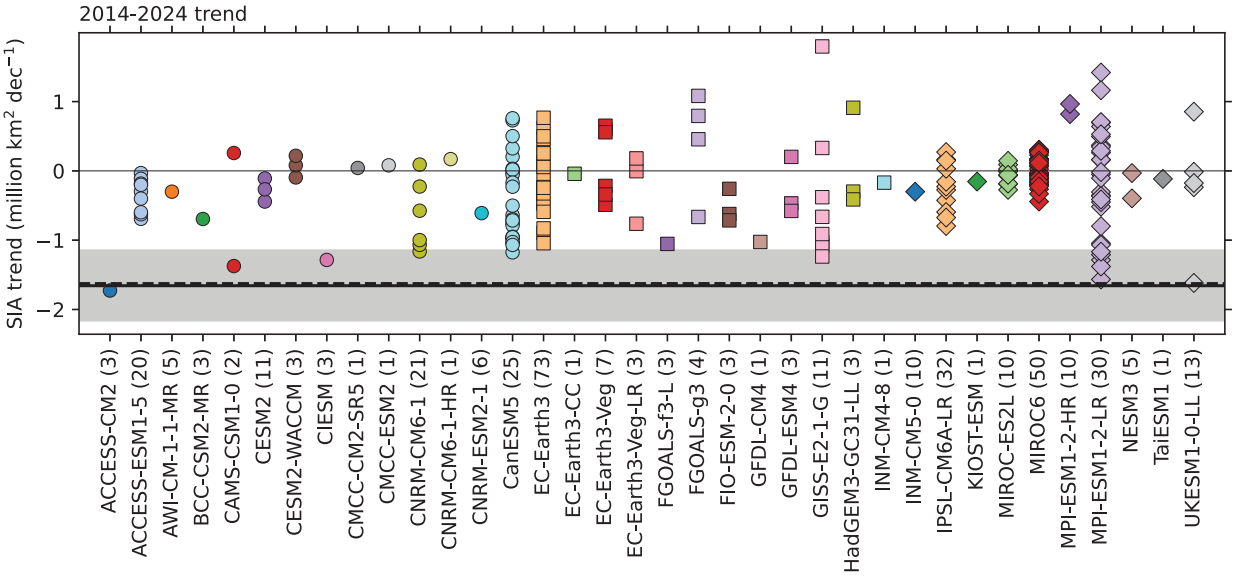

FIG. 6. As in Fig. 3, but for 2014–24, using the SSP2-4.5 scenario.

SIA over 2014–24 closest to observations. These runs, however, exhibit pronounced declines throughout the full period, again inconsistent with the observed trend structure.

Finally, we examine the spatial patterns in the models whose SIA trend over 2014–24 is consistent with observations. As for the expansion phase, one can see from Fig. S7 that February is largely ice free in most models, with the exception of CanESM5 and UKESM1-0-LL. The corresponding September 2014–24 sea ice concentration trends for these 12 ensemble members are shown in Fig. 8. Note that seven of them have very diffuse trends within the ice pack, similar to the expansion period. The remaining simulations display spatial patterns somewhat closer to observations, having a more defined band of sea ice change at the edge of the sea ice zone, although the actual locations of positive and negative trends generally fail to match observations.

4. Summary and discussion

Analyzing the same model output used in two recent studies (Holmes et al. 2024; Liu 2025) that suggested Antarctic sea ice area (SIA) trends in climate models can now be reconciled

