The GFDL-CM4X climate model hierarchy, Part II: case studies

Stephen M Griffies[1], Alistair Adcroft[2], Rebecca Lynn Beadling[3], Mitchell Bushuk[4], Chiung-Yin Chang[2], Henri Francois Drake[5], Raphael Dussin[6], Robert W. Hallberg[7], William Hurlin[8], Hemant Khatri[9], John P Krasting[10], Matthew Lobo[2], Graeme MacGilchrist[11], Brandon G Reichl[1], Aakash Sane[2], Olga V. Sergienko[12], Maike Sonnewald[13], Jacob M. Steinberg[10], Jan-Erik Tesdal[2], Matthew D Thomas[14], Katherine Elise Turner[15], Marshall L Ward[16], Michael Winton[4], Niki Zadeh[8], Laure Zanna[17], Rong Zhang[4], Wenda Zhang[2], and Ming Zhao[4]

[1]NOAA Geophysical Fluid Dynamics Laboratory
[2]Princeton University
[3]Temple University
[4]GFDL/NOAA
[5]University of California Irvine
[6]National Oceanic and Atmospheric Administration
[7]NOAA/Geophysical Fluid Dynamics Laboratory
[8]NOAA/GFDL
[9]University of Liverpool
[10]Geophysical Fluid Dynamics Laboratory
[11]University of St Andrews
[12]AOS, Princeton University/GFDL
[13]University of California, Davis
[14]Centre for Environment, Fisheries and Aquaculture Science
[15]University of Arizona
[16]Australian National University
[17]New York University

November 28, 2024

Abstract

This paper is Part II of a two-part paper that documents the CM4X (Climate Model version 4X) hierarchy of coupled climate models developed at the Geophysical Fluid Dynamics Laboratory (GFDL). Part I of this paper is presented in \citeA{CM4X_partI}. Here we present a suite of case studies that examine ocean and sea ice features that are targeted for further research, which include sea level, eastern boundary upwelling, Arctic and Southern Ocean sea ice, Southern Ocean circulation, and North Atlantic circulation. The case studies are based on experiments that follow the protocol of version 6 from the Coupled Model Intercomparison Project (CMIP6). The analysis reveals a systematic improvement in the simulation fidelity of CM4X relative to its CM4.0 predecessor, as well as an improvement when refining the ocean/sea ice horizontal grid spacing from the \$0.25ˆ{\circ}\$ of CM4X-p25 to the \$0.125ˆ{\circ}\$ of CM4X-p125. Even so, there remain many outstanding biases, thus pointing to the need for further grid refinements, enhancements to numerical methods, and/or advances in parameterizations, each of which target long-standing model biases and limitations.

The GFDL-CM4X climate model hierarchy, Part II: case studies

Stephen M. Griffies[1,2], Alistair Adcroft[2], Rebecca L. Beadling[3], Mitchell Bushuk[1], Chiung-Yin Chang[2], Henri F. Drake[4], Raphael Dussin[1], Robert W. Hallberg[1,2], William J. Hurlin[1], Hemant Khatri[5], John P. Krasting[1], Matthew Lobo[2], Graeme A. MacGilchrist[6], Brandon G. Reichl[1], Aakash Sane[2], Olga Sergienko[2], Maike Sonnewald[7], Jacob M. Steinberg[1], Jan-Erik Tesdal[2], Matthew Thomas[8], Katherine E. Turner[2], Marshall L. Ward[1], Michael Winton[1], Niki Zadeh[1], Laure Zanna[9], Rong Zhang[1,2], Wenda Zhang[2], Ming Zhao[1]

[1]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, USA
[2]Princeton University Atmospheric and Oceanic Sciences Program, Princeton, USA
[3]Department of Earth and Environmental Science, Temple University, Philadelphia, USA
[4]Department of Earth System Science, University of California, Irvine, USA
[5]Department of Earth, Ocean and Ecological Sciences, University of Liverpool, UK
[6]School of Earth and Environmental Sciences, University of St. Andrews, UK
[7]Computer Science Department, University of California, Davis, USA
[8]Centre for Environment Fisheries and Aquaculture Science, Suffolk, UK
[9]Courant Institute of Mathematical Sciences, New York University, New York, USA

November 23, 2024

Key Points:

- We present case studies of selected features of the GFDL-CM4X climate model from CMIP6 piControl, historical, and SSP5-8.5 simulations.
- Case studies include sea level, eastern boundary upwelling, sea ice, Southern Ocean circulation, and North Atlantic Ocean circulation.
- Refining ocean grid spacing from 0.25° to 0.125° has systematic improvements across a number of climate relevant features.

Corresponding author: Stephen Griffies, `Stephen.M.Griffies@gmail.com`

Abstract

This paper is Part II of a two-part paper that documents the CM4X (Climate Model version 4X) hierarchy of coupled climate models developed at the Geophysical Fluid Dynamics Laboratory (GFDL). Part I of this paper is presented in Griffies et al. (2024). Here we present a suite of case studies that examine ocean and sea ice features that are targeted for further research, which include sea level, eastern boundary upwelling, Arctic and Southern Ocean sea ice, Southern Ocean circulation, and North Atlantic circulation. The case studies are based on experiments that follow the protocol of version 6 from the Coupled Model Intercomparison Project (CMIP6). The analysis reveals a systematic improvement in the simulation fidelity of CM4X relative to its CM4.0 predecessor, as well as an improvement when refining the ocean/sea ice horizontal grid spacing from the 0.25° of CM4X-p25 to the 0.125° of CM4X-p125. Even so, there remain many outstanding biases, thus pointing to the need for further grid refinements, enhancements to numerical methods, and/or advances in parameterizations, each of which target long-standing model biases and limitations.

Plain Language Summary

We examine simulations from a new climate model hierarchy, referred to as CM4X (Climate Model version 4X). The finer grid component of the hierarchy, CM4X-p125, out shines its coarser sibling, CM4X-p25, for certain processes of interest for climate studies, though in others the results are not dramatically distinct. Each case study reveals the advances made by moving from the predecessor CM4.0 climate model to finer grid spacing in either the atmosphere or ocean. Even so, there remain many unresolved problems that help to guide further research and development goals and strategies.

1 Introduction and content of this paper

This paper is Part II of a two-part paper that documents the CM4X hierarchy of coupled climate models, with Part I presented in Griffies et al. (2024). We developed CM4X to support research into the ocean and sea ice components of the earth climate system, with CM4X comprised of two coupled climate models, CM4X-p25 and CM4X-p125. These two models are identically configured, except for their ocean (and sea ice) horizontal grid spacing and bottom topography. In Part I from Griffies et al. (2024), we documented the remarkable thermal equilibration properties of CM4X-p125, and proposed the *mesoscale dominance hypothesis* to help explain the behavior. In the present paper, we work through a suite of case studies that focus on areas of planned research with CM4X.

As detailed in Section 3.1 of Griffies et al. (2024), we present results from following CMIP6 (Eyring et al., 2016) simulations.

- <u>piControl</u>: Pre-industrial control with radiative forcing fixed at year 1850. This experiment illustrates how the models drift from their initial conditions, taken from the early 21st century, and approach thermal equilibrium under pre-industrial forcing.
- <u>Historical</u>: 01January of year 101 from the piControl is used to initialize a historical simulation that is run from 1850 to 2014. In this historical simulation, we did not account for temporal evolution in vegetation, land use, or CO2 fertilization.
- <u>SSP5-8.5</u>: 01January of year 2015 provides the initial condition for the CMIP6 SSP5-8.5 scenario experiment, which allows us to study how the CM4X models simulate climate change through to 2100.

The case studies exemplify aspects of the science going into the model and the science emerging from the model simulations. The presentation generally follows a "show and tell" approach given that our primary aim in this paper is to document features of the new CM4X hierarchy, with many of these features to be more thoroughly examined in future studies.

We begin in Section 2 with a study of the global thermosteric sea level along with statistical patterns of dynamic sea level. This analysis reveals that the historical thermosteric sea level in the CM4X models is somewhat improved relative to CM4.0, and yet the patterns of sea level skewness in CM4.0 and CM4X remain in poor agreement in comparison to ocean reanalysis. In Section 3 we examine properties of the eastern boundary upwelling zones, which are regions of particular importance for biogeochemistry. Here we find an advance arises from the refined atmospheric model grid used in CM4X relative to CM4.0 (see Section A1 of Griffies et al. (2024)), thus improving the fidelity of coastal wind patterns key to upwelling. Even so, long-standing biases in the low level clouds means that the upwelling zones remain far too warm relative to observations. Section 4 studies the Arctic Ocean and Southern Ocean sea ice properties, revealing again that the CM4X model represents an advance over the CM4.0 model, though with many longstanding biases remaining.

In Section 5 we study properties of the Southern Ocean simulation, with some focus on the region near Antarctica given its importance to ongoing studies of ice shelf melt. A particularly encouraging feature of both CM4X-p25 and CM4X-p125 concerns the absence of the unphysically large open ocean polynyas that plagued CM4.0 (Held et al., 2019) and its earth system model cousin ESM4.1 (Dunne et al., 2020). As a result, CM4X provides a versatile tool for performing perturbation experiments to examine, say, the role of fresh water melt around Antarctica such as in Beadling et al. (2022) and Tesdal et al. (2023), including examining the role of the ocean in the SST pattern effect (Armour et al., 2013; Andrews et al., 2018). We complete the case studies in Section 6 with a focus on the North Atlantic circulation, considering both horizontal and overturning circulation features in the middle and high latitudes. The CM4X simulations show some advances over CM4.0 in the overturning and supolar gyre properties, and yet there remain nontrivial biases in the overturning depth and attendant overflows (model is too shallow), as well as biases in the Gulf Stream structure (simulated jet does not penetrate far enough from the coast). We close the paper in Section 7 with concluding remarks on strategies for future ocean climate model development that are motivated by results from the CM4X model hierarchy.

2 Thermosteric and dynamic sea level

In this section we consider two aspects of sea level: global mean thermosteric sea level and patterns of dynamic sea level.

2.1 Global mean thermosteric sea level

Changes to global mean thermosteric sea level occur with seawater density changes affected by temperature changes. Although seawater density is a highly nonlinear function of temperature, salinity, and pressure, we expect the time series for thermosteric sea level to reflect that for global mean temperature, with Figure 1 supporting this expectation. To generate this figure, we computed thermosteric sea level according to Section H9.5 of Griffies et al. (2016) (CMIP variable *zostoga*), using software described in Krasting et al. (2024) and with full-depth monthly mean fields. Notably, the CM4.0 piControl drift is larger than CM4X-p25, whereas there is negligible drift in CM4X-p125.

Figure 1 suggests that differences between CM4X-p25 and CM4X-p125 are mostly due to the difference in drift seen in the piControl runs. Removing a linear trend com-

ACRONYM	MEANING	CITATION OR SECTION
AM4	GFDL Atmospheric Model 4.0	Zhao et al. (2018b, 2018a)
CM2-O	GFDL climate model hierarchy 2.0	Delworth et al. (2006), Griffies et al. (2015)
C96	AM4 with cubed-sphere (≈ 100 km)	Zhao et al. (2018b, 2018a)
C192	AM4 with cubed-sphere (≈ 50 km) in CM4X	Zhao (2020)
CMIP6	Coupled Model Intercomparison Project 6	Eyring et al. (2016)
CM4.0	GFDL Climate Model 4.0 (0.25° ocn & C96 atm)	Held et al. (2019)
CM4X	GFDL Climate Model hierarchy	this paper
CM4X-p25	CM4X w/ 0.25° ocn and C192 atm	this paper
CM4X-p125	CM4X with 0.125° ocn and C192 atm	this paper
CM4X-p25-C96	CM4X with 0.25° ocn and C96 atm	3
ESM4.1	GFDL Earth System Model 4.1	Dunne et al. (2020)
GFDL	Geophysical Fluid Dynamics Laboratory	–
MOM6	Modular Ocean Model version 6	Adcroft et al. (2019), Griffies et al. (2020)
NWA12	NorthWest Atlantic 1/12° model	6.1 and Ross et al. (2023)
OM4.0	GFDL Ocean/Sea-ice Model 4.0 (0.25°)	Adcroft et al. (2019)
SIS2	Sea Ice Simulator version 2	Delworth et al. (2006), Adcroft et al. (2019)
AABW	Antarctic Bottom Water	5
AAIW	Antarctic Intermediate Water	5
ACC	Antarctic Circumpolar Current	5
AIS	Antarctic Ice Shelf	5
AMOC	Atlantic Meridional Overturning Circulation	6
ASC	Antarctic Slope Current	5
CDW	Circumpolar Deep Water	5
EBUS	eastern boundary upwelling system	3
DSW	Dense Shelf Water	5
NADW	North Atlantic Deep Water	6
MKE	mean kinetic energy	6.1
MLD	mixed layer depth	6
OSNAP	Overturning in the North Atlantic Subpolar Program	6
RAPID	Rapid Climate Change Programme	6
RMSE	root-mean-square error	4
SAMW	Sub-Antarctic Mode Water	5
SIC	Sea Ice Concentration	4
SIE	Sea Ice Extent	4
SIT	Sea Ice Thickness	4
SIV	Sea Ice Volume	4
WMT	watermass transformation	5.2 and 6.4

Table 1. Acronyms used in this paper, their meaning, and relevant citation or section. The upper portion refers to model related acronyms and the lower portion to oceanographic and statistical related acronyms.

puted from the piControl leads to very similar global thermal expansion in the historical and SSP5-8.5 simulations (Figure 2). Evidently, the nonlinear effects noted by Hallberg et al. (2013) are not revealed by these two simulations, presumably since their piControl states have not drifted too far apart after 100 years.

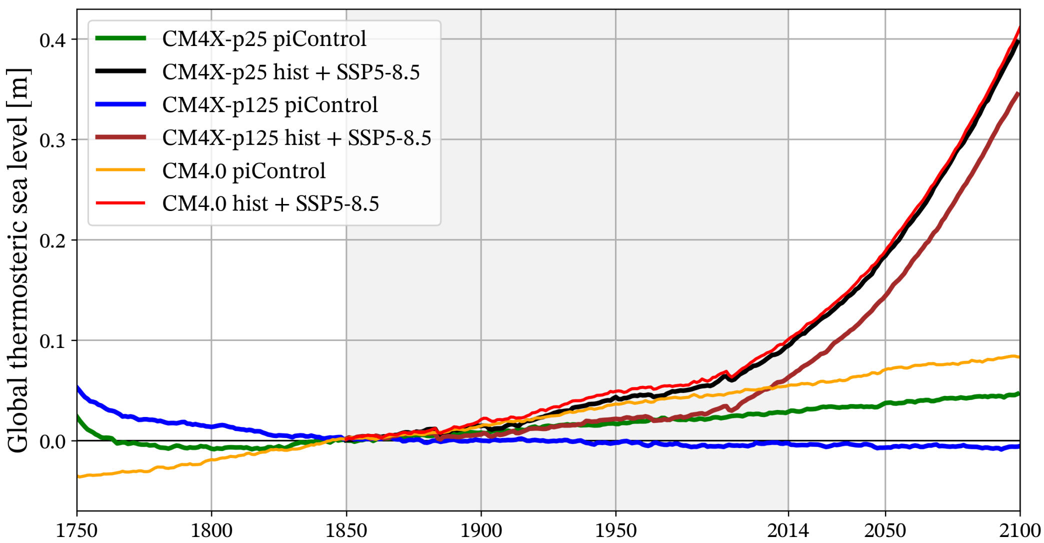

Figure 1. Global thermosteric sea level in piControl, historical (1850-2014), and SSP5-8.5 (2014-2100) simulations using CM4X and CM4.0 climate models. Historical simulations for CM4X-p25 and CM4X-p125 branch from the corresponding piControl at year 101, whereas CM4.0 is branched from its piControl at year 251. This different branching explains why the CM4.0 piControl does not line up with the CM4X piControls. Furthermore, CM4.0's later branching means that its initial cooling phase see in Part I (Griffies et al., 2024) is outside of the time range of this figure, so that the CM4.0 piControl exhibits a nearly linear drift throughout.

During the 20th and early 21st centuries, the observed global-mean sea level exhibited significant increases, with thermosteric rise becoming increasingly significant in recent decades (Frederikse et al., 2020). Figure 2 shows changes relative to the year 2002-2018 time mean, plotted over the historical simulations (from year 1850 through 2014) and eight years of the SSP5-8.5 projection (years 2015 to 2022). We accounted for model drift by removing the long term linear trend in the corresponding piControl run from each historical + SSP5-8.5 time series. We also compare model results to multiple observation based analyses.

In Figure 2 we see that CM4.0 shows a nearly flat thermosteric sea level during 1940–1990, with similar behavior found for many other CMIP6 models discussed by Jevrejeva et al. (2021). In contrast, the CM4X simulations better align with the thermosteric sea level rise found by the observations during this period. Recent decades have seen an upward acceleration of thermosteric sea level rise (Dangendorf et al., 2019), with the CM4.0 and CM4X simulations also showing an acceleration. However, the models appear to overestimate the observational trend since 1990, indicative of the large transient climate sensitivity found in CM4.0 (Winton et al., 2020). We qualify this point by noting the models are mostly within the observational product uncertainty ranges, and further updates of Ishii et al. (2017), including data up to 2022, diverge from the other observations and align more closely with the models.

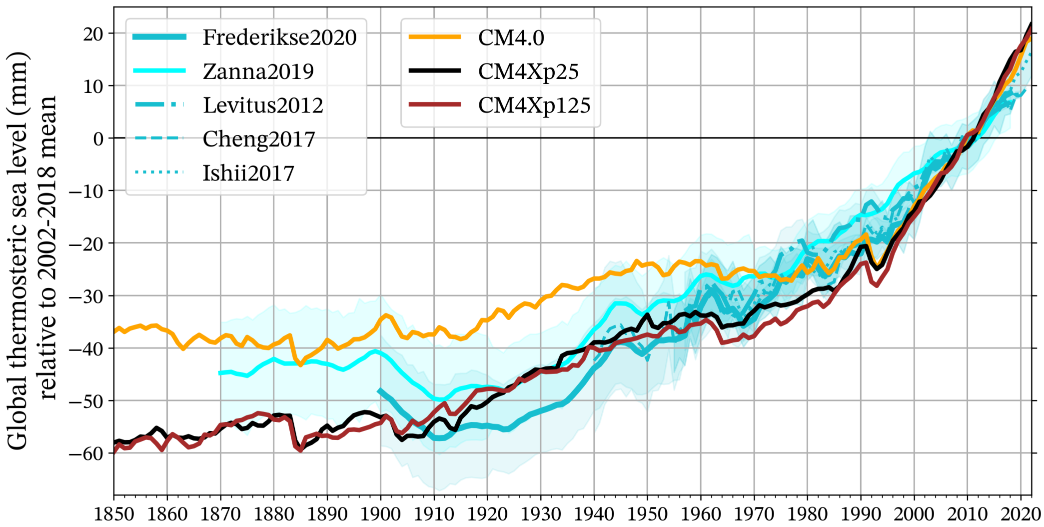

Figure 2. Historical and projected thermosteric sea level derived from CM4.0 and CM4X simulations, along with observational estimates. Models are plotted over the entire historical period (1850-2014), including years 2015-2022 from the SSP5-8.5 projections. Observational estimates are depicted by cyan-shaded lines with dark-solid (Frederikse et al., 2020), light-solid (Zanna et al., 2019), dash-dotted (Levitus et al., 2012), dashed (Cheng et al., 2017) and dotted (Ishii et al., 2017). The global mean thermosteric sea level estimates by Levitus et al. (2012), Cheng et al. (2017) and Ishii et al. (2017) do not include the deep ocean contribution (below 2000 meters). Note that the themosteric sea level time series for CM4.0 and CM4X were detrended by subtracting the long-term linear trend in the piControl from the combined historical and SSP5-8.5 scenario time series. The linear trend was derived from a linear fit over the time period of the piControl run, matching the branch-off year and duration of the historical and SSP5-8.5 scenario simulations (250 years).

2.2 Statistical measures of dynamic sea level fluctuations

We consider statistical properties of the daily mean dynamic sea level (`zos` as in Griffies et al. (2016)), thus allowing for a quantitative characterization of spatial structure of sea level fluctuations. In particular, we focus on the standard deviation and skewness computed over the 20-year segment 1995-2014 of the historical simulation. In fact, the standard deviation was already presented in Figure 3 of Part I (Griffies et al., 2024) as part of our discussion of mesoscale eddy activity. Here we present the skewness.

To compute the statistics, we generate a 20-year climatology of the daily mean dynamic sea level over the historical period 1995-2014, denoted by $\overline{\texttt{zos}}$. We then compute anomalies relative to the climatology

$$\texttt{zos}'(t_n) = \texttt{zos}(t_n) - \overline{\texttt{zos}}(t_{\text{mod}(n,365)}), \tag{1}$$

where t_n is the day within the $N = 20 * 365$ total number of days, and $t_{\text{mod}(n,365)}$ is the climatological day. The standard deviation and skewness, computed at each horizontal ocean grid cell, are given by

$$s = \left[\frac{1}{N}\sum_{n=1}^{N}[\texttt{zos}'(t_n)]^2\right]^{1/2} \quad \text{and} \quad \mathcal{S} = \frac{\sum_{n=1}^{N}[\texttt{zos}'(t_n)]^3}{N\, s^3}. \tag{2}$$

The sea level standard deviation has dimensions of length whereas skewness is dimensionless. We compare results from CM4X-p25 and CM4X-p125, and include the 1/12° GLORYS12 analysis from Lellouche et al. (2021) as a benchmark.

2.3 Skewness

Skewness is a third-order statistic that quantifies the asymmetry of a distribution. A positive skewness means that fluctuations are biased positive relative to a Gaussian distribution, and vice versa for negative skewness. K. R. Thompson and Demirov (2006) and Hughes et al. (2010) noted that sea level skewness is positive on the poleward side of strong eastward currents (e.g., Gulf Stream, Kuroshio) and negative on the equatorward side, so that strong currents are generally aligned with the zero contour.

As seen in Figure 3, CM4X-p25 contains no clear zero skewness signature of the Gulf Stream jet, contrary to that found in GLORYS12. CM4X-p125 shows some hint of a zero skewness contour, but far less coherent than in GLORYS12. The Kuroshio Current in CM4X-p25 is also poorly revealed by the CM4X-p25 skewness, whereas the skewness in CM4X-p125 resembles GLORYS12 though with a muted signature. The Agulhas region in the CM4X simulations suffers from the complement bias found in the western boundary currents. Namely, Agulhas eddies in the CM4X simulations remain somewhat more coherent than found in GLORY12, thus producing a nontrivial positive skewness signature reaching into the central portion of the South Atlantic, with this signature in the models far larger than found in GLORYS12.

As noted by K. R. Thompson and Demirov (2006), tropical sea level skewness is dominated by large patterns associated with variability such as the El Niño-Southern Oscillation and Indian Ocean variability. The CM4X models generally show a muted tropical variability relative to GLORYS12, which is consistent with a muted tropical sea surface temperature variability as revealed by the power spectra in Figure 10 of Part I (Griffies et al., 2024). Correspondingly, the positive skewness in GLORYS12 extending along the coasts of North and South America is missing in both CM4X-p25 and CM4X-p125.

2.4 Conclusions regarding sea level

Figure 2 shows that both CM4X simulations reduce thermosteric sea level biases found in CM4.0, and yet that result is possibly due to simplifications of the land model as detailed in Section 2.4 and Appendix A of Part I (Griffies et al., 2024). The patterns for dynamic sea level standard deviation (Figure 3 in Part I) and skewness expose nontrivial biases in the middle latitude boundary currents. These biases are reduced with CM4X-p125 relative to CM4X-p25, and yet they suggest a need for either improved parameterizations or, as emphasized by Chassignet and Xu (2017), substantially finer grid spacing. We find a highly muted CM4X tropical variability as revealed by the sea level skewness, thus revealing how sea level skewness complements the sea surface temperature power spectra shown in Figure 10 of Part I (Griffies et al., 2024), which also shows muted tropical sea surface temperature variability.

3 Eastern boundary upwelling systems

Eastern boundary upwelling systems (EBUS) are among the most biologically productive areas of the World Ocean (Strub et al., 2013). They are characterized by a sharp drop in the sea surface temperature near the coast, which results from upwelling of cooler interior waters through Ekman suction and lateral Ekman transport driven by equatorward wind stresses. In Figure 4 we present summer SST taken from CM4X-p125. By showing the summertime season we clearly expose the cool upwelling waters in contrast to the warm surrounding waters. As summarized by Richter (2015), many climate models exhibit large biases in the SST in eastern boundary regions due to both a lack of cloud

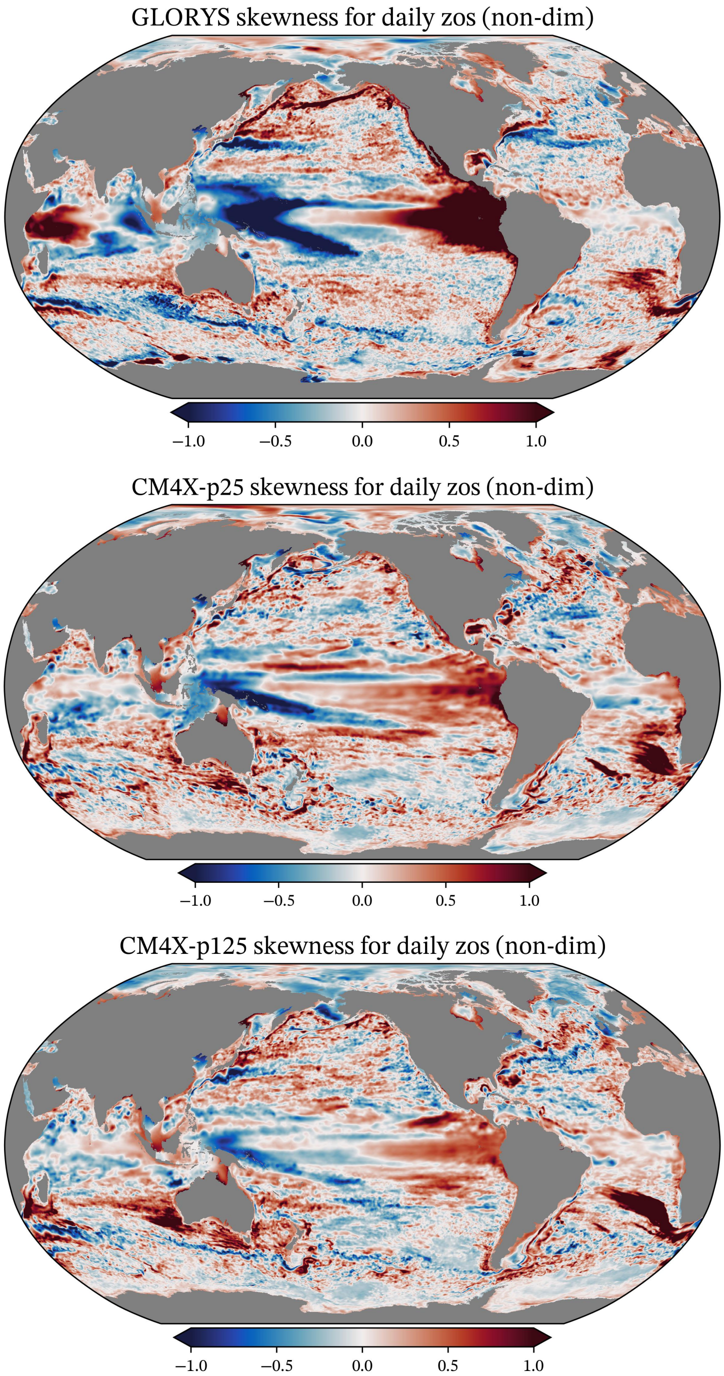

Figure 3. Skewness (non-dimensional) for the daily mean dynamic sea level from GLORYS12 (Lellouche et al., 2021) (top panel), and CM4X-p25 (middle panel) and CM4X-p125 (lower panel). Each figure is created from the deviation of the daily mean sea level relative to the climatological mean for that day of the year, as detailed in Section 2.2. Each climatology is created from years 1995-2014.

cover (in particular stratocumulus decks) and weaker than observed upwelling favorable winds (see also C. Wang et al. (2014)). The CM4X models are similarly lacking the appropriate cloud cover in EBUS. Compared to cloud cover estimates from Kaspar et al. (2009), both CM4X-p25 and CM4X-p125 have at best 20% less cloud cover over EBUS and as much as 40% less in regions in the Pacific systems, thus resulting in the warm SST biases presented in Figure 13 of Part I (Griffies et al., 2024). Gordon et al. (2000) shows that a more realistic marine stratocumulus significantly improves the annual cycle of SST and ocean dynamics in the tropics. Representation of upwelling favorable winds is the second source of SST biases in EBUS. In this section, we focus on improvements to SST biases from the 50 km atmosphere in CM4X models.

3.1 Modeling eastern boundary upwelling systems

Because of their importance to the marine-based economy, eastern boundary upwelling systems have been extensively studied with ocean and climate models. The impact of wind stress on upwelling characteristics is typically addressed in a regional context, such as the studies by Albert et al. (2010), Jacox and Edwards (2012), Junker et al. (2015), Small et al. (2015), and Sylla et al. (2022). Furthermore, sensitivity of upwelling regions to climate change is a topic of great interest, such as studied by Bakun et al. (2015), Rykaczewski et al. (2015), and Bograd et al. (2023). We here consider the representation of upwelling in the CM4X climate models during their historical simulations. We focus on the four upwelling systems shown in Figure 4: California and Peru in the Pacific and Canary and Benguela in the Atlantic.

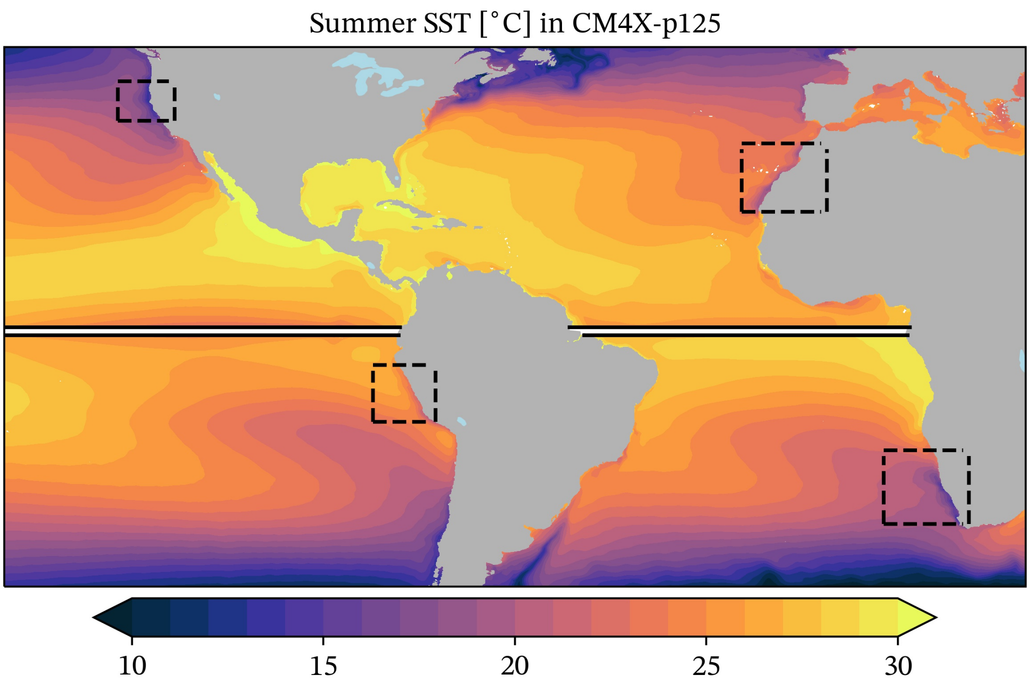

Figure 4. Sea surface temperature from CM4X-p125, during summer months (July-September in northern hemisphere, January-March in southern hemisphere) averaged over years 1980-2014. The boxes indicate the four eastern boundary upwelling systems considered in our analysis: California, Peru, Canary and Benguela. We place a white stripe at the equator since we map the summer months for both hemisphere, and so there is a jump at the equator.

Because of the fine spatial scales involved in coastal upwelling, both atmosphere and ocean models with relatively coarse grids are limited in their ability to capture the relevant dynamical processes. Fine resolution is needed in the atmosphere to represent the coastal wind stresses and their curl, and fine resolution is needed in the ocean to realize upwelling localized near the coast. Varela et al. (2022) found that refined grid spacing allows many CMIP6 climate models to improve their simulation of coastal SST relative to earlier model classes. For example, the CM4.0 climate model captures the imprint of upwelling using its C96 atmosphere (approximately 100 km) and 0.25° ocean. Further refining the atmospheric grid to C192 (approximately 50 km) in CM4X leads us to the question of its impact on the upwelling systems, as does refinement of the ocean grid from CM4X-p25 to CM4X-p125. To help address these questions we include a companion experiment, CM4X-p25-C96, which uses the C96 atmosphere along with the same ocean model configuration as CM4X-p25.

3.2 Winds and SSTs in the upwelling regions

In Figure 5 we show the alongshore wind stress and wind stress curl as a function of the distance to the coast and as averaged over the four upwelling regions. The alongshore/equatorward wind stress remains stronger nearshore with the C192 atmosphere, with a drop confined to the inner 100 km from the coast instead of 150 km found in CM4X-p25-C96. This result holds for both CM4X-p25 and CM4X-p125, which is expected since the ocean grid spacing is not a leading order effect on the wind stress near the coast. The sharper wind drop near the coast found with the C192 atmosphere results in a greater wind stress curl in the nearshore region (within 50 km from the coast) and a decrease in the 50-100 km band. The CM4X experiments are compared with the SCOW estimates (Risien & Chelton, 2008) based on QuikSCAT satellite measurements. Except for the California system, modeled alongshore wind stresses are stronger than observed in the 100-300 km band. In all regions, the modeled wind drop-off at the coast is steeper than in the SCOW estimates, resulting in a stronger wind stress curl at the coast.

The distinct alongshore wind and wind stress curl lead to differences in the resulting SST profiles as averaged over the four eastern boundary regions (Figure 6). The CM4X results are typically biased warm offshore in the Pacific, whereas they are in better agreement in the Atlantic. The C192 atmosphere leads to an SST that drops faster in the 50-150 km band in all eastern boundary regions. This result has a favorable impact in the California system where it compensates for a warm offshore bias, and yet the SST gradient is stronger than observed. In the Peru system, the SST gradient is in good agreement with observations and leads to a much improved simulation using the C192 atmosphere. In the Benguela system, the SST drop at the coast is not present using the C96 atmosphere, whereas the CM4X experiments using the C192 atmosphere are superior. Finally, the stronger SST gradient in the CM4X simulations with the C192 atmosphere overshoots the observed values in the Canary system, whereas this region is better captured using the C96 atmosphere.

3.3 Conclusions regarding eastern boundary upwelling

Ekman mechanics accounts for two key processes important for the upwelling regions: the cross-shore Ekman transport is proportional to alongshore wind stress (Smith, 1968) and Ekman suction is proportional to the wind stress curl (Enriquez & Friehe, 1995; Pickett & Paduan, 2003). As noted in Jacox and Edwards (2012), cross-shore Ekman transport dominates in a narrow coastal band (within 50 km of the coast), whereas Ekman suction creates small but important upwelling velocities in a broader area extending from outside the coastal band to 200-300 km offshore (Jacox & Edwards, 2012).

Results from the CM4X experiments suggest that reduction in wind stress curl in the 50-200 km band, and its intensification in the narrow nearshore area, favor a stronger cross-shore SST gradient. In the CM4X models using the C192 atmosphere, Ekman suction is concentrated in the nearshore area at the deficit of the broader offshore region.

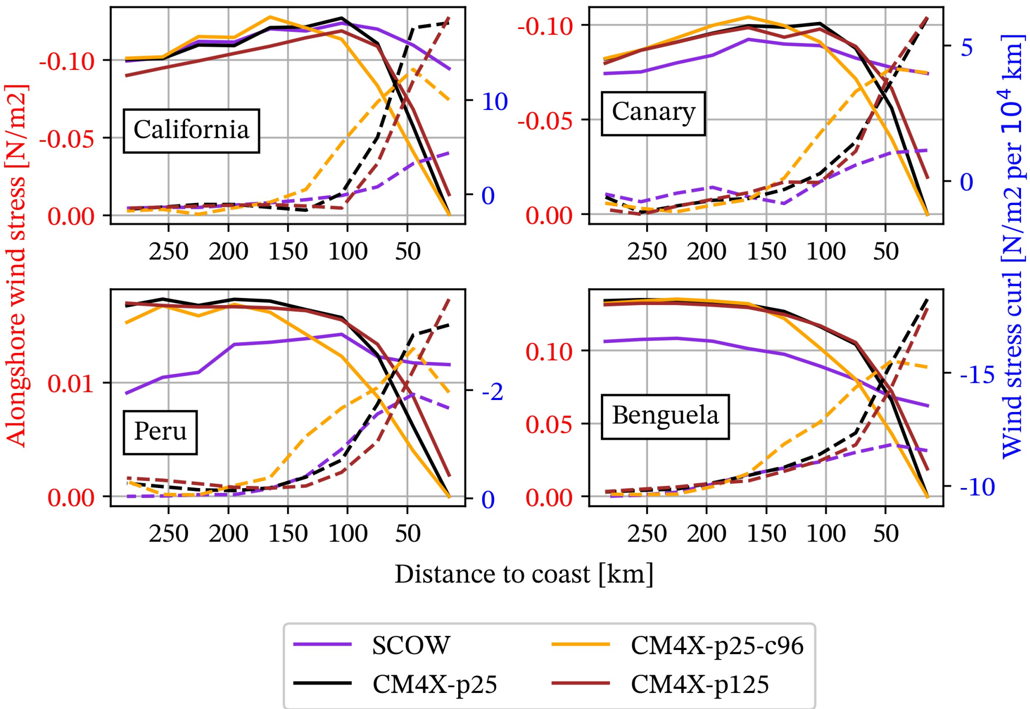

Figure 5. Alongshore wind stress (solid lines/red ticks) and wind stress curl (dashed lines/blue ticks) in the four major eastern boundary upwelling systems during the summer months of 1980-2014. For coastal upwelling at these latitudes, the β term (see equation (2) of Taboada et al. (2018)) is neglected. Gridded data is averaged over the boxes of Figure 4. Note the distinct vertical axes: signs can be reversed so that wind stress decreases at the coast and wind stress curl increases. CM4X-p25-c96 uses the C96 atmosphere model, whereas the other models us the C192 atmosphere. Satellite measurements from SCOW (Risien and Chelton (2008)) are added for reference although they cover a shorter time period (1999-2009). Note the distinct vertical axes on the panels.

The cross-shore Ekman transport with the C192 atmosphere remains strong closer to the coast than with the C96 atmosphere. This result suggests that strengthening of the Ekman transport part of the upwelling in the 50-150 km band, in conjunction with the concentration of the wind stress curl at the coast, favors the upwelling. These results are consistent with previous works from Gent et al. (2010) and Small et al. (2015) in which intensification of the coastal winds leads to stronger upwelling at the coast and overall reduction in SST bias.

The simulated cross-shore Ekman transport at the coast is weaker than the SCOW estimates but stronger offshore in all but the California system, with an intersect ranging from 50 to 100 km depending on the region and experiment considered. Since cross-shore Ekman transport is expected to dominate in the inner 50 km to the coast, this dominance should lead to decreased coastal upwelling, which contrasts to the CM4X SST gradients of Figure 6. In addition, the Ekman suction part of the upwelling is remarkably consistent with satellite estimates in the offshore part (distance $>$ 100 km) with the C192 atmosphere. The C96 atmosphere departs largely from these estimates in the 100-150 km band, which should lead to more Ekman suction upwelling velocities. Even so, we do not find this upwelling signal appears in the SST gradients. We suspect that the biases are

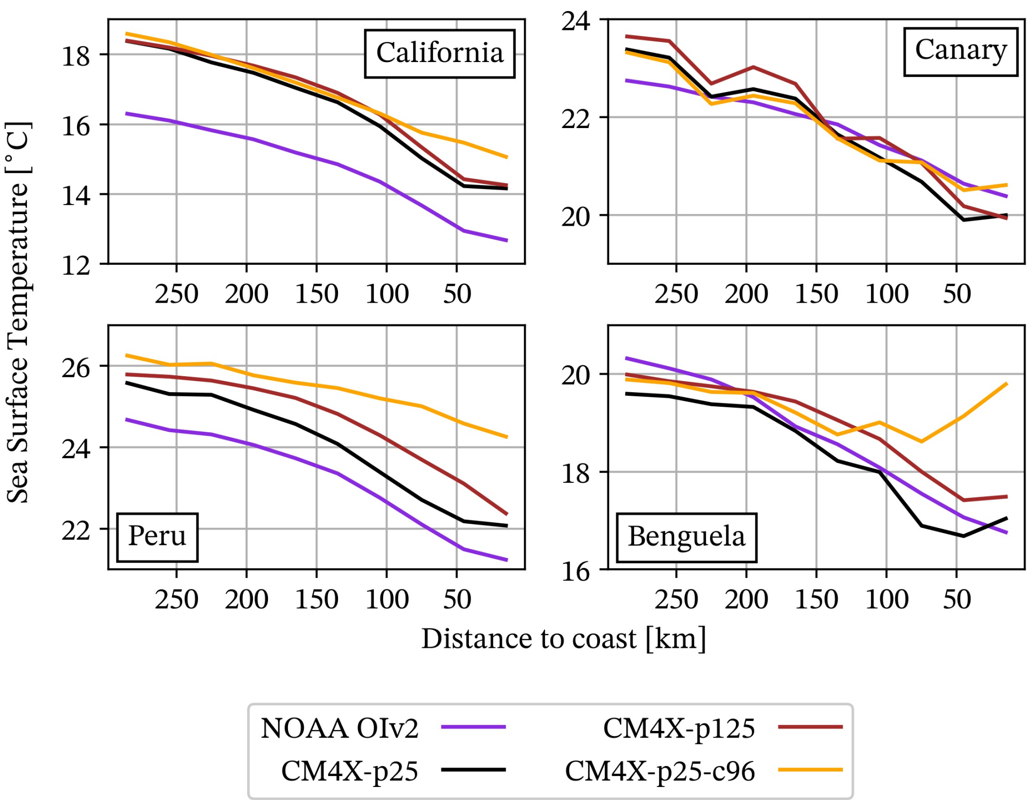

Figure 6. Sea surface temperature in the four major eastern boundary upwelling systems during the summer months of 1980-2014. Results are averaged over the boxes of Figure 4. We include the satellite estimate from NOAA OIv2 (Huang et al., 2020) over the period 1982-2014. Note the different temperature scales on the vertical axes.

modified only slightly due to the more dominant issues with atmospheric model's biases in representing low level clouds in the eastern boundary upwelling regions, with these biases a well known feature of atmospheric models (e.g., Richter (2015); Ceppi et al. (2024)) that remain a topic of ongoing research.

4 Sea ice in the Arctic Ocean and Southern Ocean

In this section, we analyze the simulated Arctic Ocean and Southern Ocean ("Antarctic") sea ice within the CM4X model hierarchy, focusing on the historical and SSP5-8.5 experiments. We also include results from CM4.0 as a point of comparison. To maximize use of the most recent observations, which include record low sea ice conditions, we compute climatologies and trends over 1979-2023 (as opposed to 1979-2014, with 2014 the end of the historical experiment). Doing so requires appending years 2015-2023 from the SSP5-8.5 scenario simulations to the historical simulations. We consider this additional model forcing uncertainty worth the benefits of an additional nine years of observational data.

4.1 Arctic Ocean sea ice

Figure 7a plots Pan-Arctic sea ice extent (SIE) climatologies computed over 1979–2023 across the model hierarchy. The seasonal cycle of Pan-Arctic SIE is well simulated

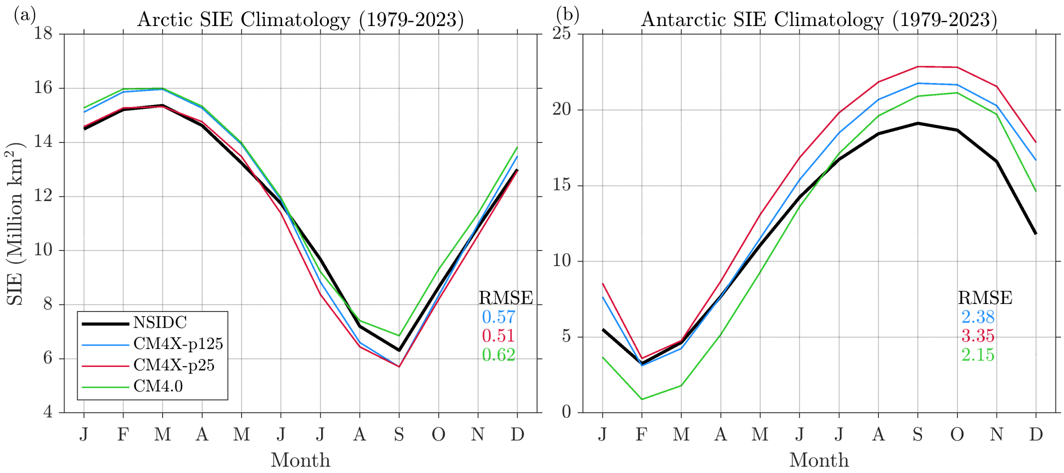

Figure 7. Pan-Arctic and Pan-Antarctic sea ice extent (SIE) climatologies (10^6 km^2) computed over the years 1979–2023 from satellite observations (black), CM4X-p125 (blue), CM4X-p25 (red), and CM4.0 (green). Pan-Arctic and Pan-Antarctic SIE are defined as the area integral of all grid cells covered by at least 15% sea ice concentration (SIC) in the northern and southern hemispheres, respectively. Root-mean-square error (RMSE) are computed between the simulated and observed seasonal cycles and shown in colored text. The observed SIE is computed using passive microwave satellite sea ice concentration observations from the NOAA/National Snow and Ice Data Center (NSIDC) Climate Data Record (CDR) of SIC, version 4 (Data Set ID: G02202; Meier et al. (2021)).

by both CM4X models and both models show improvements relative to the CM4.0 simulation. The RMS errors of the Pan-Arctic SIE climatology are 0.57×10^6 km^2 and 0.51×10^6 km^2 in the CM4X-p125 and CM4X-p25 models, respectively, which are lower than the median CMIP5 model RMSE of 1.45×10^6 km^2 (Shu et al., 2015) and the CM4.0 RMSE of 0.62×10^6 km^2. CM4X-p25 has Pan-Arctic SIE improvements in non-summer months relative to CM4.0, and has a negative summer bias which is similar in magnitude to CM4.0's positive summertime bias. CM4x-p125 is generally similar to CM4.0 in winter and spring and similar to CM4X-p25 in summer and autumn months.

The climatological sea ice concentration (SIC) biases of the models are shown in Figure 8. The spatial pattern of winter Arctic SIC bias is similar across the models (top row in Figure 8), implying that the improved winter Pan-Arctic SIE simulation of CM4X-p25 primarily results from cancellation of positive and negative SIC errors. In winter months, the models have positive biases (too much ice coverage) in the Greenland-Iceland-Norwegian (GIN), Barents, and Bering Seas and negative biases in the Labrador Sea and the Sea of Okhotsk. These winter SIC biases closely mirror the patterns of SST bias (compare with Figure 14 from Part I of Griffies et al. (2024)). The notable positive SIC bias in the GIN and Barents Seas has been persistent across earlier generations of GFDL models, such as CM2.1, ESM2M, ESM2G, and CM3 (Delworth et al., 2006; Dunne et al., 2012; Griffies et al., 2011), possibly related to a combination of too much poleward ocean heat transport and too much ice export through Fram Strait (see Figure 9). The finer grid spacing in CM4X-p125, with its enhanced mesoscale eddy activity, does not ameliorate this bias.

The CM4X models have similar patterns and magnitudes of summer SIC bias (second row of Figure 8), which differ from the SIC biases of CM4.0. The main improvement in CM4X is the reduced positive bias in the Beaufort, Chukchi, and East Siberian Seas.

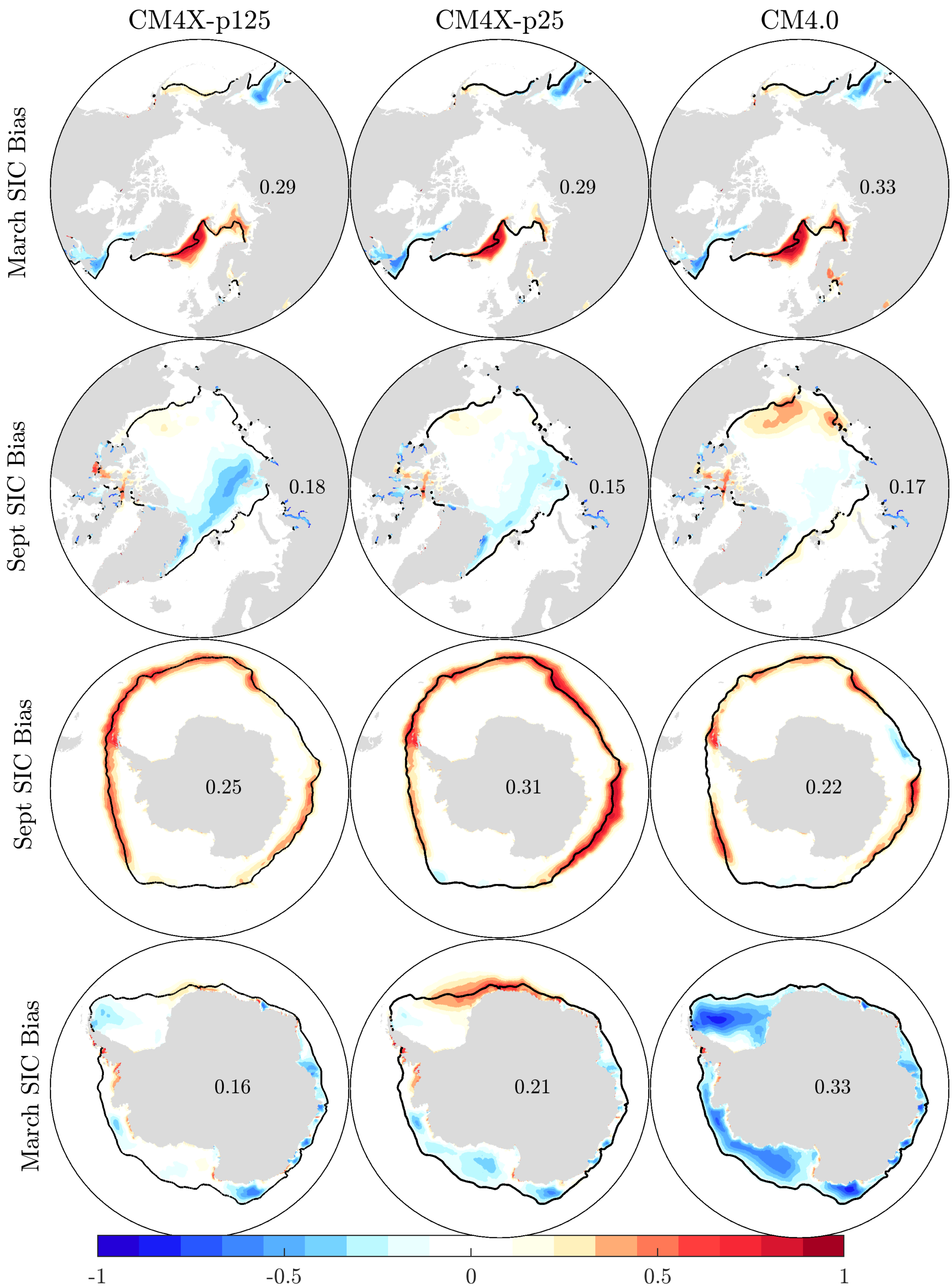

Figure 8. Arctic and Southern Ocean sea ice concentration (SIC) climatological biases (model minus NSIDC CDR observations) in March and September computed over 1979–2023. The black contours indicate the observed climatological sea ice edge position (15% SIC contour). Black text indicates the SIC RMSE area-averaged over the zone of SIC variability, defined as all grid points where the model or observed monthly SIC standard deviation exceeds 5%.

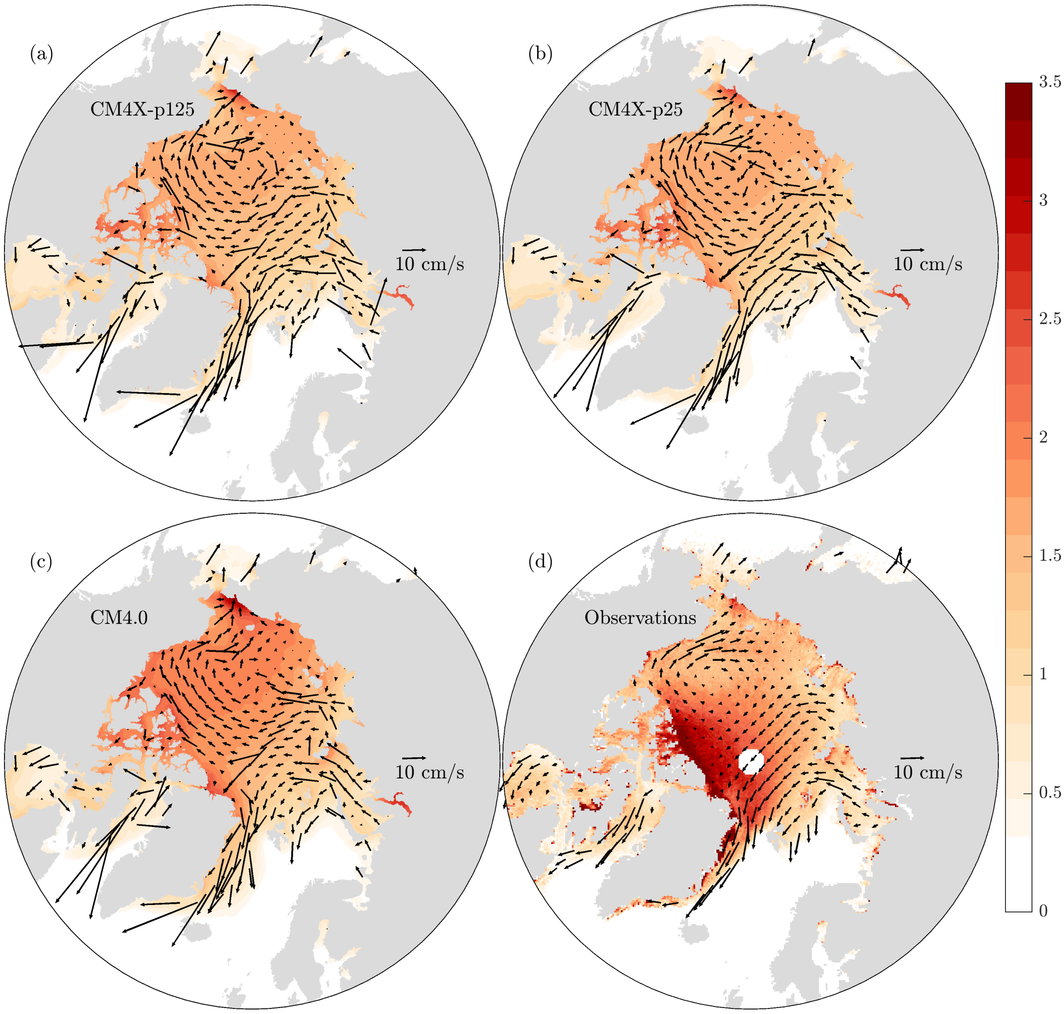

Figure 9. March Arctic sea ice thickness (SIT) climatologies (m; shading) and climatological DJFM sea ice drift (cm/s; vectors) in CM4X, CM4.0, and observations computed over the period 2011–2023. The SIT observations come from the Alfred-Wegener-Institute monthly SIT product (Ricker et al., 2014) and span 2011–2023. The sea ice drift observations are from the low-resolution daily sea ice drift product of the EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSISAF) and span 2010–2023 (Lavergne et al., 2010).

This summer SIC bias in CM4.0 is associated with an erroneous pattern of winter sea ice thickness (SIT), which has the model's thickest ice located in the Beaufort, Chukchi and East Siberian Seas rather than north of Greenland and Ellesmere Island as found in the observations (see Figures 9c,d). This anomalously thick winter ice in the Beaufort, Chukchi, and East Siberian Seas leads to delayed melt in these regions, resulting in a spatially coincident positive bias in summer SIC. We find that the CM4X models share a similar bias in SIT spatial pattern, however, their mean SIT is reduced, resulting in lower summer SIC throughout most of the Arctic (Figures 9a,b). This thinner mean state also results in an exacerbated negative SIC bias in the Greenland Sea and along the northern boundaries of the Barents, Kara, and Laptev Seas, which is a degradation relative to CM4.0. One region that is unchanged is the Canadian Arctic Archipelago which has a consistent positive summer SIC bias across CM4.0, CM4X-p25, and CM4X-p125.

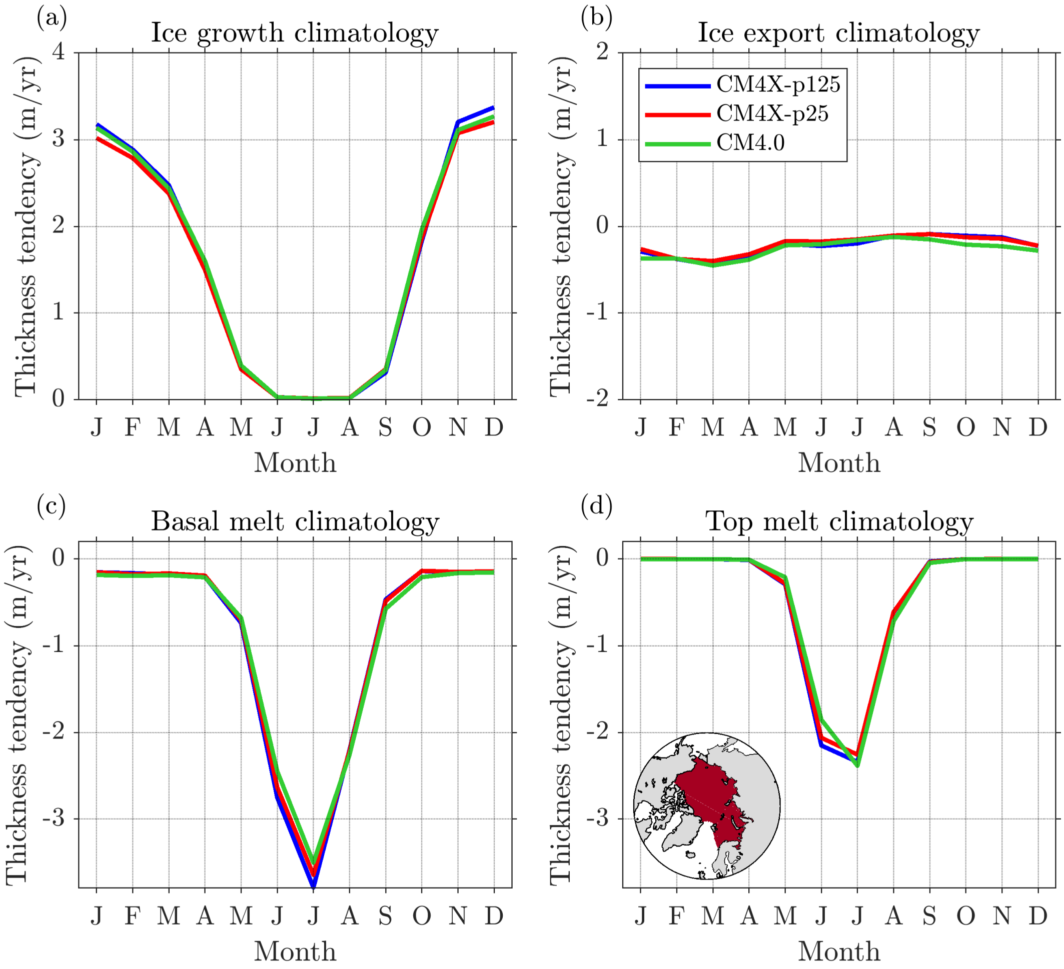

Figure 10. Arctic sea ice mass budget climatologies computed over the Central Arctic basin domain shown in panel (d) over the time period 1979–2023. The mass budget consists of sea ice thickness (SIT) tendency terms (m/yr) corresponding to congelation and frazil ice growth (a), mass transport convergence (b), basal melt (c), and top melt (d). Positive values correspond to mass gain and negative values correspond to mass loss. The y-axis range is the same across the panels, thus allowing for direct comparison of the various terms.

Figure 10 shows Arctic sea ice mass budget climatologies computed following the methodology of Keen et al. (2021), which averages mass budget terms over a Central Arctic domain (see inset of Figure 10d). This domain encompasses the region of thickest Arctic ice and its boundaries include all flux gates to the North Atlantic and North Pacific sectors. The sea ice mass budget consists of a dynamic tendency term associated with ice mass transport convergence (export) and thermodynamic tendency terms associated with congelation and frazil ice growth, basal melt, and top melt. All terms are defined such that positive values indicate mass gain and negative values indicate mass loss and are expressed as a thickness tendency in m/yr (note that the SIS2 model uses a constant sea ice density of 905 kg/m^3). The thinner Arctic ice in CM4X primarily results from increased summer basal and top melt relative to CM4.0. The CM4X models have less mass loss due to ice export in autumn. This reduced autumn ice export in CM4X is likely associated with the thinner and less extensive mean state in these models, as their ice drift patterns are similar to CM4.0 (see Fig. 9). CM4X-p125 has a similar SIT mean state to CM4X-p25, but has more winter growth, potentially associated with enhanced ice formation in sea ice leads in this higher resolution model.

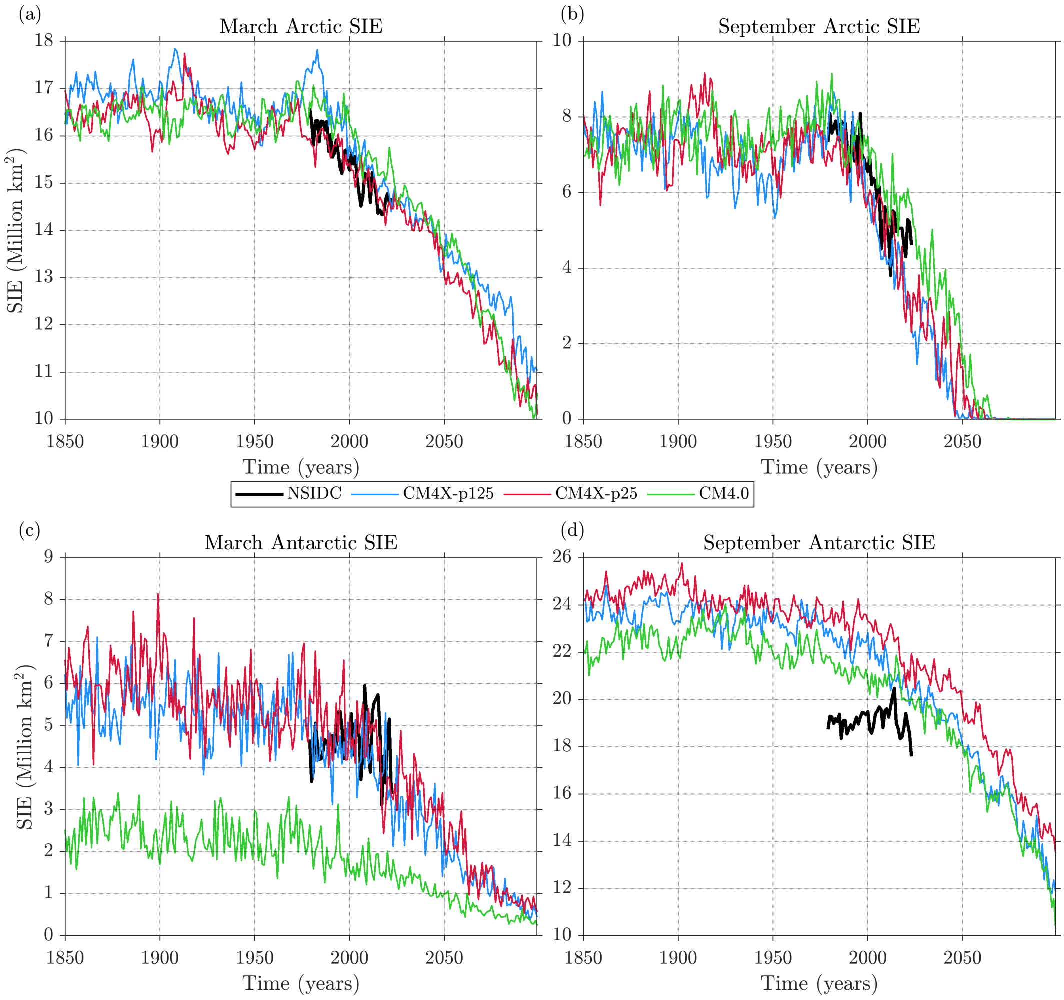

Figure 11. Time series of Arctic and Southern Ocean sea ice extent (SIE) in March and September in CM4X-p125 (blue), CM4X-p25 (red), CM4.0 (green), and NSIDC observations (black) in March and September. The simulations use CMIP6 Historical (1850–2014) and SSP5-8.5 forcings (2015–2099).

We next consider the time evolution of Arctic SIE and sea ice volume (SIV) in Figures 11 and 12. Each model simulates SIE trends in March and September in reasonable quantitative agreement with the observed SIE decline. Given the large degree of internal variability in Arctic SIE trends, we do not expect a perfect match between observations and a single model realization (Jahn et al., 2016; DeRepentigny et al., 2020). The trend differences between the CM4X and CM4.0 historical simulations and observations are smaller than the typical ranges estimated by single-model initial conditional large ensembles performed with other GCMs (Horvat, 2021), suggesting that the CM4X and CM4.0 models are not inconsistent with observed trends. There are some trend differences between the CM4X and CM4.0 models, but multi-member ensembles are necessary to determine if differences are statistically robust. CM4X and CM4.0 each simulate substantial decadal-to-multidecadal variability over the 20th century. This low-freqeuncy

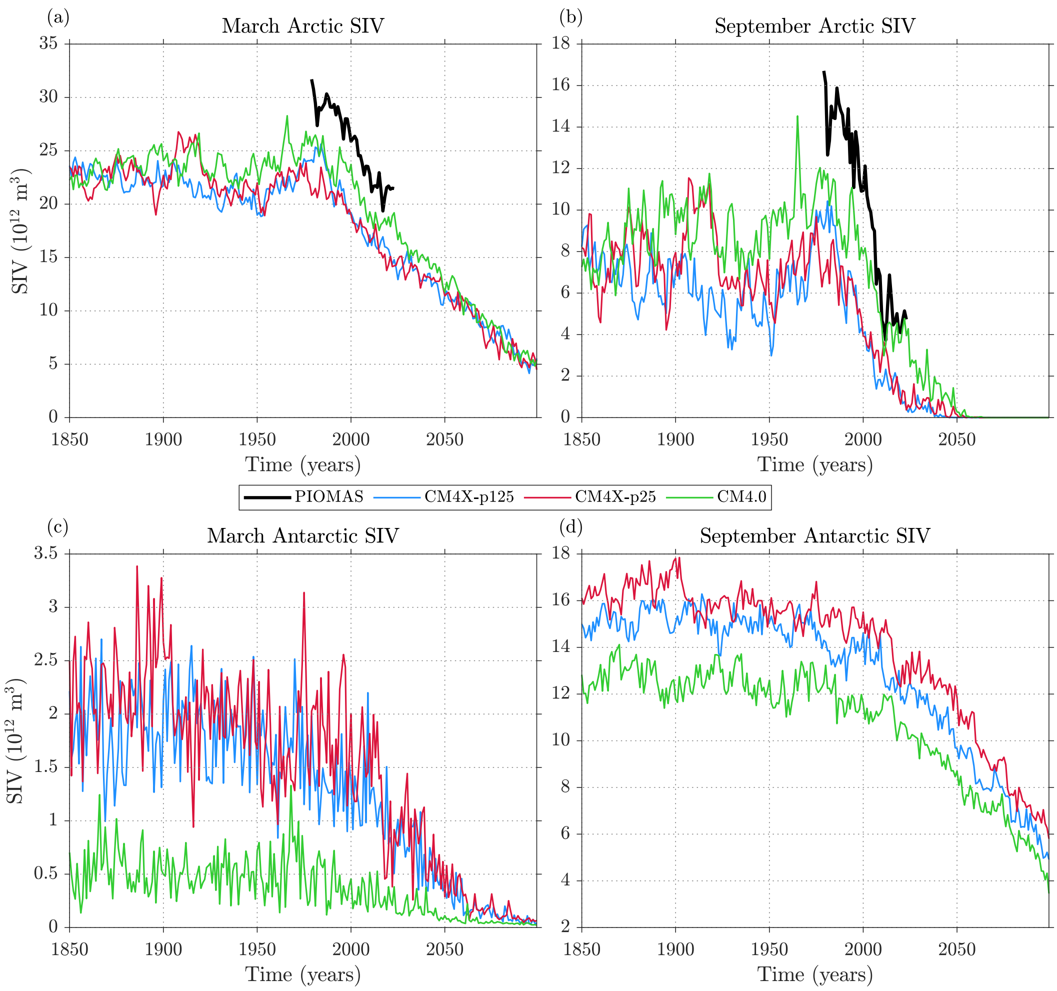

Figure 12. Time series of Arctic and Antarctic sea ice volume (SIV) in March and September in CM4X-p125 (blue), CM4X-p25 (red), CM4.0 (green), and PIOMAS sea ice thickness (SIT) reanalysis (black) in March and September. The simulations use CMIP6 Historical (1850–2014) and SSP5-8.5 forcings (2015–2099). The PIOMAS data spans 1979–2023 and is based on an assimilation system that incorporates SIC, SST, and atmopsheric reanalysis constraints (J. Zhang & Rothrock, 2003).

variability becomes notably muted under the high-forcing SSP5-8.5 scenario used over the 21st century.

Figure 12 shows that the models are biased thin relative to the Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS; J. Zhang and Rothrock (2003)) SIT reanalysis, a product that has reasonably good agreement with available in situ, aircraft, and satellite observations of SIT (X. Wang et al., 2016; Landy et al., 2022). The CM4X models have a larger thin bias than CM4.0, but similar SIV timeseries to each other. This result suggest that the thinner ice in CM4X-p125 relative to CM4.0 is not the result of refined ocean grid spacing. The models simulate a strong decline of Arctic SIV in all months of the year (Figure 12). Despite their mean state SIV biases, the models simulate similar rates of historical SIV loss to PIOMAS.

Under the SSP5-8.5 forcing scenario, the models simulate a complete loss of summer Arctic SIE and SIV over the 21st century. The first ice-free summers occur in the years 2040, 2038, and 2052 in the CM4X-p125, CM4X-p25, and CM4.0 models, respectively (defined here as SIE $< 10^6$ km^2). This ice-free timing is consistent with the ice-free range of 2015–2052 as estimated by selected CMIP6 models (SIMIP Community, 2020; Jahn et al., 2024). All three models also reach ice free conditions in the months of July–November by the year 2100 (not shown).

4.2 Southern Ocean sea ice

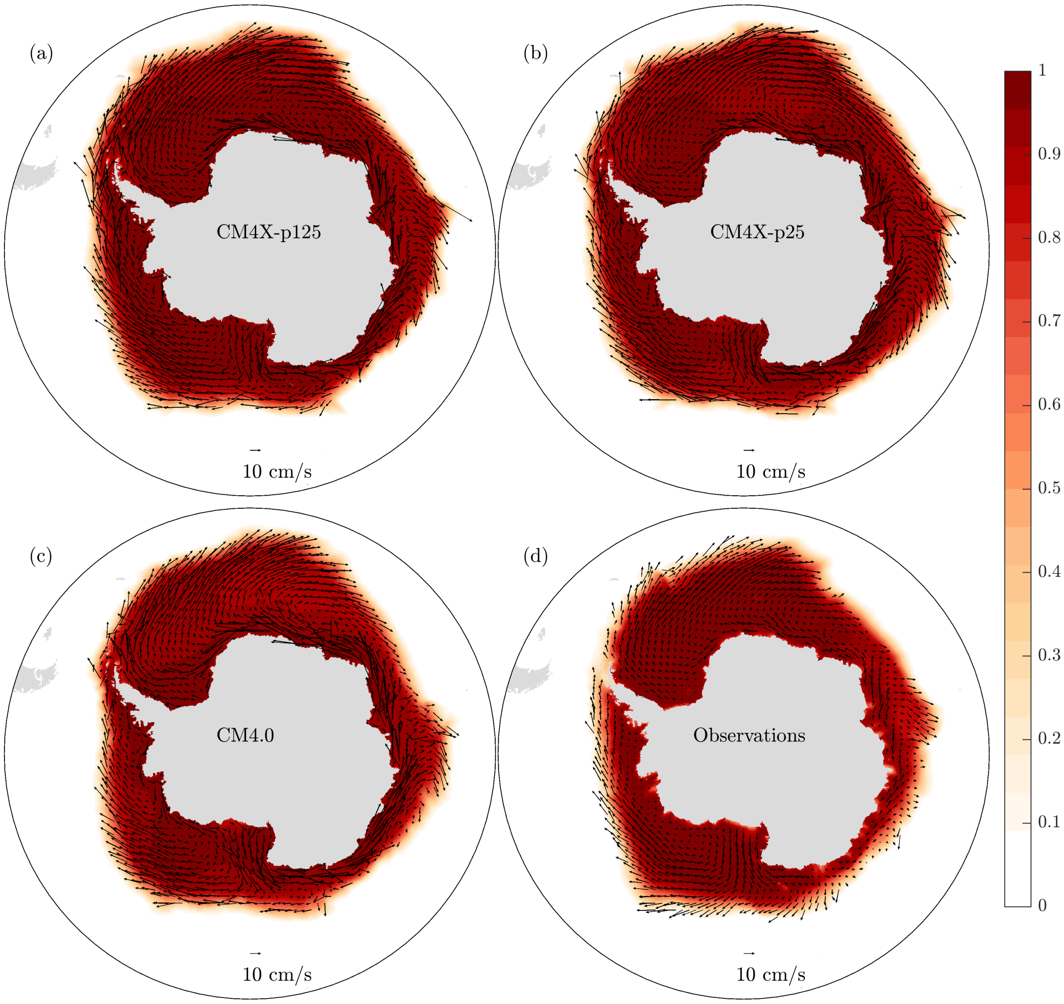

Figure 13. September Southern Ocean SIC climatologies (shading) and climatological JJAS sea ice drift (cm/s; vectors) in CM4X, CM4.0, and observations computed over the period 2011–2023. SIC and sea ice drift observations are from NSIDC and OSISF, respectively. Note that the OSISAF southern hemisphere drift observations span the period 2013–2023.

CM4.0, CM4X-p25, and CM4X-p125 each capture the asymmetric seasonal cycle of Southern Ocean SIE, with five months of ice retreat and seven months of ice advance (Figure 7b). Compared to the Arctic, the models' SIE biases are generally larger in the Southern Ocean. CM4.0 has an exaggerated Southern Ocean SIE seasonal cycle with too little sea ice in austral summer and too much sea ice in austral winter. The CM4X models have more extensive Southern Ocean sea ice than CM4.0 in all months of the year,

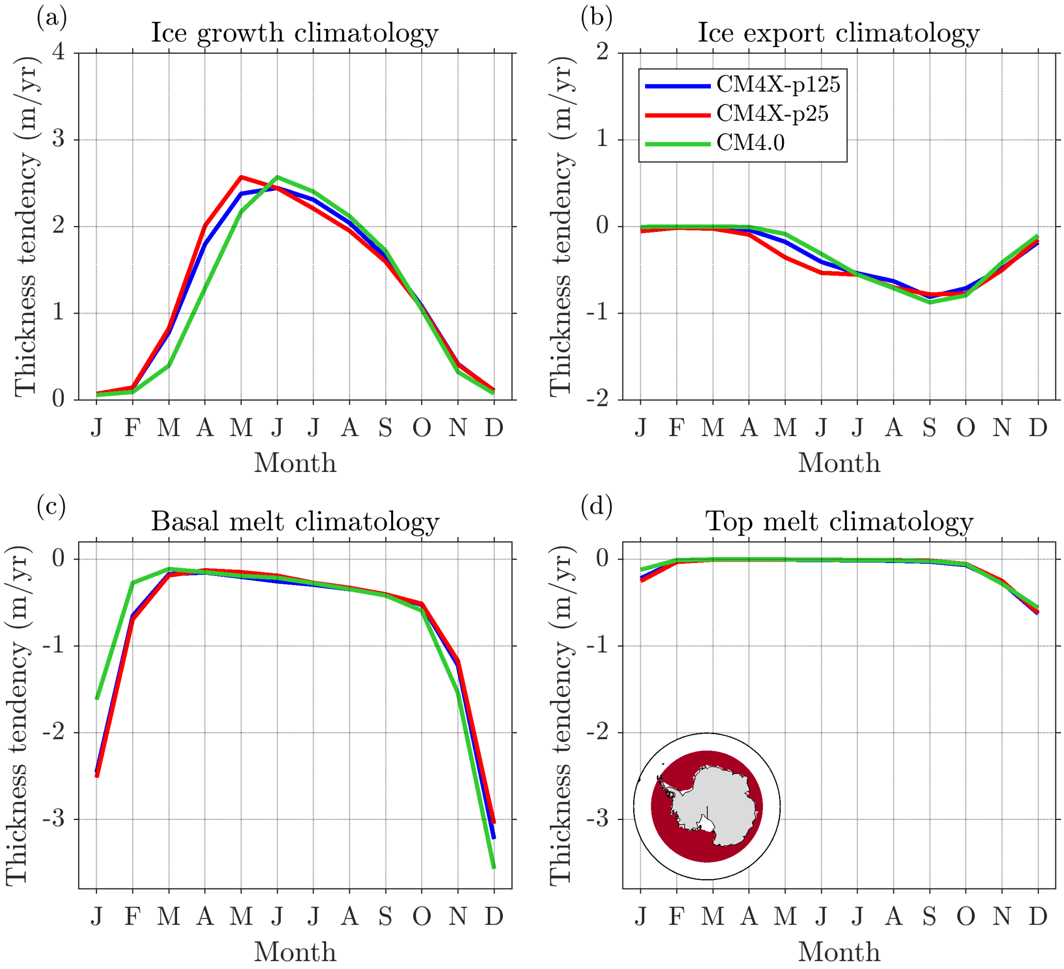

Figure 14. Southern Ocean sea ice mass budget climatologies computed over all gridpoints south of 63°S as shown in panel (d) over the period 1979–2023. The mass budget consists of SIT tendency terms (m/year) corresponding to congelation and frazil ice growth (a), mass transport convergence (b), basal melt (c), and top melt (d). Positive values correspond to mass gain and negative values correspond to mass loss. The y-axis range is the same across the panels, thus allowing for direct comparison of the various terms, and for comparing to the Arctic mass budget in Figure 10.

likely associated with the increased near-infrared land ice albedo values that were used in CM4X in order to promote Southern Ocean ventilation and production of AABW (see Appendix A4 in Part I (Griffies et al., 2024)). These higher albedos result in a cooler Southern Ocean surface climate with more sea ice than CM4.0. The increased sea ice coverage in CM4X improves upon the CM4.0 model biases in summer months yet exacerbates the winter sea ice biases. The RMS errors of the Pan-Antarctic SIE climatology are 2.38×10^6 km^2, 3.35×10^6 km^2, and 2.15×10^6 km^2 in CM4X-p125, CM4X-p25, and CM4.0, respectively, which can be compared to the CMIP5 multi-model mean RMSE of $3.42{\times}10^6$ km^2(Shu et al., 2015). A low bias in summer Southern Ocean sea ice is a ubiquitous bias across CMIP5 and CMIP6 models (Roach et al., 2020), which the CM4X models ameliorate.

Figure 8 shows climatological Southern Ocean SIC biases. We find that the spatial pattern of summer SIC is well captured by the CM4X-p125 model, whereas CM4X-p25 simulates too much sea ice in the Weddell Sea and too little in the Ross Sea (bottom row of Figure 8). All three models fail to simulate summer sea ice along the coastlines of the Indian Ocean and West Pacific sectors, which are regions with substantial landfast sea ice coverage (Fraser et al., 2023). We note that this sea ice model does not include a landfast ice parameterization. The summer SIC RMSE values of both CM4X models are reduced relative to CM4.0, which has negative SIC biases throughout the summer sea ice zone. All three models have positive winter SIC biases that are relatively circumpolar, however the biases are progressively stronger in CM4.0, CM4X-p125, and CM4X-p25 (third row of Fig. 8). CM4.0 has negative SIC biases within the sea ice pack in the Weddell Sea near Maud Rise, suggestive of too much vertical mixing and a tendency to form open-ocean polynyas in this region. These negative SIC biases are not present in the CM4X models.

The climatological winter Southern Ocean sea ice drift is shown in Figure 13. The general patterns of observed Southern Ocean sea ice drift are well captured by the models, with each model simulating northward sea ice export in the Weddell and Ross Seas, westward drift along the Antarctic coastal current, and strong eastward drift associated with the Antarctic circumpolar current. The models have drift speeds that are generally too fast relative to observations. This bias may contribute to the models' positive biases in wintertime SIC, since stronger drift implies a greater northward export of sea ice. We also find that the drift speeds along the Antarctic coastal current are notably higher than observed, especially in the zone of landfast sea ice along the eastern Antarctic coastline. This overly mobile sea ice shows that the model is unable to simulate landfast ice, and this potentially underpins the negative summer SIC biases in this region.

Figure 14 shows Southern Ocean sea ice mass budgets computed over the region south of 63°S (see inset of panel d). This region was chosen to encompass the primary zone of sea ice growth and melt while also capturing the dominant flux gates for sea ice export. Relative to CM4.0, the CM4X models show a clear shift in sea ice growth to earlier in the autumn season, consistent with the higher glacier albedos and cooler surface climate in CM4X (see Appendix A3 in Part I (Griffies et al., 2024)). The CM4X models also have more total annual sea ice growth, forming approximately an additional 0.1m of sea ice each year. The CM4X models also have more sea ice export during the autumn months, likely associated with the enhanced sea ice growth and thicker sea ice produced over these months. The dominant melt contributions come from basal melt, with the CM4X models showing a later spring onset of basal melt compared to CM4.0, consistent with the higher albedo and cooler surface climate in these runs. We also note that, compared to the Arctic, Southern Ocean sea ice has larger basal melt contributions during winter months, which tend to increase as winter mixed layers deepen.

Figures 11c,d and Figures 12c,d show time evolution of Southern Ocean SIE and SIV, respectively. Roach et al. (2020) showed that nearly every CMIP6 model simulates a negative Southern Ocean SIE trend in both summer and winter, failing to capture the observed trends which are close to zero in these seasons. This behavior is also the case for the CM4X and CM4.0 models, which simulate declines of Southern Ocean SIE in all months of the year over the period 1979–2023. This mismatch in modeled and observed trends may have contributions from missing meltwater forcing from Antarctic ice sheet and ice shelf melt (Bronselaer et al., 2018; Schmidt et al., 2023), systematic coupled model errors (Purich et al., 2016; Kay et al., 2016; Rackow et al., 2022), and internal climate variability (Meehl et al., 2016; L. Zhang et al., 2019). The CM4X and CM4.0 models simulate ice free Southern Ocean conditions (defined here as SIE $< 10^6$ km^2) in January–March by the year 2100, with CM4.0 reaching ice free states much earlier than CM4X. CM4X-p125 and CM4X-p25 have their first ice-free February in years 2049 and 2050, respectively, whereas CM4.0 simulates an ice free February states intermittently throughout the 20th century.

4.3 Conclusions regarding the sea ice simulations

Both CM4X models provide credible simulations of the Pan-Arctic sea ice mean state and trends, however the models simulate notable regional SIC errors, which are likely a combination of sea ice model physics errors and coupled model errors that originate in the atmospheric and oceanic components. It is notable that the spatial pattern of SIC and SIT model errors are very similar across the $1/4°$ and $1/8°$ CM4X configurations, and closely resemble the error patterns of the GFDL-ESM4.1 (Dunne et al., 2020) and GFDL-SPEAR (Bushuk et al., 2022) models, which have nominal horizontal grid spacings of $1/2°$ and $1°$, respectively. This similarity suggests that sea ice model errors are relatively insensitive to horizontal grid spacing across the $1/8°$–$1°$ range. Models in this range are not eddy resolving in the Arctic Ocean basin and Subpolar seas (see Figure 1 from Part I (Griffies et al., 2024)), and it is possible that a clearer impact of fine horizontal grid spacing would emerge in models that are fully eddy resolving in the Arctic. These eddy-resolving grids represent scales below the formal length scale of validity for viscous-plastic sea ice rheologies (Feltham, 2008). Even so, recent work has shown that the viscous-plastic rheology can simulate good agreement with observed sea ice drift and deformation even with 1 km grid spacing (Hutter et al., 2018). Key Arctic sea ice priorities for future model development include improving the spatial pattern of SIT, improving the magnitude and pattern of sea ice drift, and improving the persistent positive SIC bias in the GIN and Barents Seas.

Compared to the Arctic, the CM4X models have larger errors for the Southern Ocean sea ice mean state and trends, which is also a generic property across most CMIP models (Shu et al., 2015). For Southern Ocean sea ice, there appears to be a modest benefit from refined ice-ocean grid spacing, as CM4X-p125 has slightly reduced biases relative to CM4X-p25. It is notable that CM4X ice-ocean resolution does not clearly influence historical Southern Ocean sea ice trends, as CM4X-p125 and CM4X-p25 have comparable SIE trends across all months of the year. Key Southern Ocean sea ice priorities for future model development include improving the significant positive biases in wintertime SIC, reducing sea ice drift speeds, adding a representation of landfast Antarctic sea ice, improving simulated summer SIC in East Antarctica, and improving the simulation of SIE trends across all seasons.

5 Southern Ocean

The Southern Ocean plays a dominant role in anthropogenic heat and carbon uptake (Frölicher et al., 2015; Roemmich et al., 2015; DeVries et al., 2019), thus the representation of physical processes in this region is critical for accurately simulating the transient climate response. The Southern Ocean is home to the strongest current on the planet, the Antarctic Circumpolar Current (ACC), which acts as the primary pathway for inter-basin exchange of physical and biogeochemical tracers. Intimately linked to the structure of the ACC, the Southern Ocean is also home to a meridional overturning circulation whose deep branch ventilates the densest waters in the World Ocean, namely the Antarctic Bottom Water (AABW), and whose intermediate branch plays a large role in the oceanic sequestration of anthropogenic heat and carbon (see Morrison et al. (2022) for a review of physical processes). Additionally, the waters on and just offshore of the Antarctic continental shelf in the subpolar Southern Ocean directly influence the mass balance of the Antarctic Ice Sheet and thus dynamics in this region exert a strong influence on global sea level rise (Paolo et al., 2015). Indeed, such concerns about sea level rise have placed a growing appreciation for the important role of Antarctic shelf processes in the global climate system. Recent work has emphasized the need for improved model representation of ocean dynamics near to and along the Antarctic shelf, including the Antarctic Slope Current (ASC) and Antarctic Coastal Current (ACoC) (A. F. Thompson et al., 2018; Moorman et al., 2020; Purich & England, 2021; Beadling et al., 2022).

Mesoscale eddies, as well as jets and boundary currents, play a central role in the dynamics of the ACC and ASC, the meridional overturning circulation, and shelf-open-ocean exchange (e.g., Goddard et al. (2017); Stewart et al. (2018, 2019)). Hence, this region provides strong motivation to refine ocean grid spacing for studying the role of the Southern Ocean in climate. Although the CM4X models remain too coarse to resolve many key processes on the shelf (e.g., see Figure 1 from Part I (Griffies et al., 2024)), they succeed in pushing the envelope of global climate models by offering a refined representation of flows on the shelf and slope (grid spacing is roughly 7 km at 70°S in CM4X-p125), thus offering a tool to probe the role of these currents on larger scale climate.

5.1 Features of the horizontal circulation

The ACC has been a topic of study with large-scale and fine resolution numerical models since the 1990s, following the pioneering efforts of the FRAM project (FRAM Group, 1991) and further pursued across a grid resolution hierarchy by Hallberg and Gnanadesikan (2006). These studies, and many more, have helped to establish the fundamental importance of the ACC for large scale climate dynamics (Rintoul et al. (2001); Rintoul and Naveira Garabato (2013); Rintoul (2018)). Despite this fundamental importance, many coupled climate models still struggle to accurately represent the mean-state ACC strength and structure (Beadling et al., 2020). For the CM4X models, the ACC is revealed by a strong eastward zonal flow comprised of multiple jet-like structures such as seen in Figures 4 and 5 of Part I (Griffies et al., 2024), as well as Figure 15 shown here. Many details of the stronger ACC flow patterns are similar between CM4X-p25 and CM4X-p125, reflecting the deep reaching nature of Southern Ocean currents that are affected by bottom topography and thus generally follow f/H contours (f is the Coriolis parameter and H is the bottom depth). We also commented on this feature of the Southern Ocean in Figure 5 from Griffies et al. (2024), where much of the kinetic energy in the Southern Ocean is dominated by the depth averaged velocity.

Moving south towards Antarctica, we encounter the westward flowing (ASC) along the continental slope. As reviewed by A. F. Thompson et al. (2018), the ASC is present in most regions around Antarctica, with the notable exception of the West Antarctic Peninsula and westward until reaching the Amundsen Sea.

Though transporting far less mass than the ACC, there is a growing appreciation for the impacts of the ASC on regional and global climate. In particular, as reviewed by Beadling (2023), the ASC acts as a barrier to meltwater originating from Antarctic ice shelves leaving the continental shelf, and conversely as a barrier to relatively warm Circumpolar Deep Waters penetrating towards the continental shelf from the north. As encountered in modeling studies such as Goddard et al. (2017); Moorman et al. (2020); Lockwood et al. (2021); Beadling et al. (2022); Tesdal et al. (2023), a realistically strong ASC introduces a fundamentally new dynamical regime into the Southern Ocean that is absent from coarse models (roughly those ocean models with horizontal grids coarser than 0.25°). Given the extremely small Rossby deformation radius along the Antarctic continental slope/shelf region (see Figure 1 from Part I (Griffies et al., 2024)), it is likely that global models will fail to accurately represent the full dynamical impacts of the ASC until reaching toward 1 km horizontal grid spacing.

In Figure 15 we display two meridional-depth sections, one through the Drake Passage and one south of Tasmania. Here we see the deep reaching eastward jet-like flows found in both sections, along with distinct westward flows. The westward flows are generally weaker and found particularly at depth and, for the Tasmanian section, we find the strong westward flowing ASC along the Antarctic continental slope. Note the rather weak flow in the Tasmanian section for latitudes between approximately 60°S and 55°S, with the upper flow weakly westward and deep flows very weak. The topography between these regions is rather fine scale, suggesting that this "rough" topography acts to weaken

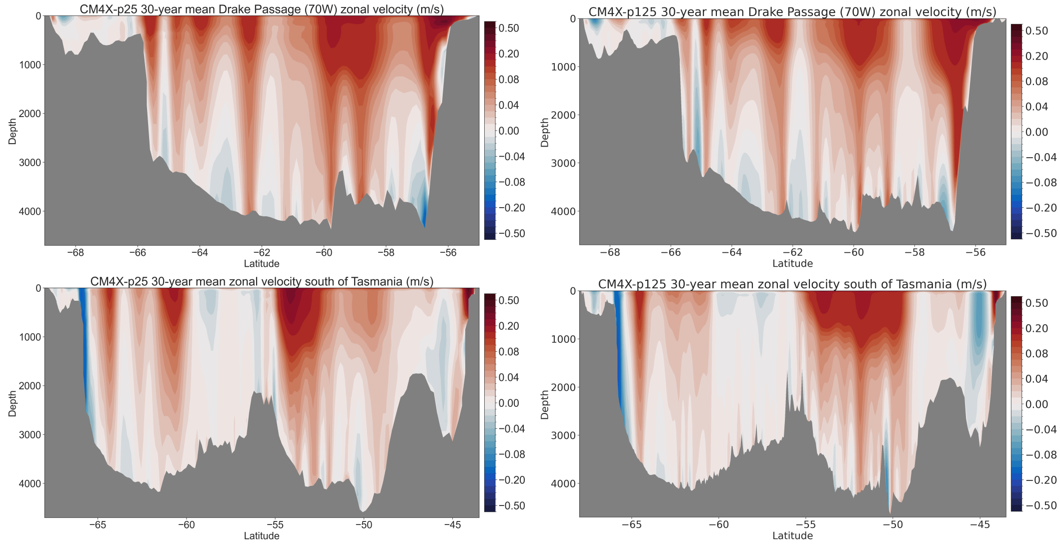

Figure 15. 30 year mean zonal velocity along a longitude within the Drake Passage (upper row) and south of Tasmania (lower row). Note the deep reaching zonal flows, with some deep reaching westward flows. Also note that strong and deep reaching westward flowing Antarctic Slope Current seen in the section south of Tasmania, whereas this current is largely absent in the Drake Passage section (see A. F. Thompson et al. (2018) for a review). The colorbars are the same for all panels, though the latitude range differs for the top row and bottom row.

the otherwise deep reaching flow. This feature of the flow over the rough topography may suggest a role for the rough bottom modes of LaCasce (2017). Even so, the horizontal flow does *not* generally vanish at the bottom, which contrasts to the assumptions of LaCasce (2017). Indeed, the Drake Passage section of CM4X-p125 is notable for its bottom enhanced westward flows. It is notable that all eastward flows are surface intensified (equivalent barotropic), whereas some of the westward flows in the open ACC, and particularly in the Drake Passage section, are bottom intensified. The presence of deep westward jets in the vicinity of the ACC have been noted in previous observational studies and high resolution simulations (Xu et al., 2020).

In Figure 16 we show the Drake Passage transport, which provides a traditional measure summarizing the zonal flow in the ACC. CM4X-p25 is roughly 10-15 Sv stronger than CM4X-p125 throughout the historical simulation, with CM4X-p125 consistent with the CMIP6 ensemble mean whereas CM4X-p25 is consistent with the lower end of the observational based estimates. After roughly 100 years of spin-up, both piControl simulations exhibit multi-decadal fluctuations of roughly 10 Sv. For the historical simulation, CM4X-p125 shows a slight decrease whereas CM4X-p25 is roughly unchanged. Both models show a decline during the SSP5-8.5 until around year 2060, at which point the strength stabilizes (CM4X-p25) or begins to rebound (CM4X-p125).

5.2 Ventilation, watermass transformation, and overturning circulation

The CM4X mixed layer depths in Figure 9 from Part I (Griffies et al., 2024) reveals a narrow band of deep mixing on the Antarctic continental shelf where dense shelf water (DSW) forms and then subsequently overflows down the continental slope, ventilating the deep ocean and leading to the formation of AABW (Rintoul & Naveira Garabato, 2013; Rintoul, 2018). In Figure 17 we display bottom temperature and salinity on the Antarctic continental shelf, defined here as the region landward of the 1000-m iso-

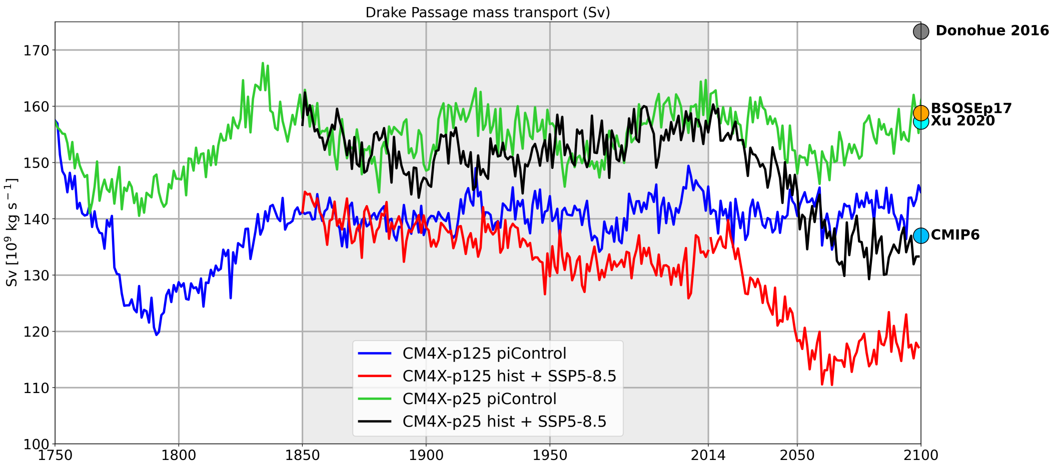

Figure 16. Time series of annual mean mass transport through the Drake Passage in units of Sverdrup (10^9 kg s^{-1}) for the CM4X simulations. The gray shaded region is the historical portion of the simulation. The colored markers on the right edge of the plot indicate the observed estimate of the total flow through the Drake Passage from the cDrake array from Donohue et al. (2016), based on an observation period of 2007 to 2011; the observationally-constrained Biogeochemical Southern Ocean State Estimate (Verdy & Mazloff, 2017) at 1/6-degree and integrated from 2013–2018; the 1/12th HYbrid Coordinate Ocean Model (Xu et al., 2020); and the CMIP6 Drake Passage transport ensemble mean of historical simulations averaged from 1986 to 2005 (Beadling et al., 2020).

bath. These shelf properties play a central role in determining the volume and formation rate of DSW and influence the stability of ice shelves ringing the continent. Both CM4X-p25 and CM4X-p125 versions produce spatial distributions of bottom temperature and salinity qualitatively consistent with observations, yet with notable differences in magnitudes in specific regions. For bottom temperatures, both the CM4X models are more consistent with observations than that simulated by most models within the CMIP6 ensemble, which exhibit significant biases (e.g., Figure S8 from Purich and England 2021). CM4X-p125 exhibits a warmer and slightly more saline West Antarctic shelf compared to CM4X-p25. The West Antarctic shelf regime is characterized as a "warm shelf" (Thompson et al., 2018) where upward sloping isopycnals and the lack of an ASC allows warm Circumpolar Deep Water (CDW) to readily access the shelf. The warmer West Antarctic bottom temperatures in CM4X-p125 are consistent with a slightly warmer mean-state off shore reservoir of CDW compared to CM4X-p25. CM4X-p125 also exhibits a much fresher East Antarctic shelf that could be related to its stronger ASC compared to CM4X-p25.

The spatial pattern of surface watermass transformation (WMT) across the densest waters confirm that both versions of CM4X simulate DSW formation and subsequent down slope flow in realistic locations along the Antarctic shelf (Jacobs, 2004; Silvano et al., 2023), highlighting the shelves around Weddell Sea, Prydz Bay, Adelie Land, and the Ross Sea (Figure 18). The red shading in Figures 18a,b shows the mean surface WMT per unit area across σ_2 (potential density referenced to 2000 dbar) classes for the period 1975-2012, focusing on the densest water classes. Choosing the densest waters available in each region captures the spatial distribution of where DSW formation occurs that contributes to the formation of AABW. The bottom age distribution (green shading in Figures 18a,b) illustrates the pathways of newly formed dense waters from the Antarctic slope to the abyssal ocean. These patterns align with the known AABW pathways iden-

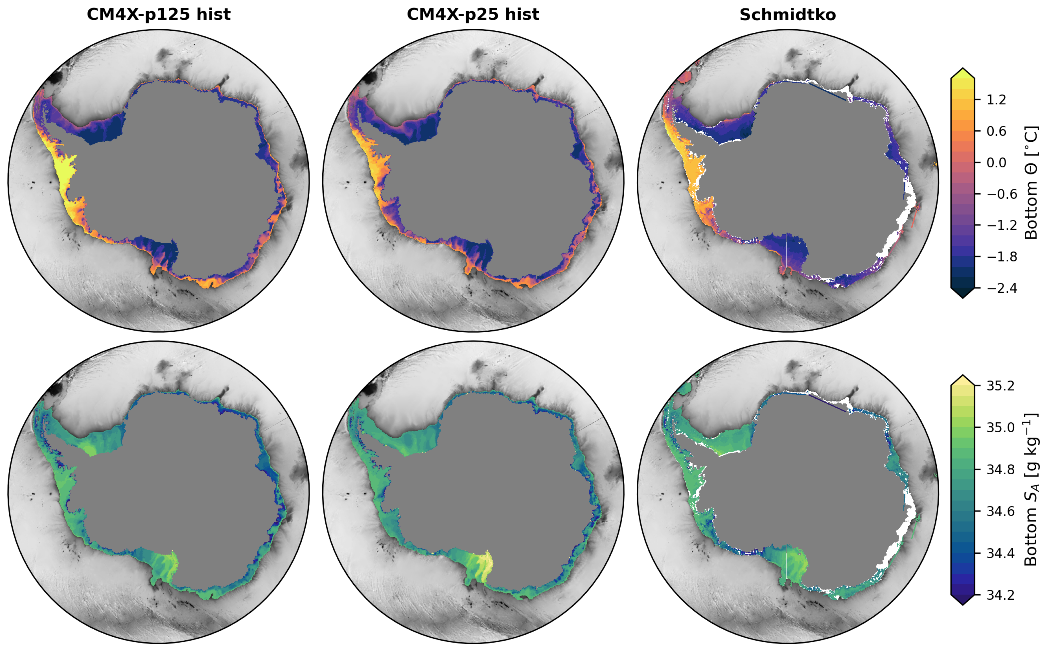

Figure 17. Temperature and salinity along the bottom of the Antarctic shelf from the CM4X models and as compared to the observational based analysis of Schmidtko et al. (2014). The fields are averages for the period 1975-2012, which is the time period used by Schmidtko et al. (2014). We define the shelf as the region with depth shallower than 1000 m.

tified from observational studies (e.g., Silvano et al. (2023)) and is consistent with passive tracers studies using a reanalysis-forced model (Solodoch et al., 2022).