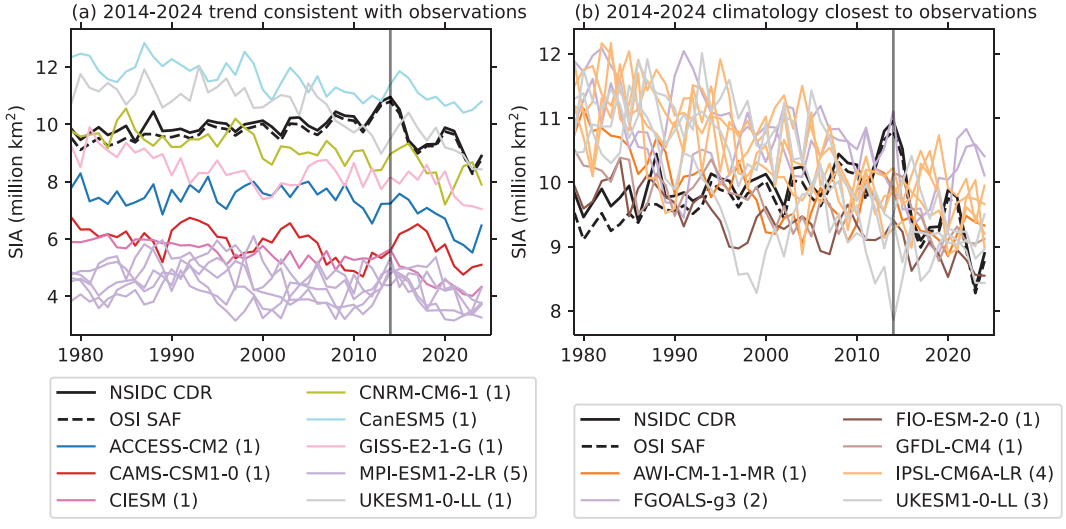

FIG. 7. Annual-mean SIA in historical simulations extended with the SSP2-4.5 scenario. (a) The 12 simulations whose 2014–24 annual-mean SIA trend is consistent with observations. (b) The 12 simulations whose 2014–24 mean SIA is closest to observations, based on RMSE. In both panels, a vertical gray line marks 2014. In the legend, the numbers in parentheses indicate the number of ensemble members included here. Both panels include the observed SIA time series from NSIDC CDR and OSI SAF.

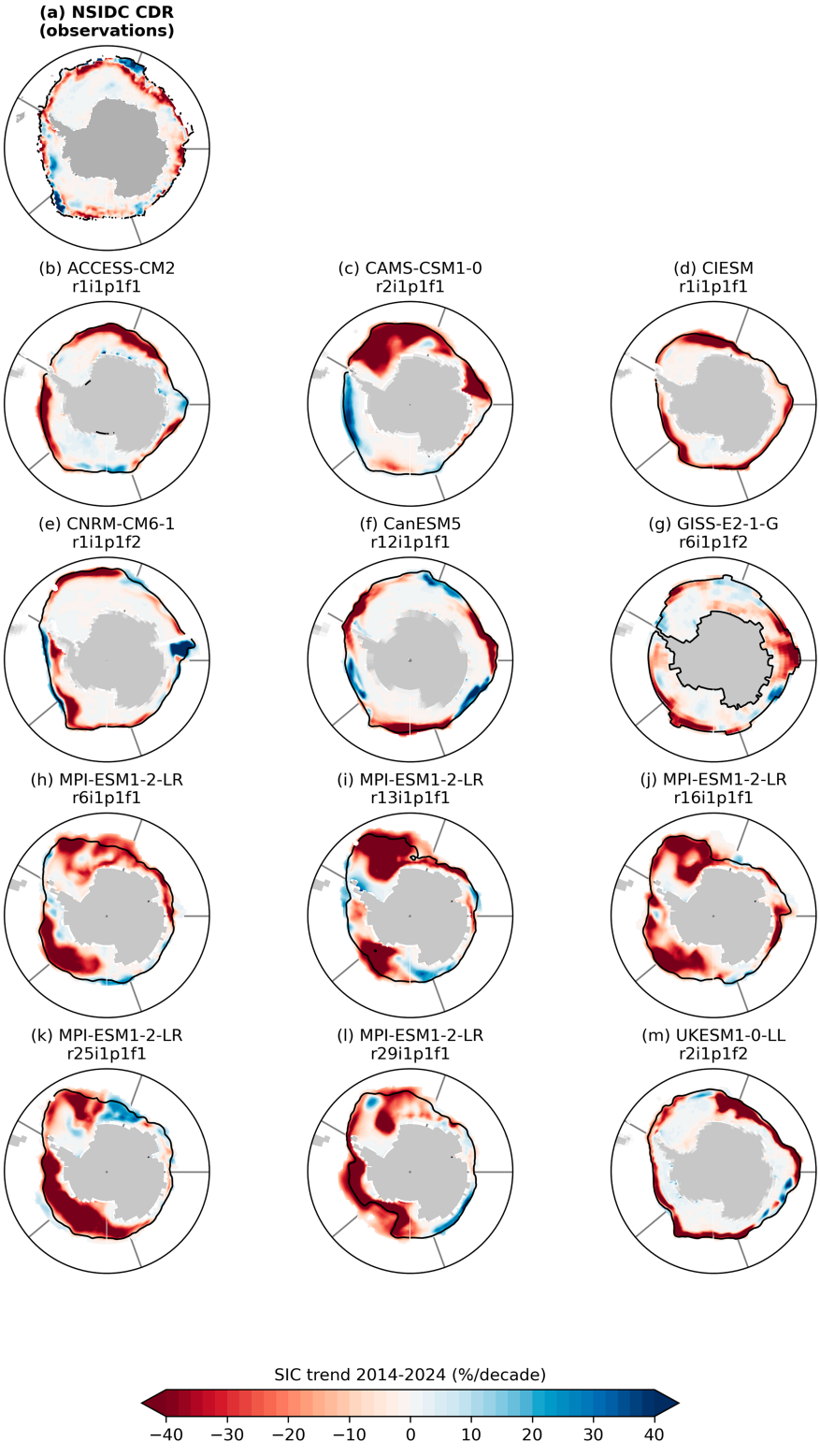

FIG. 8. September 2014–24 sea ice concentration trends in (a) NSIDC CDR observations and (b)–(m) the 12 simulations whose 2014–24 annual-mean SIA trend is consistent with observations. The black line shows the 15% sea ice concentration contour for the 2014–24 climatology. Gray lines show the sea ice sectors.

with observations, we draw different conclusions. For example, Holmes et al. (2024), computing linear trends over 1979–2023, have concluded that the models now agree with observations. The observed SIA time series, however, does not resemble a straight line, and we argue that the expansion and decline phases need to be examined separately. Regardless of whether the trends in these two phases are externally forced or internally driven, all skillful models should capture them if their ensembles are sufficiently large to adequately sample internal variability of the Antarctic climate system. And yet, not a single simulation among the several hundreds available from many models in CMIP6 resembles observations for both the 35-yr expansion phase and the 11-yr decline phase.

When we consider these two phases separately, we find that the observed sea ice expansion from 1979 to 2014 is simulated by a minuscule fraction of the CMIP6 models considered here (only around 2% of simulations). These simulations are largely the CMIP6 models identified by Liu (2025), who used them to argue that models can be reconciled with observations and to analyze associated climate conditions. However, these simulations have very biased climatological sea ice area, in many cases less than half of the observed amount. In addition, their variability does not resemble observations and neither do regional trend patterns. A larger fraction of simulations, around 6%, capture the decline phase (2014–24). However, these are also plagued by low-biased climatological sea ice area, high variability, and trend patterns that do not match the observations. Comparing short (11 yr) trends across the observational record, models are least able to capture the most recent years. Taken together, these facts lead us to conclude that the extended observational record, including the recent SIA minima, does not increase confidence in the ability of CMIP6 models to accurately simulate sea ice around Antarctica.

We have followed best practices for model–observation comparison as far as possible (Simpson et al. 2025). We conducted a like-for-like comparison, computing sea ice area on native grids and showing sea ice concentration trends regridded onto a common grid. We used two sea ice observational datasets to sample observational uncertainty. We compared observations to individual model runs, made use of many different models, and have accounted for internal variability. Furthermore, we have verified that our analysis is not sensitive to small changes in the end points of the different time periods. While we believe that the entire 1979–2024 period cannot be fitted with a single linear curve, we nonetheless applied a simple linear trend to two subperiods because they offer a simple and readily interpretable perspective and are a reasonable approximation once the sea ice time series is separated into expansion and decline phases.

An important caveat is that many models provide few ensemble members, complicating estimates of internal variability, which vary significantly across models. Applying a simple correction (based on the model mean) only slightly increases the number of models we identify as consistent with observations in terms of sea ice expansion. We also concede that the satellite record, while now 45 years in length, may still not be fully representative of the entire spectrum of time scales of variability at play in the Antarctic climate system (Fan et al. 2014). Advances in proxy reconstructions provide better context, suggesting that the more recent sea ice changes are unprecedented over the past century (Fogt et al. 2022; Raphael et al. 2025). However, we note that Antarctic sea ice evolution differs considerably across recent reconstructions, even in the latter half of the twentieth century (Goosse et al. 2024; Cooper et al. 2025). And, although we have here focused on SIA, as it is relatively well observed from satellites, accurate sea ice thickness observations are still lacking, limiting full characterization of Antarctic sea ice evolution.