The total DSW formation over the Antarctic continental shelf, estimated as the maximum surface WMT, is approximately 5 Sv and occurs for $\sigma_2 > 37$, with CM4X-p25 exhibiting a slightly greater DSW formation compared to CM4X-p125 (Figure 18c). Furthermore, the surface WMT is shifted towards higher σ_2 by around 0.1 in CM4X-p25 compared to CM4X-p125, especially in the Ross Sea sector (dashed lines in Figure 18c), which is consistent with the formation of higher salinity waters in this region (Figure 17). The Ross Sea sector shows the highest WMT rates compared to other regions, while the Weddell Sea sector (dotted lines in Figure 18c) also exhibits significant WMT, in agreement with observational estimates identifying these two regions as the major sources of DSW and AABW (Silvano et al., 2023).

To interpret Figure 18c, it is important to note at what densities positive WMT occurs. Both the Ross and Weddell sections in Figure 18c represent the majority of the WMT in the range of $\sigma_2 > 37$. Adding these two together explains the bulk contribution to dense water formation required for AABW. Other regions over the Antarctic shelf show positive WMT. However, these transformations occur at generally lighter densities ($\sigma_2 < 37$). As shown in Figure 7 from Part I (Griffies et al., 2024), the bottom cell overturning is associated with waters denser than $\sigma_2 = 37$, indicating that any surface WMT at lighter densities are not expected to contribute to AABW formation and bottom cell overturning.

The spatial pattern in bottom age (Figures 18a,b) is also consistent with the difference in surface WMT between the two CM4X models, and the dominant role of surface WMT in the Ross and Weddell sections, showing slightly younger bottom waters in CM4X-p25 along the AABW pathways due to the higher density waters formed on the shelf, emanating away from the Ross and Weddell Seas. The total DSW formation rate of approximately 5 Sv in both CM4X model configurations is on the lower end of the observational range, which spans 5-15 Sv (Silvano et al., 2023). However, this formation rate is a significant improvement compared to most CMIP6 models (Heuzé, 2021), suggesting that both CM4X simulations capture the relevant processes responsible for DSW and AABW production.

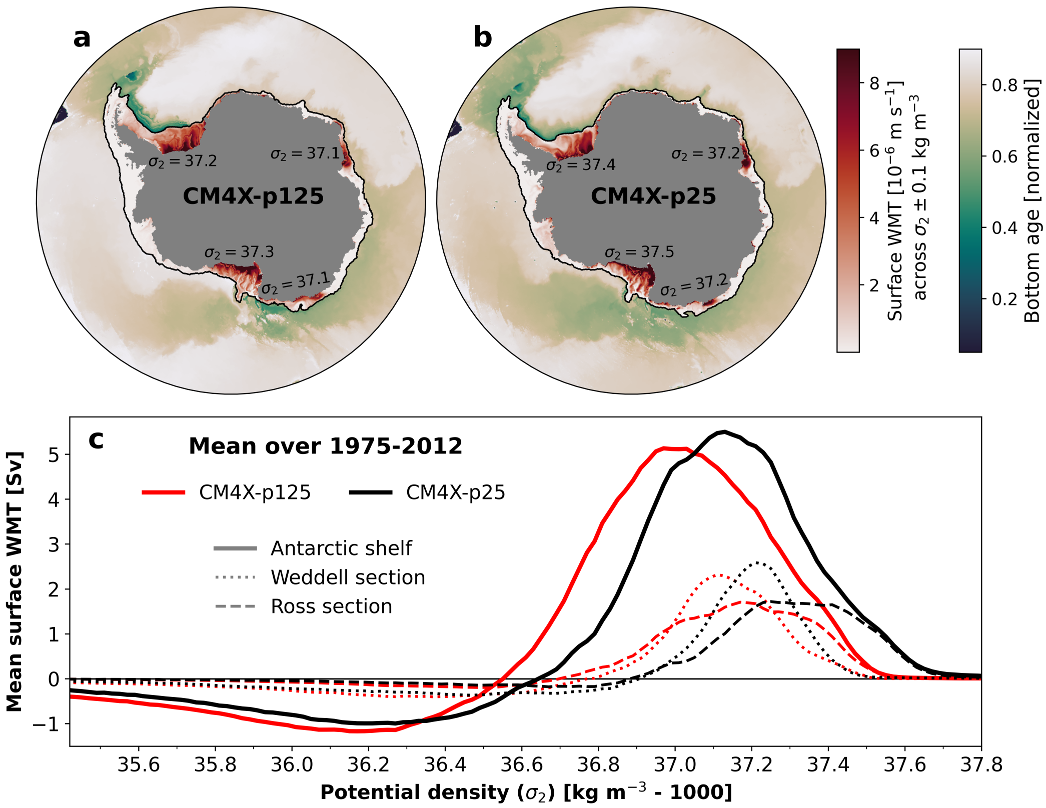

Figure 18. Surface water mass transformation (WMT) on the Antarctic continental shelf and bottom water age distribution over the abyssal Southern Ocean in (a) CM4X-p125 and (b) CM4X-p25. The red shading in panels a and b represent the time mean (1975-2012) surface WMT per unit area across σ_2 potential density classes (potential density referenced to 2000 dbar), separately determined in the four key DSW formation regions based on the densest water class. For the Weddell, Prydz/Adelie, and Ross shelves the surface WMT is mapped across σ_2 isopycnals of 37.2, 37.1 and 37.3 in CM4X-p125, and 37.4, 37,2 and 37.5 in CM4X-p25. The green shading in panels a and b represents the bottom age tracer at year 2009, normalized to the total simulation length (100 years of spinup + 1850-2009 historical = 260 years). (c) Mean (1975-2012) surface WMT in σ_2 integrated over the Antarctic shelf, Weddell (62°W-10°E) and Ross section (154°E-134°W) for CM4X-p125 (red lines) and CM4X-p25 (black lines).

The overturning circulation streamfunction offers a means to both measure and to visualize ventilation of the ocean interior. As a complement to the pole-to-pole overturning in Figure 7 from Part I (Griffies et al., 2024) that illustrates connections between the Southern Ocean and North Atlantic, in Figure 19 we focus on the Southern Ocean overturning. The AABW cell is the densest cell (blue counterclockwise cell) associated

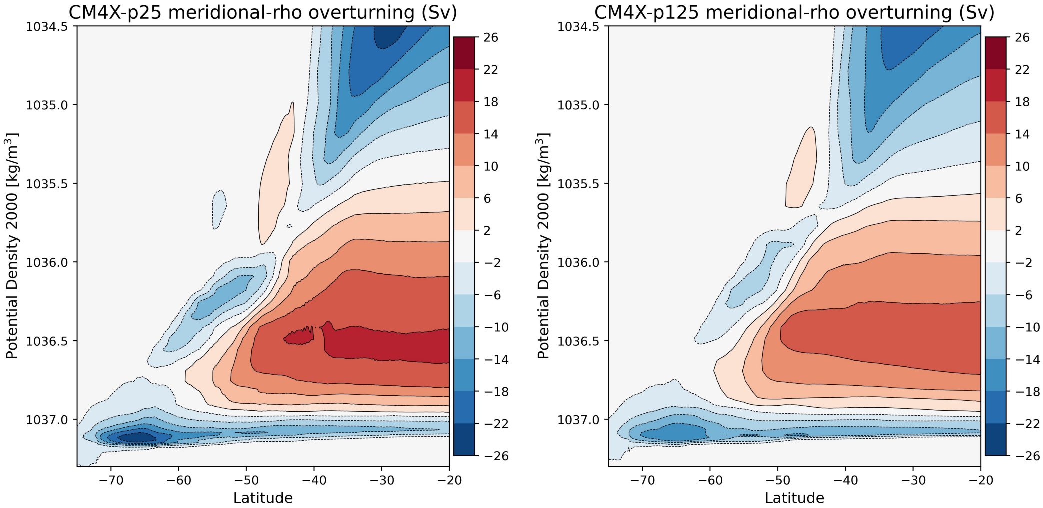

Figure 19. Meridional-density (ρ_{2000}) overturning circulation in the Southern Ocean as computed using time mean flow from years 1980-2009, with CM4X-p25 on the left and CM4X-p125 on the right.

with waters formed via DSW production on the Antarctic shelf and subsequent overflow and entrainment into the abyssal ocean (i.e., the processes shown in Figure 18). CM4X-p25 shows slightly larger formation around 65°S-70°S, and yet the AABW signal is slightly stronger in CM4X-p125 upon reaching 30°S. This disagreement between the strength of the AABW cell in the subpolar region and at 30°S indicates potentially larger interior mixing in CM4X-p25 which erodes the strong AABW transport away from the subpolar region. The other (blue) counterclockwise overturning cell is split into two sections in CM4X-p25 and CM4X-p125, though it is nearly connected in CM4X-p125. In the low latitudes, this cell is associated with subtropical mode waters. As discussed by Hallberg and Gnanadesikan (2006) (see their Section 3a), the merging of this cell southward across 45°S results from meridional mass transport from transient mesoscale eddies, with such eddy variability stronger in CM4X-p125 (e.g., East Australian Current, Agulhas Rings). The dense flow in the red clockwise cell is associated with North Atlantic Deep Water (NADW) and Circumpolar Deep Water (CDW) moving south, with a portion of this water lightened into Antarctic Intermediate Water (AAIW) and another portion densified into AABW.

5.3 Multi-decadal fluctuations in the piControl simulations

In Figure 20 we show the time series for the Drake Passage transport from the piControl simulations in CM4X as well as CM4.0. As described by Held et al. (2019), the large amplitude multi-decadal fluctuations in CM4.0 are associated with very large (i.e., super) polynyas in the Ross Sea. Such polynyas also appear in the ESM4 simulation of Dunne et al. (2020). In contrast, we do not find these super-polynyas in either CM4X-p25 or CM4X-p125, with both models exhibiting more modest multi-decadal fluctuations. One hypothesis for the absence of super-polynyas in CM4X relates to the increase in land ice albedo relative to CM4.0, resulting in a slightly cooler Antarctic climate and thus supporting more intermittent ventilation with smaller polynyas, rather than the buildup of massive subsurface heat charging the super-polynyas in CM4.0 (L. Zhang et al., 2021). This hypothesis comes with a caveat, however, since the land ice albedos used in CM4X are the same as ESM4, and yet ESM4 also suffers from the extremely strong polynyas.

We thus suspect that the full story for polynya events involves multiple factors, including winds, sea ice, and ice shelves.

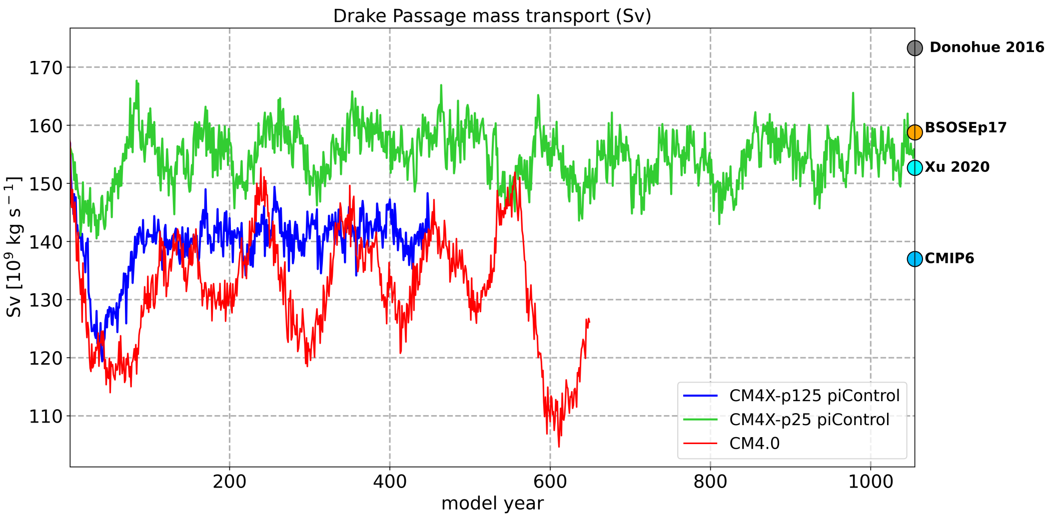

Figure 20. Time series of annual mean mass transport through the Drake Passage in units of Sverdrup (10^9 kg s^{-1}) for the piControl simulations from CM4X and CM4.0.

As a further means to distinguish the CM4X Southern Ocean simulations from CM4.0, Figure 21 provides time series for Antarctic shelf salinity, circulation strength of the bottom overturning cell, AABW transport at 30°S, as well as Hovmöller diagrams of surface water mass transformation due to heat fluxes in σ_2. Here we again see signatures of the super-polynyas in CM4.0 whereas both CM4X simulations exhibit smaller amplitude fluctuations. The variability in CM4.0 highlights how the occurrence of large open-ocean polynyas lead to a series of interconnected changes in physical processes within the Southern Ocean. The CM4.0 piControl exhibits marked multi-decadal oscillations in shelf salinity, offshore heat loss, and dense water formation, seen in Figure 21 as positive excursions of surface WMT at the densest open waters ($\sigma_2 > 37$). These episodes of enhanced dense water formation directly impact the large-scale circulation, as evidenced by the concurrent peak in the strength of the bottom overturning cell, followed by increased AABW transport at 30°S. Furthermore, the same multi-decadal fluctuations are seen in the Drake Passage transport (Figure 20), illustrating the connection between dense water formation, overturning and the strength of the ACC.

The multi-decadal oscillations are still present in the CM4X-p25 piControl simulations, but are more muted compared to CM4.0. We observe variability in the subpolar cell strength that oscillates around 20 Sv with a clear periodicity. This muted variability is also reflected in the Antarctic shelf salinity and the offshore surface WMT. Thus, there is some intrinsic variability that is still apparent in the quarter degree piControl runs, consistent with L. Zhang et al. (2021), occurring at higher frequency and with weaker amplitude when land ice albedo is increased. Figures 20 and 21 suggest that the oscillations do not have as large an impact across different parts of the Southern Ocean in CM4X-p25 as they do in CM4.0. Interestingly, these oscillations are not as clearly seen in the CM4X-p125 version, which could be due to the shorter run time.

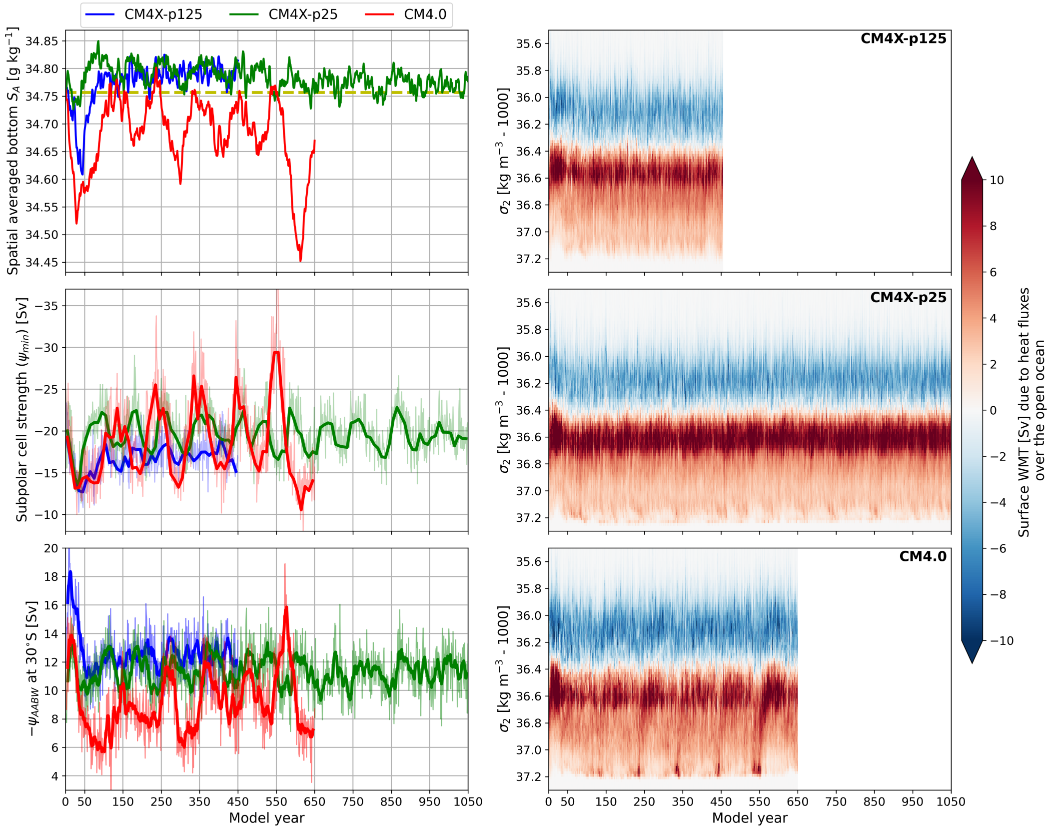

Figure 21. Time series for the annual mean salinity as area averaged around the bottom of the Antarctic shelf (upper left), circulation strength of the bottom overturning cell (middle left) and AABW transport at 30S (lower left) for CM4X-p125, CM4X-p25 and CM4.0 piControl simulations. The finer lines in the middle and lower left panels are annual means. The thick lines in the middle panel are decadal means, while the thicker lines in the lower panel are 10-year running means. The right panels show Hovmöller diagrams of surface forced water mass transformation in σ_2-space (potential density referenced to 2000 dbar) over the open Southern Ocean due to heat fluxes in CM4X-p125 (upper), CM4X-p25 (middle) and CM4.0 (lower). When calculating the area-average in the upper left panel we define the Antarctic shelf as the region with depths shallower than 1000 m. The yellow dashed line in the upper left denotes the observation-based climatological mean of Antarctic shelf salinity from Schmidtko et al. (2014).

5.4 Transient climate response

In Figure 22 we provide a suite of five time series for various properties around the Southern Ocean during the piControl spin-up, historical, and SSP5-8.5 scenario, that emphasize the coupling of properties between the shelf and open ocean. This coupling is also found in the shelf-open-ocean diagnostics presented in Figure 21. A key take away from Figure 22 is the general agreement in the transient response regardless of slight differences in the mean-state. Both models show a steep drop in shelf salinity at the start of the SSP5-8.5 simulation driven by freshening associated with an enhanced hydrological cycle and sea ice melt (and decreased sea ice formation) under increased warming. This drop also coincides with a sharp acceleration of the ASC in both models. The temporal pattern of ASC acceleration differs depending on location as the mean-state ASC is influenced by different dynamics along the shelf (Huneke et al., 2022). However, a strong ASC acceleration is consistently found from years 2014 to 2060 regardless of location.

The increased strength of the ASC also likely contributes to the consistent response found in shelf salinity via a positive feedback mechanism established between shelf freshening that enhances the ASC and thus leading to more freshwater trapping on the shelf (Moorman et al., 2020; Lockwood et al., 2021; Beadling et al., 2022). Despite initial differences in the strength of the subpolar cell, with CM4X-p25 showing stronger abyssal overturning, both models exhibit a similar magnitude of decline in bottom cell strength by the end of the 21st century under the SSP5-8.5 scenario. CM4X-p125 displays a steady weakening over time, whereas CM4X-p25 experiences a more abrupt reduction. Nevertheless, both models demonstrate comparable trends in AABW export at 30S, highlighting consistent responses to the transient conditions throughout the historical and SSP5-8.5 simulations. Investigating the differences between these two models and the mechanisms driving the response in shelf and open-ocean processes under SSP5-8.5 is a topic of further investigation.

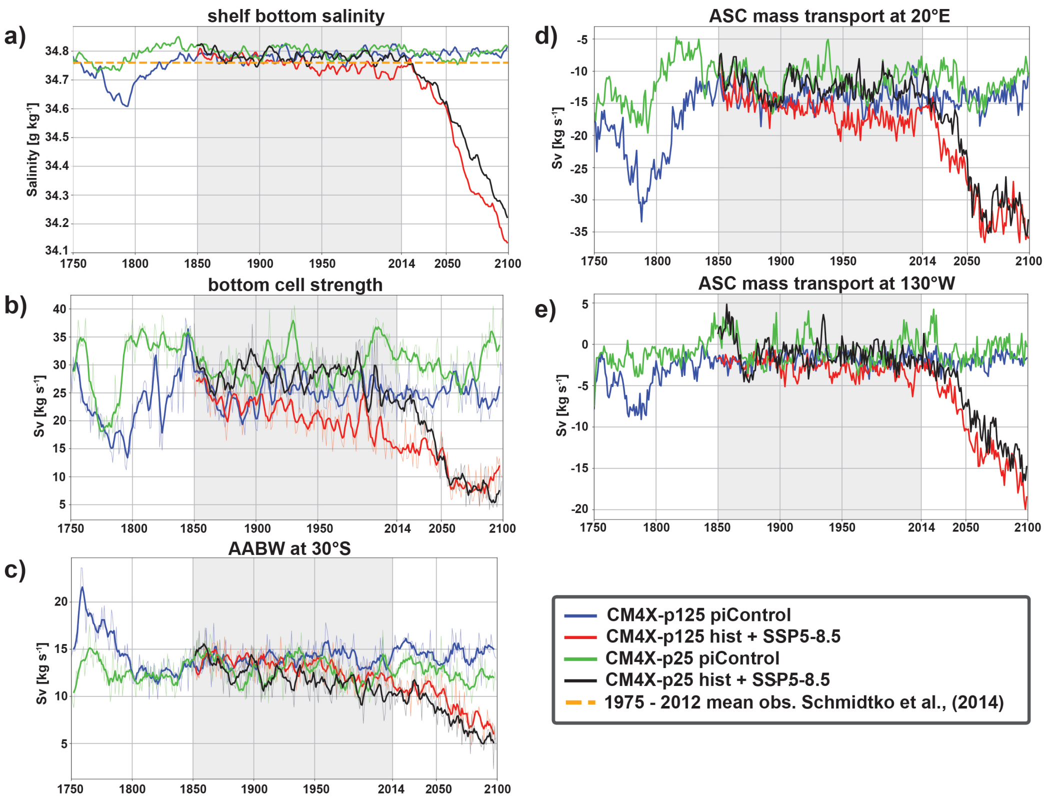

Figure 22. Time series for various properties around the Southern Ocean during the piControl spin-up, historical, and SSP5-8.5 scenario. a) Spatially averaged continental shelf bottom salinity b) circulation strength of the bottom overturning cell, c) AABW volume transport at 30°S, e) ASC strength at 20°E and f) ASC strength at 130°W. The ASC strength is shown at two different locations due to its flow characteristics being governed by different dynamics around the continental shelf (A. F. Thompson et al., 2018). The finer lines in panels b and c are annual means, while thicker lines in panels b and c are 5-year running means. When calculating the area-average in panel a, we define the Antarctic shelf as the region with depths shallower than 1000 m. The yellow dashed line in the upper left denotes the observation-based climatological mean of Antarctic shelf salinity from (Schmidtko et al., 2014).

5.5 Conclusions regarding the Southern Ocean simulations

The credible representation of Southern Ocean properties in CM4X opens up opportunities to study fully coupled Southern Ocean processes and Antarctic margin dynamics. The largest improvement in the CM4X models relative to previous generation GFDL models (CM4.0 and ESM4), is the lack of large amplitude multi-decadal oscillations associated with super polyna events. The presence of such large amplitude variability in CM4.0 prevented a suitable control run from which to branch perturbation experiments. Since the super-polynya-driven variability imprinted on the global climate and the Southern Ocean mean state, they made it nontrivial to interpret signal and "noise" in transient response to forcing perturbation experiments (Beadling et al., 2022; Tesdal et al., 2023).

Particularly notable features of the CM4X Southern Ocean simulation include the credible representations of Antarctic shelf hydrography, realistic locations of DSW production and overflow, and resolution of a strong ASC along the continental slope. A realistic representation of shelf properties and dynamics along the shelf-slope is required for examining shelf-open-ocean interactions, improving confidence in the thermal forcing of the AIS in a transient climate, and providing boundary conditions for dynamical ice sheet models. The credibility of CM4X near and along the Antarctic margin is encouraging for the utility of studying high latitude processes and for coupling with dynamical ice sheet models.

Additionally, with differing horizontal grid spacing providing for different degrees of representation of ocean mesoscale features, the CM4X suite is well positioned to probe the role of the ocean mesoscale within the climate system. Slight differences identified here between CM4X-p25 and CM4X-p125 in their velocity structure, Drake Passage transport, and Southern Ocean overturning may be linked to differing representation of mesoscale features and topography. Future work will aim at disentangling the sources of these differences between the two models and understanding whether these differences imprint on the transient climate response.

6 North Atlantic circulation

In this section we survey the horizontal and meridional overturning circulation features of the North Atlantic portion of the CM4X simulations, with comparisons made to observational estimates and other models, when available.

6.1 Gulf Stream

Here we focus on the Gulf Stream representation in CM4X as it compares to theoretical expectations, observed three dimensional structure (Todd, 2021), and the representation in a 1/12 degree regional ocean simulation also using the MOM6 dynamical core (NWA12, Ross et al. (2023)).

Expectations from previous studies

Past efforts to determine controls on the simulated Gulf Stream have focused on the latitude of separation from the coast, as well as behavior of the jet and mesoscale eddies post separation (Chassignet & Marshall, 2008). This latitude of separation and/or the mean latitude of the Gulf Stream extension is often associated with a temperature front or the zero barotropic vorticity contour (Section 6.2). In this section we consider jet coherence and jet location relative to the continental slope. This analysis finds that once leaving the coast, the CM4X simulations have a rather diffuse and meandering Gulf Stream jet that is, unfortunately, rather distinct from observational measurements.

Fundamentally, Gulf Stream fidelity depends on model horizontal grid spacing relative to the first baroclinic Rossby radius of deformation. At middle latitudes, Hallberg (2013) (see also Figure 1 from Part I (Griffies et al., 2024)) reveals a requirement of approximately 0.25° over the deep ocean (deeper than 3000 m) and approximately 0.125° closer to the coast (500 m $<$ depth $<$ 3000 m). Since the Gulf Stream 'leans' on the U.S. east coast continental shelf prior to separation, we expect accurate simulations require grid spacing to be at or finer than 0.125°.

Gulf Stream simulation features also depend on the spatial structure of wind forcing, deep western boundary current strength, bathymetric slope resolution, and model viscosity parameterizations that moderate resolved mesoscale turbulence (Hurlburt and Hogan (2000), Parsons (2006), Chassignet and Marshall (2008), Ezer (2016), Chassignet and Xu (2017), and Debreu et al. (2022)). Chassignet and Marshall (2008) show how choices of biharmonic and/or Laplacian viscosity control the latitude of separation, determine the presence or absence of standing eddies, and shape downstream instability behavior. The authors demonstrate that use of a low biharmonic viscosity can result in early separation and the presence of a standing eddy near Cape Hatteras, while use of a Laplacian can result in flow that is too laminar and a Gulf Stream that does not penetrate into the North Atlantic interior.

Plan view and vertical sections of the Gulf Stream

Qualitative plan view comparisons of surface mean kinetic energy (MKE) in CM4X simulations, NWA12, and observations reveal the initial latitude of separation from the coast as too far south (Fig. 23a-d). The observed time mean latitude of separation is approximately 37.5°N while in CM4X-p25 it is approximately 31°N and in CM4X-p125 approximately 32°N. This shift northwards with refined grid spacing continues when comparing to the 1/12° regional ocean model, where Gulf Stream separation occurs at approximately 37°N. As discussed by Chassignet and Marshall (2008), this characterization of the Gulf Stream separation and particular focus on jet coherence and the horizontal spreading of mean kinetic energy centers on the impact of viscosity choices on flow around the Charleston Bump. In both CM4X simulations use of a relatively low biharmonic viscosity value, as well representation limitations associated with relatively coarse horizontal resolution, are likely the cause of early separation.

Upstream of separation, surface MKE is greater in CM4X-p125 than in CM4X-p25 and closer to observed magnitudes. This increase in Gulf Stream strength in CM4X-p125 may be responsible for more realistic flow around the the Charleston Bump compared to CM4X-p25. The vertical extent of this relative increase is seen in a zonal cross-section at 29.5°N (Figure 24a-d). While the depth penetration of the Gulf Stream is similar in all panels, CM4X-p125 upper-ocean northward velocity (at depths shallower than 300 m) is over 50% larger than in CM4X-p25.

Differences between CM4X and observations and NWA12 are more significant post separation. In both CM4X-p25 and CM4X-p125, the eastward flowing jet rapidly dissipates post separation, resulting in a diffuse field of eddies that meander further north than observed eddy activity. Correspondingly, MKE is spread across roughly 30°N to 40°N, and dissipates moving eastward. Cross-sections at Cape Hatteras and at 70.5°W (Figure 24e-l) reveal the vertical structure of velocity to rapidly decrease with increasing depth. Compared to observations and NWA12, the CM4X Gulf Stream structure is far too surface intensified and horizontally diffuse. At 1800 m, a depth greater than that at which gliders collected measurements in Todd (2021), the spatial distribution of MKE appears similar in CM4X and NWA12. Noted differences between CM4X-p25 and CM4X-p125 include the near barotropic standing eddy in CM4X-p25 (Figures 23c,f) and a more energetic deep slope current north of 37.5°N in CM4X-p25.

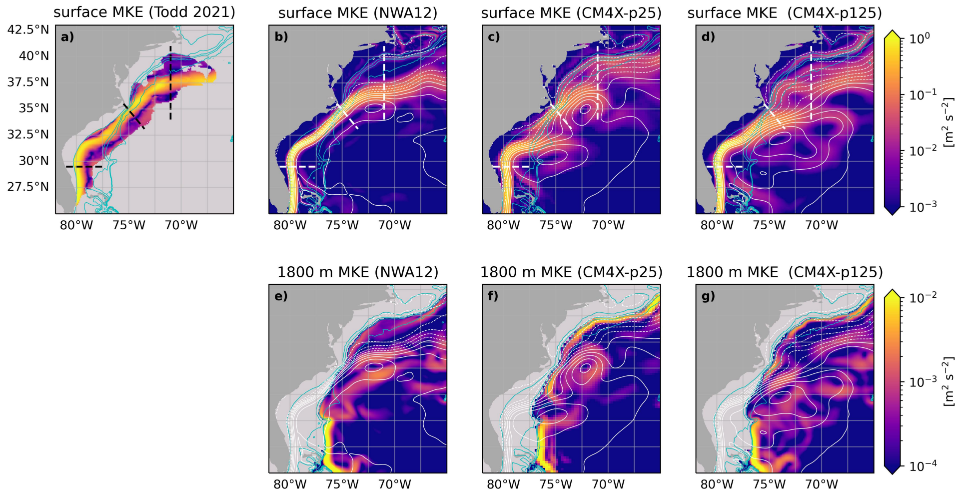

Figure 23. a) Time mean surface kinetic energy (MKE) derived from observations collected between 2015-06 and 2020-7 (Todd, 2021) (note log color scale). Blue contours identify the 100 m, 1000 m, 2000 m, and 3000 m isobaths. Black dashed lines are the locations of zonal, Cape Hatteras, and meridional cross-sections in Figure 24. b) 2010-2014 time mean surface kinetic energy from the NWA12 1/12° degree regional ocean model (Ross et al., 2023). Time mean sea surface height contours are added in white. c) Same as b) but for CM4X-p25. d) Same as b) but for CM4X-p125. e-g) Mean kinetic energy as in b-d) at 1800 m (note colormap scale change).

6.2 Barotropic vorticity

Interactions between deep ocean flows and sloping bathymetry result in bottom stretching and bottom pressure torques, influencing the western boundary currents and ocean circulation (Hughes & De Cuevas, 2001; R. Zhang & Vallis, 2007; Waldman & Giordani, 2023; Khatri et al., 2024). We investigate how spatial resolution differences between CM4X-p25 and CM4X-p125 affect the strength of the North Atlantic gyre circulations. For this purpose, we use the linear steady-state barotropic vorticity budget, which is often used to study how surface winds and bathymetry control ocean gyres (Stommel, 1948; Holland, 1967). Following Yeager (2015), we use the streamfunction, ψ, form of the barotropic vorticity budget

$$\underbrace{\int_{x_w}^{x} V \,\mathrm{d}x}_{\psi} \approx \underbrace{\frac{1}{\beta}\int_{x_w}^{x} J(p_b, H)\,\mathrm{d}x}_{\psi_{BPT}} + \underbrace{\frac{1}{\beta\rho_0}\int_{x_w}^{x} \hat{\boldsymbol{z}}\cdot(\nabla\times\tau_s)\,\mathrm{d}x.}_{\psi_{\tau_s}} \tag{3}$$

Here, V is the vertically integrated meridional velocity, β is the meridional derivative of the Coriolis parameter, p_b is bottom pressure, H is ocean depth, τ_s is the surface wind stress, x_w represents the western continental boundary and ρ_0 is the Boussinesq reference density. Note that contributions from friction and nonlinear terms are not included, as these contributions are relatively small in time-mean vorticity balances (see Khatri et al. (2024) for a complete vorticity budget analysis and diagnosis methodology).

The time-mean spatial patterns of subtropical and subpolar gyres are well captured by the combined effects of bottom pressure torque and surface wind stress (Figure 25).

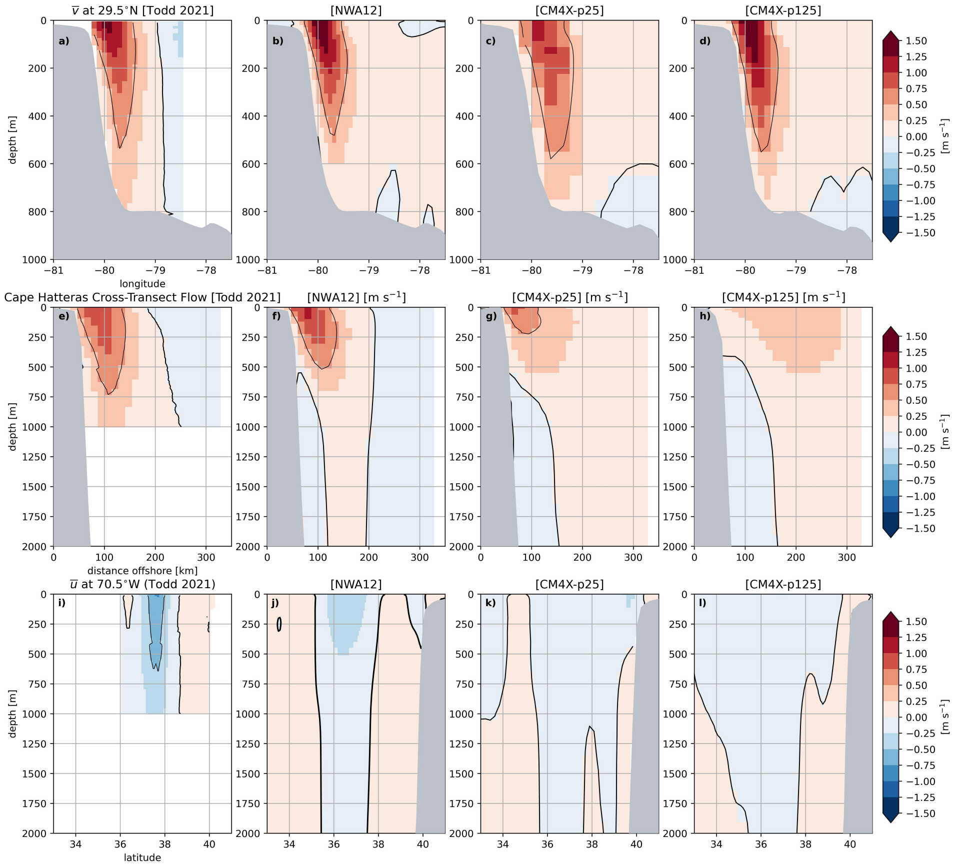

Figure 24. a-d) Time mean meridional velocity at 29.5°N as a function of depth and longitude from Todd (2021) (2015-202), NWA12 (2010-2014), CM4X-p25 (2010-2014), and CM4X-p125 (2010-2014) (note log color scale). Heavy black contour is the 0 value, light black solid/dashed contours are 0.5 and -0.5 m s^{-1}, and gray shading is the model bathymetry (NWA12 bathymetry is used in panel a). e-h) Time mean cross-transect velocity extending southwest of Cape Hatteras. i-l) Time mean zonal velocity at 70.5°W. Positive velocities are those into the page.

Surface winds control the meridional flow in open ocean gyres, while the return flow in western boundary currents is driven by bottom pressure torques (compare Figures 25a1-a2 with 25d1-d2, also see (Yeager, 2015)). However, in the Iceland basin and Nordic Seas, bottom friction and other vorticity terms are more important in guiding the meridional transport (not shown).

Comparing CM4X-p25 and CM4X-p125 reveals that refined ocean grid spacing leads to a stronger time-mean subtropical gyre transport, while the subpolar gyre transport is weaker. In CM4X-p25, the maximum transport in the subtropical gyre and minimum transport in the subpolar gyre are 25 Sv and -44 Sv, respectively (Figures 25a1, 25e).

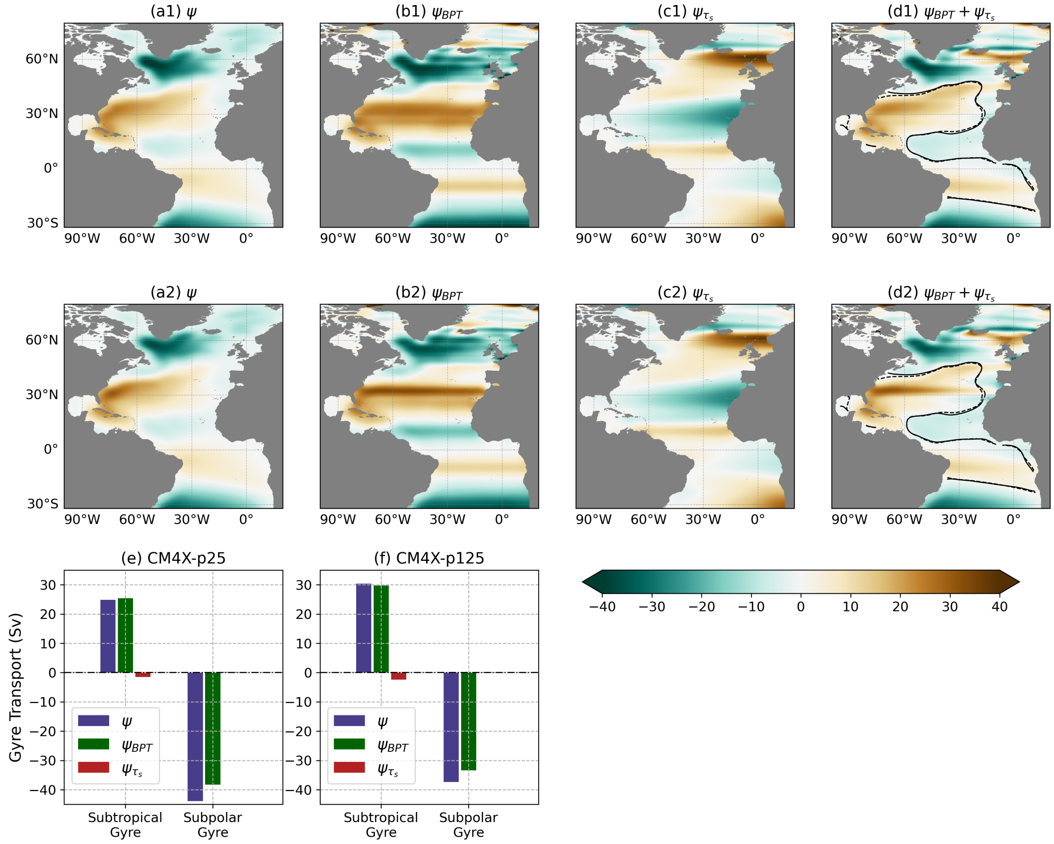

Figure 25. Barotropic streamfunction (in Sv) estimated using vorticity budget terms in equation (3) (30-year time-mean). The top row (panels a1-d1) show the CM4X-p25 piControl and the bottom row (panels a2-d2) show CM4X-p125 piControl. Panels (e,f) show the subtropical and subpolar gyre transports (estimated as the maximum or minimum of the barotropic streamfunction) and the associated contributions from ψ_{BPT} and ψ_{τ_s} (note that ψ_{τ_s} transports in the subpolar gyre are too small to see). Spatial maps in panels (d1-d2) are sums of maps shown in panels (b1-b2) and (c1-c2), and the solid (CM4X-p25) and dashed black (CM4X-p125) curves represent the zero streamfunction contour lines, which are used to infer the latitudinal position of the Gulf Stream located between the subtropical and subpolar gyres. To filter out the small-scale variability and grid-scale noise in vorticity diagnostics (Khatri et al., 2024), spatial maps are spatially smoothed to 10° resolution using the GCM-Filters from Loose et al. (2022).

On the other hand, in CM4X-p125, subtropical and subpolar gyre transports are 30 Sv and -37 Sv (Figures 25a2, 25f). These differences in gyre strength are primarily associated with differences in bottom pressure torque (compare Figures 25b1 and 25b2), which is affected by the finer scale bathymetry and better resolved mesoscale deep flows in CMX-p125 relative to CM4X-p25. These changes in gyre transports agree with changes in magnitudes of bottom pressure torques and the corresponding streamfunction transports, ψ_{BPT}.

Similar strengthening in the subtropical gyre with increased ocean model resolution has been observed in many climate models (Meccia et al., 2021). Moreover, the southward subpolar transport of 37 Sv in CM4X-p125 is close to estimates from Xu et al. (2013), who determined the southward transport in the Labrador Sea near 53°N to be approximately 39 Sv.

6.3 Atlantic meridional overturning circulation (AMOC)

Direct AMOC observations from OSNAP (Overturning in the Subpolar North Atlantic Program) over the recent periods (Lozier et al., 2019; Lim et al., 2019), as well as the reconstruction of the long-term mean AMOC over the past several decades using robust diagnostic calculations (R. Zhang & Thomas, 2021), show that the subpolar AMOC across the OSNAP section is dominated by its eastern component rather than the western component that extends across the Labrador Sea. Climate models often overestimate the maximum AMOC strength across OSNAP West (Thomas et al., 2015; Lim et al., 2019; Yeager et al., 2021). This model bias is likely related to the overestimation of the Labrador Sea open-ocean deep convection and its role in the AMOC across OSNAP West in relatively coarse models. As discussed in R. Zhang and Thomas (2021), too much open ocean Labrador Sea convection unrealistically overlaps with the less resolved boundary outflow from the Labrador Sea basin, thus leading to unrealistically high density in the less resolved boundary outflow from the Labrador Sea basin in relatively coarse models.

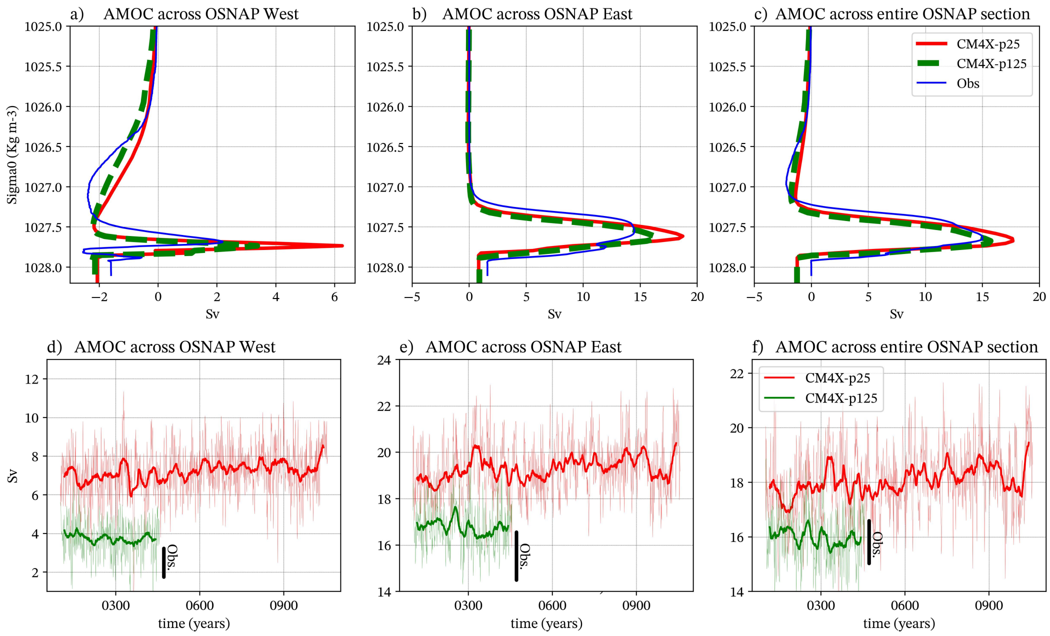

Figure 26. Top row: Climatological annual mean density space AMOC streamfunction across OSNAP West, OSNAP East, and the full OSNAP section for CM4X-p25, CM4X-p125, and the OSNAP observations over the period of 2014-2020. Bottom row: Time series of the maximum AMOC strength across OSNAP West, OSNAP East, and the full OSNAP section, in comparison with the OSNAP observations over the period of 2014-2020. The time series are smoothed with a 10-year running mean. The dark lines in bottom rows depict the ranges (mean +/- one standard deviation of rolling annual mean) of observations.

Figure 26 shows that CM4X-p125 has a maximum climatological density space AMOC strength across OSNAP West that has been reduced relative to CM4X-p25, with CM4X-p125 results closer to recent OSNAP observations of 2 Sv (Lozier et al., 2019; Li et al., 2021). The modeled maximum climatological density space AMOC strength across OSNAP East (or across the entire OSNAP section) has also been improved with CM4X-p125 relative to CM4X-p25, with Figure 26 showing CM4X-p125 values closer to the OSNAP observations. The modeled maximum climatological subpolar AMOC in density

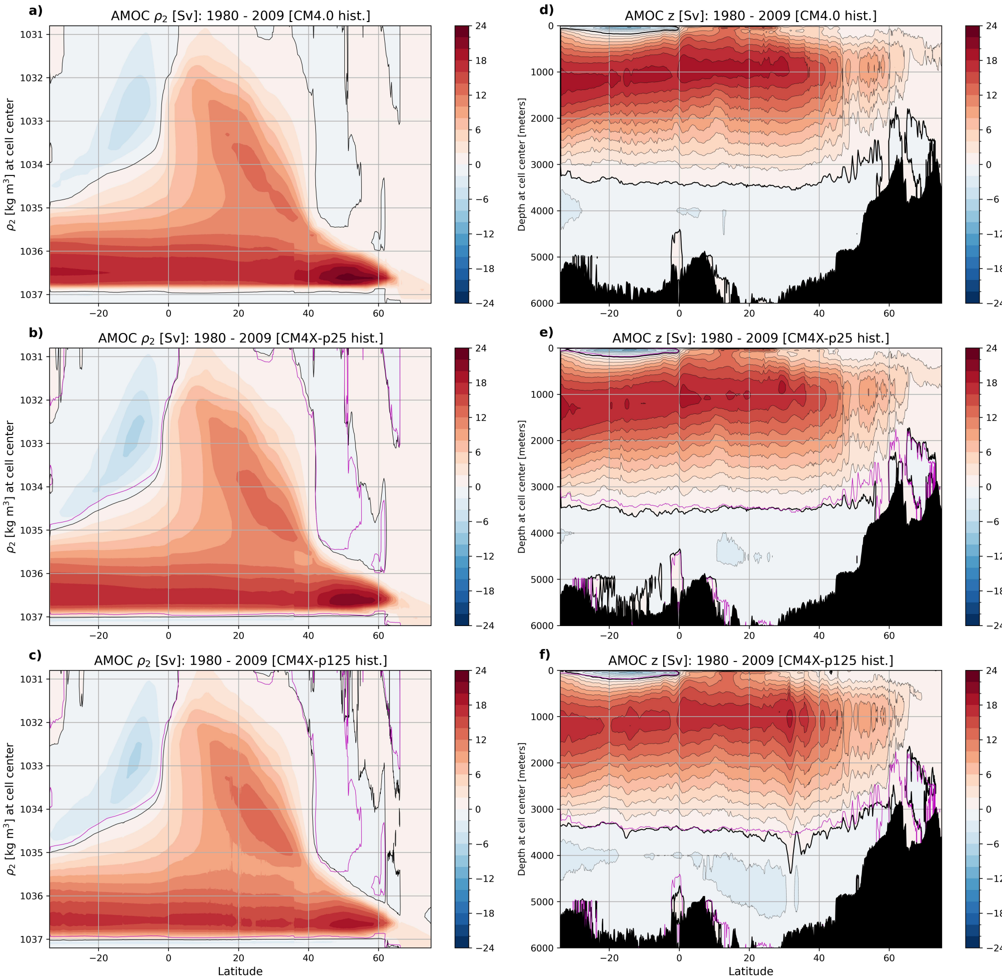

Figure 27. Climatological annual mean Atlantic meridional overturning circulation (AMOC) streamfunction as a function of latitude in potential density (ρ_2, a,b,c) and geopotential (d,e,f) for the historical simulations over the period of 1980-2009 using CM4.0, CM4X-p25, and CM4X-p125. In CM4X-p25 and CM4X-p125 panels, the magenta contour identifies the zero contour of CM4.0. Contours are shown at 2 Sv increments. The figure includes transport in the Atlantic, Arctic, Mediterranean, and Baltic basins.

space is reduced in CM4X-p125 compared to that in CM4X-p25 (Figure 27), which is consistent with the improvement across OSNAP West and East (Figure 26).