Several hypotheses have been proposed to explain why current-generation models still struggle to simulate observed Antarctic sea ice changes. These include interactions with the Antarctic ice sheet, which are not included in CMIP6 models (Schmidt et al. 2023), biases in Southern Hemisphere atmospheric circulation (Bracegirdle et al. 2020), biases in sea ice drift (Sun and Eisenman 2021), biases in the representation of Southern Ocean clouds (Zelinka et al. 2020), biases in Southern Ocean heat fluxes (Chemke 2022; Zhang et al. 2021), the low resolution of the ocean models (Rackow et al. 2022), and significant deep open-ocean convection in models that is unsupported by observations (de Lavergne et al. 2014; Heuzé 2021; Morioka et al. 2023). It is particularly noteworthy that ocean–ice hindcast simulations (Krumhardt et al. 2024; Kusahara and Tatebe 2025) and simulations with nudged winds (Blanchard-Wrigglesworth et al. 2021; Roach et al. 2023) are able to capture the observed changes, while free-running coupled models are not. This suggests that atmospheric processes and atmosphere–ocean/sea ice coupling could be the culprit. There are many processes at play in this coupled system, and more work is required to untangle the leading causes of sea ice biases.

Ultimately, we aim to understand the drivers of observed Antarctic sea ice change over the satellite record, the key question being to what extent the 1979–2014 expansion or the more recent decline is fundamentally caused by anthropogenic forcings versus merely reflecting internal (i.e., unforced) variability. To answer that question, comprehensive climate models are a crucial tool, since the anthropogenic forcings in model simulations can be included or excluded at will, and internal variability can be suppressed by averaging large ensemble to extract the forced component. However, to confidently draw conclusions from model simulations, one needs to ensure that models are able to capture observations. Unfortunately, at present, this is not the case, and this is why the observed changes in Antarctic sea ice over the last several decades remain a major puzzle.

Acknowledgments. We acknowledge the World Climate Research Programme (WRCP), which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. LAR was supported by the Initiative and

Networking Fund of the Helmholtz Association (VH-NG-20-10). LMP contributed to this work out of sheer curiosity. The authors thank the reviewers (Patricia deRepentigny, Alek Petty, and one anonymous reviewer) for their thoughtful and constructive feedback, which greatly improved the manuscript.

Data availability statement. Sea ice observations from 1979 through 2024 are from OSI SAF (EUMETSAT 2023) and computed from the NSIDC Climate Data Record version 4 sea ice concentration (Meier et al. 2021). CMIP6 model output is available online at https://esgf-node.llnl.gov/projects/esgf-llnl/, and processed sea ice area from the CMIP6 models is available from the University of Hamburg (UHH) CMIP6 SIA directory (https://www.cen.uni-hamburg.de/en/icdc/data/cryosphere/cmip6-sea-ice-area.html, accessed 7 October 2025).

REFERENCES

Armour, K. C., J. Marshall, J. R. Scott, A. Donohoe, and E. R. Newsom, 2016: Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. *Nat. Geosci.*, **9**, 549–554, https://doi.org/10.1038/ngeo2731.

Banerjee, A., J. C. Fyfe, L. M. Polvani, D. Waugh, and K.-L. Chang, 2020: A pause in Southern Hemisphere circulation trends due to the Montreal Protocol. *Nature*, **579**, 544–548, https://doi.org/10.1038/s41586-020-2120-4.

Bitz, C. M., and L. M. Polvani, 2012: Antarctic climate response to stratospheric ozone depletion in a fine resolution ocean climate model. *Geophys. Res. Lett.*, **39**, L20705, https://doi.org/10.1029/2012gl053393.

Blanchard-Wrigglesworth, E., L. A. Roach, A. Donohoe, and Q. Ding, 2021: Impact of winds and Southern Ocean SSTs on Antarctic sea ice trends and variability. *J. Climate*, **34**, 949–965, https://doi.org/10.1175/JCLI-D-20-0386.1.

Bracegirdle, T. J., C. R. Holmes, J. S. Hosking, G. J. Marshall, M. Osman, M. Patterson, and T. Rackow, 2020: Improvements in circumpolar Southern Hemisphere extratropical atmospheric circulation in CMIP6 compared to CMIP5. *Earth Space Sci.*, **7**, e2019EA001065, https://doi.org/10.1029/2019ea001065.

Chemke, R., 2022: The future poleward shift of Southern Hemisphere summer mid-latitude storm tracks stems from ocean coupling. *Nat. Commun.*, **13**, 1730, https://doi.org/10.1038/s41467-022-29392-4.

Cooper, V. T., G. J. Hakim, and K. C. Armour, 2025: Monthly sea surface temperature, sea ice, and sea level pressure over 1850–2023 from coupled data assimilation. *J. Climate*, **38**, 5461–5490, https://doi.org/10.1175/JCLI-D-25-0021.1.

de Lavergne, C., J. B. Palter, E. D. Galbraith, R. Bernardello, and I. Marinov, 2014: Cessation of deep convection in the open Southern Ocean under anthropogenic climate change. *Nat. Climate Change*, **4**, 278–282, https://doi.org/10.1038/nclimate2132.

Diamond, R., L. C. Sime, C. R. Holmes, and D. Schroeder, 2024: CMIP6 models rarely simulate Antarctic winter sea-ice anomalies as large as observed in 2023. *Geophys. Res. Lett.*, **51**, e2024GL109265, https://doi.org/10.1029/2024gl109265.

Eisenman, I., T. Schneider, D. S. Battisti, and C. M. Bitz, 2011: Consistent changes in the sea ice seasonal cycle in response to global warming. *J. Climate*, **24**, 5325–5335, https://doi.org/10.1175/2011jcli4051.1.

EUMETSAT, 2023: OSI SAF sea ice index 1978-onwards, version 2.2. EUMETSAT Ocean and Sea Ice Satellite Application Facility, accessed 25 June 2025, https://osi-saf.eumetsat.int/products/osi-420.

Fan, T., C. Deser, and D. P. Schneider, 2014: Recent Antarctic sea ice trends in the context of Southern Ocean surface climate variations since 1950. *Geophys. Res. Lett.*, **41**, 2419–2426, https://doi.org/10.1002/2014GL059239.

Ferreira, D., J. Marshall, C. M. Bitz, S. Solomon, and A. Plumb, 2015: Antarctic Ocean and sea ice response to ozone depletion: A two-time-scale problem. *J. Climate*, **28**, 1206–1226, https://doi.org/10.1175/jcli-d-14-00313.1.

Fogt, R. L., A. M. Sleinkofer, M. N. Raphael, and M. S. Handcock, 2022: A regime shift in seasonal total Antarctic sea ice extent in the twentieth century. *Nat. Climate Change*, **12**, 54–62, https://doi.org/10.1038/s41558-021-01254-9.

Fox-Kemper, B., and Coauthors, 2021: Ocean, cryosphere and sea level change. *Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte et al., Eds., Cambridge University Press, 1211–1362.

Goosse, H., Q. Dalaiden, F. Feba, B. Mezzina, and R. L. Fogt, 2024: A drop in Antarctic sea ice extent at the end of the 1970s. *Commun. Earth Environ.*, **5**, 628, https://doi.org/10.1038/s43247-024-01793-x.

Heuzé, C., 2021: Antarctic Bottom Water and North Atlantic Deep Water in CMIP6 models. *Ocean Sci.*, **17**, 59–90, https://doi.org/10.5194/os-17-59-2021.

Hobbs, W., and Coauthors, 2024: Observational evidence for a regime shift in summer Antarctic sea ice. *J. Climate*, **37**, 2263–2275, https://doi.org/10.1175/JCLI-D-23-0479.1.

Holland, P. R., and R. Kwok, 2012: Wind-driven trends in Antarctic sea-ice drift. *Nat. Geosci.*, **5**, 872–875, https://doi.org/10.1038/ngeo1627.

Holmes, C. R., T. J. Bracegirdle, P. R. Holland, J. Stroeve, and J. Wilkinson, 2024: Brief communication: New perspectives on the skill of modelled sea ice trends in light of recent Antarctic sea ice loss. *Cryosphere*, **18**, 5641–5652, https://doi.org/10.5194/tc-18-5641-2024.

Ivanova, N., O. M. Johannessen, L. T. Pedersen, and R. T. Tonboe, 2014: Retrieval of arctic sea ice parameters by satellite passive microwave sensors: A comparison of eleven sea ice concentration algorithms. *IEEE Trans. Geosci. Remote Sens.*, **52**, 7233–7246, https://doi.org/10.1109/TGRS.2014.2310136.

Krumhardt, K. M., and Coauthors, 2024: From nutrients to fish: Impacts of mesoscale processes in a global CESM-FEISTY eddying ocean model framework. *Prog. Oceanogr.*, **227**, 103314, https://doi.org/10.1016/j.pocean.2024.103314.