Since both CM4X-p125 and CM4X-p25 have the same atmosphere model, improvements in the CM4X-p125 mean state subpolar AMOC are likely due to its refined horizontal ocean grid spacing compared to CM4X-p25. Improvements in the simulated maximum AMOC strength across OSNAP West in CM4X-p125 (Figure 26a,d) is likely linked to its improved and reduced Labrador Sea winter deep convection area compared to that simulated in CM4X-p25, as reflected in the climatological winter mixed layer depth (MLD) seen in Figure 28. The Labrador Sea winter MLD in CM4X-p25 is unrealistically deep and broad, which is likely related to the under-representation of the dense overflows entering the Labrador Sea and the less resolved mesoscale eddy restratification (Tagklis et al., 2020). The winter Labrador Sea MLD is improved (shoaled) in CM4X-p125, which has a slightly better representation of the dense overflows across the OSNAP section as

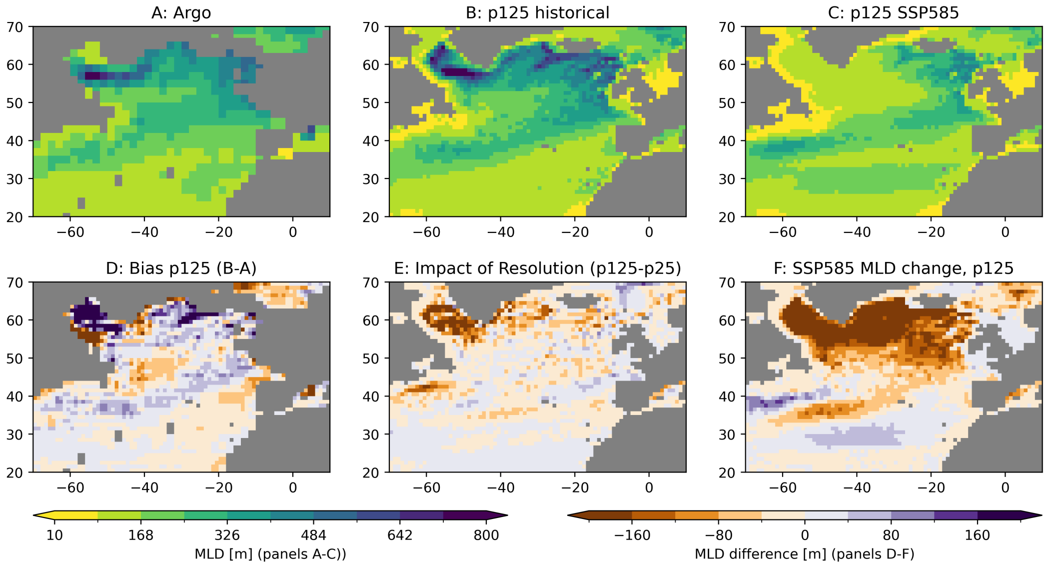

Figure 28. Winter climatology of the North Atlantic maximum mixed layer depth from Argo (years 2004-2023) (Argo, 2023), as well as years 2000-2014 (historical simulation) and years 2085-2099 (SSP5-8.5 simulation) from CM4X-p25 and CM4X-p125, computed as in Figure 9 from Part I (Griffies et al., 2024) Panel A: estimates from the Argo profiling floats; Panel B: results from CM4X-p125 historical experiment; Panel C: CM4X-p125 SSP5-8.5 simulation. Panel D shows the differences between CM4X-p125 and Argo (Panels B-A), whereas Panel E shows the difference between CM4X-p125 and CM4X-p25 (with a shoaling of MLD in CM4X-p125 relative to CM4X-p25). Finally, Panel F shows the impacts from the SSP5-8.5 climate change, showing years 2084-2099 minus years 2000-2014 (Panels C-B) from CM4X-p125. Note that differences documented in Panels D and E are robust to a longer time average over model years 1955-2014.

seen in Figure 29, as well as a stronger resolved mesoscale eddy restratification. However, neither CM4X-p125 nor CM4X-p25 simulate the observed winter deep convection in the Greenland Sea.

Over the subpolar North Atlantic, horizontal circulations across sloping isopycnals play an important role in density-space AMOC, and a Sigma-Z diagram was designed to illustrate the role of horizontal circulations in density-space (R. Zhang & Thomas, 2021). Across OSNAP East, both CM4X-p125 and CM4X-p25 simulate the horizontal circulation contribution (positive inflow and negative outflow canceled at the same depth level, but not at the same density level), which corresponds to that observed in the upper ocean (Figure 30). However, both models underestimate the horizontal circulation contribution compared to that observed in the deep ocean (Figure 30). This modeling bias in the deep ocean is related to the underestimation of the Nordic Sea overflows (i.e. Denmark Strait overflow and Iceland-Scotland overflow) and associated recirculation in both CM4X-p125 and CM4X-p25 (Figure 29). With refined horizontal grid spacing, CM4X-p125 has a slightly better representation of the Nordic Sea overflows and associated recirculation (Figure 29), and thus a slightly better representation of the horizontal circulation contribution in the deep ocean (Figure 30). Although the difference between CM4X simulations and the observation in the density space AMOC streamfunction across OSNAP East does not appear pronounced (Figure 26), the Sigma-Z diagram clearly reveals differences in the simulated deep ocean AMOC structure in Figure 30.

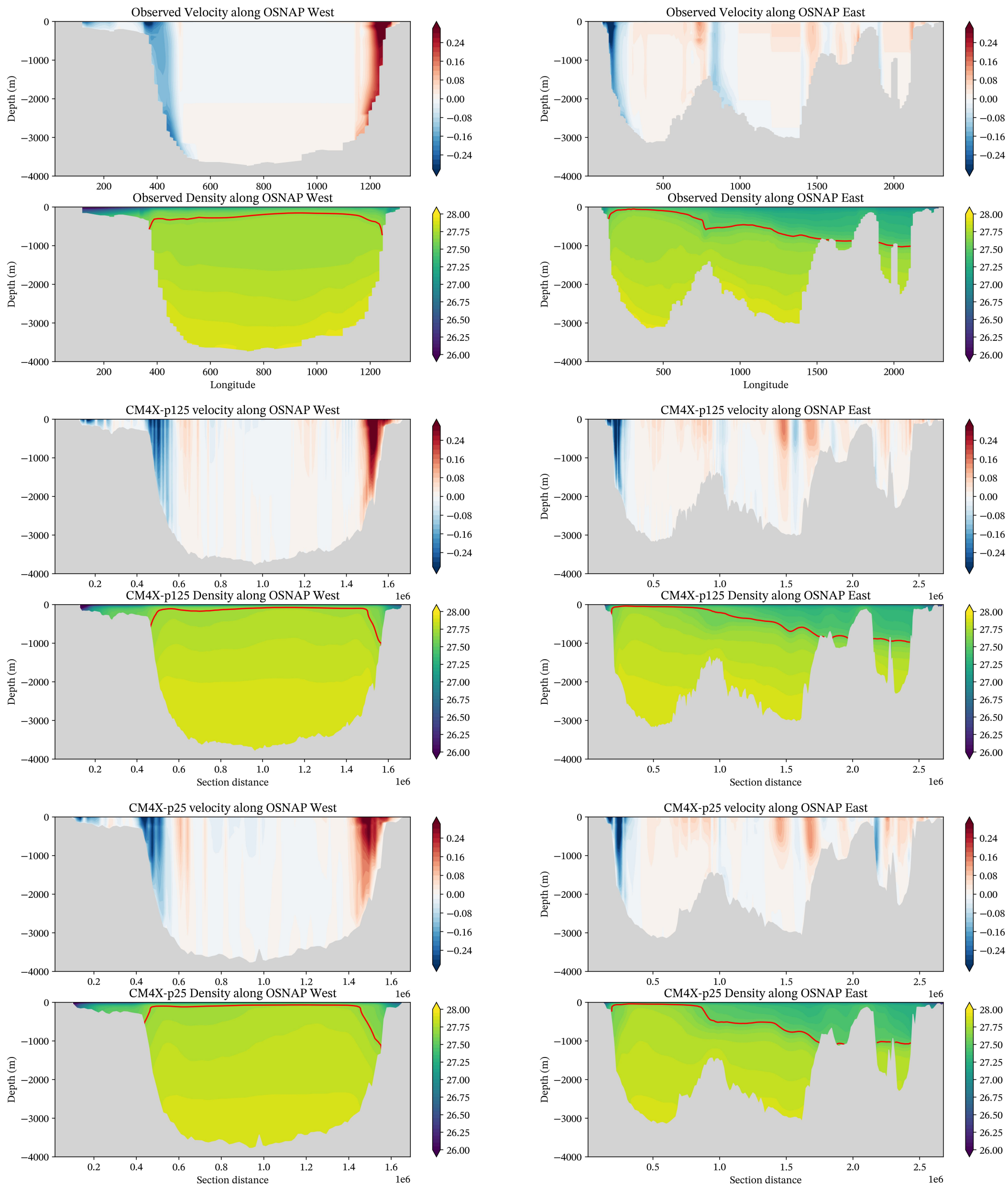

Figure 29. Mean velocity and potential density across OSNAP West (left) and OSNAP East (right). The top two rows show the OSNAP observations averaged over years 2014-2020 from Fu et al. (2023). The middle two row shows CM4X-p125 as averaged over years 2014-2020. The bottom two rows show CM4X-p25 as averaged over years 2014-2020.

Across OSNAP West, the horizontal circulation contribution to the density space AMOC in the upper ocean is too strong and too deep in CM4X-p25 compared to that estimated from observations (Figure 31). However, in CM4X-p125, the horizontal circulation contribution becomes less strong and shallower, which then compares better with the field measurements than CM4X-p25 (Figure 31). This improvement in CM4X-p125 is related to the improved (reduced) Labrador Sea winter deep convection area (Figure 28) and the better resolved Labrador Sea boundary current in CM4X-p125 (Figure 29). Consistently, the density contrast between the Labrador Sea boundary outflow and in-

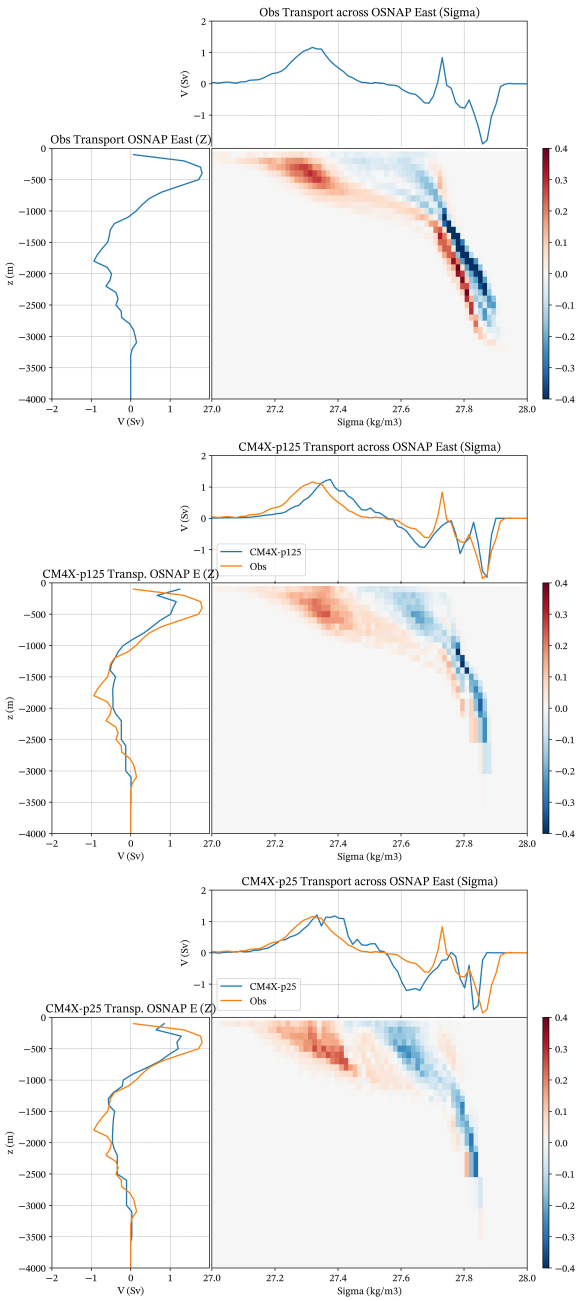

Figure 30. Sigma-z diagram of volume transport across OSNAP East for OSNAP observations (Fu et al., 2023) (top), CM4X-p125 (middle), and CM4X-p25 (bottom), all computed over years 2014-2020. The color shading in each panel shows the integrated volume transport (Sv) across each subsection over each potential density (σ_0; potential density referenced to 0 dbar) bin (x-axis) and depth (z) bin (y-axis); the integrated transport across each subsection over each potential density bin summed over the entire depth range is shown in the blue curve above; the integrated transport across each subsection at each depth bin summed over the entire potential density range is shown in the blue curve on the left. The line plots for the CM4X results have the OSNAP measurements plotted for easier comparison.

flow is reduced and thus the maximum density space AMOC across OSNAP West (mainly contributed by the horizontal circulation) becomes weaker in CM4X-p125 compared to that simulated in CM4X-p25 (Figure 26a,d). The observed density structure between the Labrador Sea boundary outflow and inflow is almost symmetric (Figure 29), consistent with a very weak horizontal circulation contribution (Figure 31) and thus a very weak maximum density space AMOC strength across OSNAP West (Figure 26a,d) in observations.

Across the RAPID section in the subtropical North Atlantic, both CM4X-p125 and CM4X-p25 simulate a similar maximum AMOC strength as that estimated from the RAPID program (Cunningham et al., 2007; McCarthy et al., 2015; Smeed et al., 2018), but the penetration depth of the simulated AMOC is much shallower than that observed (Figure 32). The simulated shallow AMOC across the RAPID section is a typical bias found in many models due to deficiencies representing the dense and deep-penetrating Nordic Sea overflows (Danabasoglu et al., 2010; R. Zhang et al., 2011; Danabasoglu et al., 2014; H. Wang et al., 2015). The penetration depth of the simulated AMOC across the RAPID section is slightly deeper in CM4X-p125 than that in CM4X-p25 (Figure 27d,e,f and Figure 32) due to slightly better representation of the Nordic Sea overflows in CM4X-p125.

The simulated multidecadal AMOC variability across the RAPID section in both CM4X-p125 and CM4X-p25 piControl simulations (Figure 33) is much weaker than the observationally-based estimate (Yan et al., 2018). The simulated historical AMOC across the RAPID section in both CM4X-p125 and CM4X-p25 has less pronounced multidecadal variations than simulated in CM4.0 (Figure 34). The pronounced multidecadal AMOC variations in CM4.0 (e.g., an increase up to 19 Sv and decline afterwards) are mainly induced by multidecadal changes in external radiative forcing (e.g. anthropogenic aerosols) as also found in many CMIP6 models (Menary et al., 2020), which are unrealistic and opposite to those indicated by the observed AMOC fingerprint (Yan et al., 2019) and inconsistent with further observational measures (Held et al., 2019; Menary et al., 2020; Robson et al., 2022). The improved historical AMOC changes simulated in CM4X are likely related to the refined horizontal atmospheric grid spacing used by CM4X (50 km) compared to that employed in CM4.0 (100 km) (see Appendix A1 in Part I (Griffies et al., 2024)). Previous studies have shown that climate models with coarse horizontal atmospheric grid spacing overestimate aerosol indirect effect (Donner et al., 2016; Sato et al., 2018; Zhao et al., 2018a). The impact of the horizontal atmospheric grid spacing on the aerosol indirect effect deserves more future investigation. The static vegetation approach (no land use change, no CO2 fertilization effect, and no demography change) employed in CM4X historical simulations might also contribute to the different simulated historical AMOC changes compared to those simulated in CM4.0 (again, see Appendix A2 in Part I (Griffies et al., 2024)). Additional experiments in future studies are needed to fully understand the improvement in simulated historical AMOC changes in CM4X. In Figure 34 we also show the AMOC strength for the SSP5-8.5 scenario experiment, whereby the AMOC weakens starting around year 2000 in CM4.0 and CM4X, with the year 2100 strength in each model less than half their pre-industrial strength.

6.4 Watermass transformation in the subpolar North Atlantic

Figure 35 shows the time-mean of the water mass budget (diagnosed following Drake et al. (2024)) in potential density coordinates,

$$\partial_t \mathcal{M}_{\geq} = \Psi_{\geq} + \mathcal{S}_{\geq} + \mathcal{G} \tag{4}$$

for two key regions of the Subpolar North Atlantic. $\mathcal{M}_{\geq} = \mathcal{M}_{\geq}(\sigma_2, t)$ is the mass of water denser than σ_2 at time t within a given region; $\Psi_{\geq}$ is the total transport into the region for waters denser than σ_2; $\mathcal{S}_{\geq}$ is direct addition of seawater mass from boundary fluxes; and $\mathcal{G}$ is the total water mass transformation rate (positive when it tends to increase the mass of denser water). We further decompose $\Psi_{\geq} = \sum \Psi_{\geq}^{(\text{boundary})}$ into con-

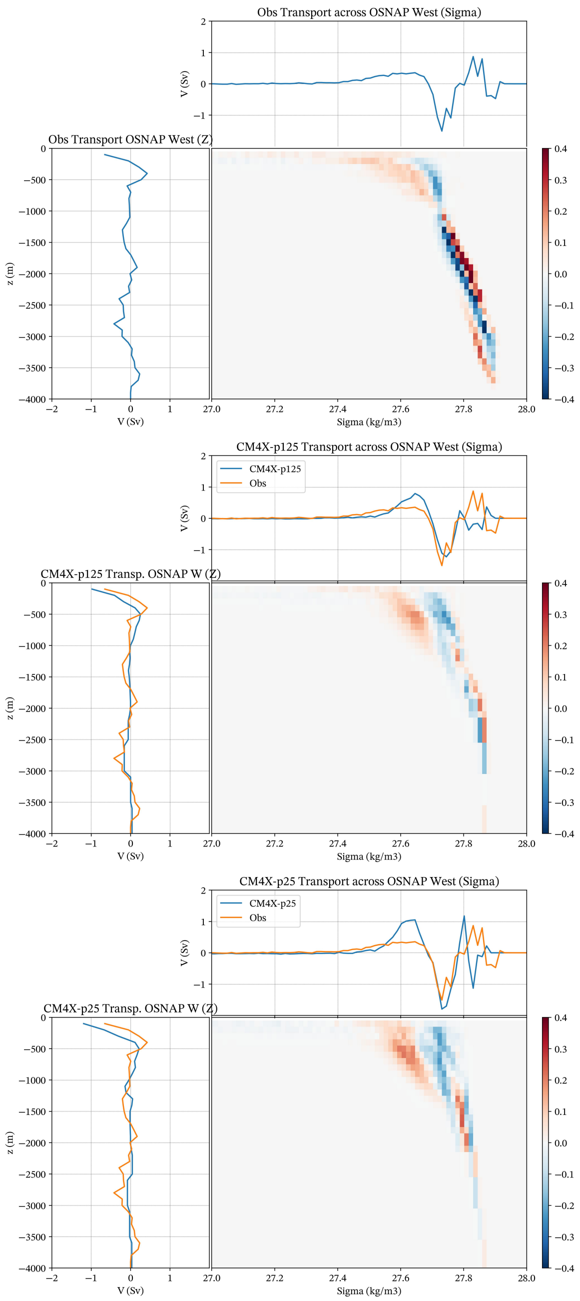

Figure 31. The Sigma-z diagram of volume transport across OSNAP West for OSNAP observations (Fu et al., 2023) (top), CM4X-p125 (middle), and CM4X-p25 (bottom), all computed over years 2014-2020. The color shading in each panel shows the integrated volume transport (Sv) across each subsection over each potential density (σ_0) bin (x-axis) and depth (z) bin (y-axis); the integrated transport across each subsection over each potential density bin summed over the entire depth range is shown in the blue curve above; the integrated transport across each subsection at each depth bin summed over the entire potential density range is shown in the blue curve on the left. The line plots for the CM4X results have the OSNAP measurements plotted for easier comparison.

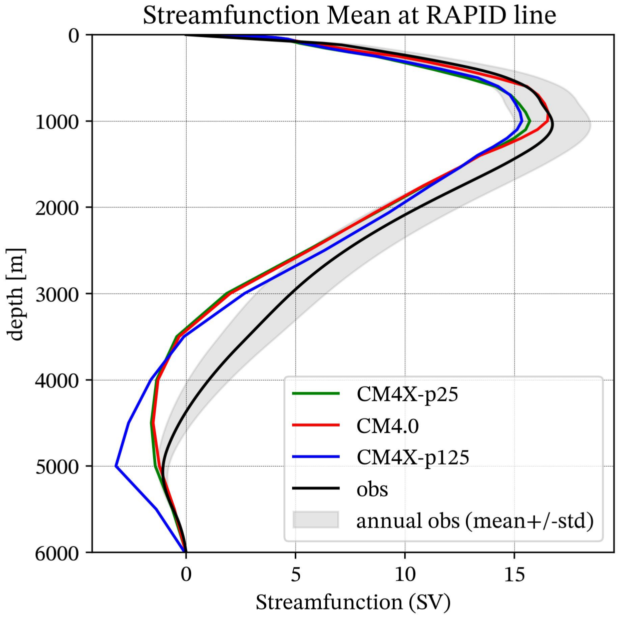

Figure 32. Climatological annual mean AMOC streamfunction strength across the RAPID section (26.5°N) for the models computed over years 2004-2022 using CM4.0, CM4X-p25, and CM4X-p125, in comparison with the RAPID observations over the period of 2004-2022.

tributions from different sections along the regions' boundaries and $\mathcal{G} \equiv \sum \mathcal{G}^{(\text{process})}$ to distinguish contributions from boundary fluxes $\mathcal{G}^{(\text{BF})}$ (primarily due to air-sea fluxes of heat, salt, and freshwater), parameterized mixing processes $\mathcal{G}^{(\text{mix})}$, and spurious numerical mixing $\mathcal{G}^{(\text{Spu})}$. All terms are diagnosed directly from mass, heat, and salt tendencies, except for the spurious numerical mixing $\mathcal{G}^{(\text{Spu})}$, which we identify as the budget residual (see Drake et al. (2024) for details). Because we use monthly-mean budget tendencies that are binned online into depth coordinates, however, some error may be introduced by the omission of sub-monthly correlations between the tendencies and σ_2.

In the Labrador Sea (Figure 35a,b), there is a nearly perfect time-mean balance between surface-forced water mass transformations (peaking at $-7\,\text{Sv}$ at $\sigma_2 = 36.6\,\text{kg/m}^3$) and export across the OSNAP-West section. Contributions from other terms, such as the inflow of dense water through Davis Strait and local transformation by mixing processes, are negligible. In the Irminger Sea and Iceland Basin (Figure 35c,d), by contrast, the peak surface-forced water mass transformation is weaker (10 Sv) and occurs at a lighter density ($36.2\,\text{kg/m}^3$) than what is exported southward across the OSNAP-East section, which peaks at 15 Sv for waters denser than $\sigma_2 = 36.4\,\text{kg/m}^3$. The difference between the transformation and overturning transport is due to 1) relatively dense Nordic Sea waters overflowing across the Greenland Scotland Ridge (peaking at $36.6\,\text{kg/m}^3$) and 2) spurious numerical entrainment of the lighter surface-forced waters (as opposed to parameterized entrainment, which is relatively weak). These model results are qualitatively consistent with the observation-based analysis of (Evans et al., 2023), in that the total surface-forced water mass transformation peaks at a lighter density than the overturning across OSNAP and that interior mixing processes play a non-negligible role in this transformation. In both cases, the differences between CM4Xp25 and CM4Xp125 are relatively modest: due to slightly less spurious numerical mixing, CM4Xp125 appears to export less deep water across OSNAP-East at the peak density but more at the highest densities.

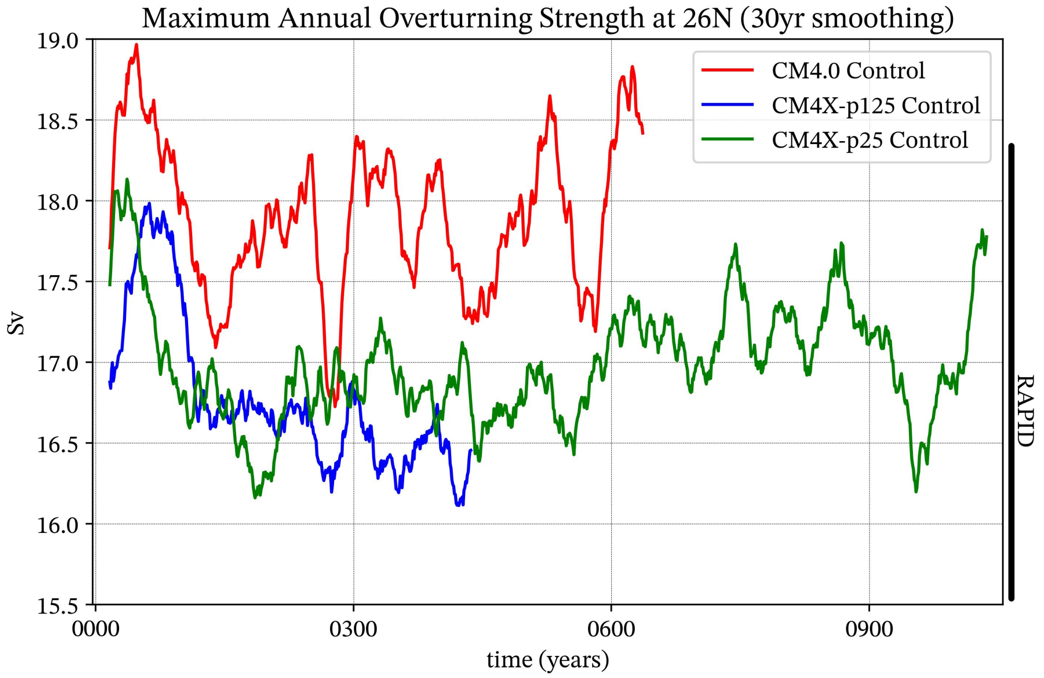

Figure 33. Time series of the maximum annual mean AMOC strength across the RAPID section (26.5°N) for the preindustrial control simulations using CM4.0, CM4X-25, and CM4X-p125. The time series are smoothed with a 30-year running mean. The dark line on the right depicts the range (mean +/- one standard deviation of annual mean) of the observed maximum annual mean AMOC strength across the RAPID section over the period of 2004-2022.

Figure 35e shows the spatial distribution of surface-forced water mass transformations at the isopycnal of peak export across the OSNAP arrays in each region. In the Irminger Sea and Iceland Basin, transformations are concentrated immediately downstream of the Denmark Strait, further downstream along the East Greenland Current, along the Reykjanes Ridge, and on the southern edge of the Iceland Faroe Ridge, qualitatively consistent with observation-based estimates (Petit et al., 2020). In the Labrador Sea, transformations are concentrated along the northwestern continental slope, also qualitatively consistent with observation-based estimates (Zou et al., 2024).

As discussed in Section 6.3, the deep water export across OSNAP-West is unrealistically strong in both CM4X-p25 and CM4X-p125. Our water mass analysis suggests this result is a direct result of too-high surface-forced water mass transformations, supporting the previous section's hypothesis that this is related to a bias in dense layer outcrop areas or mixed-layer depths. Furthermore, our analysis suggests that the light and shallow biases of the CM4X AMOC is in part attributable to spurious numerical mixing, which transforms a few Sv of inflowing dense Nordic waters (with $36.8\,\mathrm{kg/m^3} \leq \sigma_2 \leq 37.2\,\mathrm{kg/m^3}$) towards lighter density classes.

6.5 Conclusions regarding the North Atlantic simulations

While representation of the deep western boundary current is expected to play a role in altering Gulf Stream vertical structure, Ezer (2016), Schoonover et al. (2017) and Debreu et al. (2022) stress the importance of realistically resolving a smooth continental slope so to allow for realistic flow-topography interactions. Resolving flow-topography

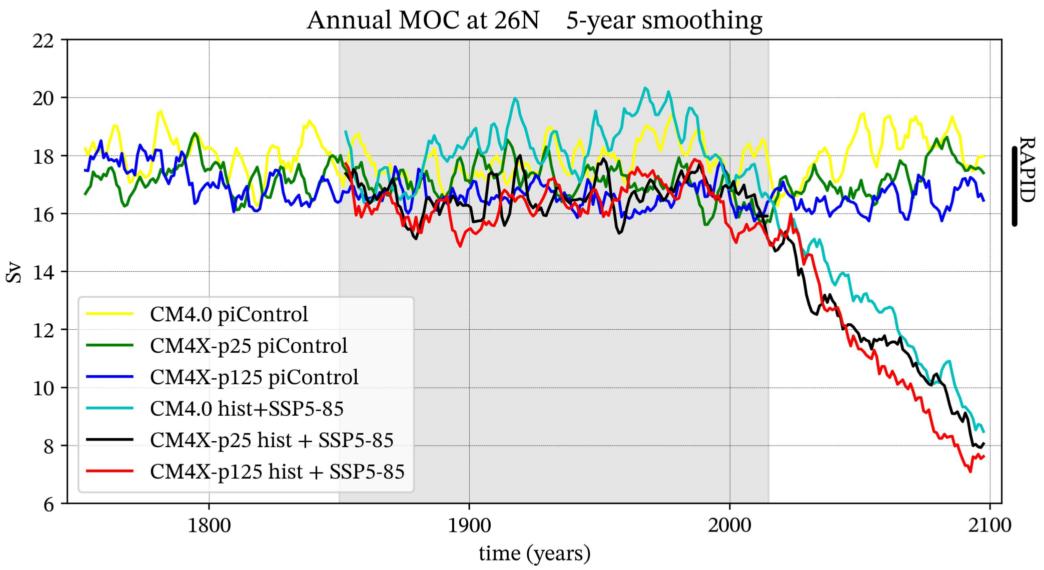

Figure 34. Time series from CM4.0, CM4X-p25, and CM4X-p125 of the maximum annual mean AMOC strength across the RAPID section (26.5°N) for the piControl, historical over years 1850-2014 (shaded region), and the SSP5-8.5 scenario experiment from 2014-2100. The time series are smoothed with a 5-year running mean. The dark line on the right depicts the range (mean +/- one standard deviation of annual mean) of the observed maximum annual mean AMOC strength across the RAPID section over the period of 2004-2022 for comparison.

interactions is a function of model horizontal and vertical grid spacing. Likely the result of improvements on many fronts, Chassignet and Xu (2017) conclude that a transition from unrealistic to realistic representation of the Gulf Stream and offshore mesoscale turbulence should be expected once reaching 1/12° and finer. Hence, we withhold expectation that the Gulf Stream in CM4X should match that in Todd (2021), with the slight improvements in CM4X-p125 relative to CM4X-p25 suggesting a trend in the right direction.

When compared to CM4.0, the CM4X simulations generally have a better representation of the maximum AMOC strength across both the subpolar and the subtropical sections, and multidecadal AMOC changes in the CM4X historical simulations are also improved. Comparing CM4X-p125 and CM4X-p25, the refined grid spacing in CM4X-p125 leads to a slightly better AMOC representation across the OSNAP section, especially across the OSNAP West subsection enclosing the Labrador Sea with improved (reduced) density contrast between the Labrador Sea boundary outflow and inflow. A linear relationship between subpolar gyre strength and the maximum subpolar overturning strength is generally observed in many climate models (Meccia et al., 2021). Thus, the relatively weaker subpolar gyre in CM4X-p125 compared to CM4X-p25 is consistent with the reduction in the maximum AMOC strength across the OSNAP section.

Many North Atlantic modeling biases in CM4.0 also exist in CM4X, such as the overestimation of winter deep convection in the Labrador Sea, the missing of winter deep convection in the Greenland Sea, the underestimation of the strength of the dense Nordic Sea overflows across the OSNAP section, the shallow AMOC depth, and the weak low-frequency AMOC variability in the control simulation. These long-standing modeling biases in the North Atlantic, which often appear in coupled climate models, are key areas for future improvements and more experiments are needed to address the processes

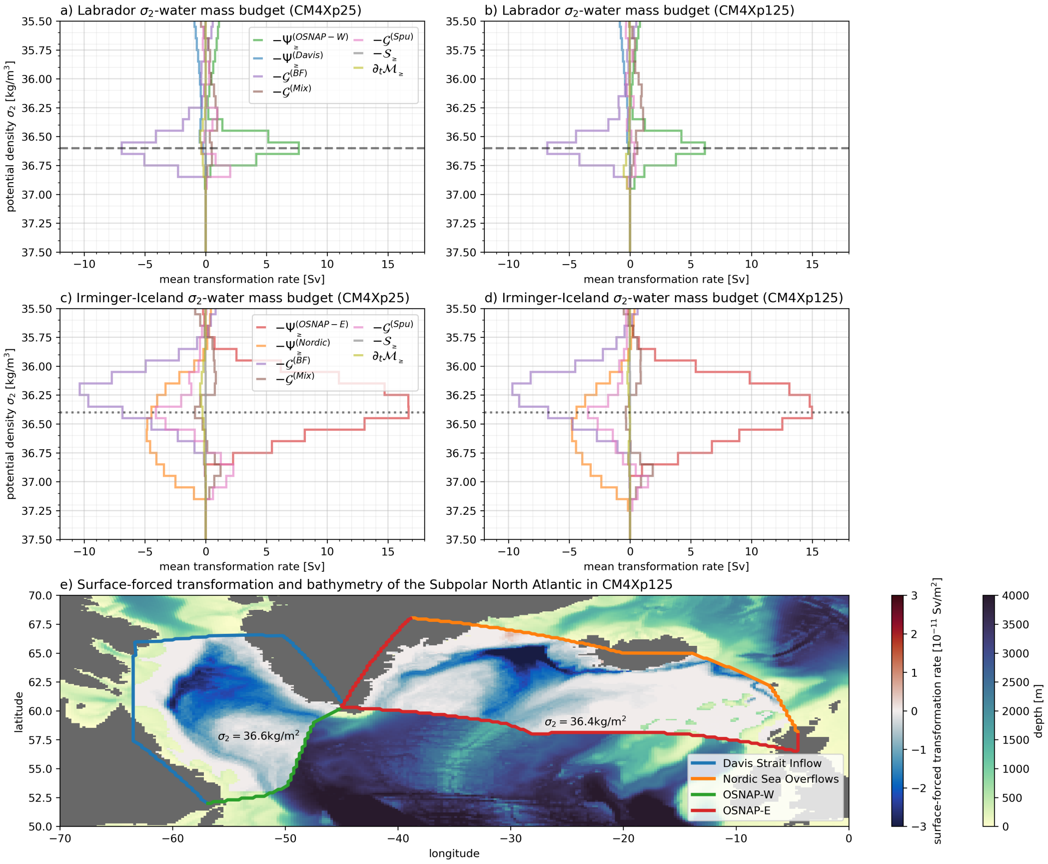

Figure 35. Water mass budgets in potential density (σ_2) coordinates in in CM4Xp25 (a,c) and CM4Xp125 (b,d,e) for a Labrador Sea region (a,b) and a Irminger Sea and Iceland Basin region (c,d). In panel (e), the Labrador region is bounded to the northwest by a Davis Strait section (blue) and to the southeast by the OSNAP-West section (green; see Fu et al. (2023)) while the Irminger-Iceland region is bounded to the northeast by the Greenland-Scotland Ridge (orange) and the to south by the OSNAP-East section (red). Within the two regions, colors represent the rates of water mass transformation per unit area across the isopycnal of peak transport across the OSNAP array (grey lines in a-d) due to boundary fluxes (i.e. the integrand of $\mathcal{G}^{(\mathrm{BF})}$, which is dominated by air-sea heat and freshwater fluxes)Outside of the two regions, colors represent seafloor depth. All terms in the water mass budget are computed following Drake et al. (2024) and represent the 2010-2024 time-mean of monthly-mean transformation rates diagnosed from the forced (historical + SSP5-8.5) scenarios. [Overturning transport values differ from those reported in Section 6.3 because we use different density coordinates, averaging intervals, and forcing scenarios.]

involved in reducing these biases. In particular, future improvements in the strength of the dense Nordic Sea overflows (including both the Denmark Strait overflow and the Iceland-Scotland overflow) across the subpolar section will enhance the vertical stratification in the Labrador Sea and thus reduce the Labrador Sea winter deep convection strength, as well as deepen the downstream AMOC. Future improvements in the Arctic dense water formation might also contribute to the improvement of the dense Nordic Sea overflows and the downstream AMOC.

7 Strategies for ocean climate model development

CM4X exemplifies the value of a hierarchy of coupled climate models where the only difference is the grid spacing. Such hierarchies provide a means to expose the importance, or lack thereof, for the enhanced representation of dynamical processes. Having two or more model configurations among a hierarchy provokes questions that go unasked with a single model. We propose that the expansion of model phase space to include carefully built hierarchies, such as the CM4X whose members represent a vigorous ocean mesoscale, is a useful, if not necessary, step to furthering the science going into climate models and the science emerging from simulations.

Although CM4X-p125 reaches thermal equilibrium in a remarkably short time (see the Part I analysis in Griffies et al. (2024)), the analysis in this paper revealed that there remain many familiar biases in need of being addressed in future development projects. Particular biases revealed in our analysis include: weak interannual variability in the tropics, as revealed by the skewness of the sea level (Figure 3 in Section 2.2); a poor representation of the eastern boundary upwelling zones, in part due to under-resolved ocean and atmosphere dynamics as well as the representation of low level clouds (Section 3); biases in the sea ice seasonal cycle, reflecting a number of possible process biases (Section 4); overly strong ventilation properties of mode and intermediate waters of the Southern Ocean, likely related to under-resolved mesoscale eddy processes at the high latitudes (Figure 9 of Part I); an overly diffuse Gulf Stream as it leaves the American coast, possibly due to over-dissipation (Sections 2.2 and 6.1); overly shallow overflows in the North Atlantic (Figure 27) and overly deep mixed layers in the Labrador Sea (Figure 28), likely due to under-resolved mesoscale processes and too much entrainment in the deep overflows. Even with these remaining issues, each of the biases are generally reduced in CM4X-p125 relative to CM4X-p25. From an oceanographic perspective, the key advantage of rapid thermal equilibration is that it allows one to examine ocean properties within a thermally equilibrated climate model, with those properties not having drastically drifted outside their physically sensible range. So although CM4X-p125 has many biases in need of further reduction, we propose that it provides a powerful venue for studying climate dynamics.

There are compelling arguments that, for purposes of centennial climate studies and climate change projections, the community should prioritize advances in numerical and computational methods to facilitate the direct simulation of the ocean mesoscale (Silvestri et al., 2024). The order of magnitude reduction in thermal equilibration time found in the CM4X-p125 simulation, relative to CM4X-p25, bolsters that argument. That is, the finer ocean grid used in CM4X-p125 has four times the number of grid points and uses half the time step; however, this factor of eight added expense for CM4X-p125 is offset by its factor of ten shorter thermal equilibration time. Advances signaled by CM4X-p125 were supported by advances in the numerical methods related to the MOM6 vertical Lagrangian dynamical core (Griffies et al., 2020). In particular, the quasi-isopycnal vertical coordinate used in the ocean interior greatly reduces spurious diapycnal mixing relative to a geopotential-like coordinate (Adcroft et al., 2019). Spurious numerical mixing is very difficult to minimize using standard Eulerian based numerical methods in the presence of a strong downscale cascade of tracer variance enabled by mesoscale eddies (Griffies et al., 2000). Developments leading to the MOM6 dynamical core were partly motivated to address this difficulty, particularly when confronted with the hundreds to thousands of mesoscale eddy turnover timescales accessed by climate simulations, thus allowing for seemingly small numerical errors to accumulate to have nontrivial detrimental impacts on stratification. Coupling advances in numerical methods to advances in computational hardware and software could render CM4X-p125 the coarsest, not the finest, member of a future climate model hierarchy.

8 Open Research

Software comprising the model as well as the software used for creating the figures will be placed on Zenodo at the revision stage of this work.

Observation-based datasets used in this paper are cited locally. We are indebted to the many efforts of the various programs providing observational-based data used to help evaluate these simulations, including the following.

- The Argo program provides data that were collected and made freely available by the International Argo Program and the national programs that contribute to it, with access available from

  http://www.argo.ucsd.edu and http://argo.jcommops.org

  The Argo Program is part of the Global Ocean Observing System.
- OSNAP data were collected and made freely available by the OSNAP (Overturning in the Subpolar North Atlantic Program) project and all the national programs that contribute to it (www.o-snap.org). The DOI for this data set is

  https://doi.org/10.35090/gatech/70342

Acknowledgments
This project started in May 2020, during the early stages of the Covid-19 pandemic shutdown. We are grateful to the GFDL computer operations team for keeping the computational resources reliable during the shutdown. We thank the GFDL management for providing the computer resources needed for the development and analysis documented here. We thank John Dunne and Matthew Harrison for very helpful comments on early drafts of this manuscript. Krista Dunne, Sergey Malyshev and Chris Milly kindly provided expertise in helping to update the rivers and lakes for use with the C192 atmosphere coupled to the p125 ocean. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce.

- The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce.
- A.A. was supported by Award NA18OAR4320123 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce.
- R.B. was supported under NSF Division of Polar Programs Grant NSF2319828.
- C.Y.C. was supported by Award NA19OAR4310365 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce.
- H.F.D. was supported by the NOAA Climate and Global Change Postdoctoral Fellowship Program, administered by UCAR's Cooperative Programs for the Advancement of Earth System Science (CPAESS) under Award NA18NWS4620043B.
- H.K. acknowledges the support from Natural Environment Research Council grants NE/T013494/1 and NE/W001543/1.
- M.L. was supported by award NA18OAR4320123 and NA23OAR4320198 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce.
- G.A.M was supported by NSF (PLR-1425989) and UKRI (MR/W013835/1).
- A.S. was supported by Schmidt Sciences, LLC under the M^{2}LInES project.
- K.E.T acknowledges support from the Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) Project under NSF Awards PLR-1425989 and OPP-1936222 and 2332379.
- L.Z. was supported by Schmidt Sciences, LLC under the M^{2}LInES project, NSF grant OCE 1912357 and NOAA CVP NA19OAR4310364.

- W.Z. was supported by the National Science Foundation under Grant Number F1240-01(NSF OCE 1912357). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Adcroft, A., Anderson, W., Blanton, C., Bushuk, M., Dufour, C. O., Dunne, J. P., ... Zhang, R. (2019). The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *Journal of Advances in Modeling the Earth System*, *JAMES*. doi: 10.1029/2019MS001726

Albert, A., Echevin, V., Lévy, M., & Aumont, O. (2010). Impact of nearshore wind stress curl on coastal circulation and primary productivity in the Peru upwelling system. *Journal of Geophysical Research: Oceans*, *115*(C12). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JC006569` doi: https://doi.org/10.1029/2010JC006569

Andrews, T., Gregory, J. M., Paynter, D., Silvers, L. G., Zhou, C., Mauritsen, T., ... Titchner, H. (2018). Accounting for changing temperature patterns increases historical estimates of climate sensitivity. *Geophysical Research Letters*, *45*, 8490–8499. doi: 10.1029/2018GL078887

Argo. (2023). *Argo float data and metadata from Global Data Assembly Centre (Argo GDAC)* [dataset]. SEANOE. Retrieved from `https://www.seanoe.org/data/00311/42182/` doi: 10.17882/42182

Armour, K. C., Bitz, C. M., & Roe, G. H. (2013). Time-varying climate sensitivity from regional feedbacks. *Journal of Climate*, *26*, 4518–4534. doi: 10.1175/JCLI-D-12-00544.1

Bakun, A., Black, B. A., Bograd, S. J., Garcia-Reyes, M., Miller, A. J., Rykaczewski, R. R., & Sydeman, W. J. (2015). Anticipated effects of climate change on coastal upwelling ecosystems. *Current Climate Change Reports*, *1*, 85–93.

Beadling, R. L. (2023). Global consequences of regional connectivity along the Antarctic margin. *Journal of Geophysical Research: Oceans*, *128*. doi: 10.1029/2023JC019908

Beadling, R. L., Krasting, J. P., Griffies, S. M., Hurlin, W. J., Bronselear, B., Russell, J. L., ... Winton, M. (2022). Importance of the Antarctic Slope Current in the Southern Ocean response to ice sheet melt and wind stress change. *Journal of Geophysical Research - Oceans*, *127*. doi: 10.1029/2021jc017608

Beadling, R. L., Russell, J. L., Stouffer, R. J., Mazloff, M., Talley, L. D., Goodman, P. J., ... Pandde, A. (2020). Representation of Southern Ocean properties across coupled model intercomparison project generations: CMIP3 to CMIP6. *Journal of Climate*, *33*, 6555–6581. doi: 10.1175/JCLI-D-19-0970.1

Bograd, S. J., Jacox, M. G., Hazen, E., Lovecchio, E., Montes, I., Buil, P. M., ... Rykaczewski, R. R. (2023). Climate change impacts on eastern boundary upwelling systems. *Annual Review of Marine Science*, *15*(1), 303-328. Retrieved from `https://doi.org/10.1146/annurev-marine-032122-021945` (PMID: 35850490) doi: 10.1146/annurev-marine-032122-021945

Bronselaer, B., Winton, M., Griffies, S. M., Hurlin, W. J., Rodgers, K. B., Sergienko, O. V., ... Russell, J. L. (2018). Change in future climate due to Antarctic meltwater. *Nature*, *564*(7734), 53.

Bushuk, M., Zhang, Y., Winton, M., Hurlin, B., Delworth, T., Lu, F., ... others (2022). Mechanisms of regional Arctic sea ice predictability in two dynamical seasonal forecast systems. *J. Climate*, *35*(13), 4207–4231.

Ceppi, P., Myers, T. A., Nowack, P., Wall, C. J., & Zelinka, M. D. (2024). Implications of a pervasive climate model bias for low-cloud feedback. *Geophysical Research Letters*, *51*. doi: 10.1029/2024GL110525

Chassignet, E. P., & Marshall, D. P. (2008). Gulf Stream separation in numerical ocean models. *Geophysical Monograph Series*, *177*. doi: https://doi.org/10.1029/177GM05

Chassignet, E. P., & Xu, X. (2017). Impact of horizontal resolution (1/12° to 1/50°) on Gulf Stream separation, penetration, and variability. *Journal of Physical Oceanography*, *47*(10), 1999-2021. doi: 10.1175/JPO-D-17-0031.1

Cheng, L., Trenberth, K. E., Fasullo, J., Boyer, T., Abraham, J., & Zhu, J. (2017). Improved estimates of ocean heat content from 1960 to 2015. *Science Advances*, *3*(3), e1601545. doi: 10.1126/sciadv.1601545

Cunningham, S. A., Kanzow, T., Rayner, D., Baringer, M. O., Johns, W. E., Marotzke, J., ... Bryden, H. L. (2007). Temporal variability of the Atlantic meridional overturning circulation at 26.5° n. *Science*, *317*. doi: 10.1126/science.1141304

Danabasoglu, G., Large, W., & Briegleb, B. (2010). Climate impacts of parameterized Nordic Sea overflows. *Journal of Geophysical Research*, *115, C11005*. doi: 10.1029/2010JC006243

Danabasoglu, G., Yeager, S. G., Bailey, D., Behrens, E., Bentsen, M., Bi, D., ... Wang, Q. (2014). North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part I: Mean states. *Ocean Modelling*, *73*(0), 76–107. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1463500313001868` doi: 10.1016/j.ocemod.2013.10.005

Dangendorf, S., Hay, C., Calafat, F. M., Marcos, M., Piecuch, C. G., Berk, K., & Jensen, J. (2019). Steric sea level rise in GFDL CM4 and ESM4: Insights into model drift and water mass representation. *Nature Climate Change*, *9*, 705–710. doi: 10.1038/s41558-019-0531-8

Debreu, L., Kevlahan, N., & Marchesiello, P. (2022). Improved Gulf Stream separation through Brinkman penalization. *Ocean Modeling*, *179*. doi: https://doi.org/10.1016/j.ocemod.2022.102121

Delworth, T. L., Broccoli, A. J., Rosati, A., Stouffer, R. J., Balaji, V., Beesley, J. A., ... Zhang, R. (2006). GFDL's CM2 global coupled climate models - Part 1: Formulation and simulation characteristics. *Journal of Climate*, *19*, 643–674. doi: 10.1175/JCLI3629.1

DeRepentigny, P., Jahn, A., Holland, M. M., & Smith, A. (2020). Arctic sea ice in two configurations of the CESM2 during the 20th and 21st centuries. *J. Geophys. Res.: Oceans*, *125*(9), e2020JC016133.