Kusahara, K., and H. Tatebe, 2025: Causes of the abrupt and sustained 2016–2023 Antarctic sea-ice decline: A sea ice–ocean model perspective. *Geophys. Res. Lett.*, **52**, e2025GL115256, https://doi.org/10.1029/2025gl115256.

Landrum, L. L., M. M. Holland, M. N. Raphael, and L. M. Polvani, 2017: Stratospheric ozone depletion: An unlikely driver of the regional trends in Antarctic sea ice in austral fall in the late twentieth century. *Geophys. Res. Lett.*, **44**, 11 062–11 070, https://doi.org/10.1002/2017gl075618.

Liu, W., 2025: Simulated Antarctic sea ice expansion reconciles climate model with observation. *npj Climate Atmos. Sci.*, **8**, 4, https://doi.org/10.1038/s41612-024-00881-1.

Marshall, J., K. C. Armour, J. R. Scott, Y. Kostov, U. Hausmann, D. Ferreira, T. G. Shepherd, and C. M. Bitz, 2014: The ocean's role in polar climate change: Asymmetric Arctic and Antarctic responses to greenhouse gas and ozone forcing. *Philos. Trans. Roy. Soc.*, **A372**, 20130040, https://doi.org/10.1098/rsta.2013.0040.

Meier, W., F. Fetterer, A. Windnagel, and S. Stewart, 2021: NOAA/NSIDC climate data record of passive microwave sea ice concentration, version 4. NSIDC, accessed 8 April 2025, https://doi.org/10.7265/EFMZ-2T65.

Merrifield, A. L., L. Brunner, R. Lorenz, V. Humphrey, and R. Knutti, 2023: Climate model Selection by Independence, Performance, and Spread (ClimSIPS v1.0.1) for regional applications. *Geosci. Model Dev.*, **16**, 4715–4747, https://doi.org/10.5194/gmd-16-4715-2023.

Morioka, Y., L. Zhang, T. L. Delworth, X. Yang, F. Zeng, M. Nonaka, and S. K. Behera, 2023: Multidecadal variability and predictability of Antarctic sea ice in the GFDL SPEAR_LO model. *Cryosphere*, **17**, 5219–5240, https://doi.org/10.5194/tc-17-5219-2023.

Notz, D., 2014: Sea-ice extent and its trend provide limited metrics of model performance. *Cryosphere*, **8**, 229–243, https://doi.org/10.5194/tc-8-229-2014.

Polvani, L. M., and K. L. Smith, 2013: Can natural variability explain observed Antarctic sea ice trends? New modeling evidence from CMIP5. *Geophys. Res. Lett.*, **40**, 3195–3199, https://doi.org/10.1002/grl.50578.

——, and Coauthors, 2021: Interannual SAM modulation of Antarctic sea ice extent does not account for its long-term trends, pointing to a limited role for ozone depletion. *Geophys. Res. Lett.*, **48**, e2021GL094871, https://doi.org/10.1029/2021gl094871.

Purich, A., and M. H. England, 2019: Tropical teleconnections to Antarctic sea ice during austral spring 2016 in coupled pacemaker experiments. *Geophys. Res. Lett.*, **46,** 6848–6858, https://doi.org/10.1029/2019gl082671.

——, and E. W. Doddridge, 2023: Record low Antarctic sea ice coverage indicates a new sea ice state. *Commun. Earth Environ.*, **4**, 314, https://doi.org/10.1038/s43247-023-00961-9.

Rackow, T., S. Danilov, H. F. Goessling, H. H. Hellmer, D. V. Sein, T. Semmler, D. Sidorenko, and T. Jung, 2022: Delayed Antarctic sea-ice decline in high-resolution climate change simulations. *Nat. Commun.*, **13**, 637, https://doi.org/10.1038/s41467-022-28259-y.

Raphael, M. N., T. J. Maierhofer, R. L. Fogt, W. R. Hobbs, and M. S. Handcock, 2025: A twenty-first century structural change in Antarctica's sea ice system. *Commun. Earth Environ.*, **6**, 131, https://doi.org/10.1038/s43247-025-02107-5.

Roach, L. A., and Coauthors, 2020: Antarctic sea ice area in CMIP6. *Geophys. Res. Lett.*, **47**, e2019GL086729, https://doi.org/10.1029/2019GL086729.

——, K. D. Mankoff, A. Romanou, E. Blanchard-Wrigglesworth, T. W. N. Haine, and G. A. Schmidt, 2023: Winds and meltwater together lead to Southern Ocean surface cooling and sea ice expansion. *Geophys. Res. Lett.*, **50**, e2023GL105948, https://doi.org/10.1029/2023gl105948.

Schmidt, G. A., and Coauthors, 2023: Anomalous meltwater from ice sheets and ice shelves is a historical forcing. *Geophys. Res. Lett.*, **50**, e2023GL106530, https://doi.org/10.1029/2023GL106530.

Seviour, W. J. M., and Coauthors, 2019: The Southern Ocean sea surface temperature response to ozone depletion: A multimodel comparison. *J. Climate*, **32**, 5107–5121, https://doi.org/10.1175/jcli-d-19-0109.1.

Shu, Q., Q. Wang, Z. Song, F. Qiao, J. Zhao, M. Chu, and X. Li, 2020: Assessment of sea ice extent in CMIP6 with comparison to observations and CMIP5. *Geophys. Res. Lett.*, **47**, e2020GL087965, https://doi.org/10.1029/2020gl087965.

Simpson, I. R., and Coauthors, 2025: Confronting Earth System Model trends with observations. *Sci. Adv.*, **11**, eadt8035, https://doi.org/10.1126/sciadv.adt8035.

Singh, H. A., L. M. Polvani, and P. J. Rasch, 2019: Antarctic sea ice expansion, driven by internal variability, in the presence of increasing atmospheric CO_2. *Geophys. Res. Lett.*, **46**, 14 762–14 771, https://doi.org/10.1029/2019gl083758.

Sun, S., and I. Eisenman, 2021: Observed Antarctic sea ice expansion reproduced in a climate model after correcting biases in sea ice drift velocity. *Nat. Commun.*, **12**, 1060, https://doi.org/10.1038/s41467-021-21412-z.

Swart, N. C., and Coauthors, 2023: The Southern Ocean Freshwater Input from Antarctica (SOFIA) initiative: Scientific objectives and experimental design. *Geosci. Model Dev.*, **16**, 7289–7309, https://doi.org/10.5194/gmd-16-7289-2023.

Turner, J., T. J. Bracegirdle, T. Phillips, G. J. Marshall, and J. S. Hosking, 2013: An initial assessment of Antarctic sea ice extent in the CMIP5 models. *J. Climate*, **26**, 1473–1484, https://doi.org/10.1175/jcli-d-12-00068.1.

——, J. S. Hosking, G. J. Marshall, T. Phillips, and T. J. Bracegirdle, 2016: Antarctic sea ice increase consistent with intrinsic variability of the Amundsen Sea Low. *Climate Dyn.*, **46**, 2391–2402, https://doi.org/10.1007/s00382-015-2708-9.

——, T. Phillips, G. J. Marshall, J. S. Hosking, J. O. Pope, T. J. Bracegirdle, and P. Deb, 2017: Unprecedented springtime retreat of Antarctic sea ice in 2016. *Geophys. Res. Lett.*, **44**, 6868–6875, https://doi.org/10.1002/2017gl073656.

Wilson, E. A., D. B. Bonan, A. F. Thompson, N. Armstrong, and S. C. Riser, 2023: Mechanisms for abrupt summertime circumpolar surface warming in the Southern Ocean. *J. Climate*, **36**, 7025–7039, https://doi.org/10.1175/JCLI-D-22-0501.1.

Zelinka, M. D., T. A. Myers, D. T. McCoy, S. Po-Chedley, P. M. Caldwell, P. Ceppi, S. A. Klein, and K. E. Taylor, 2020: Causes of higher climate sensitivity in CMIP6 models. *Geophys. Res. Lett.*, **47**, e2019GL085782, https://doi.org/10.1029/2019gl085782.

Zhang, X., C. Deser, and L. Sun, 2021: Is there a tropical response to recent observed Southern Ocean cooling? *Geophys. Res. Lett.*, **48**, e2020GL091235, https://doi.org/10.1029/2020gl091235.

Zunz, V., H. Goosse, and F. Massonnet, 2013: How does internal variability influence the ability of CMIP5 models to reproduce the recent trend in Southern Ocean sea ice extent? *Cryosphere*, **7**, 451–468, https://doi.org/10.5194/tc-7-451-2013.