DeVries, T., Le Quéré, C., Berthet, S., Hauck, J., Ilyina, T., Landschützer, P., ... Séférian, R. (2019). Decadal trends in the ocean carbon sink. *Proceedings of the National Academy of Sciences*, *116*, 11646–11651. doi: 10.1073/pnas.1900371116

Donner, L. J., O'Brien, T. A., Rieger, D., Vogel, B., & Cooke, W. F. (2016). Are atmospheric updrafts a key to unlocking climate forcing and sensitivity? *Atmospheric Chemistry and Physics*, *16*. doi: 10.5194/acp-16-12983-2016

Donohue, K. A., Tracey, K. L., Watts, D. R., Chidichimo, M. P., & Chereskin, T. K. (2016). Mean Antarctic Circumpolar Current transport measured in Drake Passage. *Geophysical Research Letters*, *43*, 11,760–11,767. doi: 10.1002/2016GL070319

Drake, H. F., Bailey, S., Dussin, R., Griffies, S. M., Krasting, J. P., MacGilchrist, G. A., ... Zika, J. D. (2024, April). *Water mass transformation budgets in finite-volume generalized vertical coordinate ocean models.* Retrieved 2024-05-31, from `https://essopenarchive.org/users/242115/articles/788549-water-mass-transformation-budgets-in-finite-volume-generalized-vertical-coordinate-ocean-models?commit=e6178873ac99d67c98963b54ba60699d57b7703f` doi: 10.22541/essoar.171284935.57181910/v1

Dunne, J. P., Horowitz, L. W., Adcroft, A., Ginoux, P., Held, I. M., John, J. G., ... Zhao, M. (2020). The GFDL Earth system model version 4.1 (GFDL-ESM4.1):

Model description and simulation characteristics. *Journal of Advances in Modeling Earth Systems*, *12*. doi: 10.1029/2019MS002015

Dunne, J. P., John, J. G., Hallberg, R. W., Griffies, S. M., Shevliakova, E. N., Stouffer, R. J., … Zadeh, N. (2012). GFDL's ESM2 global coupled climate-carbon Earth System Models Part I: Physical formulation and baseline simulation characteristics. *Journal of Climate*, *25*, 6646-–6665. doi: 10.1175/JCLI-D-11-00560.1

Enriquez, A. G., & Friehe, C. A. (1995). Effects of wind stress and wind stress curl variability on coastal upwelling. *Journal of Physical Oceanography*, *25*(7), 1651 - 1671. Retrieved from `https://journals.ametsoc.org/view/journals/phoc/25/7/1520-0485_1995_025_1651_eowsaw_2_0_co_2.xml` doi: https://doi.org/10.1175/1520-0485(1995)025⟨1651:EOWSAW⟩2.0.CO;2

Evans, D. G., Holliday, N. P., Bacon, S., & Le Bras, I. (2023, June). Mixing and air–sea buoyancy fluxes set the time-mean overturning circulation in the subpolar North Atlantic and Nordic Seas. *Ocean Science*, *19*(3), 745–768. Retrieved 2023-09-18, from `https://os.copernicus.org/articles/19/745/2023/` (Publisher: Copernicus GmbH) doi: 10.5194/os-19-745-2023

Eyring, V., Bony, S., Meehl, J. A., Senior, C., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organisation. *Geoscientific Model Development*, *9*, 1937–1958. Retrieved from `http://www.geosci-model-dev.net/9/1937/2016/` doi: 10.5194/gmd-9-1937-2016

Ezer, T. (2016). Revisiting the problem of the Gulf Stream separation: on the represenation of topography in ocean models with different types of vertical grids. *Ocean Modelling*, *104*. doi: https://doi.org/10.1016/j.ocemod.2016.05.008

Feltham, D. L. (2008). Sea ice rheology. *Annu. Rev. Fluid Mech.*, *40*, 91–112.

FRAM Group. (1991). An eddy-resolving model of the Southern Ocean. *EOS Transactions of the American Geophysical Union*, *72*, 169–175. doi: 10.1029/90EO00128

Fraser, A., Wongpan, P., Langhorne, P., Klekociuk, A., Kusahara, K., Lannuzel, D., … others (2023). Antarctic landfast sea ice: A review of its physics, biogeochemistry and ecology. *Reviews of Geophysics*, *61*(2), e2022RG000770.

Frederikse, T., Landerer, F., Caron, L., Adhikari, S., Parkes, D., Humphrey, V. W., … Wu, Y.-H. (2020). The causes of sea-level rise since 1900. *Nature*, *584*(7821), 393–397. doi: 10.1038/s41586-020-2591-3

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., & Winton, M. (2015). Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *Journal of Climate*, *28*, 862-886. doi: 10.1175/JCLI-D-14-00117.1

Fu, Y., Lozier, M. S., Biló, T. C., Bower, A. S., Cunningham, S. A., Cyr, F., … Yashayaev, I. (2023). Seasonality of the meridional overturning circulation in the subpolar north atlantic [dataset]. *Communications Earth & Environment*, *4*(1), 181.

Gent, P. R., Yeager, S. G., Neale, R. B., Levis, S., & Bailey, D. A. (2010). Improvements in a half degree atmosphere/land version of the ccsm. *Climate Dynamics*, *34*, 819–833.

Goddard, P., Dufour, C., Yin, J., Griffies, S. M., & Winton, M. (2017). CO2-induced ocean warming of the Antarctic continental shelf in an eddying global climate model. *Journal of Geophysical Research - Oceans*, *122*. doi: 10.1002/2017JC012849

Gordon, C., Rosati, A., & Gudgel, R. (2000). Tropical sensitivity of a coupled model to specified isccp low clouds. *Journal of Climate*, *13*(13), 2239–2260.

Griffies, S. M., Adcroft, A., Alberty, M. S., Beadling, R. L., Bushuk, M., Chang, C.-Y., … Zika, J. D. (2024). The GFDL-CM4X climate model hierarchy, Part I: thermal properties and the mesoscale dominance hypothesis. *Journal of*

Advances in Modeling Earth Systems.

Griffies, S. M., Adcroft, A., & Hallberg, R. W. (2020). A primer on the vertical Lagrangian-remap method in ocean models based on finite volume generalized vertical coordinates. *Journal of Advances in Modeling Earth Systems*, *12*. doi: 10.1029/2019MS001954

Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Böning, C. W., ... Yeager, S. G. (2016). OMIP contribution to CMIP6: experimental and diagnostic protocol for the physical component of the ocean model intercomparison project. *Geoscientific Model Development*, *9*, 3231–3296. doi: 10.5194/gmd-9-3231-2016

Griffies, S. M., Pacanowski, R. C., & Hallberg, R. W. (2000). Spurious diapycnal mixing associated with advection in a z-coordinate ocean model. *Monthly Weather Review*, *128*, 538–564. doi: 10.1175/1520-0493(2000)128⟨0538:SDMAWA⟩2.0.CO;2

Griffies, S. M., Winton, M., Anderson, W. G., Benson, R., Delworth, T. L., Dufour, C., ... Zhang, R. (2015). Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *Journal of Climate*, *28*, 952–977. doi: 10.1175/JCLI-D-14-00353.1

Griffies, S. M., Winton, M., Donner, L. J., Horowitz, L. W., Downes, S. M., Farneti, R., ... others (2011). The GFDL CM3 coupled climate model: Characteristics of the ocean and sea ice simulations. *J. Climate*, *24*(13), 3520–3544.

Hallberg, R. W. (2013). Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects. *Ocean Modelling*, *72*, 92-103. doi: 10.1016/j.ocemod.2013.08.007

Hallberg, R. W., Adcroft, A. J., Dunne, J. P., Krasting, J. P., & Stouffer, R. (2013). Sensitivity of twenty-first-century global-mean steric sea level rise to ocean model formulation. *Journal of Climate*, *26*, 2947-–2956. doi: 10.1175/JCLI-D-12-00506.1

Hallberg, R. W., & Gnanadesikan, A. (2006). On the role of eddies in determining the structure and response of the wind-driven southern hemisphere overturning: Results from the Modeling Eddies in the Southern Ocean (MESO) project. *Journal of Physical Oceanography*, *36*, 2232–2252. doi: 10.1175/JPO2980.1

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., ... Zadeh, N. (2019). Structure and performance of GFDL's CM4.0 climate model. *Journal of Advances in Modeling the Earth System*, *11*, 3691–3726. doi: 10.1029/2019MS001829

Heuzé, C. (2021). Antarctic Bottom Water and North Atlantic Deep Water in CMIP6 models. *Ocean Science*, *17*. doi: 10.5194/os-17-59-2021

Holland, W. R. (1967). On the wind-driven circulation in an ocean with bottom topography. *Tellus*, *19*(4), 582–600. doi: 10.3402/tellusa.v19i4.9825

Horvat, C. (2021). Marginal ice zone fraction benchmarks sea ice and climate model skill. *Nature Communications*, *12*(1), 2221.

Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., ... Zhang, H.-M. (2020). Improvements of the daily optimum interpolation sea surface temperature (DOISST) Version 2.1. *Journal of Climate*, *34*, 2923–2939. doi: 10.1175/JCLI-D-20-0166.1

Hughes, C. W., & De Cuevas, B. A. (2001). Why western boundary currents in realistic oceans are inviscid: A link between form stress and bottom pressure torques. *Journal of Physical Oceanography*, *31*(10), 2871–2885. doi: 10.1175/1520-0485(2001)031⟨2871:WWBCIR⟩2.0.CO;2

Hughes, C. W., Thompson, A. F., & Wilson, C. (2010). Identification of jets and mixing barriers from sea level and vorticity measurements using simple statistics. *Ocean Modelling*, *32*, 44–57. doi: 10.1016/j.ocemod.2009.10.004

Huneke, W. G. C., Morrison, A. K., & Hogg, A. M. (2022). Spatial and subannual variability of the Antarctic Slope Current in an eddying ocean sea ice model. *Journal of Physical Oceanograph*, *52*, 347–361. doi: 10.1175/jpo-d-21-0143.1

Hurlburt, H. E., & Hogan, P. J. (2000). Impact of 1/8 to 1/64 resolution on Gulf Stream model-data comparisons in basin-scale subtropical atlantic ocean models. *Dynamics of Atmospheres and Oceans*, *32*. doi: https://doi.org/10.1016/S0377-0265(00)00050-6

Hutter, N., Losch, M., & Menemenlis, D. (2018). Scaling properties of Arctic sea ice deformation in a high-resolution viscous-plastic sea ice model and in satellite observations. *Journal of Geophysical Research: Oceans*, *123*(1), 672–687.

Ishii, M., Fukuda, Y., Hirahara, S., Yasui, S., Suzuki, T., & Sato, K. (2017). Accuracy of global upper ocean heat content estimation expected from present observational data sets. *SOLA*, *13*(0), 163–167. doi: 10.2151/sola.2017-030

Jacobs, S. S. (2004). Bottom water production and its links with the thermohaline circulation. *Antarctic Science*, *16*(4), 427–437. doi: 10.1017/S095410200400224X

Jacox, M. G., & Edwards, C. A. (2012). Upwelling source depth in the presence of nearshore wind stress curl. *Journal of Geophysical Research: Oceans*, *117*(C5). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JC007856` doi: https://doi.org/10.1029/2011JC007856

Jahn, A., Holland, M. M., & Kay, J. E. (2024). Projections of an ice-free Arctic Ocean. *Nature Reviews Earth & Environment*, 1–13.

Jahn, A., Kay, J. E., Holland, M. M., & Hall, D. M. (2016). How predictable is the timing of a summer ice-free Arctic? *Geophys. Res. Lett.*, *43*(17), 9113–9120.

Jevrejeva, S., H, P., & Jackson, L. P. (2021). Global mean thermosteric sea level projections by 2100 in CMIP6 climate models. *Environmental Research Letters*, *16*(1), 705–710. doi: 10.1088/1748-9326/abceea

Junker, T., Schmidt, M., & Mohrholz, V. (2015). The relation of wind stress curl and meridional transport in the Benguela upwelling system. *Journal of Marine Systems*, *143*, 1-6. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0924796314002425` doi: https://doi.org/10.1016/j.jmarsys.2014.10.006

Kaspar, F., Hollmann, R., Lockhoff, M., Karlsson, K.-G., Dybbroe, A., Fuchs, P., ... Schulz, J. (2009). Operational generation of AVHRR-based cloud products for Europe and the Arctic at EUMETSAT's Satellite Application Facility on Climate Monitoring (CM-SAF). *Advances in Science and Research*, *3*(1), 45–51.

Kay, J. E., Wall, C., Yettella, V., Medeiros, B., Hannay, C., Caldwell, P., & Bitz, C. (2016). Global climate impacts of fixing the Southern Ocean shortwave radiation bias in the Community Earth System Model (CESM). *Journal of Climate*, *29*(12), 4617–4636.

Keen, A., Blockley, E., Bailey, D., Debernard, J. B., Bushuk, M., Delhaye, S., ... Wyser, K. (2021). An inter-comparison of the mass budget of the Arctic sea ice in CMIP6 models. *The Cryosphere*, *15*(2), 951–982.

Khatri, H., Griffies, S. M., Storer, B. A., Buzzicotti, M., Aluie, H., Sonnewald, M., ... Shao, A. (2024). A scale-dependent analysis of the barotropic vorticity budget in a global ocean simulation. *Journal of Advances in Modeling Earth Systems*, *16*(6), e2023MS003813. doi: 10.1029/2023MS003813

Krasting, J. P., Griffies, S. M., Tesdal, J.-E., MacGilchrist, G., Beadling, R. L., & Little, C. M. (2024). Steric sea level rise and relationships with model drift and water mass representation in GFDL CM4 and ESM4. *Journal of Climate*, *accepted*.

LaCasce, J. (2017). The prevalence of oceanic surface modes. *Geophysical Research Letters*, *44*, 11,097–11,105. doi: 10.1002/2017GL075430

Landy, J. C., Dawson, G. J., Tsamados, M., Bushuk, M., Stroeve, J. C., Howell, S. E., ... others (2022). A year-round satellite sea-ice thickness record from

CryoSat-2. *Nature*, *609*(7927), 517–522.
Lavergne, T., Eastwood, S., Teffah, Z., Schyberg, H., & Breivik, L.-A. (2010). Sea ice motion from low-resolution satellite sensors: An alternative method and its validation in the Arctic. *J. Geophys. Res.: Oceans*, *115*(C10).
Lellouche, J.-M., Greiner, E., Bourdallé-Badie, R., Garric, G., Melet, A., Drévillon, M., ... Le Traon, P.-Y. (2021). The Copernicus global 1/12° oceanic and sea ice GLORYS12 reanalysis. *Frontiers in Earth Science*, *9*. doi: 10.3389/feart.2021.698876
Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., Locarnini, R. A., ... Zweng, M. M. (2012, 05). World ocean heat content and thermosteric sea level change (0-2000 m), 1955-2010. *Geophysical Research Letters*, *39*(10), n/a – n/a. doi: 10.1029/2012gl051106
Li, F., Lozier, M. S., Bacon, S., Bower, A. S., Cunningham, S. A., de Jong, M. F., ... Zhou, C. (2021). Subpolar North Atlantic western boundary density anomalies and the meridional overturning circulation. *Nature Communications*, *12*. doi: 10.1038/s41467-021-23350-2
Lim, F., Lozier, M. S., Danagasoglu, G., Holliday, N. P., Kwon, Y.-O., Romanou, A., ... Zhang, R. (2019). Local and downstream relationships between Labrador Sea Water volume and North Atlantic meridional overturning circulation variability. *Journal of Climate*, *32*, 3883-3898. doi: 10.1175/JCLI-D-18-0735.1
Lockwood, J. W., Griffies, C. O. D. S. M., & Winton, M. (2021). On the role of the Antarctic Slope Front on the occurrence of the Weddell Sea polynya under climate change. *Journal of Climate*, *34*, 2529–2548. doi: 10.1175/JCLI-D-20-0069.1
Loose, N., Abernathey, R., Grooms, I., Busecke, J., Guillaumin, A., Yankovsky, E., ... others (2022). GCM-filters: A Python package for diffusion-based spatial filtering of gridded data. *Journal of Open Source Software*, *7*(70).
Lozier, M. S., Li, F., Bacon, S., Bahr, F., Bower, A. S., Cunningham, S. A., ... Zhao, J. (2019). A sea change in our view of overturning in the subpolar North Atlantic. *Science*, *363*(6426), 516-521. Retrieved from `https://www.science.org/doi/abs/10.1126/science.aau6592` doi: 10.1126/science.aau6592
McCarthy, G. D., Smeed, D. A., Johns, W. E., Frajka-Williams, E., Moat, B. I., Rayner, D., ... Bryden, H. L. (2015). Measuring the Atlantic meridional overturning circulation at 26°N. *Progress in Oceanography*, *130*, 91-111. doi: 10.1016/j.pocean.2014.10.006
Meccia, V. L., Iovino, D., & Bellucci, A. (2021). North Atlantic gyre circulation in PRIMAVERA models. *Climate Dynamics*, *56*(11), 4075–4090. doi: 10.1007/s00382-021-05686-z
Meehl, G. A., Arblaster, J. M., Bitz, C. M., Chung, C. T., & Teng, H. (2016). Antarctic sea-ice expansion between 2000 and 2014 driven by tropical Pacific decadal climate variability. *Nat. Geosci.*, *9*(8), 590–595.
Meier, W. N., Fetterer, F., Windnagel, A. K., & Stewart, S. (2021). *NOAA/NSIDC climate data record of passive microwave sea ice concentration, version 4.* National Snow and Ice Data Center. Retrieved from `https://nsidc.org/data/g02202` doi: 10.7265/efmz-2t65
Menary, M. B., Robson, J., Allan, R. P., Booth, B. B. B., Cassou, C., Gastineau, G., ... Zhang, R. (2020). Aerosol-forced AMOC changes in CMIP6 historical simulations. *Geophysical Research Letters*, *47*. doi: 10.1029/2020GL088166
Moorman, R., Morrison, A. K., & Hogg, A. M. C. (2020). Thermal responses to Antarctic ice shelf melt in an eddy-rich global ocean-sea ice model. *Journal of Climate*, *33*. doi: 0.1175/jcli-d-19-0846.1
Morrison, A. K., Waugh, D. W., Hogg, A. M., Jones, D. C., & Abernathey, R. P. (2022). Ventilation of the Southern Ocean pycnocline. *Annual Review of Marine Science*, *14*, 15.1–15.26. doi: 10.1146/annurev-marine-010419-011012

Paolo, F. S., Fricker, H. A., & Padman, L. (2015). Volume loss from Antarctic ice shelves is accelerating. *Science*, *348*, 327–331. doi: 10.1126/science.aaa0940

Parsons, A. (2006). A two-layer model of Gulf Stream separation. *Journal of Fluid Mechanics*, *39*. doi: https://doi.org/10.1017/S0022112069002308

Petit, T., Lozier, M. S., Josey, S. A., & Cunningham, S. A. (2020). Atlantic Deep Water Formation Occurs Primarily in the Iceland Basin and Irminger Sea by Local Buoyancy Forcing. *Geophysical Research Letters*, *47*(22), e2020GL091028. Retrieved 2024-07-03, from `https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL091028` (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL091028) doi: 10.1029/2020GL091028

Pickett, M. H., & Paduan, J. D. (2003). Ekman transport and pumping in the California Current based on the U.S. Navy's high-resolution atmospheric model (COAMPS). *Journal of Geophysical Research: Oceans*, *108*(C10). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003JC001902` doi: https://doi.org/10.1029/2003JC001902

Purich, A., Cai, W., England, M. H., & Cowan, T. (2016). Evidence for link between modelled trends in Antarctic sea ice and underestimated westerly wind changes. *Nat. Comms.*, *7*, 10409.

Purich, A., & England, M. H. (2021). Historical and future projected warming of Antarctic Shelf Bottom Water in CMIP6 models. *Geophysical Research Letters*, *48*. doi: 10.1029/2021GL092752

Rackow, T., Danilov, S., Goessling, H. F., Hellmer, H. H., Sein, D. V., Semmler, T., … Jung, T. (2022). Delayed Antarctic sea-ice decline in high-resolution climate change simulations. *Nature communications*, *13*(1), 637.

Richter, I. (2015). Climate model biases in the eastern tropical oceans: Causes, impacts and ways forward. *Wiley Interdisciplinary Reviews: Climate Change*, *6*(3), 345–358.

Ricker, R., Hendricks, S., Helm, V., Skourup, H., & Davidson, M. (2014). Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation. *The Cryosphere*, *8*(4), 1607–1622.

Rintoul, S. R. (2018). The global influence of localized dynamics in the Southern Ocean. *Nature*, *558*, 209–218. doi: 10.1038/s41586-018-0182-3

Rintoul, S. R., Hughes, C. W., & Olbers, D. (2001). The Antarctic Circumpolar Current system. In G. Siedler, J. Gould, & J. Church (Eds.), *Ocean circulation and climate, 1st edition* (Vol. 77, pp. 271–301). Academic Press.

Rintoul, S. R., & Naveira Garabato, A. C. (2013). Dynamics of the Southern Ocean circulation. In G. Siedler, S. M. Griffies, J. Gould, & J. Church (Eds.), *Ocean circulation and climate, 2nd edition: A 21st century perspective* (Vol. 103, pp. 471–492). Academic Press. doi: 10.1016/B978-0-12-391851-2.00018-0

Risien, C. M., & Chelton, D. B. (2008). A global climatology of surface wind and wind stress fields from eight years of QuikSCAT scatterometer data. *Journal of Physical Oceanography*, *38*(11), 2379 - 2413. Retrieved from `https://journals.ametsoc.org/view/journals/phoc/38/11/2008jpo3881.1.xml` doi: https://doi.org/10.1175/2008JPO3881.1

Roach, L. A., Dörr, J., Holmes, C. R., Massonnet, F., Blockley, E. W., Notz, D., … others (2020). Antarctic sea ice area in CMIP6. *Geophys. Res. Lett.*, *47*(9), e2019GL086729.

Robson, J., Menary, M. B., Sutton, R. T., Mecking, J., Gregory, J. M., Jones, C., … Wilcox, L. J. (2022). The role of anthropogenic aerosol forcing in the 1850–1985 strengthening of the AMOC in CMIP6 historical simulations. *Journal of Climate*, *35*, 6843–6863. doi: 10.1175/JCLI-D-22-0124.1

Roemmich, D., Church, J., Gilson, J., Monselesan, D., Sutton, P., & Wijffels, S. (2015). Unabated planetary warming and its ocean structure since 2006. *Nature Climate Change*, *5*, 240–245. doi: 10.1038/nclimate2513

Ross, A. C., Stock, C. A., Adcroft, A., Curchitser, E., Hallberg, R., Harrison, M. J., ... Simkins, J. (2023). A high-resolution physical-biogeochemical model for marine resource applications in the northwest atlantic (mom6-cobalt-nwa12 v1.0). *Geoscientific Model Development*, *16*. doi: https://doi.org/10.5194/gmd-16-6943-2023

Rykaczewski, R. R., Dunne, J. P., Sydeman, W. J., García-Reyes, M., Black, B. A., & Bograd, S. J. (2015). Poleward displacement of coastal upwelling-favorable winds in the ocean's eastern boundary currents through the 21st century. *Geophysical Research Letters*, *42*(15), 6424-6431. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL064694` doi: https://doi.org/10.1002/2015GL064694

Sato, Y., Goto, D., Michibata, T., Suzuki, K., Takemura, T., Tomita, H., & Nakajima, T. M. (2018). Aerosol effects on cloud water amounts were successfully simulated by a global cloud-system resolving model. *Nature Communications*, *9*. doi: 10.1038/s41467-018-03379-6

Schmidt, G. A., Romanou, A., Roach, L. A., Mankoff, K. D., Li, Q., Rye, C. D., ... Busecke, J. J. (2023). Anomalous meltwater from ice sheets and ice shelves is a historical forcing. *Geophys. Res. Lett.*, *50*(24), e2023GL106530.

Schmidtko, S., Heywood, K. J., Thompson, A. F., & Aoki, S. (2014). Multidecadal warming of Antarctic waters. *Science*, *346*, 1227–1231. doi: 10.1126/science.1256117

Schoonover, J., Dewar, W. K., Wienders, N., & Deremble, B. (2017). Local sensitivities of the Gulf Stream separation. *Journal of Physical Oceanography*, *47*. doi: https://doi.org/10.1175/JPO-D-16-0195.1

Shu, Q., Song, Z., & Qiao, F. (2015). Assessment of sea ice simulations in the CMIP5 models. *The Cryosphere*, *9*(1), 399–409.

Silvano, A., Purkey, S., Gordon, A. L., Castagno, P., Stewart, A. L., & S. R. Rintoul, e. a. (2023). Observing Antarctic Bottom Water in the Southern Ocean. *Frontiers in Marine Science*, *10*. doi: 10.3389/fmars.2023.1221701

Silvestri, S., Wagner, G. L., Campin, J.-M., Constantinou, N. C., Hill, C. N., Souza, A., & Ferrari, R. (2024). A new WENO-based momentum advection scheme for simulations of ocean mesoscale turbulence. *Journal of Climate*, *16*. doi: 10.1029/2023MS004130

SIMIP Community, T. (2020). Arctic sea ice in CMIP6. *Geophys. Res. Lett.*, *47*(10), e2019GL086749.

Small, R. J., Curchitser, E., Hedstrom, K., Kauffman, B., & Large, W. G. (2015). The benguela upwelling system: Quantifying the sensitivity to resolution and coastal wind representation in a global climate model. *Journal of Climate*, *28*(23), 9409 - 9432. Retrieved from `https://journals.ametsoc.org/view/journals/clim/28/23/jcli-d-15-0192.1.xml` doi: https://doi.org/10.1175/JCLI-D-15-0192.1

Smeed, D. A., Josey, S. A., Beaulieu, C., Johns, W. E., & amd E. Frajka-Williams., B. I. M. (2018). The North Atlantic Ocean is in a state of reduced overturning. *Geophysical Research Letters*, *45*, 1527–1533. doi: 10.1002/2017GL076350

Smith, R. L. (1968). Upwelling. *Oceanography and Marine Biology: an Annual Review*, *6*, 11–46.

Solodoch, A., Stewart, A. L., Hogg, A. M., Morrison, A. K., Kiss, A. E., Thompson, A. F., ... Cimoli, L. (2022). How does Antarctic Bottom Water cross the Southern Ocean? *Geophysical Research Letters*, *49*. doi: 10.1029/2021gl097211

Stewart, A., Klocker, A., & Menemenlis, D. (2018). Circum-Antarctic shoreward heat transport derived from an eddy- and tide-resolving simulation. *Geophysical Research Letters*, *45*. doi: 10.1002/2017GL075677

Stewart, A., Klocker, A., & Menemenlis, D. (2019). Acceleration and overturning of the Antarctic Slope Current by winds, eddies, and tides. *Journal of Physical*

Oceanograph, *49*, 2043–2074. doi: 10.1175/JPO-D-18-0221.1
Stommel, H. (1948). The westward intensification of wind-driven ocean currents. *EOS, Transactions American Geophysical Union*, *29*(2), 202–206. doi: 10.1029/TR029i002p00202
Strub, P. T., Combes, V., Shillington, F. A., & Pizarro, O. (2013). Currents and processes along the eastern boundaries. In G. Siedler, S. M. Griffies, J. Gould, & J. Church (Eds.), *Ocean circulation and climate, 2nd edition: A 21st century perspective* (Vol. 103, pp. 339–384). Academic Press. doi: 10.1016/B978-0-12-391851-2.00014-3
Sylla, A., Sanchez Gomez, E., Mignot, J., & López-Parages, J. (2022). Impact of increased resolution on the representation of the Canary upwelling system in climate models. *Geoscientific Model Development*, *15*(22), 8245–8267. Retrieved from https://gmd.copernicus.org/articles/15/8245/2022/ doi: 10.5194/gmd-15-8245-2022
Taboada, F. G., Stock, C. A., Griffies, S. M., Dunne, J., John, J. G., Small, R. J., & Tsujino, H. (2018). Surface winds from atmospheric reanalysis lead to contrasting oceanic forcing and coastal upwelling patterns. *Ocean Modelling*, *133*, 79–111. doi: 10.1016/j.ocemod.2018.11.003
Tagklis, F., Bracco, A., Ito, T., & Castelao, R. M. (2020). Submesoscale modulation of deep water formation in the Labrador Sea. *Scientific Reports*, *10*. doi: 10.1038/s41598-020-74345-w
Tesdal, J.-E., MacGilchrist, G. A., Beadling, R. L., Griffies, S. M., Krasting, J. P., & Durack, P. J. (2023). Revisiting interior water mass responses to surface forcing changes and the subsequent effects on overturning in the Southern Ocean. *Journal of Geophysical Research - Oceans*, *128*. doi: 10.1029/2022JC019105
Thomas, M. D., Tréguier, A. M., Blanke, B., Deshayes, J., & Voldoire, A. A. (2015). A Lagrangian method to isolate the impacts of mixed layer subduction on the meridional overturning circulation in a numerical model. *Journal of Climate*, *28*, 7503–7517. doi: 10.1175/JCLI-D-14-00631.1
Thompson, A. F., Stewart, A. L., Spence, P., & Heywood, K. J. (2018). The Antarctic Slope Current in a changing climate. *Reviews of Geophysics*, *56*, 741–770.
Thompson, K. R., & Demirov, E. (2006). Skewness of sea level variability of the world's oceans. *Journal of Geophysical Research: Oceans*, *111*. doi: 10.1029/2004JC002839
Todd, R. E. (2021). Gulf Stream mean and eddy kinetic energy: Three-dimensional estimates from underwater glider observations. *Geophysical Research Letters*, *48*. doi: https://doi.org/10.1029/2020GL090281
Varela, R., DeCastro, M., Rodriguez-Diaz, L., Dias, J. M., & Gómez-Gesteira, M. (2022). Examining the ability of CMIP6 models to reproduce the upwelling SST imprint in the Eastern Boundary Upwelling Systems. *Journal of Marine Science and Engineering*, *10*(12). Retrieved from https://www.mdpi.com/2077-1312/10/12/1970 doi: 10.3390/jmse10121970
Verdy, A., & Mazloff, M. R. (2017). A data assimilating model for estimating Southern Ocean biogeochemistry. *Journal of Geophysical Research: Oceans*, *122*, 6968–6988. doi: 10.1002/2016JC012650
Waldman, R., & Giordani, H. (2023). Ocean barotropic vorticity balances: theory and application to numerical models. *Journal of Advances in Modeling Earth Systems*, *15*(4), e2022MS003276. doi: 10.1029/2022MS003276
Wang, C., Zhang, L., Lee, S.-K., Wu, L., & Mechoso, C. R. (2014). A global perspective on CMIP5 climate model biases. *Nature Climate Change*, *4*. doi: 10.1038/NCLIMATE2118
Wang, H., Legg, S. A., & Hallberg, R. W. (2015). Representations of the Nordic Seas overflows and their large scale climate impact in coupled models. *Ocean Modelling*, *86*, 76–92. doi: 10.1016/j.ocemod.2014.12.005
Wang, X., Key, J., Kwok, R., & Zhang, J. (2016). Comparison of Arctic sea ice

thickness from satellites, aircraft, and PIOMAS data. *Remote Sensing*, *8*(9), 713.

Winton, M., Adcroft, A., Dunne, J. P., Held, I. M., Shevliakova, E., Zhao, M., ... Zhang, R. (2020). Climate sensitivity of GFDL's CM4.0. *Journal of Advances in Modeling the Earth System*, *12*. doi: 10.1029/2019MS001838

Xu, X., Chassignet, E. P., Firing, Y. L., & Donohue, K. (2020). Antarctic Circumpolar Current transport through Drake Passage: What can we learn from comparing high-resolution model results to observations? *Journal of Geophysical Research: Oceans*, *125*. doi: 10.1029/2020jc016365

Xu, X., Hurlburt, H., Schmitz Jr, W., Zantopp, R., Fischer, J., & Hogan, P. (2013). On the currents and transports connected with the Atlantic meridional overturning circulation in the subpolar North Atlantic. *Journal of Geophysical Research: Oceans*, *118*(1), 502–516. doi: 10.1002/jgrc.20065

Yan, X., Zhang, R., & Knutson, T. R. (2018). Underestimated AMOC variability and implications for AMV and predictability in CMIP models. *Geophysical Research Letters*, *45*. doi: 10.1029/2018GL077378

Yan, X., Zhang, R., & Knutson, T. R. (2019). A multivariate AMV index and associated discrepancies between observed and CMIP5 externally forced amv. *Geophysical Research Letters*, *46*. doi: 10.1029/2019GL082787

Yeager, S. (2015). Topographic coupling of the Atlantic overturning and gyre circulations. *Journal of Physical Oceanography*, *45*(5), 1258–1284. doi: 10.1175/JPO-D-14-0100.1

Yeager, S., Castruccio, F., Chang, P., Danabasoglu, G., Maroon, E., Small, J., ... Zhang, S. (2021). An outsized role for the Labrador Sea in the multidecadal variability of the Atlantic overturning circulation. *Science Advances*, *7*. doi: 10.1126/sciadv.abh3592

Zanna, L., Khatiwala, S., Gregory, J., Ison, J., & Heimbach, P. (2019). Global reconstruction of historical ocean heat storage and transport. *Proceedings of the National Academy of Science*. doi: 10.1073/pnas.1808838115

Zhang, J., & Rothrock, D. (2003). Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates. *Mon. Wea. Rev.*, *131*(5), 845–861.

Zhang, L., Delworth, T. L., Cooke, W., Goosse, H., Bushuk, M., Morioka, Y., & Yang, X. (2021). The dependence of internal multidecadal variability in the Southern Ocean on the ocean background mean state. *Journal of Climate*, *34*(3), 1061–1080. doi: 0.1175/JCLI-D-20-0049.1

Zhang, L., Delworth, T. L., Cooke, W., & Yang, X. (2019). Natural variability of Southern Ocean convection as a driver of observed climate trends. *Nature Climate Change*, *9*(1), 59–65.

Zhang, R., Delworth, T. L., Rosati, A., Anderson, W. G., Dixon, K. W., Lee, H.-C., & Zeng, F. (2011). Sensitivity of the North Atlantic Ocean circulation to an abrupt change in the Nordic Sea overflow in a high resolution global coupled climate model. *Journal of Geophysical Research*, *116*. doi: 10.1029/2011JC007240

Zhang, R., & Thomas, M. (2021). Horizontal circulation across density surfaces contributes substantially to the long-term mean northern Atlantic meridional overturning circulation. *Communications Earth & Environment*, *2*. doi: 10.1038/s43247-021-00182-y

Zhang, R., & Vallis, G. K. (2007). The role of bottom vortex stretching on the path of the North Atlantic Western Boundary Current and on the Northern Recirculation Gyre. *Journal of Physical Oceanography*, *37*, 2053–2080. doi: 10.1175/JPO3102.1

Zhao, M. (2020). Simulations of atmospheric rivers, their variability, and response to global warming using GFDL's new high-resolution general circulation model. *Journal of Climate*, *33*, 10287–10303. doi: 10.1175/JCLI-D-20-0241.1

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., ... Xiang, B. (2018a). The GFDL Global Atmosphere and Land Model AM4.0/LM4.0 - Part II: Model Description, Sensitivity Studies, and Tuning Strategies. *Journal of Advances in Modeling Earth Systems*, *10*. doi: 10.1002/2017MS001209

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., ... Xiang, B. (2018b). The GFDL Global Atmosphere and Land Model AM4.0/LM4.0 - Part I: Simulation Characteristics with Prescribed SSTs. *Journal of Advances in Modeling Earth Systems*, *10*. doi: 10.1002/2017MS001208

Zou, S., Petit, T., Li, F., & Lozier, M. S. (2024, April). Observation-based estimates of water mass transformation and formation in the Labrador Sea. *Journal of Physical Oceanography*, *-1*(aop). Retrieved 2024-04-28, from `https://journals.ametsoc.org/view/journals/phoc/aop/JPO-D-23-0235.1/JPO-D-23-0235.1.xml` (Publisher: American Meteorological Society Section: Journal of Physical Oceanography) doi: 10.1175/JPO-D-23-0235.1

The GFDL-CM4X climate model hierarchy, Part II: case studies

Stephen M. Griffies[1,2], Alistair Adcroft[2], Rebecca L. Beadling[3], Mitchell Bushuk[1], Chiung-Yin Chang[2], Henri F. Drake[4], Raphael Dussin[1], Robert W. Hallberg[1,2], William J. Hurlin[1], Hemant Khatri[5], John P. Krasting[1], Matthew Lobo[2], Graeme A. MacGilchrist[6], Brandon G. Reichl[1], Aakash Sane[2], Olga Sergienko[2], Maike Sonnewald[7], Jacob M. Steinberg[1], Jan-Erik Tesdal[2], Matthew Thomas[8], Katherine E. Turner[2], Marshall L. Ward[1], Michael Winton[1], Niki Zadeh[1], Laure Zanna[9], Rong Zhang[1,2], Wenda Zhang[2], Ming Zhao[1]

[1]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, USA
[2]Princeton University Atmospheric and Oceanic Sciences Program, Princeton, USA
[3]Department of Earth and Environmental Science, Temple University, Philadelphia, USA
[4]Department of Earth System Science, University of California, Irvine, USA
[5]Department of Earth, Ocean and Ecological Sciences, University of Liverpool, UK
[6]School of Earth and Environmental Sciences, University of St. Andrews, UK
[7]Computer Science Department, University of California, Davis, USA
[8]Centre for Environment Fisheries and Aquaculture Science, Suffolk, UK
[9]Courant Institute of Mathematical Sciences, New York University, New York, USA

November 23, 2024

Key Points:

- We present case studies of selected features of the GFDL-CM4X climate model from CMIP6 piControl, historical, and SSP5-8.5 simulations.
- Case studies include sea level, eastern boundary upwelling, sea ice, Southern Ocean circulation, and North Atlantic Ocean circulation.
- Refining ocean grid spacing from 0.25° to 0.125° has systematic improvements across a number of climate relevant features.

Corresponding author: Stephen Griffies, `Stephen.M.Griffies@gmail.com`

Abstract

This paper is Part II of a two-part paper that documents the CM4X (Climate Model version 4X) hierarchy of coupled climate models developed at the Geophysical Fluid Dynamics Laboratory (GFDL). Part I of this paper is presented in Griffies et al. (2024). Here we present a suite of case studies that examine ocean and sea ice features that are targeted for further research, which include sea level, eastern boundary upwelling, Arctic and Southern Ocean sea ice, Southern Ocean circulation, and North Atlantic circulation. The case studies are based on experiments that follow the protocol of version 6 from the Coupled Model Intercomparison Project (CMIP6). The analysis reveals a systematic improvement in the simulation fidelity of CM4X relative to its CM4.0 predecessor, as well as an improvement when refining the ocean/sea ice horizontal grid spacing from the 0.25° of CM4X-p25 to the 0.125° of CM4X-p125. Even so, there remain many outstanding biases, thus pointing to the need for further grid refinements, enhancements to numerical methods, and/or advances in parameterizations, each of which target long-standing model biases and limitations.

Plain Language Summary

We examine simulations from a new climate model hierarchy, referred to as CM4X (Climate Model version 4X). The finer grid component of the hierarchy, CM4X-p125, out shines its coarser sibling, CM4X-p25, for certain processes of interest for climate studies, though in others the results are not dramatically distinct. Each case study reveals the advances made by moving from the predecessor CM4.0 climate model to finer grid spacing in either the atmosphere or ocean. Even so, there remain many unresolved problems that help to guide further research and development goals and strategies.

1 Introduction and content of this paper

This paper is Part II of a two-part paper that documents the CM4X hierarchy of coupled climate models, with Part I presented in Griffies et al. (2024). We developed CM4X to support research into the ocean and sea ice components of the earth climate system, with CM4X comprised of two coupled climate models, CM4X-p25 and CM4X-p125. These two models are identically configured, except for their ocean (and sea ice) horizontal grid spacing and bottom topography. In Part I from Griffies et al. (2024), we documented the remarkable thermal equilibration properties of CM4X-p125, and proposed the *mesoscale dominance hypothesis* to help explain the behavior. In the present paper, we work through a suite of case studies that focus on areas of planned research with CM4X.

As detailed in Section 3.1 of Griffies et al. (2024), we present results from following CMIP6 (Eyring et al., 2016) simulations.

- piControl: Pre-industrial control with radiative forcing fixed at year 1850. This experiment illustrates how the models drift from their initial conditions, taken from the early 21st century, and approach thermal equilibrium under pre-industrial forcing.
- Historical: 01January of year 101 from the piControl is used to initialize a historical simulation that is run from 1850 to 2014. In this historical simulation, we did not account for temporal evolution in vegetation, land use, or CO2 fertilization.
- SSP5-8.5: 01January of year 2015 provides the initial condition for the CMIP6 SSP5-8.5 scenario experiment, which allows us to study how the CM4X models simulate climate change through to 2100.

The case studies exemplify aspects of the science going into the model and the science emerging from the model simulations. The presentation generally follows a "show and tell" approach given that our primary aim in this paper is to document features of the new CM4X hierarchy, with many of these features to be more thoroughly examined in future studies.

We begin in Section 2 with a study of the global thermosteric sea level along with statistical patterns of dynamic sea level. This analysis reveals that the historical thermosteric sea level in the CM4X models is somewhat improved relative to CM4.0, and yet the patterns of sea level skewness in CM4.0 and CM4X remain in poor agreement in comparison to ocean reanalysis. In Section 3 we examine properties of the eastern boundary upwelling zones, which are regions of particular importance for biogeochemistry. Here we find an advance arises from the refined atmospheric model grid used in CM4X relative to CM4.0 (see Section A1 of Griffies et al. (2024)), thus improving the fidelity of coastal wind patterns key to upwelling. Even so, long-standing biases in the low level clouds means that the upwelling zones remain far too warm relative to observations. Section 4 studies the Arctic Ocean and Southern Ocean sea ice properties, revealing again that the CM4X model represents an advance over the CM4.0 model, though with many longstanding biases remaining.

In Section 5 we study properties of the Southern Ocean simulation, with some focus on the region near Antarctica given its importance to ongoing studies of ice shelf melt. A particularly encouraging feature of both CM4X-p25 and CM4X-p125 concerns the absence of the unphysically large open ocean polynyas that plagued CM4.0 (Held et al., 2019) and its earth system model cousin ESM4.1 (Dunne et al., 2020). As a result, CM4X provides a versatile tool for performing perturbation experiments to examine, say, the role of fresh water melt around Antarctica such as in Beadling et al. (2022) and Tesdal et al. (2023), including examining the role of the ocean in the SST pattern effect (Armour et al., 2013; Andrews et al., 2018). We complete the case studies in Section 6 with a focus on the North Atlantic circulation, considering both horizontal and overturning circulation features in the middle and high latitudes. The CM4X simulations show some advances over CM4.0 in the overturning and supolar gyre properties, and yet there remain nontrivial biases in the overturning depth and attendant overflows (model is too shallow), as well as biases in the Gulf Stream structure (simulated jet does not penetrate far enough from the coast). We close the paper in Section 7 with concluding remarks on strategies for future ocean climate model development that are motivated by results from the CM4X model hierarchy.

2 Thermosteric and dynamic sea level

In this section we consider two aspects of sea level: global mean thermosteric sea level and patterns of dynamic sea level.

2.1 Global mean thermosteric sea level

Changes to global mean thermosteric sea level occur with seawater density changes affected by temperature changes. Although seawater density is a highly nonlinear function of temperature, salinity, and pressure, we expect the time series for thermosteric sea level to reflect that for global mean temperature, with Figure 1 supporting this expectation. To generate this figure, we computed thermosteric sea level according to Section H9.5 of Griffies et al. (2016) (CMIP variable *zostoga*), using software described in Krasting et al. (2024) and with full-depth monthly mean fields. Notably, the CM4.0 piControl drift is larger than CM4X-p25, whereas there is negligible drift in CM4X-p125.

Figure 1 suggests that differences between CM4X-p25 and CM4X-p125 are mostly due to the difference in drift seen in the piControl runs. Removing a linear trend com-

ACRONYM	MEANING	CITATION OR SECTION
AM4	GFDL Atmospheric Model 4.0	Zhao et al. (2018b, 2018a)
CM2-O	GFDL climate model hierarchy 2.0	Delworth et al. (2006), Griffies et al. (2015)
C96	AM4 with cubed-sphere ($\approx$ 100 km)	Zhao et al. (2018b, 2018a)
C192	AM4 with cubed-sphere ($\approx$ 50 km) in CM4X	Zhao (2020)
CMIP6	Coupled Model Intercomparison Project 6	Eyring et al. (2016)
CM4.0	GFDL Climate Model 4.0 (0.25° ocn & C96 atm)	Held et al. (2019)
CM4X	GFDL Climate Model hierarchy	this paper
CM4X-p25	CM4X w/ 0.25° ocn and C192 atm	this paper
CM4X-p125	CM4X with 0.125° ocn and C192 atm	this paper
CM4X-p25-C96	CM4X with 0.25° ocn and C96 atm	3
ESM4.1	GFDL Earth System Model 4.1	Dunne et al. (2020)
GFDL	Geophysical Fluid Dynamics Laboratory	–
MOM6	Modular Ocean Model version 6	Adcroft et al. (2019), Griffies et al. (2020)
NWA12	NorthWest Atlantic 1/12° model	6.1 and Ross et al. (2023)
OM4.0	GFDL Ocean/Sea-ice Model 4.0 (0.25°)	Adcroft et al. (2019)
SIS2	Sea Ice Simulator version 2	Delworth et al. (2006), Adcroft et al. (2019)
AABW	Antarctic Bottom Water	5
AAIW	Antarctic Intermediate Water	5
ACC	Antarctic Circumpolar Current	5
AIS	Antarctic Ice Shelf	5
AMOC	Atlantic Meridional Overturning Circulation	6
ASC	Antarctic Slope Current	5
CDW	Circumpolar Deep Water	5
EBUS	eastern boundary upwelling system	3
DSW	Dense Shelf Water	5
NADW	North Atlantic Deep Water	6
MKE	mean kinetic energy	6.1
MLD	mixed layer depth	6
OSNAP	Overturning in the North Atlantic Subpolar Program	6
RAPID	Rapid Climate Change Programme	6
RMSE	root-mean-square error	4
SAMW	Sub-Antarctic Mode Water	5
SIC	Sea Ice Concentration	4
SIE	Sea Ice Extent	4
SIT	Sea Ice Thickness	4
SIV	Sea Ice Volume	4
WMT	watermass transformation	5.2 and 6.4

Table 1. Acronyms used in this paper, their meaning, and relevant citation or section. The upper portion refers to model related acronyms and the lower portion to oceanographic and statistical related acronyms.

puted from the piControl leads to very similar global thermal expansion in the historical and SSP5-8.5 simulations (Figure 2). Evidently, the nonlinear effects noted by Hallberg et al. (2013) are not revealed by these two simulations, presumably since their piControl states have not drifted too far apart after 100 years.

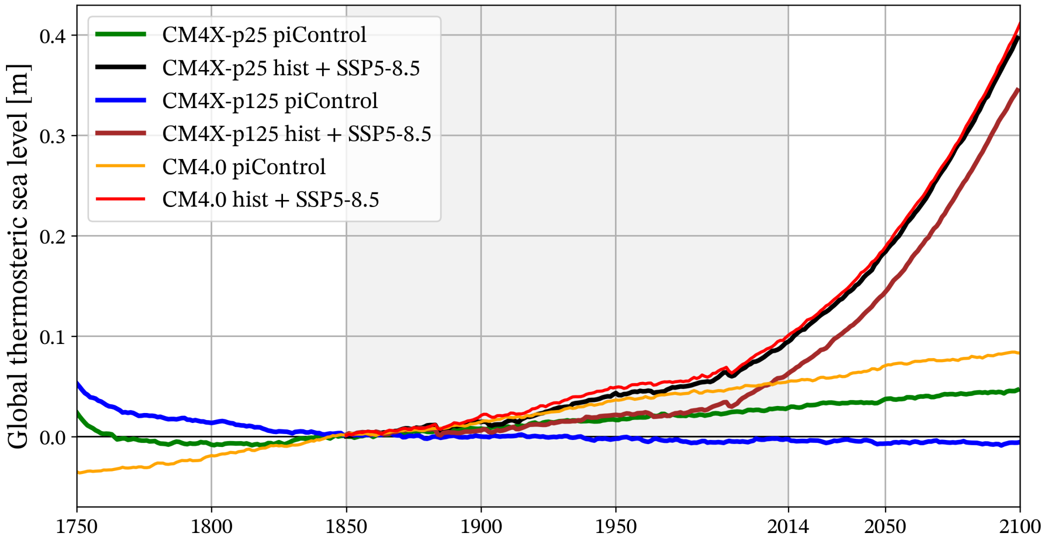

Figure 1. Global thermosteric sea level in piControl, historical (1850-2014), and SSP5-8.5 (2014-2100) simulations using CM4X and CM4.0 climate models. Historical simulations for CM4X-p25 and CM4X-p125 branch from the corresponding piControl at year 101, whereas CM4.0 is branched from its piControl at year 251. This different branching explains why the CM4.0 piControl does not line up with the CM4X piControls. Furthermore, CM4.0's later branching means that its initial cooling phase see in Part I (Griffies et al., 2024) is outside of the time range of this figure, so that the CM4.0 piControl exhibits a nearly linear drift throughout.

During the 20th and early 21st centuries, the observed global-mean sea level exhibited significant increases, with thermosteric rise becoming increasingly significant in recent decades (Frederikse et al., 2020). Figure 2 shows changes relative to the year 2002-2018 time mean, plotted over the historical simulations (from year 1850 through 2014) and eight years of the SSP5-8.5 projection (years 2015 to 2022). We accounted for model drift by removing the long term linear trend in the corresponding piControl run from each historical + SSP5-8.5 time series. We also compare model results to multiple observation based analyses.

In Figure 2 we see that CM4.0 shows a nearly flat thermosteric sea level during 1940–1990, with similar behavior found for many other CMIP6 models discussed by Jevrejeva et al. (2021). In contrast, the CM4X simulations better align with the thermosteric sea level rise found by the observations during this period. Recent decades have seen an upward acceleration of thermosteric sea level rise (Dangendorf et al., 2019), with the CM4.0 and CM4X simulations also showing an acceleration. However, the models appear to overestimate the observational trend since 1990, indicative of the large transient climate sensitivity found in CM4.0 (Winton et al., 2020). We qualify this point by noting the models are mostly within the observational product uncertainty ranges, and further updates of Ishii et al. (2017), including data up to 2022, diverge from the other observations and align more closely with the models.

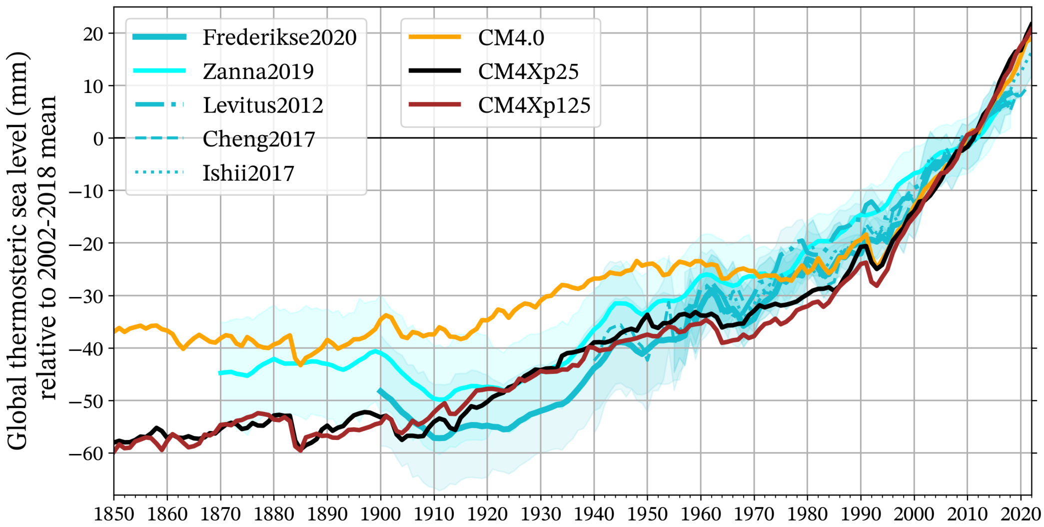

Figure 2. Historical and projected thermosteric sea level derived from CM4.0 and CM4X simulations, along with observational estimates. Models are plotted over the entire historical period (1850-2014), including years 2015-2022 from the SSP5-8.5 projections. Observational estimates are depicted by cyan-shaded lines with dark-solid (Frederikse et al., 2020), light-solid (Zanna et al., 2019), dash-dotted (Levitus et al., 2012), dashed (Cheng et al., 2017) and dotted (Ishii et al., 2017). The global mean thermosteric sea level estimates by Levitus et al. (2012), Cheng et al. (2017) and Ishii et al. (2017) do not include the deep ocean contribution (below 2000 meters). Note that the themosteric sea level time series for CM4.0 and CM4X were detrended by subtracting the long-term linear trend in the piControl from the combined historical and SSP5-8.5 scenario time series. The linear trend was derived from a linear fit over the time period of the piControl run, matching the branch-off year and duration of the historical and SSP5-8.5 scenario simulations (250 years).

2.2 Statistical measures of dynamic sea level fluctuations

We consider statistical properties of the daily mean dynamic sea level ($\texttt{zos}$ as in Griffies et al. (2016)), thus allowing for a quantitative characterization of spatial structure of sea level fluctuations. In particular, we focus on the standard deviation and skewness computed over the 20-year segment 1995-2014 of the historical simulation. In fact, the standard deviation was already presented in Figure 3 of Part I (Griffies et al., 2024) as part of our discussion of mesoscale eddy activity. Here we present the skewness.

To compute the statistics, we generate a 20-year climatology of the daily mean dynamic sea level over the historical period 1995-2014, denoted by $\overline{\texttt{zos}}$. We then compute anomalies relative to the climatology

$$\texttt{zos}'(t_n) = \texttt{zos}(t_n) - \overline{\texttt{zos}}(t_{\mathrm{mod}(n,365)}), \tag{1}$$

where t_n is the day within the $N = 20 * 365$ total number of days, and $t_{\mathrm{mod}(n,365)}$ is the climatological day. The standard deviation and skewness, computed at each horizontal ocean grid cell, are given by

$$s = \left[\frac{1}{N}\sum_{n=1}^{N}[\texttt{zos}'(t_n)]^2\right]^{1/2} \quad \text{and} \quad \mathcal{S} = \frac{\sum_{n=1}^{N}[\texttt{zos}'(t_n)]^3}{N\, s^3}. \tag{2}$$

The sea level standard deviation has dimensions of length whereas skewness is dimensionless. We compare results from CM4X-p25 and CM4X-p125, and include the 1/12° GLORYS12 analysis from Lellouche et al. (2021) as a benchmark.

2.3 Skewness

Skewness is a third-order statistic that quantifies the asymmetry of a distribution. A positive skewness means that fluctuations are biased positive relative to a Gaussian distribution, and vice versa for negative skewness. K. R. Thompson and Demirov (2006) and Hughes et al. (2010) noted that sea level skewness is positive on the poleward side of strong eastward currents (e.g., Gulf Stream, Kuroshio) and negative on the equatorward side, so that strong currents are generally aligned with the zero contour.

As seen in Figure 3, CM4X-p25 contains no clear zero skewness signature of the Gulf Stream jet, contrary to that found in GLORYS12. CM4X-p125 shows some hint of a zero skewness contour, but far less coherent than in GLORYS12. The Kuroshio Current in CM4X-p25 is also poorly revealed by the CM4X-p25 skewness, whereas the skewness in CM4X-p125 resembles GLORYS12 though with a muted signature. The Agulhas region in the CM4X simulations suffers from the complement bias found in the western boundary currents. Namely, Agulhas eddies in the CM4X simulations remain somewhat more coherent than found in GLORY12, thus producing a nontrivial positive skewness signature reaching into the central portion of the South Atlantic, with this signature in the models far larger than found in GLORYS12.

As noted by K. R. Thompson and Demirov (2006), tropical sea level skewness is dominated by large patterns associated with variability such as the El Niño-Southern Oscillation and Indian Ocean variability. The CM4X models generally show a muted tropical variability relative to GLORYS12, which is consistent with a muted tropical sea surface temperature variability as revealed by the power spectra in Figure 10 of Part I (Griffies et al., 2024). Correspondingly, the positive skewness in GLORYS12 extending along the coasts of North and South America is missing in both CM4X-p25 and CM4X-p125.

2.4 Conclusions regarding sea level

Figure 2 shows that both CM4X simulations reduce thermosteric sea level biases found in CM4.0, and yet that result is possibly due to simplifications of the land model as detailed in Section 2.4 and Appendix A of Part I (Griffies et al., 2024). The patterns for dynamic sea level standard deviation (Figure 3 in Part I) and skewness expose nontrivial biases in the middle latitude boundary currents. These biases are reduced with CM4X-p125 relative to CM4X-p25, and yet they suggest a need for either improved parameterizations or, as emphasized by Chassignet and Xu (2017), substantially finer grid spacing. We find a highly muted CM4X tropical variability as revealed by the sea level skewness, thus revealing how sea level skewness complements the sea surface temperature power spectra shown in Figure 10 of Part I (Griffies et al., 2024), which also shows muted tropical sea surface temperature variability.

3 Eastern boundary upwelling systems

Eastern boundary upwelling systems (EBUS) are among the most biologically productive areas of the World Ocean (Strub et al., 2013). They are characterized by a sharp drop in the sea surface temperature near the coast, which results from upwelling of cooler interior waters through Ekman suction and lateral Ekman transport driven by equatorward wind stresses. In Figure 4 we present summer SST taken from CM4X-p125. By showing the summertime season we clearly expose the cool upwelling waters in contrast to the warm surrounding waters. As summarized by Richter (2015), many climate models exhibit large biases in the SST in eastern boundary regions due to both a lack of cloud

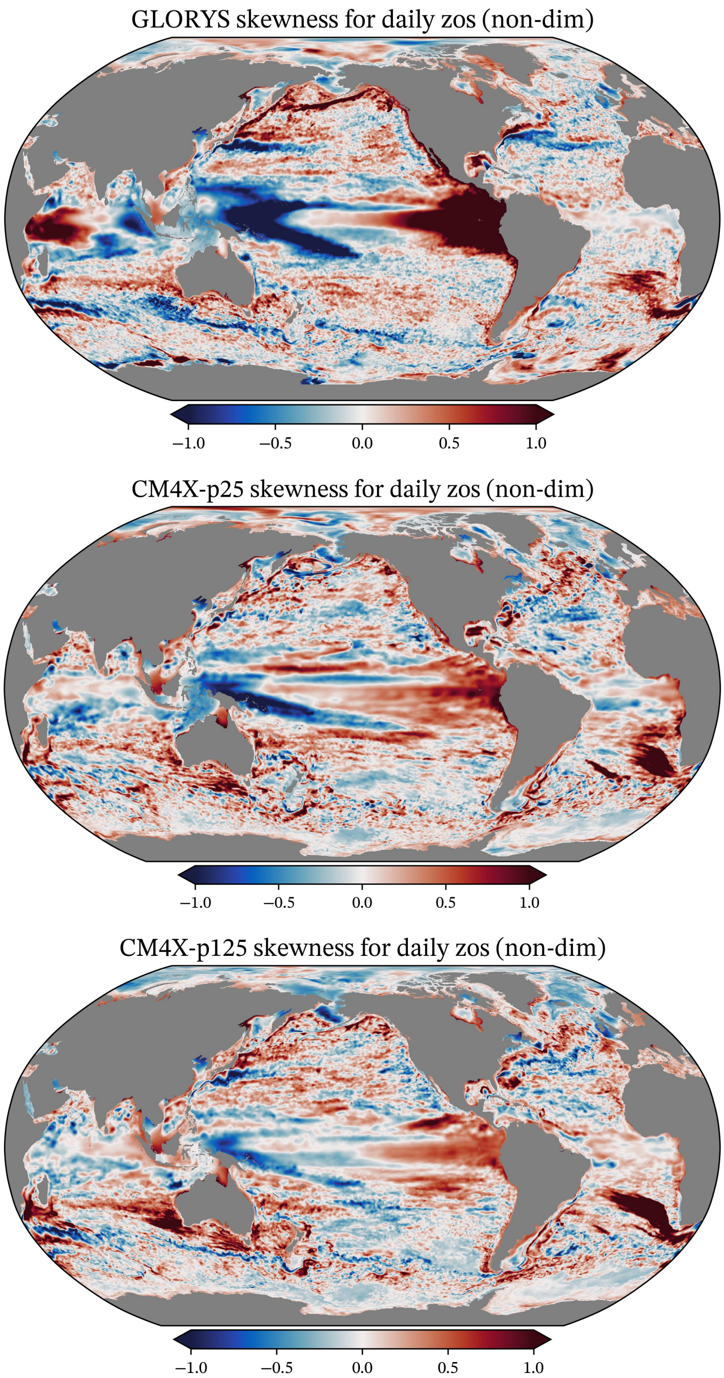

Figure 3. Skewness (non-dimensional) for the daily mean dynamic sea level from GLORYS12 (Lellouche et al., 2021) (top panel), and CM4X-p25 (middle panel) and CM4X-p125 (lower panel). Each figure is created from the deviation of the daily mean sea level relative to the climatological mean for that day of the year, as detailed in Section 2.2. Each climatology is created from years 1995-2014.

cover (in particular stratocumulus decks) and weaker than observed upwelling favorable winds (see also C. Wang et al. (2014)). The CM4X models are similarly lacking the appropriate cloud cover in EBUS. Compared to cloud cover estimates from Kaspar et al. (2009), both CM4X-p25 and CM4X-p125 have at best 20% less cloud cover over EBUS and as much as 40% less in regions in the Pacific systems, thus resulting in the warm SST biases presented in Figure 13 of Part I (Griffies et al., 2024). Gordon et al. (2000) shows that a more realistic marine stratocumulus significantly improves the annual cycle of SST and ocean dynamics in the tropics. Representation of upwelling favorable winds is the second source of SST biases in EBUS. In this section, we focus on improvements to SST biases from the 50 km atmosphere in CM4X models.

3.1 Modeling eastern boundary upwelling systems

Because of their importance to the marine-based economy, eastern boundary upwelling systems have been extensively studied with ocean and climate models. The impact of wind stress on upwelling characteristics is typically addressed in a regional context, such as the studies by Albert et al. (2010), Jacox and Edwards (2012), Junker et al. (2015), Small et al. (2015), and Sylla et al. (2022). Furthermore, sensitivity of upwelling regions to climate change is a topic of great interest, such as studied by Bakun et al. (2015), Rykaczewski et al. (2015), and Bograd et al. (2023). We here consider the representation of upwelling in the CM4X climate models during their historical simulations. We focus on the four upwelling systems shown in Figure 4: California and Peru in the Pacific and Canary and Benguela in the Atlantic.

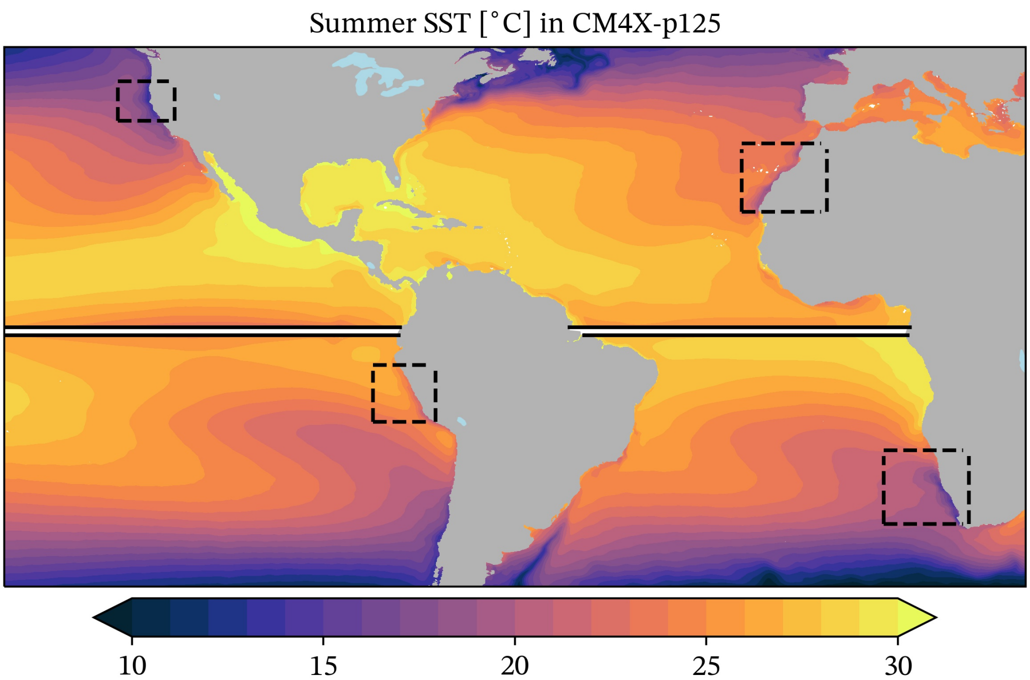

Figure 4. Sea surface temperature from CM4X-p125, during summer months (July-September in northern hemisphere, January-March in southern hemisphere) averaged over years 1980-2014. The boxes indicate the four eastern boundary upwelling systems considered in our analysis: California, Peru, Canary and Benguela. We place a white stripe at the equator since we map the summer months for both hemisphere, and so there is a jump at the equator.

Because of the fine spatial scales involved in coastal upwelling, both atmosphere and ocean models with relatively coarse grids are limited in their ability to capture the relevant dynamical processes. Fine resolution is needed in the atmosphere to represent the coastal wind stresses and their curl, and fine resolution is needed in the ocean to realize upwelling localized near the coast. Varela et al. (2022) found that refined grid spacing allows many CMIP6 climate models to improve their simulation of coastal SST relative to earlier model classes. For example, the CM4.0 climate model captures the imprint of upwelling using its C96 atmosphere (approximately 100 km) and 0.25° ocean. Further refining the atmospheric grid to C192 (approximately 50 km) in CM4X leads us to the question of its impact on the upwelling systems, as does refinement of the ocean grid from CM4X-p25 to CM4X-p125. To help address these questions we include a companion experiment, CM4X-p25-C96, which uses the C96 atmosphere along with the same ocean model configuration as CM4X-p25.

3.2 Winds and SSTs in the upwelling regions

In Figure 5 we show the alongshore wind stress and wind stress curl as a function of the distance to the coast and as averaged over the four upwelling regions. The alongshore/equatorward wind stress remains stronger nearshore with the C192 atmosphere, with a drop confined to the inner 100 km from the coast instead of 150 km found in CM4X-p25-C96. This result holds for both CM4X-p25 and CM4X-p125, which is expected since the ocean grid spacing is not a leading order effect on the wind stress near the coast. The sharper wind drop near the coast found with the C192 atmosphere results in a greater wind stress curl in the nearshore region (within 50 km from the coast) and a decrease in the 50-100 km band. The CM4X experiments are compared with the SCOW estimates (Risien & Chelton, 2008) based on QuikSCAT satellite measurements. Except for the California system, modeled alongshore wind stresses are stronger than observed in the 100-300 km band. In all regions, the modeled wind drop-off at the coast is steeper than in the SCOW estimates, resulting in a stronger wind stress curl at the coast.

The distinct alongshore wind and wind stress curl lead to differences in the resulting SST profiles as averaged over the four eastern boundary regions (Figure 6). The CM4X results are typically biased warm offshore in the Pacific, whereas they are in better agreement in the Atlantic. The C192 atmosphere leads to an SST that drops faster in the 50-150 km band in all eastern boundary regions. This result has a favorable impact in the California system where it compensates for a warm offshore bias, and yet the SST gradient is stronger than observed. In the Peru system, the SST gradient is in good agreement with observations and leads to a much improved simulation using the C192 atmosphere. In the Benguela system, the SST drop at the coast is not present using the C96 atmosphere, whereas the CM4X experiments using the C192 atmosphere are superior. Finally, the stronger SST gradient in the CM4X simulations with the C192 atmosphere overshoots the observed values in the Canary system, whereas this region is better captured using the C96 atmosphere.

3.3 Conclusions regarding eastern boundary upwelling

Ekman mechanics accounts for two key processes important for the upwelling regions: the cross-shore Ekman transport is proportional to alongshore wind stress (Smith, 1968) and Ekman suction is proportional to the wind stress curl (Enriquez & Friehe, 1995; Pickett & Paduan, 2003). As noted in Jacox and Edwards (2012), cross-shore Ekman transport dominates in a narrow coastal band (within 50 km of the coast), whereas Ekman suction creates small but important upwelling velocities in a broader area extending from outside the coastal band to 200-300 km offshore (Jacox & Edwards, 2012).

Results from the CM4X experiments suggest that reduction in wind stress curl in the 50-200 km band, and its intensification in the narrow nearshore area, favor a stronger cross-shore SST gradient. In the CM4X models using the C192 atmosphere, Ekman suction is concentrated in the nearshore area at the deficit of the broader offshore region.

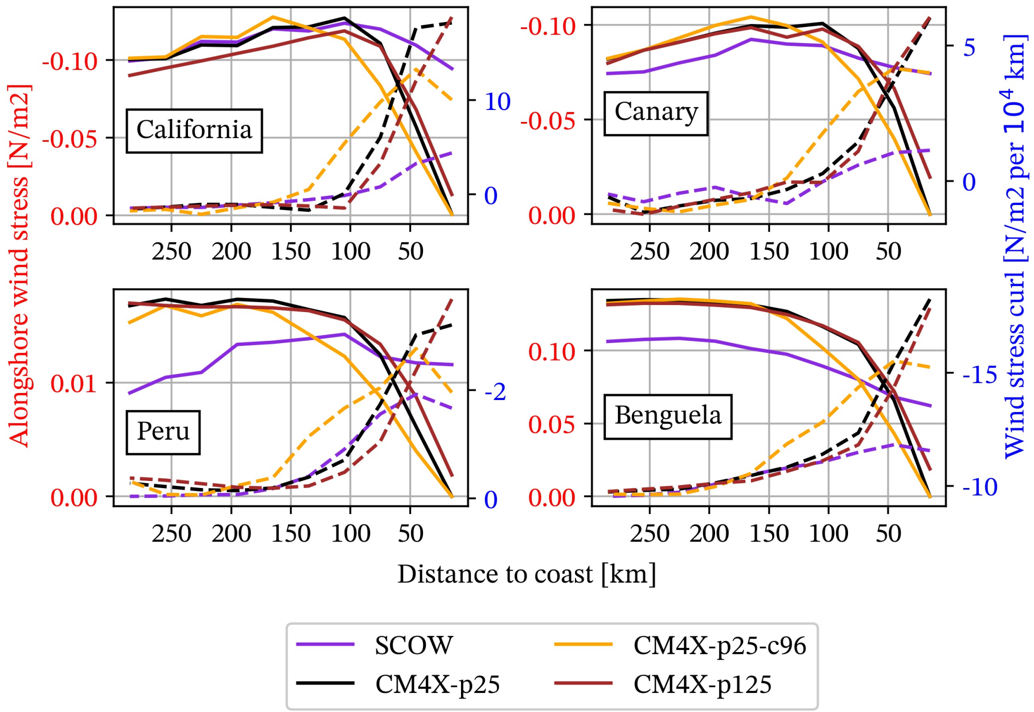

Figure 5. Alongshore wind stress (solid lines/red ticks) and wind stress curl (dashed lines/blue ticks) in the four major eastern boundary upwelling systems during the summer months of 1980-2014. For coastal upwelling at these latitudes, the β term (see equation (2) of Taboada et al. (2018)) is neglected. Gridded data is averaged over the boxes of Figure 4. Note the distinct vertical axes: signs can be reversed so that wind stress decreases at the coast and wind stress curl increases. CM4X-p25-c96 uses the C96 atmosphere model, whereas the other models us the C192 atmosphere. Satellite measurements from SCOW (Risien and Chelton (2008)) are added for reference although they cover a shorter time period (1999-2009). Note the distinct vertical axes on the panels.

The cross-shore Ekman transport with the C192 atmosphere remains strong closer to the coast than with the C96 atmosphere. This result suggests that strengthening of the Ekman transport part of the upwelling in the 50-150 km band, in conjunction with the concentration of the wind stress curl at the coast, favors the upwelling. These results are consistent with previous works from Gent et al. (2010) and Small et al. (2015) in which intensification of the coastal winds leads to stronger upwelling at the coast and overall reduction in SST bias.

The simulated cross-shore Ekman transport at the coast is weaker than the SCOW estimates but stronger offshore in all but the California system, with an intersect ranging from 50 to 100 km depending on the region and experiment considered. Since cross-shore Ekman transport is expected to dominate in the inner 50 km to the coast, this dominance should lead to decreased coastal upwelling, which contrasts to the CM4X SST gradients of Figure 6. In addition, the Ekman suction part of the upwelling is remarkably consistent with satellite estimates in the offshore part (distance $>$ 100 km) with the C192 atmosphere. The C96 atmosphere departs largely from these estimates in the 100-150 km band, which should lead to more Ekman suction upwelling velocities. Even so, we do not find this upwelling signal appears in the SST gradients. We suspect that the biases are

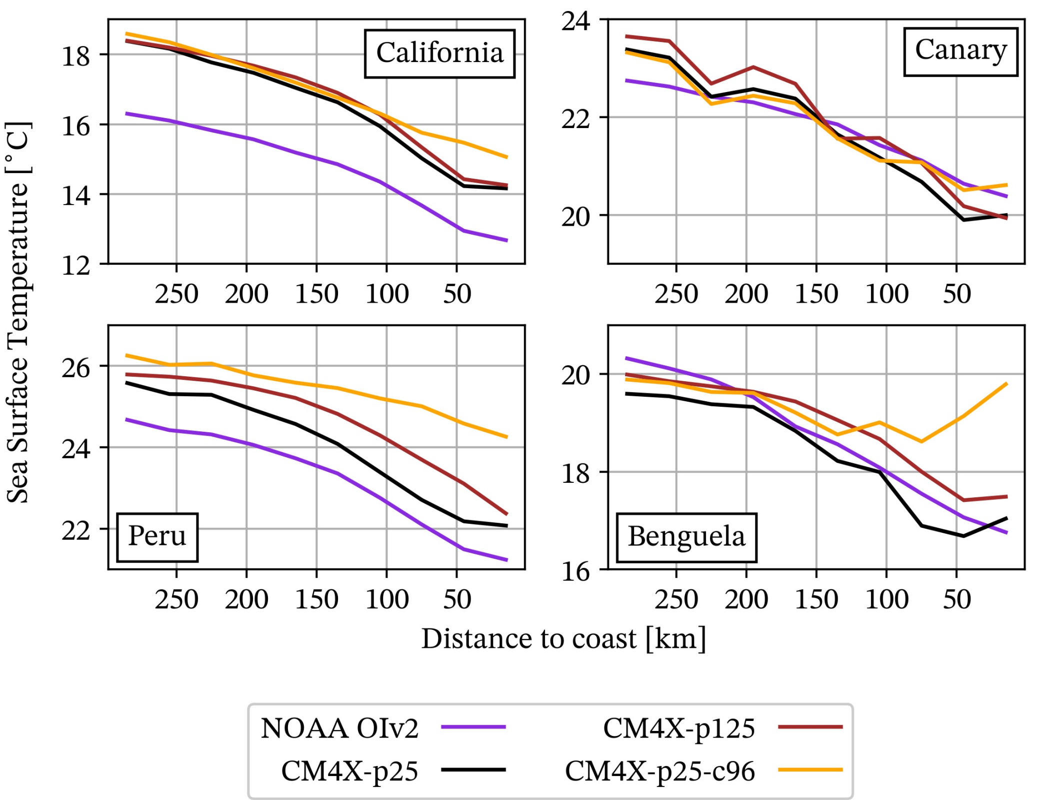

Figure 6. Sea surface temperature in the four major eastern boundary upwelling systems during the summer months of 1980-2014. Results are averaged over the boxes of Figure 4. We include the satellite estimate from NOAA OIv2 (Huang et al., 2020) over the period 1982-2014. Note the different temperature scales on the vertical axes.

modified only slightly due to the more dominant issues with atmospheric model's biases in representing low level clouds in the eastern boundary upwelling regions, with these biases a well known feature of atmospheric models (e.g., Richter (2015); Ceppi et al. (2024)) that remain a topic of ongoing research.

4 Sea ice in the Arctic Ocean and Southern Ocean

In this section, we analyze the simulated Arctic Ocean and Southern Ocean ("Antarctic") sea ice within the CM4X model hierarchy, focusing on the historical and SSP5-8.5 experiments. We also include results from CM4.0 as a point of comparison. To maximize use of the most recent observations, which include record low sea ice conditions, we compute climatologies and trends over 1979-2023 (as opposed to 1979-2014, with 2014 the end of the historical experiment). Doing so requires appending years 2015-2023 from the SSP5-8.5 scenario simulations to the historical simulations. We consider this additional model forcing uncertainty worth the benefits of an additional nine years of observational data.

4.1 Arctic Ocean sea ice

Figure 7a plots Pan-Arctic sea ice extent (SIE) climatologies computed over 1979–2023 across the model hierarchy. The seasonal cycle of Pan-Arctic SIE is well simulated

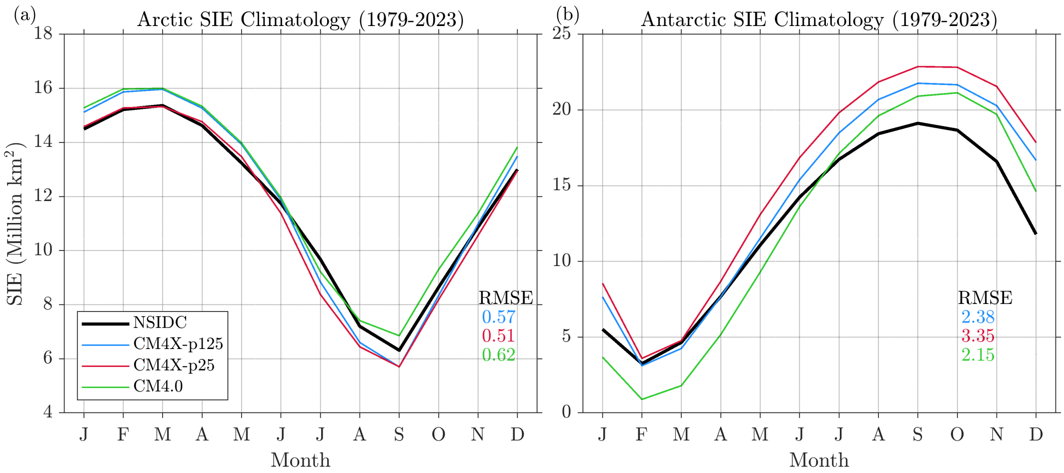

Figure 7. Pan-Arctic and Pan-Antarctic sea ice extent (SIE) climatologies (10^6 km^2) computed over the years 1979–2023 from satellite observations (black), CM4X-p125 (blue), CM4X-p25 (red), and CM4.0 (green). Pan-Arctic and Pan-Antarctic SIE are defined as the area integral of all grid cells covered by at least 15% sea ice concentration (SIC) in the northern and southern hemispheres, respectively. Root-mean-square error (RMSE) are computed between the simulated and observed seasonal cycles and shown in colored text. The observed SIE is computed using passive microwave satellite sea ice concentration observations from the NOAA/National Snow and Ice Data Center (NSIDC) Climate Data Record (CDR) of SIC, version 4 (Data Set ID: G02202; Meier et al. (2021)).

by both CM4X models and both models show improvements relative to the CM4.0 simulation. The RMS errors of the Pan-Arctic SIE climatology are 0.57×10^6 km^2 and 0.51×10^6 km^2 in the CM4X-p125 and CM4X-p25 models, respectively, which are lower than the median CMIP5 model RMSE of 1.45×10^6 km^2 (Shu et al., 2015) and the CM4.0 RMSE of 0.62×10^6 km^2. CM4X-p25 has Pan-Arctic SIE improvements in non-summer months relative to CM4.0, and has a negative summer bias which is similar in magnitude to CM4.0's positive summertime bias. CM4x-p125 is generally similar to CM4.0 in winter and spring and similar to CM4X-p25 in summer and autumn months.

The climatological sea ice concentration (SIC) biases of the models are shown in Figure 8. The spatial pattern of winter Arctic SIC bias is similar across the models (top row in Figure 8), implying that the improved winter Pan-Arctic SIE simulation of CM4X-p25 primarily results from cancellation of positive and negative SIC errors. In winter months, the models have positive biases (too much ice coverage) in the Greenland-Iceland-Norwegian (GIN), Barents, and Bering Seas and negative biases in the Labrador Sea and the Sea of Okhotsk. These winter SIC biases closely mirror the patterns of SST bias (compare with Figure 14 from Part I of Griffies et al. (2024)). The notable positive SIC bias in the GIN and Barents Seas has been persistent across earlier generations of GFDL models, such as CM2.1, ESM2M, ESM2G, and CM3 (Delworth et al., 2006; Dunne et al., 2012; Griffies et al., 2011), possibly related to a combination of too much poleward ocean heat transport and too much ice export through Fram Strait (see Figure 9). The finer grid spacing in CM4X-p125, with its enhanced mesoscale eddy activity, does not ameliorate this bias.

The CM4X models have similar patterns and magnitudes of summer SIC bias (second row of Figure 8), which differ from the SIC biases of CM4.0. The main improvement in CM4X is the reduced positive bias in the Beaufort, Chukchi, and East Siberian Seas.

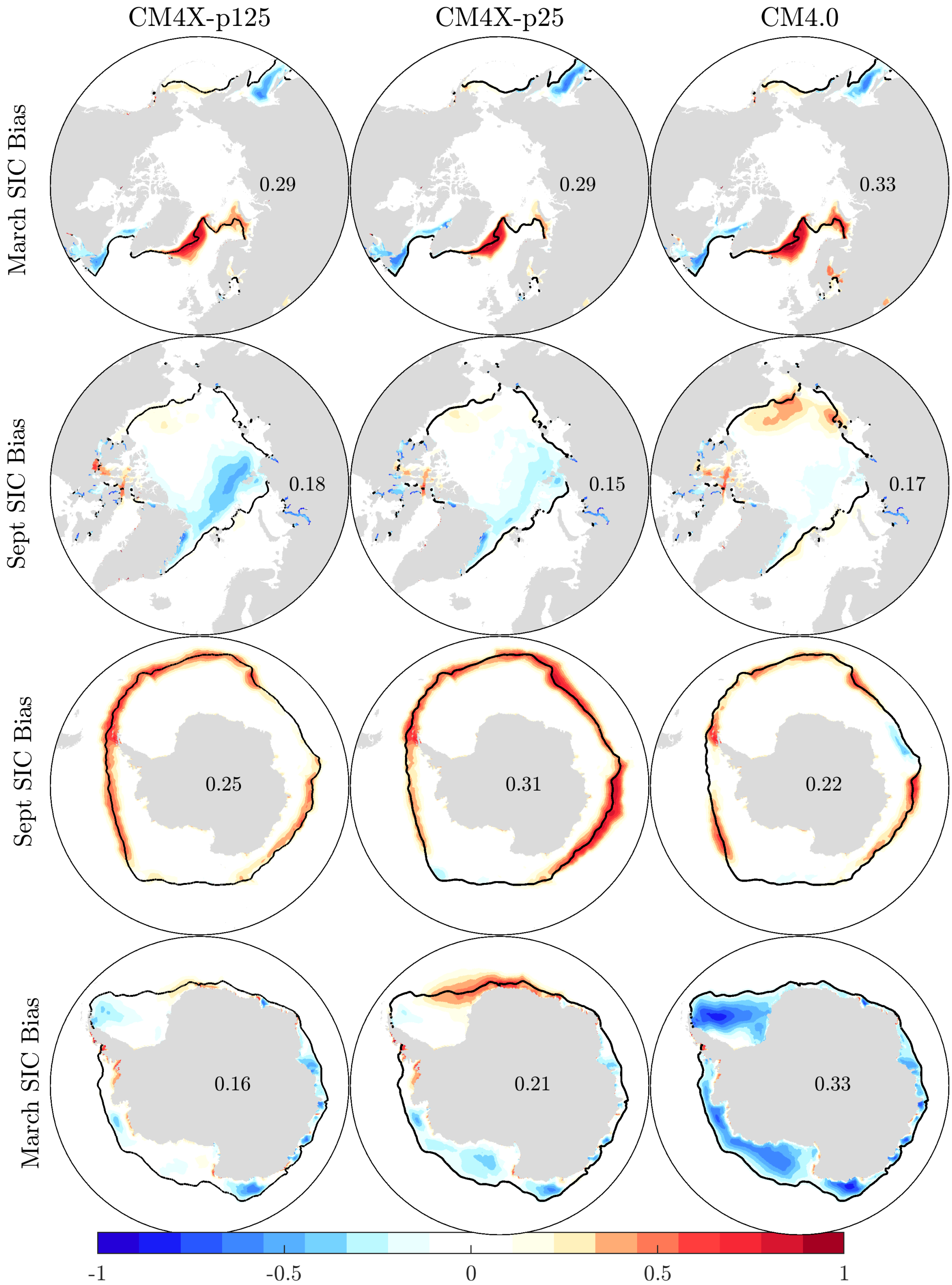

Figure 8. Arctic and Southern Ocean sea ice concentration (SIC) climatological biases (model minus NSIDC CDR observations) in March and September computed over 1979–2023. The black contours indicate the observed climatological sea ice edge position (15% SIC contour). Black text indicates the SIC RMSE area-averaged over the zone of SIC variability, defined as all grid points where the model or observed monthly SIC standard deviation exceeds 5%.

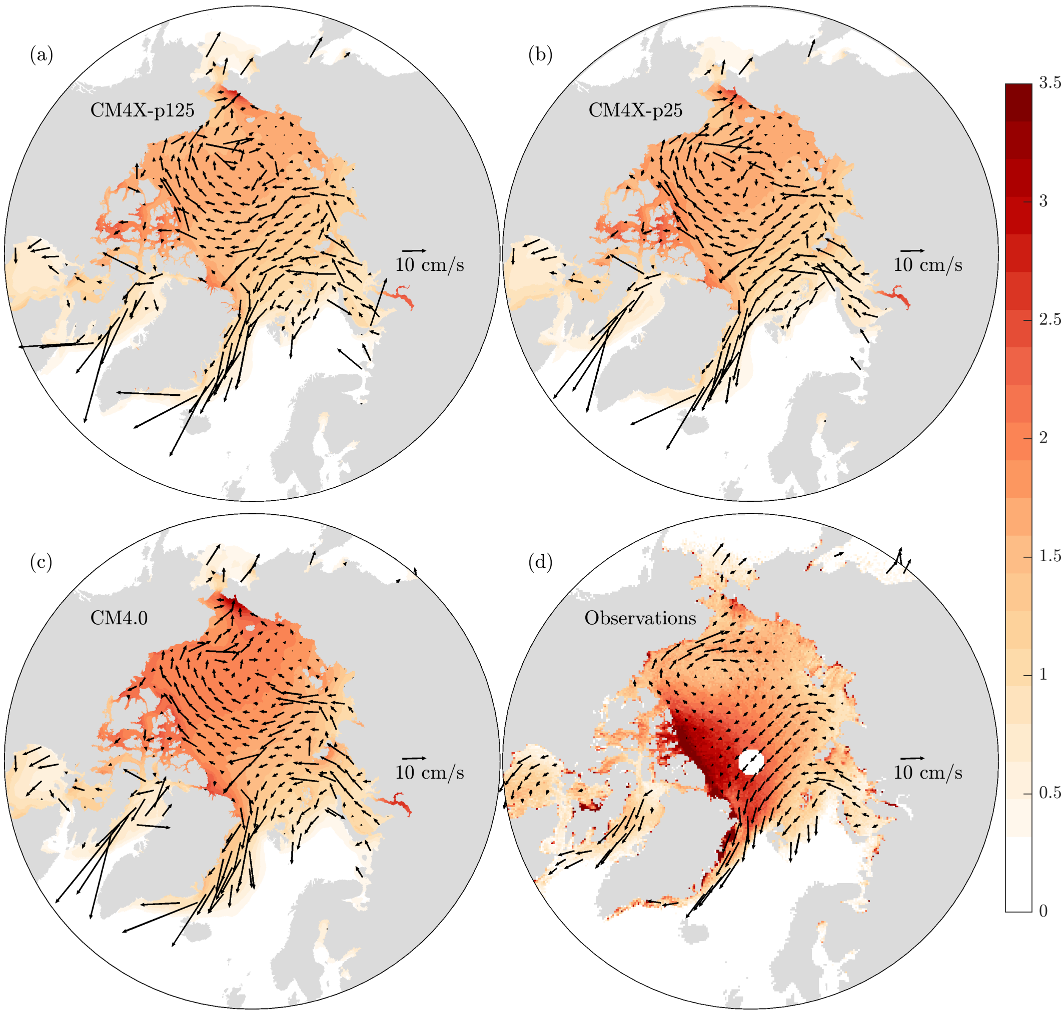

Figure 9. March Arctic sea ice thickness (SIT) climatologies (m; shading) and climatological DJFM sea ice drift (cm/s; vectors) in CM4X, CM4.0, and observations computed over the period 2011–2023. The SIT observations come from the Alfred-Wegener-Institute monthly SIT product (Ricker et al., 2014) and span 2011–2023. The sea ice drift observations are from the low-resolution daily sea ice drift product of the EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSISAF) and span 2010–2023 (Lavergne et al., 2010).

This summer SIC bias in CM4.0 is associated with an erroneous pattern of winter sea ice thickness (SIT), which has the model's thickest ice located in the Beaufort, Chukchi and East Siberian Seas rather than north of Greenland and Ellesmere Island as found in the observations (see Figures 9c,d). This anomalously thick winter ice in the Beaufort, Chukchi, and East Siberian Seas leads to delayed melt in these regions, resulting in a spatially coincident positive bias in summer SIC. We find that the CM4X models share a similar bias in SIT spatial pattern, however, their mean SIT is reduced, resulting in lower summer SIC throughout most of the Arctic (Figures 9a,b). This thinner mean state also results in an exacerbated negative SIC bias in the Greenland Sea and along the northern boundaries of the Barents, Kara, and Laptev Seas, which is a degradation relative to CM4.0. One region that is unchanged is the Canadian Arctic Archipelago which has a consistent positive summer SIC bias across CM4.0, CM4X-p25, and CM4X-p125.

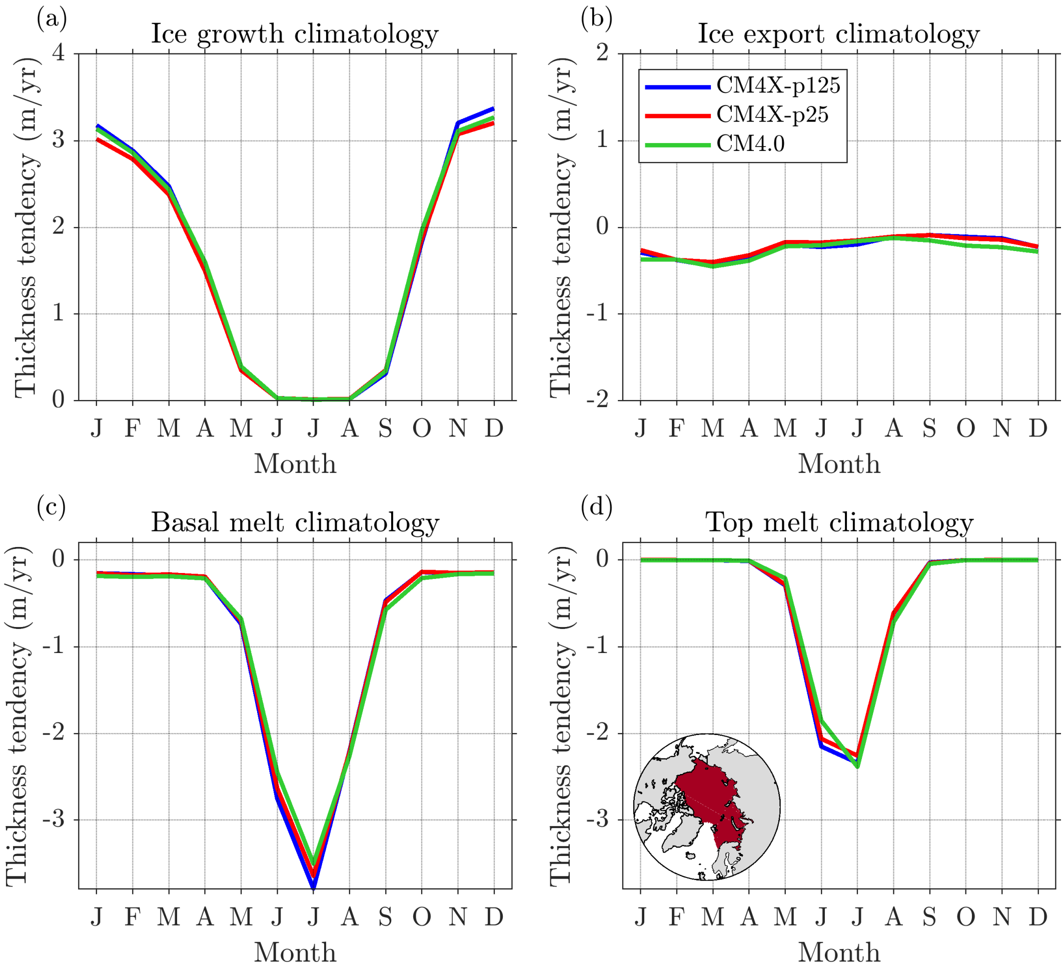

Figure 10. Arctic sea ice mass budget climatologies computed over the Central Arctic basin domain shown in panel (d) over the time period 1979–2023. The mass budget consists of sea ice thickness (SIT) tendency terms (m/yr) corresponding to congelation and frazil ice growth (a), mass transport convergence (b), basal melt (c), and top melt (d). Positive values correspond to mass gain and negative values correspond to mass loss. The y-axis range is the same across the panels, thus allowing for direct comparison of the various terms.

Figure 10 shows Arctic sea ice mass budget climatologies computed following the methodology of Keen et al. (2021), which averages mass budget terms over a Central Arctic domain (see inset of Figure 10d). This domain encompasses the region of thickest Arctic ice and its boundaries include all flux gates to the North Atlantic and North Pacific sectors. The sea ice mass budget consists of a dynamic tendency term associated with ice mass transport convergence (export) and thermodynamic tendency terms associated with congelation and frazil ice growth, basal melt, and top melt. All terms are defined such that positive values indicate mass gain and negative values indicate mass loss and are expressed as a thickness tendency in m/yr (note that the SIS2 model uses a constant sea ice density of 905 kg/m^3). The thinner Arctic ice in CM4X primarily results from increased summer basal and top melt relative to CM4.0. The CM4X models have less mass loss due to ice export in autumn. This reduced autumn ice export in CM4X is likely associated with the thinner and less extensive mean state in these models, as their ice drift patterns are similar to CM4.0 (see Fig. 9). CM4X-p125 has a similar SIT mean state to CM4X-p25, but has more winter growth, potentially associated with enhanced ice formation in sea ice leads in this higher resolution model.

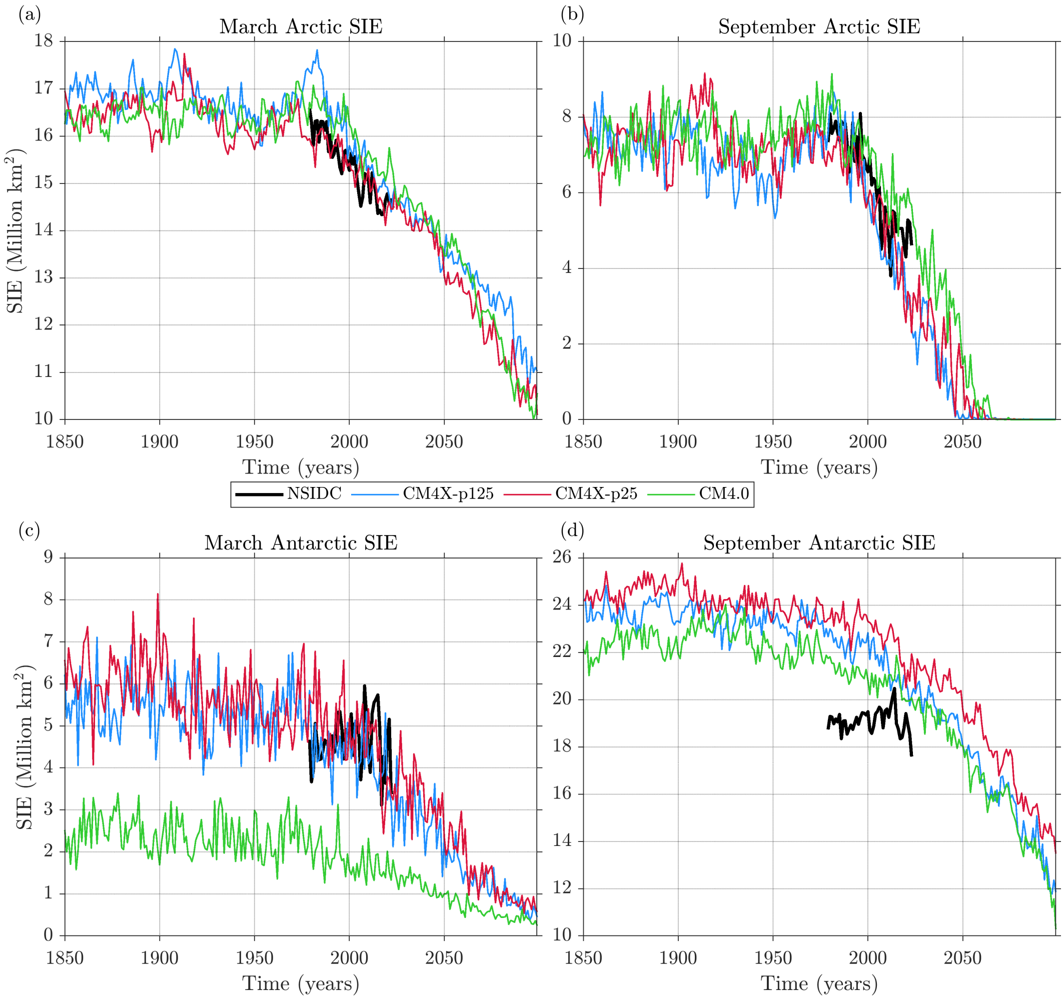

Figure 11. Time series of Arctic and Southern Ocean sea ice extent (SIE) in March and September in CM4X-p125 (blue), CM4X-p25 (red), CM4.0 (green), and NSIDC observations (black) in March and September. The simulations use CMIP6 Historical (1850–2014) and SSP5-8.5 forcings (2015–2099).

We next consider the time evolution of Arctic SIE and sea ice volume (SIV) in Figures 11 and 12. Each model simulates SIE trends in March and September in reasonable quantitative agreement with the observed SIE decline. Given the large degree of internal variability in Arctic SIE trends, we do not expect a perfect match between observations and a single model realization (Jahn et al., 2016; DeRepentigny et al., 2020). The trend differences between the CM4X and CM4.0 historical simulations and observations are smaller than the typical ranges estimated by single-model initial conditional large ensembles performed with other GCMs (Horvat, 2021), suggesting that the CM4X and CM4.0 models are not inconsistent with observed trends. There are some trend differences between the CM4X and CM4.0 models, but multi-member ensembles are necessary to determine if differences are statistically robust. CM4X and CM4.0 each simulate substantial decadal-to-multidecadal variability over the 20th century. This low-freqeuncy

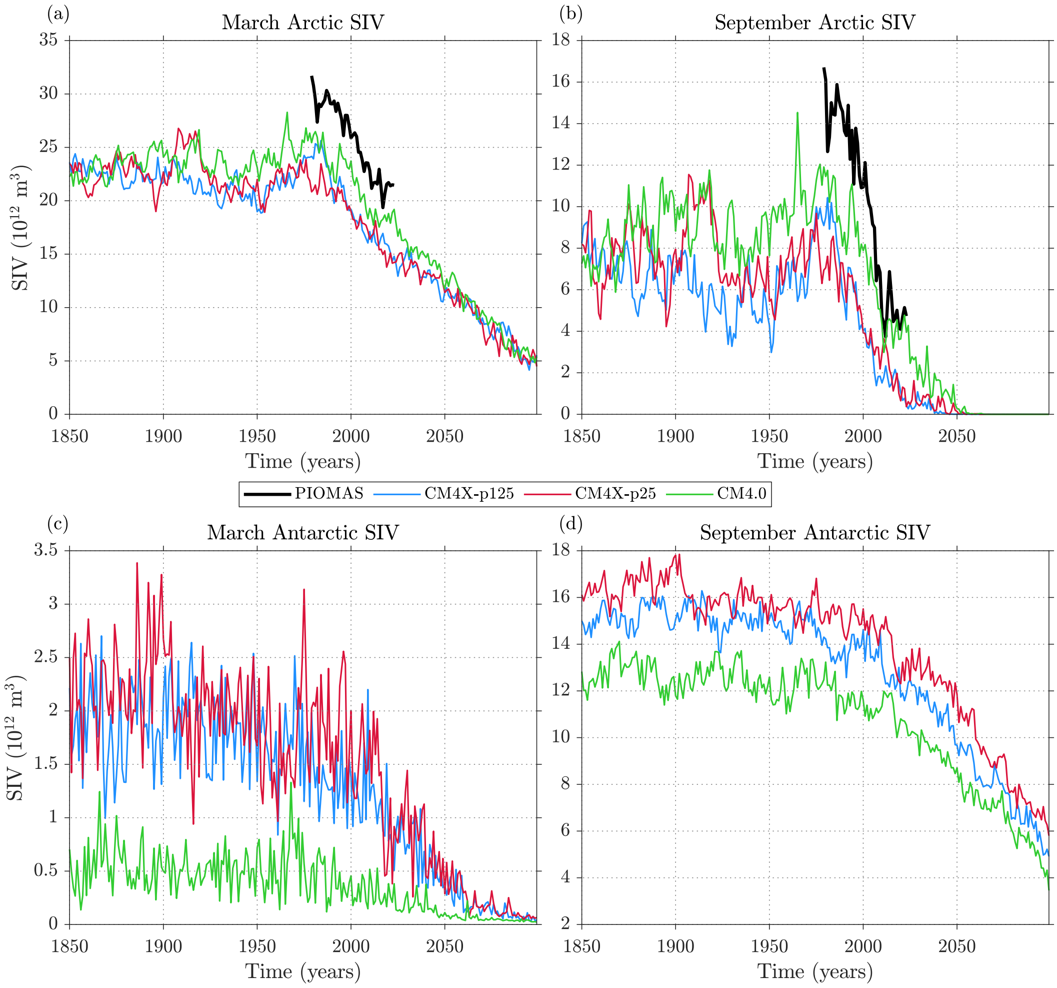

Figure 12. Time series of Arctic and Antarctic sea ice volume (SIV) in March and September in CM4X-p125 (blue), CM4X-p25 (red), CM4.0 (green), and PIOMAS sea ice thickness (SIT) reanalysis (black) in March and September. The simulations use CMIP6 Historical (1850–2014) and SSP5-8.5 forcings (2015–2099). The PIOMAS data spans 1979–2023 and is based on an assimilation system that incorporates SIC, SST, and atmopsheric reanalysis constraints (J. Zhang & Rothrock, 2003).

variability becomes notably muted under the high-forcing SSP5-8.5 scenario used over the 21st century.

Figure 12 shows that the models are biased thin relative to the Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS; J. Zhang and Rothrock (2003)) SIT reanalysis, a product that has reasonably good agreement with available in situ, aircraft, and satellite observations of SIT (X. Wang et al., 2016; Landy et al., 2022). The CM4X models have a larger thin bias than CM4.0, but similar SIV timeseries to each other. This result suggest that the thinner ice in CM4X-p125 relative to CM4.0 is not the result of refined ocean grid spacing. The models simulate a strong decline of Arctic SIV in all months of the year (Figure 12). Despite their mean state SIV biases, the models simulate similar rates of historical SIV loss to PIOMAS.

Under the SSP5-8.5 forcing scenario, the models simulate a complete loss of summer Arctic SIE and SIV over the 21st century. The first ice-free summers occur in the years 2040, 2038, and 2052 in the CM4X-p125, CM4X-p25, and CM4.0 models, respectively (defined here as SIE $< 10^6$ km^2). This ice-free timing is consistent with the ice-free range of 2015–2052 as estimated by selected CMIP6 models (SIMIP Community, 2020; Jahn et al., 2024). All three models also reach ice free conditions in the months of July–November by the year 2100 (not shown).

4.2 Southern Ocean sea ice

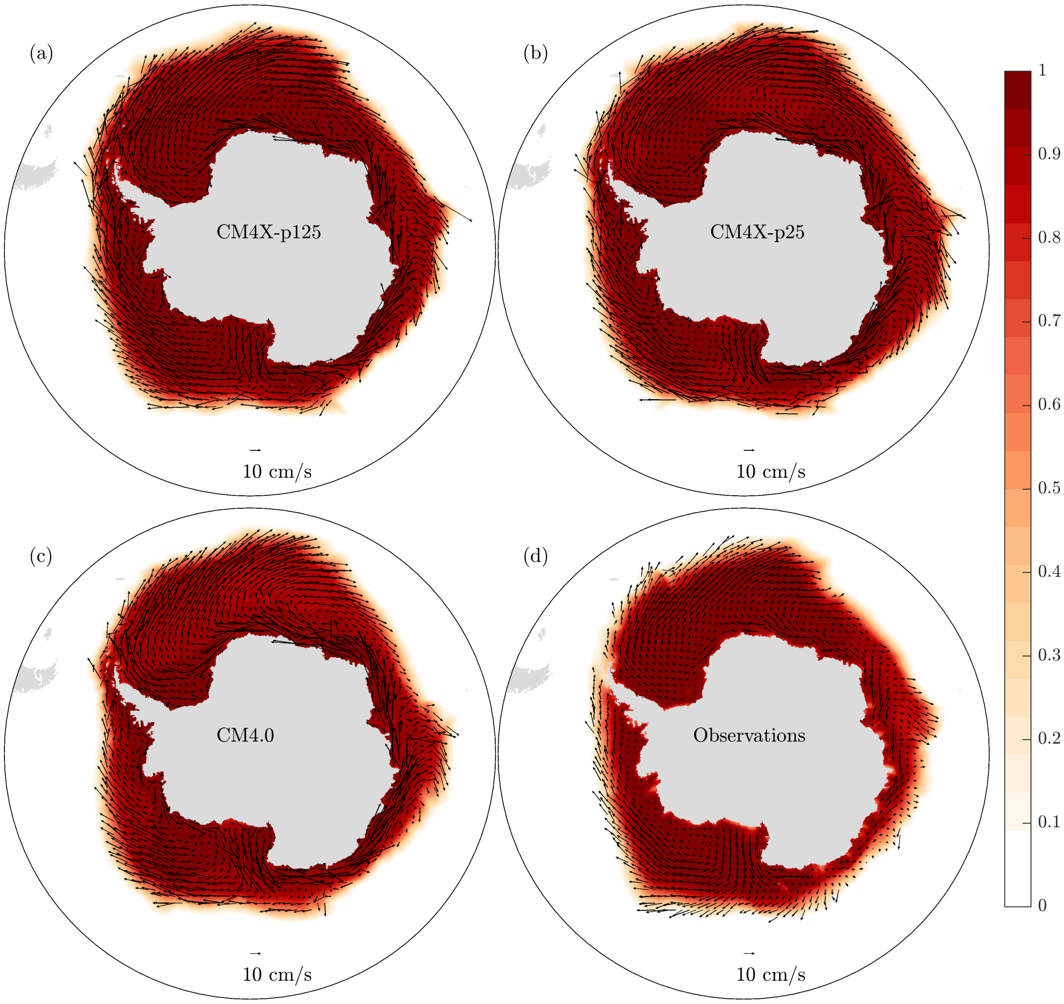

Figure 13. September Southern Ocean SIC climatologies (shading) and climatological JJAS sea ice drift (cm/s; vectors) in CM4X, CM4.0, and observations computed over the period 2011–2023. SIC and sea ice drift observations are from NSIDC and OSISF, respectively. Note that the OSISAF southern hemisphere drift observations span the period 2013–2023.

CM4.0, CM4X-p25, and CM4X-p125 each capture the asymmetric seasonal cycle of Southern Ocean SIE, with five months of ice retreat and seven months of ice advance (Figure 7b). Compared to the Arctic, the models' SIE biases are generally larger in the Southern Ocean. CM4.0 has an exaggerated Southern Ocean SIE seasonal cycle with too little sea ice in austral summer and too much sea ice in austral winter. The CM4X models have more extensive Southern Ocean sea ice than CM4.0 in all months of the year,

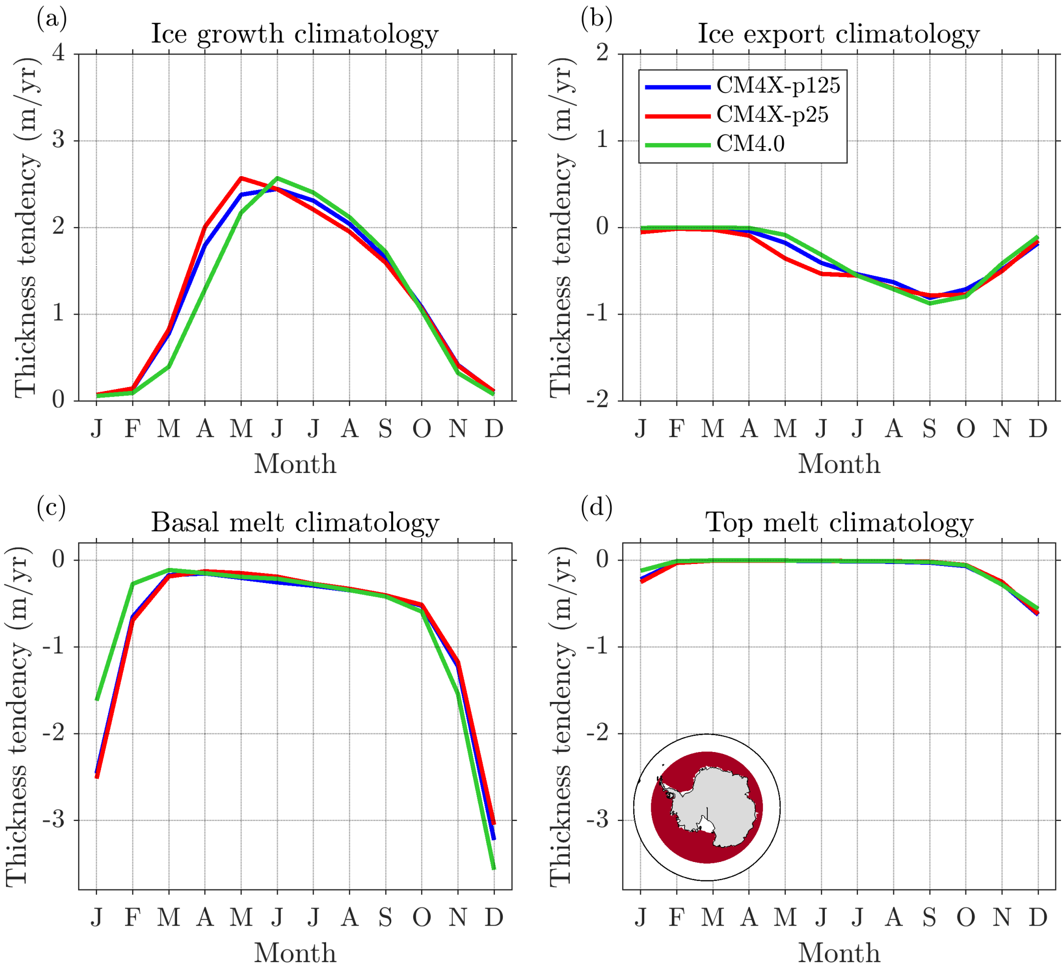

Figure 14. Southern Ocean sea ice mass budget climatologies computed over all gridpoints south of 63°S as shown in panel (d) over the period 1979–2023. The mass budget consists of SIT tendency terms (m/year) corresponding to congelation and frazil ice growth (a), mass transport convergence (b), basal melt (c), and top melt (d). Positive values correspond to mass gain and negative values correspond to mass loss. The y-axis range is the same across the panels, thus allowing for direct comparison of the various terms, and for comparing to the Arctic mass budget in Figure 10.

likely associated with the increased near-infrared land ice albedo values that were used in CM4X in order to promote Southern Ocean ventilation and production of AABW (see Appendix A4 in Part I (Griffies et al., 2024)). These higher albedos result in a cooler Southern Ocean surface climate with more sea ice than CM4.0. The increased sea ice coverage in CM4X improves upon the CM4.0 model biases in summer months yet exacerbates the winter sea ice biases. The RMS errors of the Pan-Antarctic SIE climatology are 2.38×10^6 km^2, 3.35×10^6 km^2, and 2.15×10^6 km^2 in CM4X-p125, CM4X-p25, and CM4.0, respectively, which can be compared to the CMIP5 multi-model mean RMSE of $3.42{\times}10^6$ km^2(Shu et al., 2015). A low bias in summer Southern Ocean sea ice is a ubiquitous bias across CMIP5 and CMIP6 models (Roach et al., 2020), which the CM4X models ameliorate.

Figure 8 shows climatological Southern Ocean SIC biases. We find that the spatial pattern of summer SIC is well captured by the CM4X-p125 model, whereas CM4X-p25 simulates too much sea ice in the Weddell Sea and too little in the Ross Sea (bottom row of Figure 8). All three models fail to simulate summer sea ice along the coastlines of the Indian Ocean and West Pacific sectors, which are regions with substantial landfast sea ice coverage (Fraser et al., 2023). We note that this sea ice model does not include a landfast ice parameterization. The summer SIC RMSE values of both CM4X models are reduced relative to CM4.0, which has negative SIC biases throughout the summer sea ice zone. All three models have positive winter SIC biases that are relatively circumpolar, however the biases are progressively stronger in CM4.0, CM4X-p125, and CM4X-p25 (third row of Fig. 8). CM4.0 has negative SIC biases within the sea ice pack in the Weddell Sea near Maud Rise, suggestive of too much vertical mixing and a tendency to form open-ocean polynyas in this region. These negative SIC biases are not present in the CM4X models.

The climatological winter Southern Ocean sea ice drift is shown in Figure 13. The general patterns of observed Southern Ocean sea ice drift are well captured by the models, with each model simulating northward sea ice export in the Weddell and Ross Seas, westward drift along the Antarctic coastal current, and strong eastward drift associated with the Antarctic circumpolar current. The models have drift speeds that are generally too fast relative to observations. This bias may contribute to the models' positive biases in wintertime SIC, since stronger drift implies a greater northward export of sea ice. We also find that the drift speeds along the Antarctic coastal current are notably higher than observed, especially in the zone of landfast sea ice along the eastern Antarctic coastline. This overly mobile sea ice shows that the model is unable to simulate landfast ice, and this potentially underpins the negative summer SIC biases in this region.

Figure 14 shows Southern Ocean sea ice mass budgets computed over the region south of 63°S (see inset of panel d). This region was chosen to encompass the primary zone of sea ice growth and melt while also capturing the dominant flux gates for sea ice export. Relative to CM4.0, the CM4X models show a clear shift in sea ice growth to earlier in the autumn season, consistent with the higher glacier albedos and cooler surface climate in CM4X (see Appendix A3 in Part I (Griffies et al., 2024)). The CM4X models also have more total annual sea ice growth, forming approximately an additional 0.1m of sea ice each year. The CM4X models also have more sea ice export during the autumn months, likely associated with the enhanced sea ice growth and thicker sea ice produced over these months. The dominant melt contributions come from basal melt, with the CM4X models showing a later spring onset of basal melt compared to CM4.0, consistent with the higher albedo and cooler surface climate in these runs. We also note that, compared to the Arctic, Southern Ocean sea ice has larger basal melt contributions during winter months, which tend to increase as winter mixed layers deepen.

Figures 11c,d and Figures 12c,d show time evolution of Southern Ocean SIE and SIV, respectively. Roach et al. (2020) showed that nearly every CMIP6 model simulates a negative Southern Ocean SIE trend in both summer and winter, failing to capture the observed trends which are close to zero in these seasons. This behavior is also the case for the CM4X and CM4.0 models, which simulate declines of Southern Ocean SIE in all months of the year over the period 1979–2023. This mismatch in modeled and observed trends may have contributions from missing meltwater forcing from Antarctic ice sheet and ice shelf melt (Bronselaer et al., 2018; Schmidt et al., 2023), systematic coupled model errors (Purich et al., 2016; Kay et al., 2016; Rackow et al., 2022), and internal climate variability (Meehl et al., 2016; L. Zhang et al., 2019). The CM4X and CM4.0 models simulate ice free Southern Ocean conditions (defined here as SIE $< 10^6$ km^2) in January–March by the year 2100, with CM4.0 reaching ice free states much earlier than CM4X. CM4X-p125 and CM4X-p25 have their first ice-free February in years 2049 and 2050, respectively, whereas CM4.0 simulates an ice free February states intermittently throughout the 20th century.

4.3 Conclusions regarding the sea ice simulations

Both CM4X models provide credible simulations of the Pan-Arctic sea ice mean state and trends, however the models simulate notable regional SIC errors, which are likely a combination of sea ice model physics errors and coupled model errors that originate in the atmospheric and oceanic components. It is notable that the spatial pattern of SIC and SIT model errors are very similar across the $1/4°$ and $1/8°$ CM4X configurations, and closely resemble the error patterns of the GFDL-ESM4.1 (Dunne et al., 2020) and GFDL-SPEAR (Bushuk et al., 2022) models, which have nominal horizontal grid spacings of $1/2°$ and $1°$, respectively. This similarity suggests that sea ice model errors are relatively insensitive to horizontal grid spacing across the $1/8°$–$1°$ range. Models in this range are not eddy resolving in the Arctic Ocean basin and Subpolar seas (see Figure 1 from Part I (Griffies et al., 2024)), and it is possible that a clearer impact of fine horizontal grid spacing would emerge in models that are fully eddy resolving in the Arctic. These eddy-resolving grids represent scales below the formal length scale of validity for viscous-plastic sea ice rheologies (Feltham, 2008). Even so, recent work has shown that the viscous-plastic rheology can simulate good agreement with observed sea ice drift and deformation even with 1 km grid spacing (Hutter et al., 2018). Key Arctic sea ice priorities for future model development include improving the spatial pattern of SIT, improving the magnitude and pattern of sea ice drift, and improving the persistent positive SIC bias in the GIN and Barents Seas.

Compared to the Arctic, the CM4X models have larger errors for the Southern Ocean sea ice mean state and trends, which is also a generic property across most CMIP models (Shu et al., 2015). For Southern Ocean sea ice, there appears to be a modest benefit from refined ice-ocean grid spacing, as CM4X-p125 has slightly reduced biases relative to CM4X-p25. It is notable that CM4X ice-ocean resolution does not clearly influence historical Southern Ocean sea ice trends, as CM4X-p125 and CM4X-p25 have comparable SIE trends across all months of the year. Key Southern Ocean sea ice priorities for future model development include improving the significant positive biases in wintertime SIC, reducing sea ice drift speeds, adding a representation of landfast Antarctic sea ice, improving simulated summer SIC in East Antarctica, and improving the simulation of SIE trends across all seasons.

5 Southern Ocean

The Southern Ocean plays a dominant role in anthropogenic heat and carbon uptake (Frölicher et al., 2015; Roemmich et al., 2015; DeVries et al., 2019), thus the representation of physical processes in this region is critical for accurately simulating the transient climate response. The Southern Ocean is home to the strongest current on the planet, the Antarctic Circumpolar Current (ACC), which acts as the primary pathway for inter-basin exchange of physical and biogeochemical tracers. Intimately linked to the structure of the ACC, the Southern Ocean is also home to a meridional overturning circulation whose deep branch ventilates the densest waters in the World Ocean, namely the Antarctic Bottom Water (AABW), and whose intermediate branch plays a large role in the oceanic sequestration of anthropogenic heat and carbon (see Morrison et al. (2022) for a review of physical processes). Additionally, the waters on and just offshore of the Antarctic continental shelf in the subpolar Southern Ocean directly influence the mass balance of the Antarctic Ice Sheet and thus dynamics in this region exert a strong influence on global sea level rise (Paolo et al., 2015). Indeed, such concerns about sea level rise have placed a growing appreciation for the important role of Antarctic shelf processes in the global climate system. Recent work has emphasized the need for improved model representation of ocean dynamics near to and along the Antarctic shelf, including the Antarctic Slope Current (ASC) and Antarctic Coastal Current (ACoC) (A. F. Thompson et al., 2018; Moorman et al., 2020; Purich & England, 2021; Beadling et al., 2022).

Mesoscale eddies, as well as jets and boundary currents, play a central role in the dynamics of the ACC and ASC, the meridional overturning circulation, and shelf-open-ocean exchange (e.g., Goddard et al. (2017); Stewart et al. (2018, 2019)). Hence, this region provides strong motivation to refine ocean grid spacing for studying the role of the Southern Ocean in climate. Although the CM4X models remain too coarse to resolve many key processes on the shelf (e.g., see Figure 1 from Part I (Griffies et al., 2024)), they succeed in pushing the envelope of global climate models by offering a refined representation of flows on the shelf and slope (grid spacing is roughly 7 km at 70°S in CM4X-p125), thus offering a tool to probe the role of these currents on larger scale climate.

5.1 Features of the horizontal circulation

The ACC has been a topic of study with large-scale and fine resolution numerical models since the 1990s, following the pioneering efforts of the FRAM project (FRAM Group, 1991) and further pursued across a grid resolution hierarchy by Hallberg and Gnanadesikan (2006). These studies, and many more, have helped to establish the fundamental importance of the ACC for large scale climate dynamics (Rintoul et al. (2001); Rintoul and Naveira Garabato (2013); Rintoul (2018)). Despite this fundamental importance, many coupled climate models still struggle to accurately represent the mean-state ACC strength and structure (Beadling et al., 2020). For the CM4X models, the ACC is revealed by a strong eastward zonal flow comprised of multiple jet-like structures such as seen in Figures 4 and 5 of Part I (Griffies et al., 2024), as well as Figure 15 shown here. Many details of the stronger ACC flow patterns are similar between CM4X-p25 and CM4X-p125, reflecting the deep reaching nature of Southern Ocean currents that are affected by bottom topography and thus generally follow f/H contours (f is the Coriolis parameter and H is the bottom depth). We also commented on this feature of the Southern Ocean in Figure 5 from Griffies et al. (2024), where much of the kinetic energy in the Southern Ocean is dominated by the depth averaged velocity.

Moving south towards Antarctica, we encounter the westward flowing (ASC) along the continental slope. As reviewed by A. F. Thompson et al. (2018), the ASC is present in most regions around Antarctica, with the notable exception of the West Antarctic Peninsula and westward until reaching the Amundsen Sea.

Though transporting far less mass than the ACC, there is a growing appreciation for the impacts of the ASC on regional and global climate. In particular, as reviewed by Beadling (2023), the ASC acts as a barrier to meltwater originating from Antarctic ice shelves leaving the continental shelf, and conversely as a barrier to relatively warm Circumpolar Deep Waters penetrating towards the continental shelf from the north. As encountered in modeling studies such as Goddard et al. (2017); Moorman et al. (2020); Lockwood et al. (2021); Beadling et al. (2022); Tesdal et al. (2023), a realistically strong ASC introduces a fundamentally new dynamical regime into the Southern Ocean that is absent from coarse models (roughly those ocean models with horizontal grids coarser than 0.25°). Given the extremely small Rossby deformation radius along the Antarctic continental slope/shelf region (see Figure 1 from Part I (Griffies et al., 2024)), it is likely that global models will fail to accurately represent the full dynamical impacts of the ASC until reaching toward 1 km horizontal grid spacing.

In Figure 15 we display two meridional-depth sections, one through the Drake Passage and one south of Tasmania. Here we see the deep reaching eastward jet-like flows found in both sections, along with distinct westward flows. The westward flows are generally weaker and found particularly at depth and, for the Tasmanian section, we find the strong westward flowing ASC along the Antarctic continental slope. Note the rather weak flow in the Tasmanian section for latitudes between approximately 60°S and 55°S, with the upper flow weakly westward and deep flows very weak. The topography between these regions is rather fine scale, suggesting that this "rough" topography acts to weaken

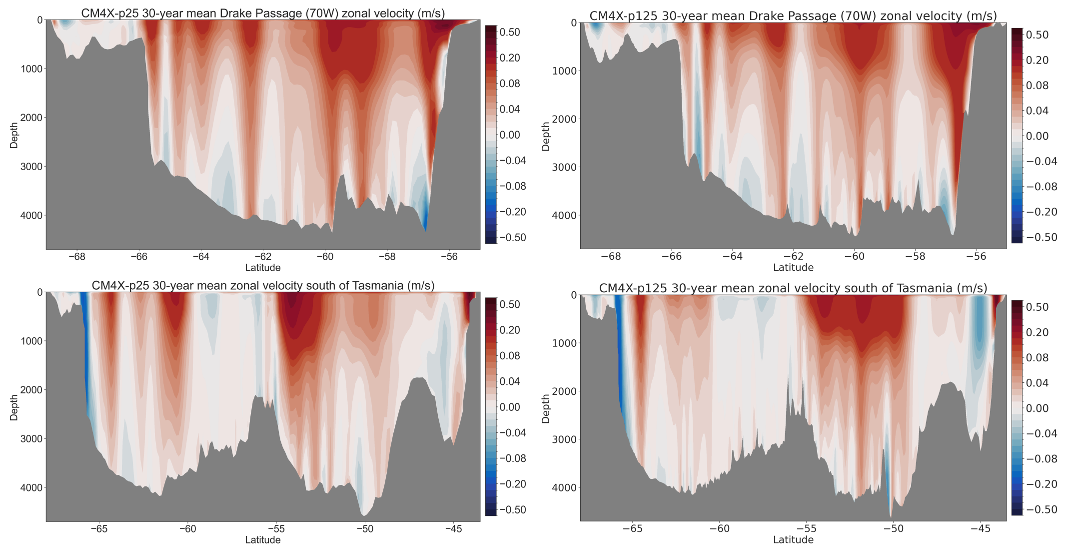

Figure 15. 30 year mean zonal velocity along a longitude within the Drake Passage (upper row) and south of Tasmania (lower row). Note the deep reaching zonal flows, with some deep reaching westward flows. Also note that strong and deep reaching westward flowing Antarctic Slope Current seen in the section south of Tasmania, whereas this current is largely absent in the Drake Passage section (see A. F. Thompson et al. (2018) for a review). The colorbars are the same for all panels, though the latitude range differs for the top row and bottom row.

the otherwise deep reaching flow. This feature of the flow over the rough topography may suggest a role for the rough bottom modes of LaCasce (2017). Even so, the horizontal flow does *not* generally vanish at the bottom, which contrasts to the assumptions of LaCasce (2017). Indeed, the Drake Passage section of CM4X-p125 is notable for its bottom enhanced westward flows. It is notable that all eastward flows are surface intensified (equivalent barotropic), whereas some of the westward flows in the open ACC, and particularly in the Drake Passage section, are bottom intensified. The presence of deep westward jets in the vicinity of the ACC have been noted in previous observational studies and high resolution simulations (Xu et al., 2020).

In Figure 16 we show the Drake Passage transport, which provides a traditional measure summarizing the zonal flow in the ACC. CM4X-p25 is roughly 10-15 Sv stronger than CM4X-p125 throughout the historical simulation, with CM4X-p125 consistent with the CMIP6 ensemble mean whereas CM4X-p25 is consistent with the lower end of the observational based estimates. After roughly 100 years of spin-up, both piControl simulations exhibit multi-decadal fluctuations of roughly 10 Sv. For the historical simulation, CM4X-p125 shows a slight decrease whereas CM4X-p25 is roughly unchanged. Both models show a decline during the SSP5-8.5 until around year 2060, at which point the strength stabilizes (CM4X-p25) or begins to rebound (CM4X-p125).

5.2 Ventilation, watermass transformation, and overturning circulation

The CM4X mixed layer depths in Figure 9 from Part I (Griffies et al., 2024) reveals a narrow band of deep mixing on the Antarctic continental shelf where dense shelf water (DSW) forms and then subsequently overflows down the continental slope, ventilating the deep ocean and leading to the formation of AABW (Rintoul & Naveira Garabato, 2013; Rintoul, 2018). In Figure 17 we display bottom temperature and salinity on the Antarctic continental shelf, defined here as the region landward of the 1000-m iso-

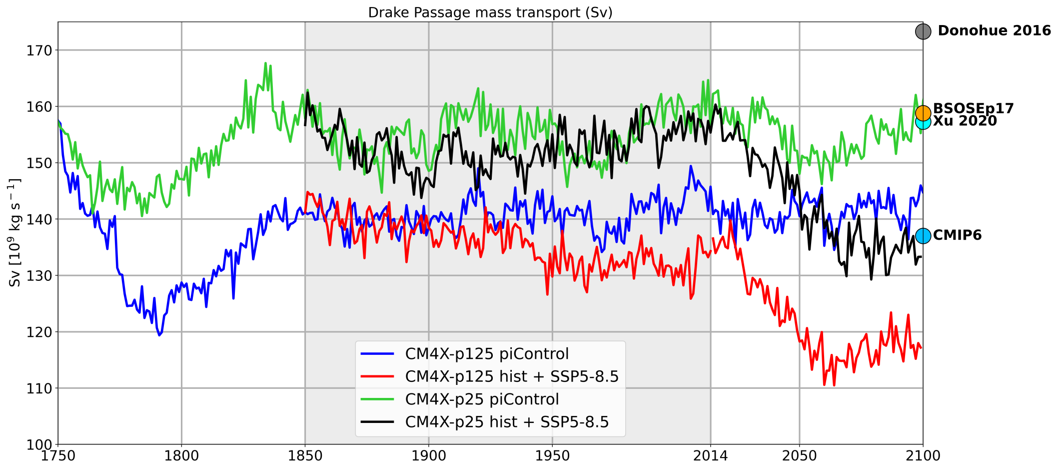

Figure 16. Time series of annual mean mass transport through the Drake Passage in units of Sverdrup (10^9 kg s^{-1}) for the CM4X simulations. The gray shaded region is the historical portion of the simulation. The colored markers on the right edge of the plot indicate the observed estimate of the total flow through the Drake Passage from the cDrake array from Donohue et al. (2016), based on an observation period of 2007 to 2011; the observationally-constrained Biogeochemical Southern Ocean State Estimate (Verdy & Mazloff, 2017) at 1/6-degree and integrated from 2013–2018; the 1/12th HYbrid Coordinate Ocean Model (Xu et al., 2020); and the CMIP6 Drake Passage transport ensemble mean of historical simulations averaged from 1986 to 2005 (Beadling et al., 2020).

bath. These shelf properties play a central role in determining the volume and formation rate of DSW and influence the stability of ice shelves ringing the continent. Both CM4X-p25 and CM4X-p125 versions produce spatial distributions of bottom temperature and salinity qualitatively consistent with observations, yet with notable differences in magnitudes in specific regions. For bottom temperatures, both the CM4X models are more consistent with observations than that simulated by most models within the CMIP6 ensemble, which exhibit significant biases (e.g., Figure S8 from Purich and England 2021). CM4X-p125 exhibits a warmer and slightly more saline West Antarctic shelf compared to CM4X-p25. The West Antarctic shelf regime is characterized as a "warm shelf" (Thompson et al., 2018) where upward sloping isopycnals and the lack of an ASC allows warm Circumpolar Deep Water (CDW) to readily access the shelf. The warmer West Antarctic bottom temperatures in CM4X-p125 are consistent with a slightly warmer mean-state off shore reservoir of CDW compared to CM4X-p25. CM4X-p125 also exhibits a much fresher East Antarctic shelf that could be related to its stronger ASC compared to CM4X-p25.

The spatial pattern of surface watermass transformation (WMT) across the densest waters confirm that both versions of CM4X simulate DSW formation and subsequent down slope flow in realistic locations along the Antarctic shelf (Jacobs, 2004; Silvano et al., 2023), highlighting the shelves around Weddell Sea, Prydz Bay, Adelie Land, and the Ross Sea (Figure 18). The red shading in Figures 18a,b shows the mean surface WMT per unit area across σ_2 (potential density referenced to 2000 dbar) classes for the period 1975-2012, focusing on the densest water classes. Choosing the densest waters available in each region captures the spatial distribution of where DSW formation occurs that contributes to the formation of AABW. The bottom age distribution (green shading in Figures 18a,b) illustrates the pathways of newly formed dense waters from the Antarctic slope to the abyssal ocean. These patterns align with the known AABW pathways iden-

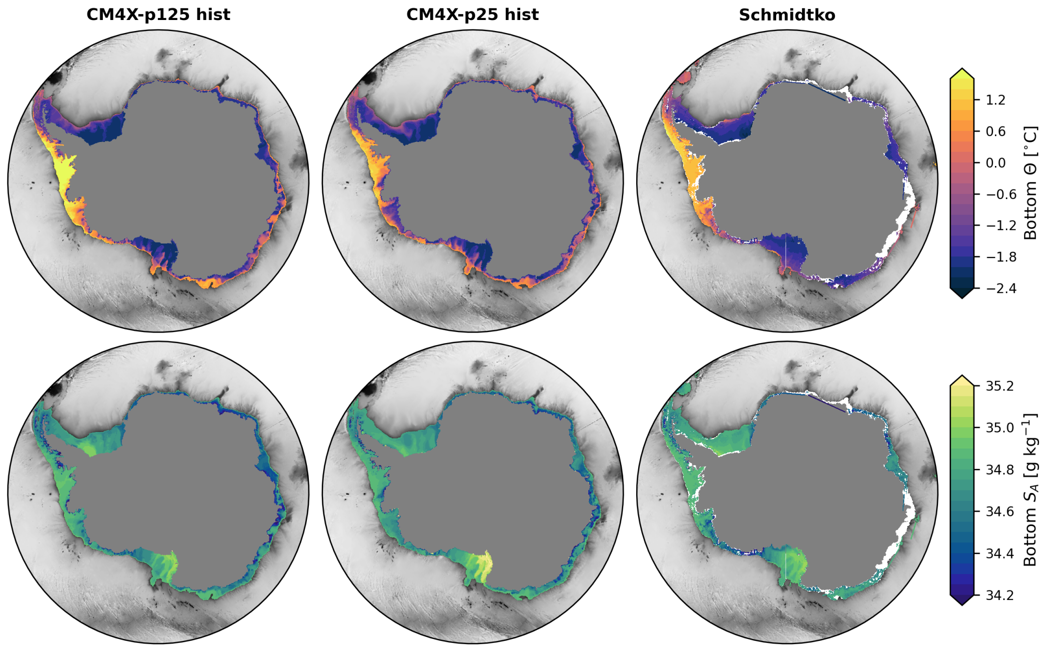

Figure 17. Temperature and salinity along the bottom of the Antarctic shelf from the CM4X models and as compared to the observational based analysis of Schmidtko et al. (2014). The fields are averages for the period 1975-2012, which is the time period used by Schmidtko et al. (2014). We define the shelf as the region with depth shallower than 1000 m.

tified from observational studies (e.g., Silvano et al. (2023)) and is consistent with passive tracers studies using a reanalysis-forced model (Solodoch et al., 2022).

The total DSW formation over the Antarctic continental shelf, estimated as the maximum surface WMT, is approximately 5 Sv and occurs for $\sigma_2 > 37$, with CM4X-p25 exhibiting a slightly greater DSW formation compared to CM4X-p125 (Figure 18c). Furthermore, the surface WMT is shifted towards higher σ_2 by around 0.1 in CM4X-p25 compared to CM4X-p125, especially in the Ross Sea sector (dashed lines in Figure 18c), which is consistent with the formation of higher salinity waters in this region (Figure 17). The Ross Sea sector shows the highest WMT rates compared to other regions, while the Weddell Sea sector (dotted lines in Figure 18c) also exhibits significant WMT, in agreement with observational estimates identifying these two regions as the major sources of DSW and AABW (Silvano et al., 2023).

To interpret Figure 18c, it is important to note at what densities positive WMT occurs. Both the Ross and Weddell sections in Figure 18c represent the majority of the WMT in the range of $\sigma_2 > 37$. Adding these two together explains the bulk contribution to dense water formation required for AABW. Other regions over the Antarctic shelf show positive WMT. However, these transformations occur at generally lighter densities ($\sigma_2 < 37$). As shown in Figure 7 from Part I (Griffies et al., 2024), the bottom cell overturning is associated with waters denser than $\sigma_2 = 37$, indicating that any surface WMT at lighter densities are not expected to contribute to AABW formation and bottom cell overturning.

The spatial pattern in bottom age (Figures 18a,b) is also consistent with the difference in surface WMT between the two CM4X models, and the dominant role of surface WMT in the Ross and Weddell sections, showing slightly younger bottom waters in CM4X-p25 along the AABW pathways due to the higher density waters formed on the shelf, emanating away from the Ross and Weddell Seas. The total DSW formation rate of approximately 5 Sv in both CM4X model configurations is on the lower end of the observational range, which spans 5-15 Sv (Silvano et al., 2023). However, this formation rate is a significant improvement compared to most CMIP6 models (Heuzé, 2021), suggesting that both CM4X simulations capture the relevant processes responsible for DSW and AABW production.

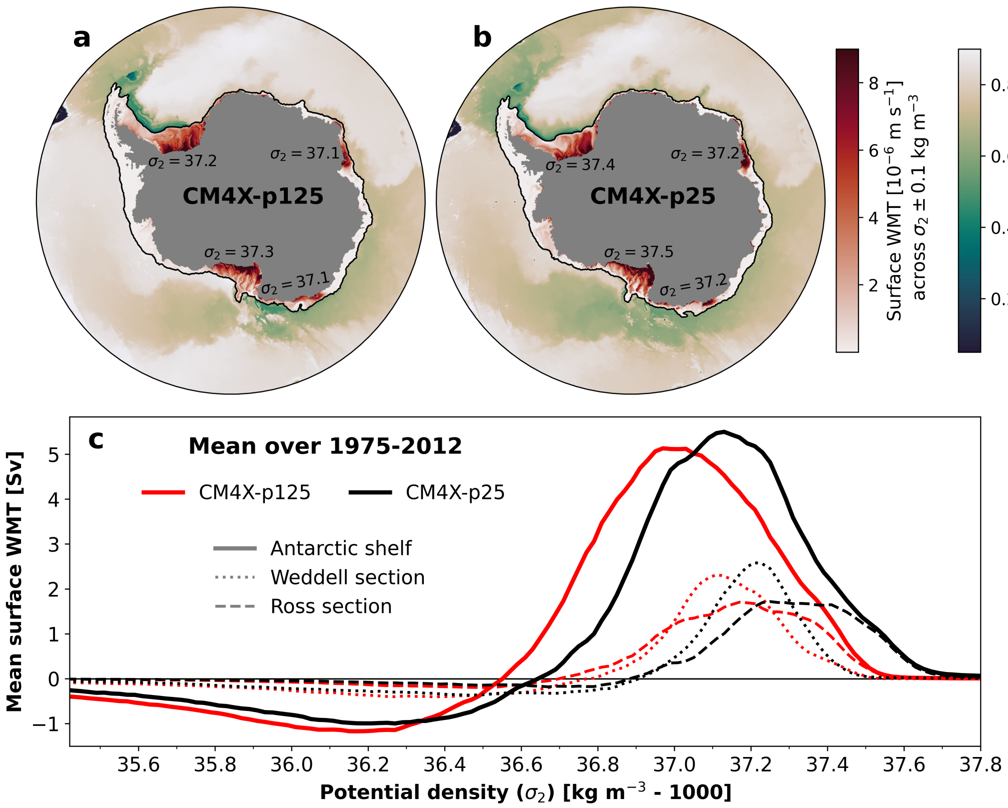

Figure 18. Surface water mass transformation (WMT) on the Antarctic continental shelf and bottom water age distribution over the abyssal Southern Ocean in (a) CM4X-p125 and (b) CM4X-p25. The red shading in panels a and b represent the time mean (1975-2012) surface WMT per unit area across σ_2 potential density classes (potential density referenced to 2000 dbar), separately determined in the four key DSW formation regions based on the densest water class. For the Weddell, Prydz/Adelie, and Ross shelves the surface WMT is mapped across σ_2 isopycnals of 37.2, 37.1 and 37.3 in CM4X-p125, and 37.4, 37,2 and 37.5 in CM4X-p25. The green shading in panels a and b represents the bottom age tracer at year 2009, normalized to the total simulation length (100 years of spinup + 1850-2009 historical = 260 years). (c) Mean (1975-2012) surface WMT in σ_2 integrated over the Antarctic shelf, Weddell (62°W-10°E) and Ross section (154°E-134°W) for CM4X-p125 (red lines) and CM4X-p25 (black lines).

The overturning circulation streamfunction offers a means to both measure and to visualize ventilation of the ocean interior. As a complement to the pole-to-pole overturning in Figure 7 from Part I (Griffies et al., 2024) that illustrates connections between the Southern Ocean and North Atlantic, in Figure 19 we focus on the Southern Ocean overturning. The AABW cell is the densest cell (blue counterclockwise cell) associated

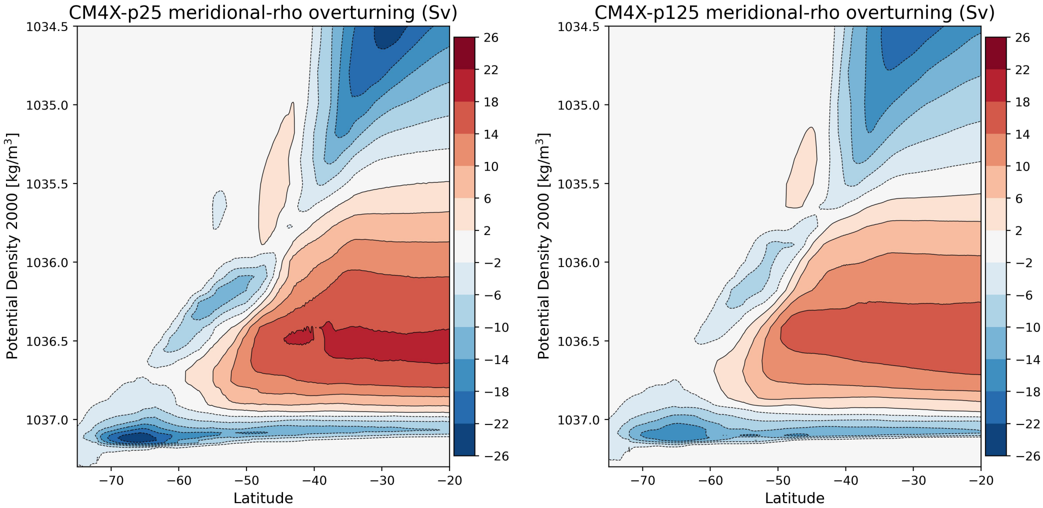

Figure 19. Meridional-density (ρ_{2000}) overturning circulation in the Southern Ocean as computed using time mean flow from years 1980-2009, with CM4X-p25 on the left and CM4X-p125 on the right.

with waters formed via DSW production on the Antarctic shelf and subsequent overflow and entrainment into the abyssal ocean (i.e., the processes shown in Figure 18). CM4X-p25 shows slightly larger formation around 65°S-70°S, and yet the AABW signal is slightly stronger in CM4X-p125 upon reaching 30°S. This disagreement between the strength of the AABW cell in the subpolar region and at 30°S indicates potentially larger interior mixing in CM4X-p25 which erodes the strong AABW transport away from the subpolar region. The other (blue) counterclockwise overturning cell is split into two sections in CM4X-p25 and CM4X-p125, though it is nearly connected in CM4X-p125. In the low latitudes, this cell is associated with subtropical mode waters. As discussed by Hallberg and Gnanadesikan (2006) (see their Section 3a), the merging of this cell southward across 45°S results from meridional mass transport from transient mesoscale eddies, with such eddy variability stronger in CM4X-p125 (e.g., East Australian Current, Agulhas Rings). The dense flow in the red clockwise cell is associated with North Atlantic Deep Water (NADW) and Circumpolar Deep Water (CDW) moving south, with a portion of this water lightened into Antarctic Intermediate Water (AAIW) and another portion densified into AABW.

5.3 Multi-decadal fluctuations in the piControl simulations

In Figure 20 we show the time series for the Drake Passage transport from the piControl simulations in CM4X as well as CM4.0. As described by Held et al. (2019), the large amplitude multi-decadal fluctuations in CM4.0 are associated with very large (i.e., super) polynyas in the Ross Sea. Such polynyas also appear in the ESM4 simulation of Dunne et al. (2020). In contrast, we do not find these super-polynyas in either CM4X-p25 or CM4X-p125, with both models exhibiting more modest multi-decadal fluctuations. One hypothesis for the absence of super-polynyas in CM4X relates to the increase in land ice albedo relative to CM4.0, resulting in a slightly cooler Antarctic climate and thus supporting more intermittent ventilation with smaller polynyas, rather than the buildup of massive subsurface heat charging the super-polynyas in CM4.0 (L. Zhang et al., 2021). This hypothesis comes with a caveat, however, since the land ice albedos used in CM4X are the same as ESM4, and yet ESM4 also suffers from the extremely strong polynyas.

We thus suspect that the full story for polynya events involves multiple factors, including winds, sea ice, and ice shelves.

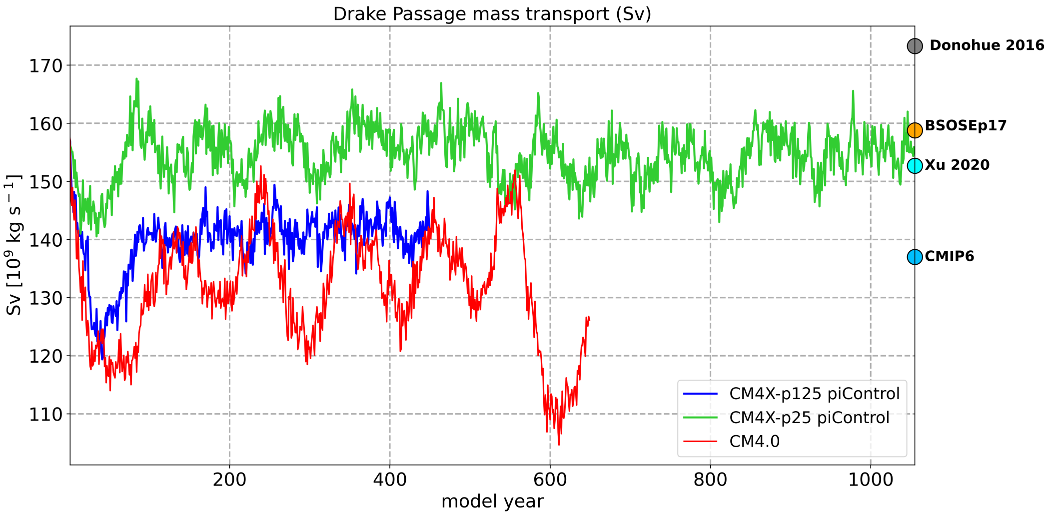

Figure 20. Time series of annual mean mass transport through the Drake Passage in units of Sverdrup (10^9 kg s^{-1}) for the piControl simulations from CM4X and CM4.0.

As a further means to distinguish the CM4X Southern Ocean simulations from CM4.0, Figure 21 provides time series for Antarctic shelf salinity, circulation strength of the bottom overturning cell, AABW transport at 30°S, as well as Hovmöller diagrams of surface water mass transformation due to heat fluxes in σ_2. Here we again see signatures of the super-polynyas in CM4.0 whereas both CM4X simulations exhibit smaller amplitude fluctuations. The variability in CM4.0 highlights how the occurrence of large open-ocean polynyas lead to a series of interconnected changes in physical processes within the Southern Ocean. The CM4.0 piControl exhibits marked multi-decadal oscillations in shelf salinity, offshore heat loss, and dense water formation, seen in Figure 21 as positive excursions of surface WMT at the densest open waters ($\sigma_2 > 37$). These episodes of enhanced dense water formation directly impact the large-scale circulation, as evidenced by the concurrent peak in the strength of the bottom overturning cell, followed by increased AABW transport at 30°S. Furthermore, the same multi-decadal fluctuations are seen in the Drake Passage transport (Figure 20), illustrating the connection between dense water formation, overturning and the strength of the ACC.

The multi-decadal oscillations are still present in the CM4X-p25 piControl simulations, but are more muted compared to CM4.0. We observe variability in the subpolar cell strength that oscillates around 20 Sv with a clear periodicity. This muted variability is also reflected in the Antarctic shelf salinity and the offshore surface WMT. Thus, there is some intrinsic variability that is still apparent in the quarter degree piControl runs, consistent with L. Zhang et al. (2021), occurring at higher frequency and with weaker amplitude when land ice albedo is increased. Figures 20 and 21 suggest that the oscillations do not have as large an impact across different parts of the Southern Ocean in CM4X-p25 as they do in CM4.0. Interestingly, these oscillations are not as clearly seen in the CM4X-p125 version, which could be due to the shorter run time.

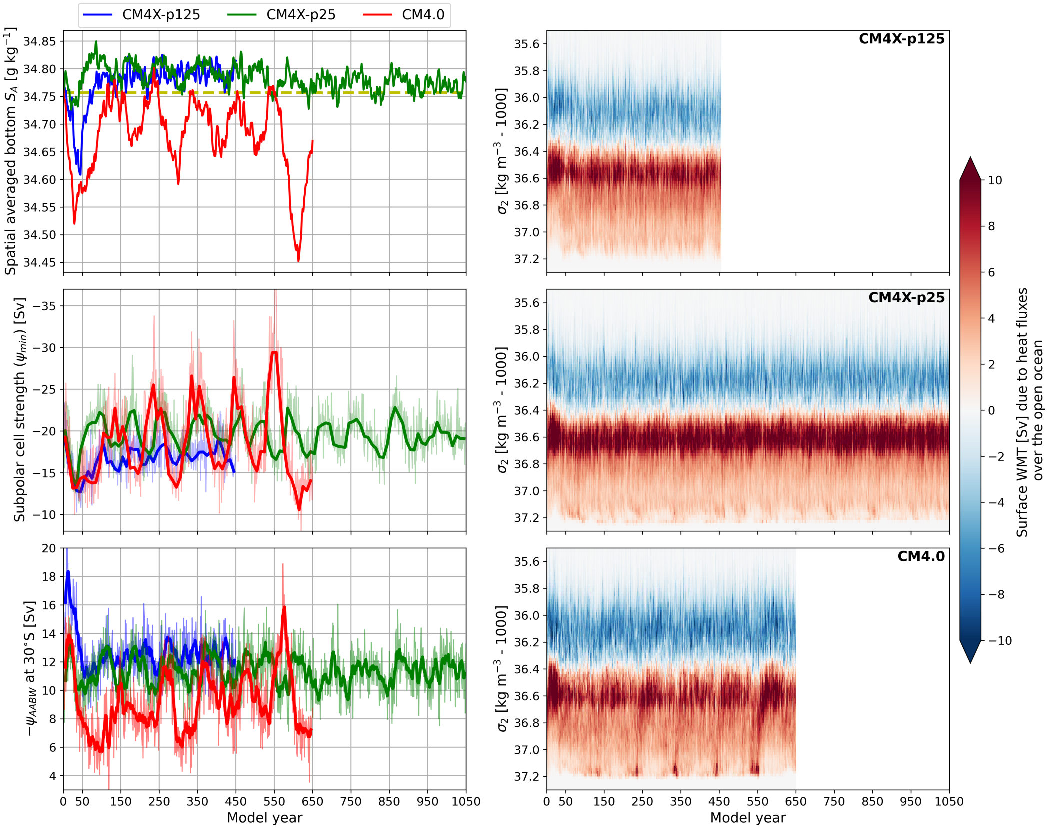

Figure 21. Time series for the annual mean salinity as area averaged around the bottom of the Antarctic shelf (upper left), circulation strength of the bottom overturning cell (middle left) and AABW transport at 30S (lower left) for CM4X-p125, CM4X-p25 and CM4.0 piControl simulations. The finer lines in the middle and lower left panels are annual means. The thick lines in the middle panel are decadal means, while the thicker lines in the lower panel are 10-year running means. The right panels show Hovmöller diagrams of surface forced water mass transformation in σ_2-space (potential density referenced to 2000 dbar) over the open Southern Ocean due to heat fluxes in CM4X-p125 (upper), CM4X-p25 (middle) and CM4.0 (lower). When calculating the area-average in the upper left panel we define the Antarctic shelf as the region with depths shallower than 1000 m. The yellow dashed line in the upper left denotes the observation-based climatological mean of Antarctic shelf salinity from Schmidtko et al. (2014).

5.4 Transient climate response

In Figure 22 we provide a suite of five time series for various properties around the Southern Ocean during the piControl spin-up, historical, and SSP5-8.5 scenario, that emphasize the coupling of properties between the shelf and open ocean. This coupling is also found in the shelf-open-ocean diagnostics presented in Figure 21. A key take away from Figure 22 is the general agreement in the transient response regardless of slight differences in the mean-state. Both models show a steep drop in shelf salinity at the start of the SSP5-8.5 simulation driven by freshening associated with an enhanced hydrological cycle and sea ice melt (and decreased sea ice formation) under increased warming. This drop also coincides with a sharp acceleration of the ASC in both models. The temporal pattern of ASC acceleration differs depending on location as the mean-state ASC is influenced by different dynamics along the shelf (Huneke et al., 2022). However, a strong ASC acceleration is consistently found from years 2014 to 2060 regardless of location.

The increased strength of the ASC also likely contributes to the consistent response found in shelf salinity via a positive feedback mechanism established between shelf freshening that enhances the ASC and thus leading to more freshwater trapping on the shelf (Moorman et al., 2020; Lockwood et al., 2021; Beadling et al., 2022). Despite initial differences in the strength of the subpolar cell, with CM4X-p25 showing stronger abyssal overturning, both models exhibit a similar magnitude of decline in bottom cell strength by the end of the 21st century under the SSP5-8.5 scenario. CM4X-p125 displays a steady weakening over time, whereas CM4X-p25 experiences a more abrupt reduction. Nevertheless, both models demonstrate comparable trends in AABW export at 30S, highlighting consistent responses to the transient conditions throughout the historical and SSP5-8.5 simulations. Investigating the differences between these two models and the mechanisms driving the response in shelf and open-ocean processes under SSP5-8.5 is a topic of further investigation.

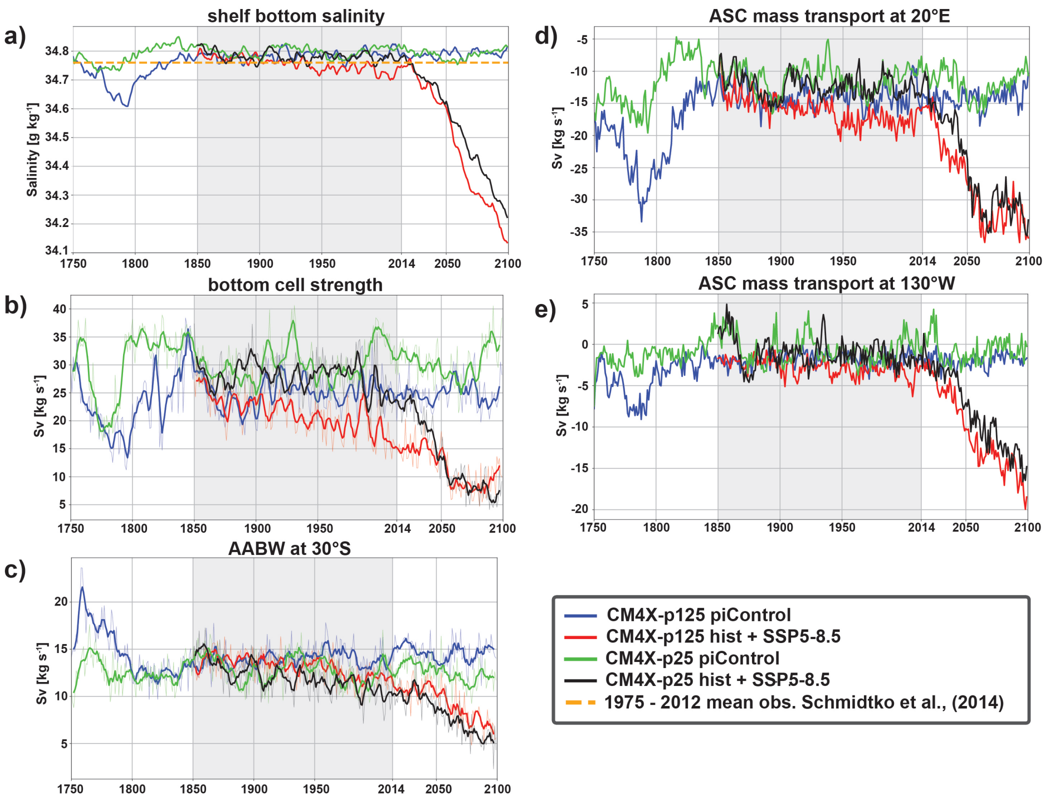

Figure 22. Time series for various properties around the Southern Ocean during the piControl spin-up, historical, and SSP5-8.5 scenario. a) Spatially averaged continental shelf bottom salinity b) circulation strength of the bottom overturning cell, c) AABW volume transport at 30°S, e) ASC strength at 20°E and f) ASC strength at 130°W. The ASC strength is shown at two different locations due to its flow characteristics being governed by different dynamics around the continental shelf (A. F. Thompson et al., 2018). The finer lines in panels b and c are annual means, while thicker lines in panels b and c are 5-year running means. When calculating the area-average in panel a, we define the Antarctic shelf as the region with depths shallower than 1000 m. The yellow dashed line in the upper left denotes the observation-based climatological mean of Antarctic shelf salinity from (Schmidtko et al., 2014).

5.5 Conclusions regarding the Southern Ocean simulations

The credible representation of Southern Ocean properties in CM4X opens up opportunities to study fully coupled Southern Ocean processes and Antarctic margin dynamics. The largest improvement in the CM4X models relative to previous generation GFDL models (CM4.0 and ESM4), is the lack of large amplitude multi-decadal oscillations associated with super polyna events. The presence of such large amplitude variability in CM4.0 prevented a suitable control run from which to branch perturbation experiments. Since the super-polynya-driven variability imprinted on the global climate and the Southern Ocean mean state, they made it nontrivial to interpret signal and "noise" in transient response to forcing perturbation experiments (Beadling et al., 2022; Tesdal et al., 2023).

Particularly notable features of the CM4X Southern Ocean simulation include the credible representations of Antarctic shelf hydrography, realistic locations of DSW production and overflow, and resolution of a strong ASC along the continental slope. A realistic representation of shelf properties and dynamics along the shelf-slope is required for examining shelf-open-ocean interactions, improving confidence in the thermal forcing of the AIS in a transient climate, and providing boundary conditions for dynamical ice sheet models. The credibility of CM4X near and along the Antarctic margin is encouraging for the utility of studying high latitude processes and for coupling with dynamical ice sheet models.

Additionally, with differing horizontal grid spacing providing for different degrees of representation of ocean mesoscale features, the CM4X suite is well positioned to probe the role of the ocean mesoscale within the climate system. Slight differences identified here between CM4X-p25 and CM4X-p125 in their velocity structure, Drake Passage transport, and Southern Ocean overturning may be linked to differing representation of mesoscale features and topography. Future work will aim at disentangling the sources of these differences between the two models and understanding whether these differences imprint on the transient climate response.

6 North Atlantic circulation

In this section we survey the horizontal and meridional overturning circulation features of the North Atlantic portion of the CM4X simulations, with comparisons made to observational estimates and other models, when available.

6.1 Gulf Stream

Here we focus on the Gulf Stream representation in CM4X as it compares to theoretical expectations, observed three dimensional structure (Todd, 2021), and the representation in a 1/12 degree regional ocean simulation also using the MOM6 dynamical core (NWA12, Ross et al. (2023)).

Expectations from previous studies

Past efforts to determine controls on the simulated Gulf Stream have focused on the latitude of separation from the coast, as well as behavior of the jet and mesoscale eddies post separation (Chassignet & Marshall, 2008). This latitude of separation and/or the mean latitude of the Gulf Stream extension is often associated with a temperature front or the zero barotropic vorticity contour (Section 6.2). In this section we consider jet coherence and jet location relative to the continental slope. This analysis finds that once leaving the coast, the CM4X simulations have a rather diffuse and meandering Gulf Stream jet that is, unfortunately, rather distinct from observational measurements.

Fundamentally, Gulf Stream fidelity depends on model horizontal grid spacing relative to the first baroclinic Rossby radius of deformation. At middle latitudes, Hallberg (2013) (see also Figure 1 from Part I (Griffies et al., 2024)) reveals a requirement of approximately 0.25° over the deep ocean (deeper than 3000 m) and approximately 0.125° closer to the coast (500 m $<$ depth $<$ 3000 m). Since the Gulf Stream 'leans' on the U.S. east coast continental shelf prior to separation, we expect accurate simulations require grid spacing to be at or finer than 0.125°.

Gulf Stream simulation features also depend on the spatial structure of wind forcing, deep western boundary current strength, bathymetric slope resolution, and model viscosity parameterizations that moderate resolved mesoscale turbulence (Hurlburt and Hogan (2000), Parsons (2006), Chassignet and Marshall (2008), Ezer (2016), Chassignet and Xu (2017), and Debreu et al. (2022)). Chassignet and Marshall (2008) show how choices of biharmonic and/or Laplacian viscosity control the latitude of separation, determine the presence or absence of standing eddies, and shape downstream instability behavior. The authors demonstrate that use of a low biharmonic viscosity can result in early separation and the presence of a standing eddy near Cape Hatteras, while use of a Laplacian can result in flow that is too laminar and a Gulf Stream that does not penetrate into the North Atlantic interior.

Plan view and vertical sections of the Gulf Stream

Qualitative plan view comparisons of surface mean kinetic energy (MKE) in CM4X simulations, NWA12, and observations reveal the initial latitude of separation from the coast as too far south (Fig. 23a-d). The observed time mean latitude of separation is approximately 37.5°N while in CM4X-p25 it is approximately 31°N and in CM4X-p125 approximately 32°N. This shift northwards with refined grid spacing continues when comparing to the 1/12° regional ocean model, where Gulf Stream separation occurs at approximately 37°N. As discussed by Chassignet and Marshall (2008), this characterization of the Gulf Stream separation and particular focus on jet coherence and the horizontal spreading of mean kinetic energy centers on the impact of viscosity choices on flow around the Charleston Bump. In both CM4X simulations use of a relatively low biharmonic viscosity value, as well representation limitations associated with relatively coarse horizontal resolution, are likely the cause of early separation.

Upstream of separation, surface MKE is greater in CM4X-p125 than in CM4X-p25 and closer to observed magnitudes. This increase in Gulf Stream strength in CM4X-p125 may be responsible for more realistic flow around the the Charleston Bump compared to CM4X-p25. The vertical extent of this relative increase is seen in a zonal cross-section at 29.5°N (Figure 24a-d). While the depth penetration of the Gulf Stream is similar in all panels, CM4X-p125 upper-ocean northward velocity (at depths shallower than 300 m) is over 50% larger than in CM4X-p25.

Differences between CM4X and observations and NWA12 are more significant post separation. In both CM4X-p25 and CM4X-p125, the eastward flowing jet rapidly dissipates post separation, resulting in a diffuse field of eddies that meander further north than observed eddy activity. Correspondingly, MKE is spread across roughly 30°N to 40°N, and dissipates moving eastward. Cross-sections at Cape Hatteras and at 70.5°W (Figure 24e-l) reveal the vertical structure of velocity to rapidly decrease with increasing depth. Compared to observations and NWA12, the CM4X Gulf Stream structure is far too surface intensified and horizontally diffuse. At 1800 m, a depth greater than that at which gliders collected measurements in Todd (2021), the spatial distribution of MKE appears similar in CM4X and NWA12. Noted differences between CM4X-p25 and CM4X-p125 include the near barotropic standing eddy in CM4X-p25 (Figures 23c,f) and a more energetic deep slope current north of 37.5°N in CM4X-p25.

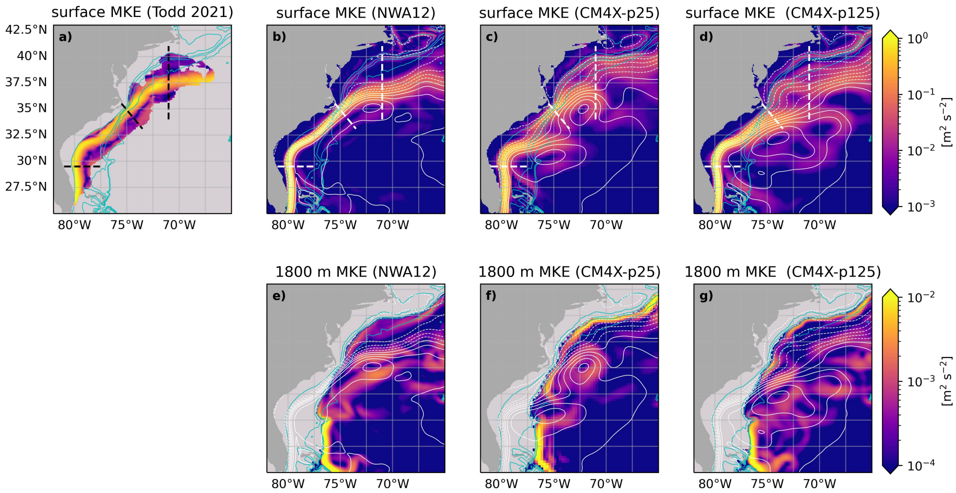

Figure 23. a) Time mean surface kinetic energy (MKE) derived from observations collected between 2015-06 and 2020-7 (Todd, 2021) (note log color scale). Blue contours identify the 100 m, 1000 m, 2000 m, and 3000 m isobaths. Black dashed lines are the locations of zonal, Cape Hatteras, and meridional cross-sections in Figure 24. b) 2010-2014 time mean surface kinetic energy from the NWA12 1/12° degree regional ocean model (Ross et al., 2023). Time mean sea surface height contours are added in white. c) Same as b) but for CM4X-p25. d) Same as b) but for CM4X-p125. e-g) Mean kinetic energy as in b-d) at 1800 m (note colormap scale change).

6.2 Barotropic vorticity

Interactions between deep ocean flows and sloping bathymetry result in bottom stretching and bottom pressure torques, influencing the western boundary currents and ocean circulation (Hughes & De Cuevas, 2001; R. Zhang & Vallis, 2007; Waldman & Giordani, 2023; Khatri et al., 2024). We investigate how spatial resolution differences between CM4X-p25 and CM4X-p125 affect the strength of the North Atlantic gyre circulations. For this purpose, we use the linear steady-state barotropic vorticity budget, which is often used to study how surface winds and bathymetry control ocean gyres (Stommel, 1948; Holland, 1967). Following Yeager (2015), we use the streamfunction, ψ, form of the barotropic vorticity budget

$$\underbrace{\int_{x_w}^{x} V \,\mathrm{d}x}_{\psi} \approx \underbrace{\frac{1}{\beta}\int_{x_w}^{x} J(p_b, H)\,\mathrm{d}x}_{\psi_{BPT}} + \underbrace{\frac{1}{\beta\rho_0}\int_{x_w}^{x} \hat{\boldsymbol{z}}\cdot(\nabla\times\tau_s)\,\mathrm{d}x.}_{\psi_{\tau_s}} \tag{3}$$

Here, V is the vertically integrated meridional velocity, β is the meridional derivative of the Coriolis parameter, p_b is bottom pressure, H is ocean depth, τ_s is the surface wind stress, x_w represents the western continental boundary and ρ_0 is the Boussinesq reference density. Note that contributions from friction and nonlinear terms are not included, as these contributions are relatively small in time-mean vorticity balances (see Khatri et al. (2024) for a complete vorticity budget analysis and diagnosis methodology).

The time-mean spatial patterns of subtropical and subpolar gyres are well captured by the combined effects of bottom pressure torque and surface wind stress (Figure 25).

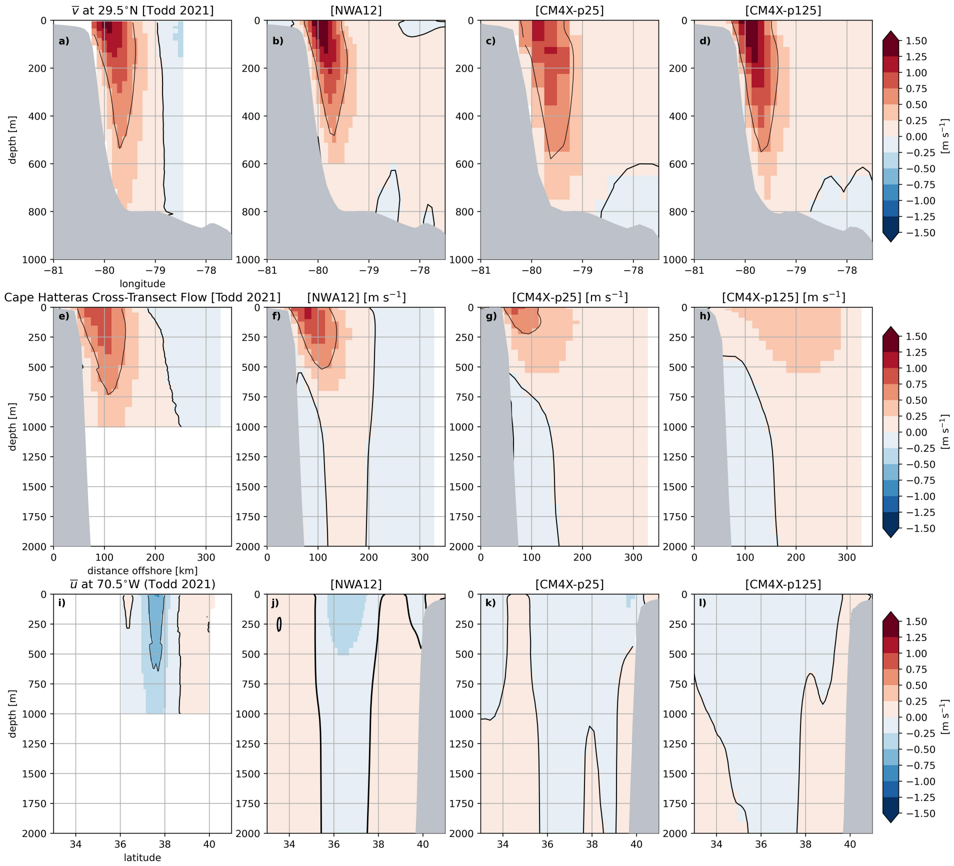

Figure 24. a-d) Time mean meridional velocity at 29.5°N as a function of depth and longitude from Todd (2021) (2015-202), NWA12 (2010-2014), CM4X-p25 (2010-2014), and CM4X-p125 (2010-2014) (note log color scale). Heavy black contour is the 0 value, light black solid/dashed contours are 0.5 and -0.5 m s^{-1}, and gray shading is the model bathymetry (NWA12 bathymetry is used in panel a). e-h) Time mean cross-transect velocity extending southwest of Cape Hatteras. i-l) Time mean zonal velocity at 70.5°W. Positive velocities are those into the page.

Surface winds control the meridional flow in open ocean gyres, while the return flow in western boundary currents is driven by bottom pressure torques (compare Figures 25a1-a2 with 25d1-d2, also see (Yeager, 2015)). However, in the Iceland basin and Nordic Seas, bottom friction and other vorticity terms are more important in guiding the meridional transport (not shown).

Comparing CM4X-p25 and CM4X-p125 reveals that refined ocean grid spacing leads to a stronger time-mean subtropical gyre transport, while the subpolar gyre transport is weaker. In CM4X-p25, the maximum transport in the subtropical gyre and minimum transport in the subpolar gyre are 25 Sv and -44 Sv, respectively (Figures 25a1, 25e).

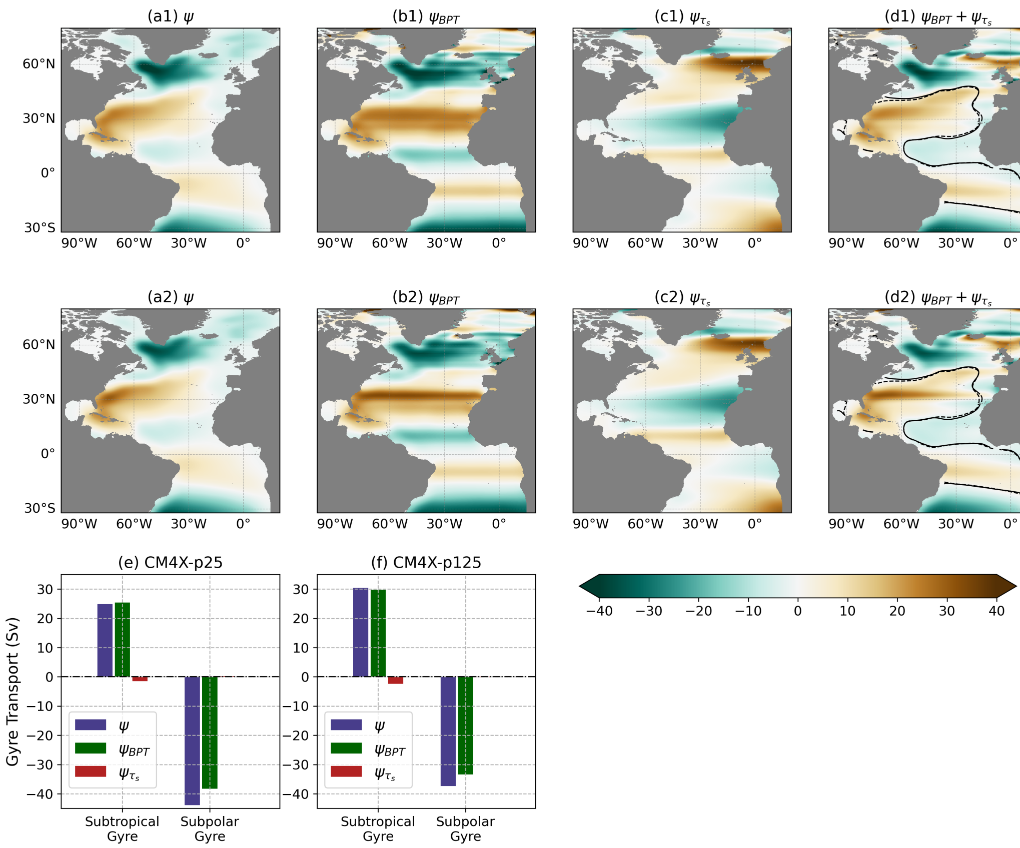

Figure 25. Barotropic streamfunction (in Sv) estimated using vorticity budget terms in equation (3) (30-year time-mean). The top row (panels a1-d1) show the CM4X-p25 piControl and the bottom row (panels a2-d2) show CM4X-p125 piControl. Panels (e,f) show the subtropical and subpolar gyre transports (estimated as the maximum or minimum of the barotropic streamfunction) and the associated contributions from ψ_{BPT} and ψ_{τ_s} (note that ψ_{τ_s} transports in the subpolar gyre are too small to see). Spatial maps in panels (d1-d2) are sums of maps shown in panels (b1-b2) and (c1-c2), and the solid (CM4X-p25) and dashed black (CM4X-p125) curves represent the zero streamfunction contour lines, which are used to infer the latitudinal position of the Gulf Stream located between the subtropical and subpolar gyres. To filter out the small-scale variability and grid-scale noise in vorticity diagnostics (Khatri et al., 2024), spatial maps are spatially smoothed to 10° resolution using the GCM-Filters from Loose et al. (2022).

On the other hand, in CM4X-p125, subtropical and subpolar gyre transports are 30 Sv and -37 Sv (Figures 25a2, 25f). These differences in gyre strength are primarily associated with differences in bottom pressure torque (compare Figures 25b1 and 25b2), which is affected by the finer scale bathymetry and better resolved mesoscale deep flows in CMX-p125 relative to CM4X-p25. These changes in gyre transports agree with changes in magnitudes of bottom pressure torques and the corresponding streamfunction transports, ψ_{BPT}.

Similar strengthening in the subtropical gyre with increased ocean model resolution has been observed in many climate models (Meccia et al., 2021). Moreover, the southward subpolar transport of 37 Sv in CM4X-p125 is close to estimates from Xu et al. (2013), who determined the southward transport in the Labrador Sea near 53°N to be approximately 39 Sv.

6.3 Atlantic meridional overturning circulation (AMOC)

Direct AMOC observations from OSNAP (Overturning in the Subpolar North Atlantic Program) over the recent periods (Lozier et al., 2019; Lim et al., 2019), as well as the reconstruction of the long-term mean AMOC over the past several decades using robust diagnostic calculations (R. Zhang & Thomas, 2021), show that the subpolar AMOC across the OSNAP section is dominated by its eastern component rather than the western component that extends across the Labrador Sea. Climate models often overestimate the maximum AMOC strength across OSNAP West (Thomas et al., 2015; Lim et al., 2019; Yeager et al., 2021). This model bias is likely related to the overestimation of the Labrador Sea open-ocean deep convection and its role in the AMOC across OSNAP West in relatively coarse models. As discussed in R. Zhang and Thomas (2021), too much open ocean Labrador Sea convection unrealistically overlaps with the less resolved boundary outflow from the Labrador Sea basin, thus leading to unrealistically high density in the less resolved boundary outflow from the Labrador Sea basin in relatively coarse models.

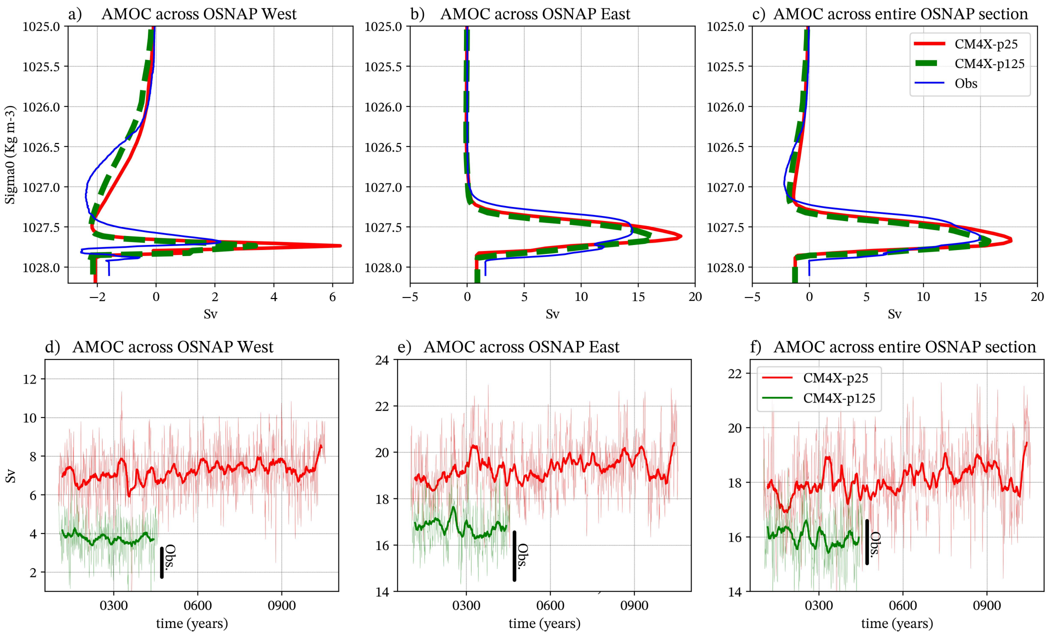

Figure 26. Top row: Climatological annual mean density space AMOC streamfunction across OSNAP West, OSNAP East, and the full OSNAP section for CM4X-p25, CM4X-p125, and the OSNAP observations over the period of 2014-2020. Bottom row: Time series of the maximum AMOC strength across OSNAP West, OSNAP East, and the full OSNAP section, in comparison with the OSNAP observations over the period of 2014-2020. The time series are smoothed with a 10-year running mean. The dark lines in bottom rows depict the ranges (mean +/- one standard deviation of rolling annual mean) of observations.

Figure 26 shows that CM4X-p125 has a maximum climatological density space AMOC strength across OSNAP West that has been reduced relative to CM4X-p25, with CM4X-p125 results closer to recent OSNAP observations of 2 Sv (Lozier et al., 2019; Li et al., 2021). The modeled maximum climatological density space AMOC strength across OSNAP East (or across the entire OSNAP section) has also been improved with CM4X-p125 relative to CM4X-p25, with Figure 26 showing CM4X-p125 values closer to the OSNAP observations. The modeled maximum climatological subpolar AMOC in density

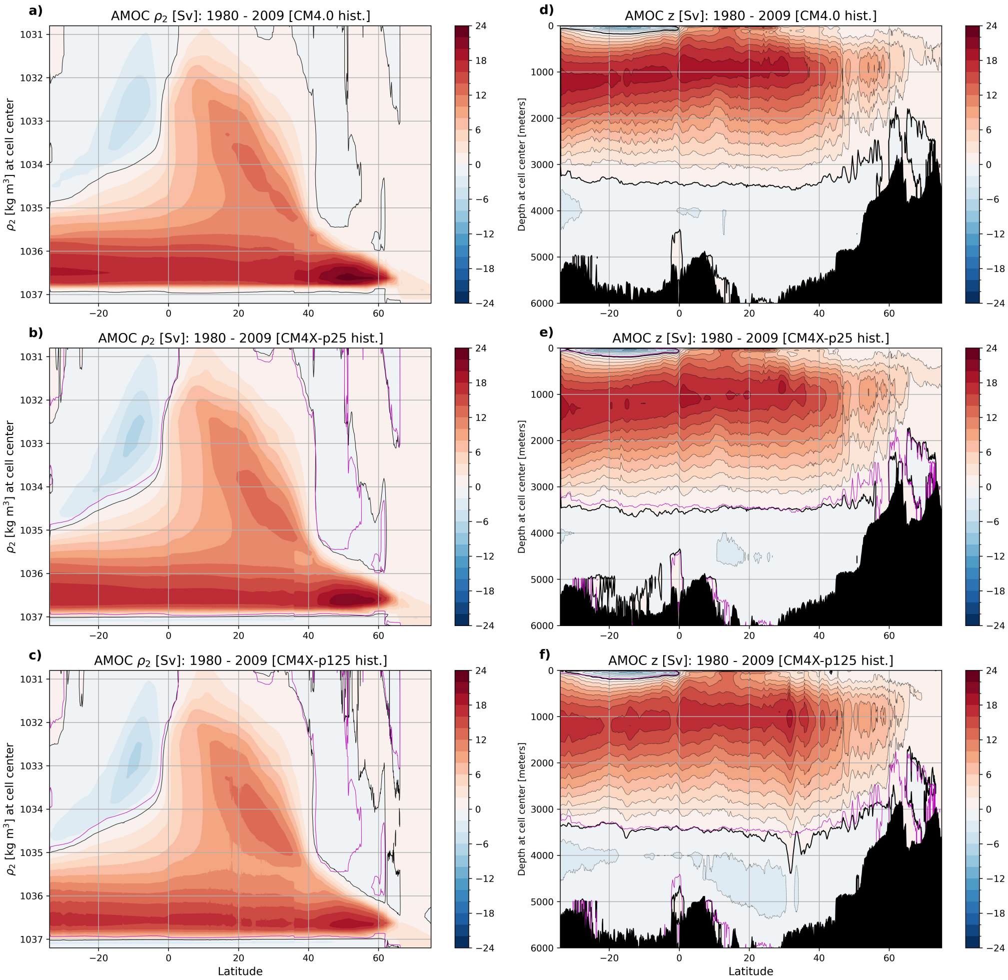

Figure 27. Climatological annual mean Atlantic meridional overturning circulation (AMOC) streamfunction as a function of latitude in potential density (ρ_2, a,b,c) and geopotential (d,e,f) for the historical simulations over the period of 1980-2009 using CM4.0, CM4X-p25, and CM4X-p125. In CM4X-p25 and CM4X-p125 panels, the magenta contour identifies the zero contour of CM4.0. Contours are shown at 2 Sv increments. The figure includes transport in the Atlantic, Arctic, Mediterranean, and Baltic basins.

space is reduced in CM4X-p125 compared to that in CM4X-p25 (Figure 27), which is consistent with the improvement across OSNAP West and East (Figure 26).

Since both CM4X-p125 and CM4X-p25 have the same atmosphere model, improvements in the CM4X-p125 mean state subpolar AMOC are likely due to its refined horizontal ocean grid spacing compared to CM4X-p25. Improvements in the simulated maximum AMOC strength across OSNAP West in CM4X-p125 (Figure 26a,d) is likely linked to its improved and reduced Labrador Sea winter deep convection area compared to that simulated in CM4X-p25, as reflected in the climatological winter mixed layer depth (MLD) seen in Figure 28. The Labrador Sea winter MLD in CM4X-p25 is unrealistically deep and broad, which is likely related to the under-representation of the dense overflows entering the Labrador Sea and the less resolved mesoscale eddy restratification (Tagklis et al., 2020). The winter Labrador Sea MLD is improved (shoaled) in CM4X-p125, which has a slightly better representation of the dense overflows across the OSNAP section as

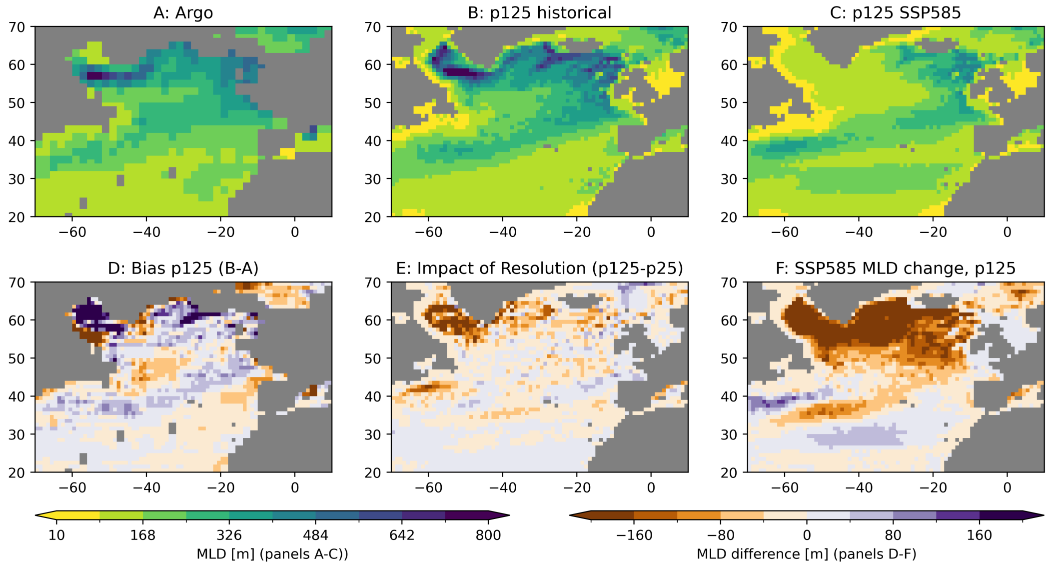

Figure 28. Winter climatology of the North Atlantic maximum mixed layer depth from Argo (years 2004-2023) (Argo, 2023), as well as years 2000-2014 (historical simulation) and years 2085-2099 (SSP5-8.5 simulation) from CM4X-p25 and CM4X-p125, computed as in Figure 9 from Part I (Griffies et al., 2024) Panel A: estimates from the Argo profiling floats; Panel B: results from CM4X-p125 historical experiment; Panel C: CM4X-p125 SSP5-8.5 simulation. Panel D shows the differences between CM4X-p125 and Argo (Panels B-A), whereas Panel E shows the difference between CM4X-p125 and CM4X-p25 (with a shoaling of MLD in CM4X-p125 relative to CM4X-p25). Finally, Panel F shows the impacts from the SSP5-8.5 climate change, showing years 2084-2099 minus years 2000-2014 (Panels C-B) from CM4X-p125. Note that differences documented in Panels D and E are robust to a longer time average over model years 1955-2014.

seen in Figure 29, as well as a stronger resolved mesoscale eddy restratification. However, neither CM4X-p125 nor CM4X-p25 simulate the observed winter deep convection in the Greenland Sea.

Over the subpolar North Atlantic, horizontal circulations across sloping isopycnals play an important role in density-space AMOC, and a Sigma-Z diagram was designed to illustrate the role of horizontal circulations in density-space (R. Zhang & Thomas, 2021). Across OSNAP East, both CM4X-p125 and CM4X-p25 simulate the horizontal circulation contribution (positive inflow and negative outflow canceled at the same depth level, but not at the same density level), which corresponds to that observed in the upper ocean (Figure 30). However, both models underestimate the horizontal circulation contribution compared to that observed in the deep ocean (Figure 30). This modeling bias in the deep ocean is related to the underestimation of the Nordic Sea overflows (i.e. Denmark Strait overflow and Iceland-Scotland overflow) and associated recirculation in both CM4X-p125 and CM4X-p25 (Figure 29). With refined horizontal grid spacing, CM4X-p125 has a slightly better representation of the Nordic Sea overflows and associated recirculation (Figure 29), and thus a slightly better representation of the horizontal circulation contribution in the deep ocean (Figure 30). Although the difference between CM4X simulations and the observation in the density space AMOC streamfunction across OSNAP East does not appear pronounced (Figure 26), the Sigma-Z diagram clearly reveals differences in the simulated deep ocean AMOC structure in Figure 30.

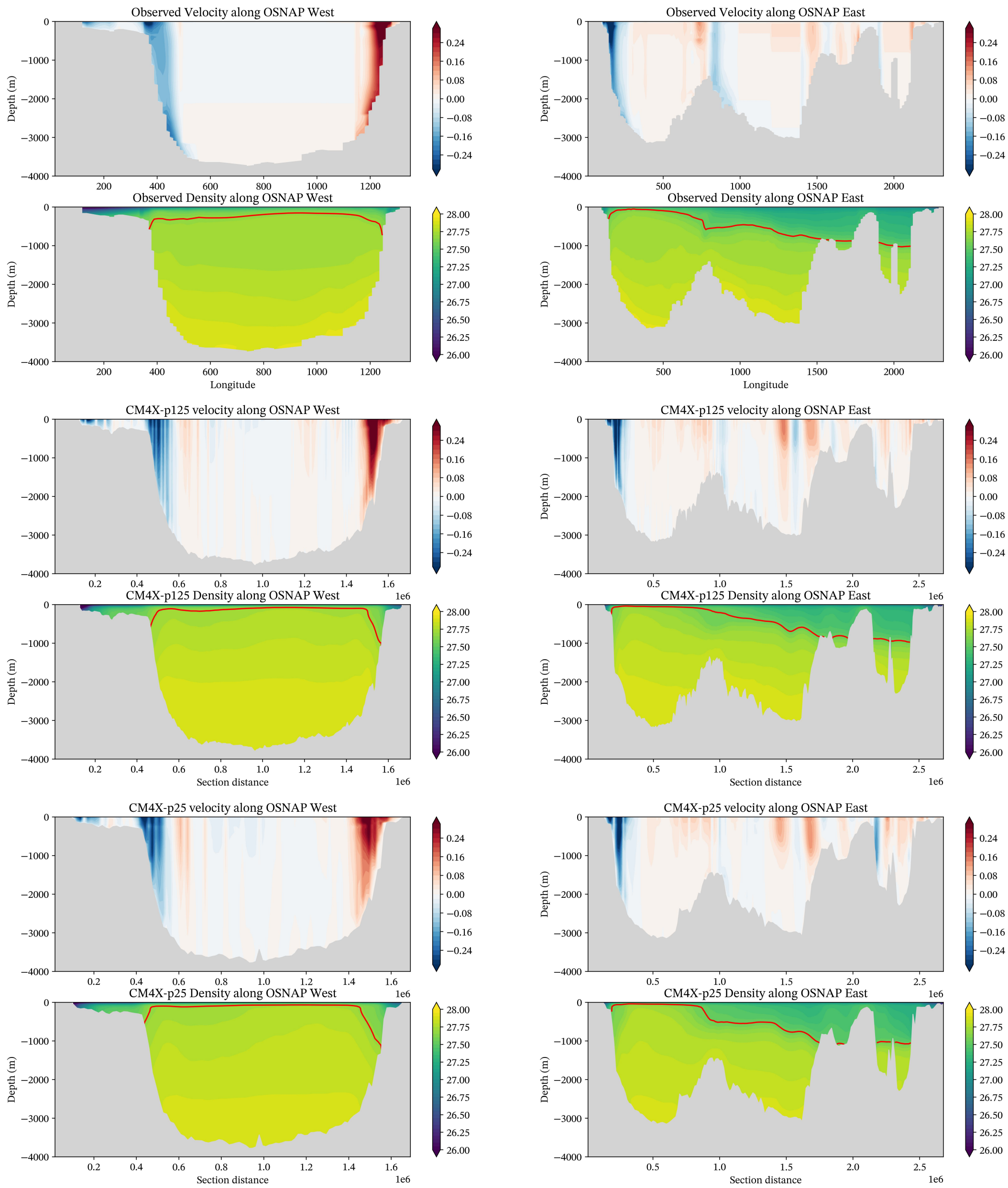

Figure 29. Mean velocity and potential density across OSNAP West (left) and OSNAP East (right). The top two rows show the OSNAP observations averaged over years 2014-2020 from Fu et al. (2023). The middle two row shows CM4X-p125 as averaged over years 2014-2020. The bottom two rows show CM4X-p25 as averaged over years 2014-2020.

Across OSNAP West, the horizontal circulation contribution to the density space AMOC in the upper ocean is too strong and too deep in CM4X-p25 compared to that estimated from observations (Figure 31). However, in CM4X-p125, the horizontal circulation contribution becomes less strong and shallower, which then compares better with the field measurements than CM4X-p25 (Figure 31). This improvement in CM4X-p125 is related to the improved (reduced) Labrador Sea winter deep convection area (Figure 28) and the better resolved Labrador Sea boundary current in CM4X-p125 (Figure 29). Consistently, the density contrast between the Labrador Sea boundary outflow and in-

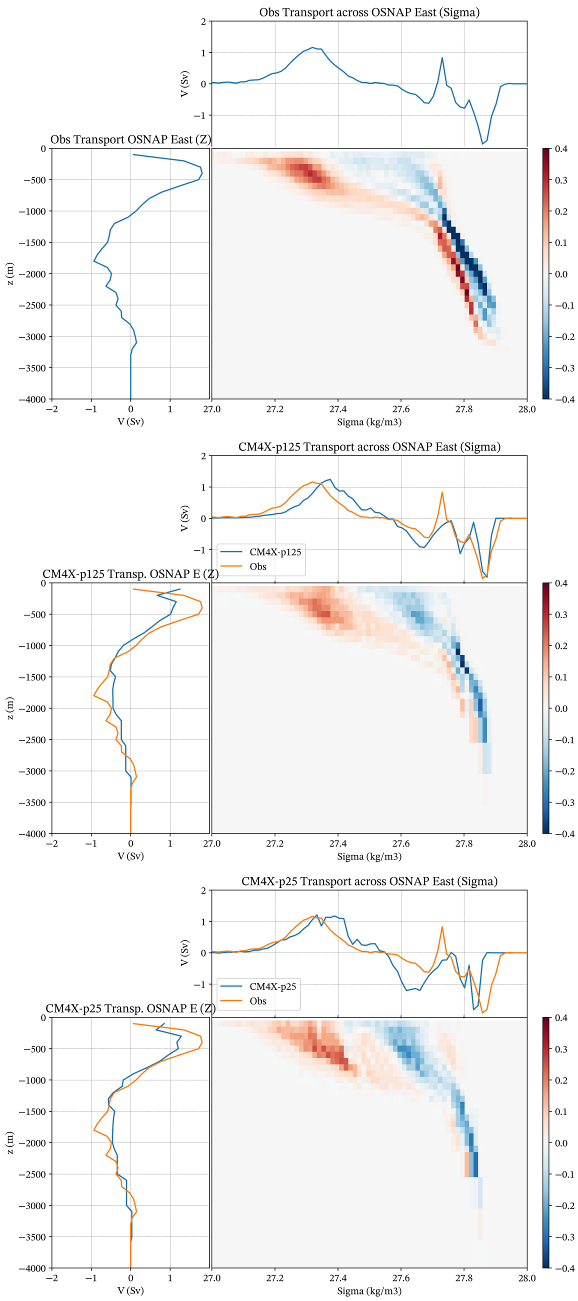

Figure 30. Sigma-z diagram of volume transport across OSNAP East for OSNAP observations (Fu et al., 2023) (top), CM4X-p125 (middle), and CM4X-p25 (bottom), all computed over years 2014-2020. The color shading in each panel shows the integrated volume transport (Sv) across each subsection over each potential density (σ_0; potential density referenced to 0 dbar) bin (x-axis) and depth (z) bin (y-axis); the integrated transport across each subsection over each potential density bin summed over the entire depth range is shown in the blue curve above; the integrated transport across each subsection at each depth bin summed over the entire potential density range is shown in the blue curve on the left. The line plots for the CM4X results have the OSNAP measurements plotted for easier comparison.

flow is reduced and thus the maximum density space AMOC across OSNAP West (mainly contributed by the horizontal circulation) becomes weaker in CM4X-p125 compared to that simulated in CM4X-p25 (Figure 26a,d). The observed density structure between the Labrador Sea boundary outflow and inflow is almost symmetric (Figure 29), consistent with a very weak horizontal circulation contribution (Figure 31) and thus a very weak maximum density space AMOC strength across OSNAP West (Figure 26a,d) in observations.

Across the RAPID section in the subtropical North Atlantic, both CM4X-p125 and CM4X-p25 simulate a similar maximum AMOC strength as that estimated from the RAPID program (Cunningham et al., 2007; McCarthy et al., 2015; Smeed et al., 2018), but the penetration depth of the simulated AMOC is much shallower than that observed (Figure 32). The simulated shallow AMOC across the RAPID section is a typical bias found in many models due to deficiencies representing the dense and deep-penetrating Nordic Sea overflows (Danabasoglu et al., 2010; R. Zhang et al., 2011; Danabasoglu et al., 2014; H. Wang et al., 2015). The penetration depth of the simulated AMOC across the RAPID section is slightly deeper in CM4X-p125 than that in CM4X-p25 (Figure 27d,e,f and Figure 32) due to slightly better representation of the Nordic Sea overflows in CM4X-p125.

The simulated multidecadal AMOC variability across the RAPID section in both CM4X-p125 and CM4X-p25 piControl simulations (Figure 33) is much weaker than the observationally-based estimate (Yan et al., 2018). The simulated historical AMOC across the RAPID section in both CM4X-p125 and CM4X-p25 has less pronounced multidecadal variations than simulated in CM4.0 (Figure 34). The pronounced multidecadal AMOC variations in CM4.0 (e.g., an increase up to 19 Sv and decline afterwards) are mainly induced by multidecadal changes in external radiative forcing (e.g. anthropogenic aerosols) as also found in many CMIP6 models (Menary et al., 2020), which are unrealistic and opposite to those indicated by the observed AMOC fingerprint (Yan et al., 2019) and inconsistent with further observational measures (Held et al., 2019; Menary et al., 2020; Robson et al., 2022). The improved historical AMOC changes simulated in CM4X are likely related to the refined horizontal atmospheric grid spacing used by CM4X (50 km) compared to that employed in CM4.0 (100 km) (see Appendix A1 in Part I (Griffies et al., 2024)). Previous studies have shown that climate models with coarse horizontal atmospheric grid spacing overestimate aerosol indirect effect (Donner et al., 2016; Sato et al., 2018; Zhao et al., 2018a). The impact of the horizontal atmospheric grid spacing on the aerosol indirect effect deserves more future investigation. The static vegetation approach (no land use change, no CO2 fertilization effect, and no demography change) employed in CM4X historical simulations might also contribute to the different simulated historical AMOC changes compared to those simulated in CM4.0 (again, see Appendix A2 in Part I (Griffies et al., 2024)). Additional experiments in future studies are needed to fully understand the improvement in simulated historical AMOC changes in CM4X. In Figure 34 we also show the AMOC strength for the SSP5-8.5 scenario experiment, whereby the AMOC weakens starting around year 2000 in CM4.0 and CM4X, with the year 2100 strength in each model less than half their pre-industrial strength.

6.4 Watermass transformation in the subpolar North Atlantic

Figure 35 shows the time-mean of the water mass budget (diagnosed following Drake et al. (2024)) in potential density coordinates,

$$\partial_t \mathcal{M}_{\geq} = \Psi_{\geq} + \mathcal{S}_{\geq} + \mathcal{G} \tag{4}$$

for two key regions of the Subpolar North Atlantic. $\mathcal{M}_{\geq} = \mathcal{M}_{\geq}(\sigma_2, t)$ is the mass of water denser than σ_2 at time t within a given region; $\Psi_{\geq}$ is the total transport into the region for waters denser than σ_2; $\mathcal{S}_{\geq}$ is direct addition of seawater mass from boundary fluxes; and $\mathcal{G}$ is the total water mass transformation rate (positive when it tends to increase the mass of denser water). We further decompose $\Psi_{\geq} = \sum \Psi_{\geq}^{(\text{boundary})}$ into con-

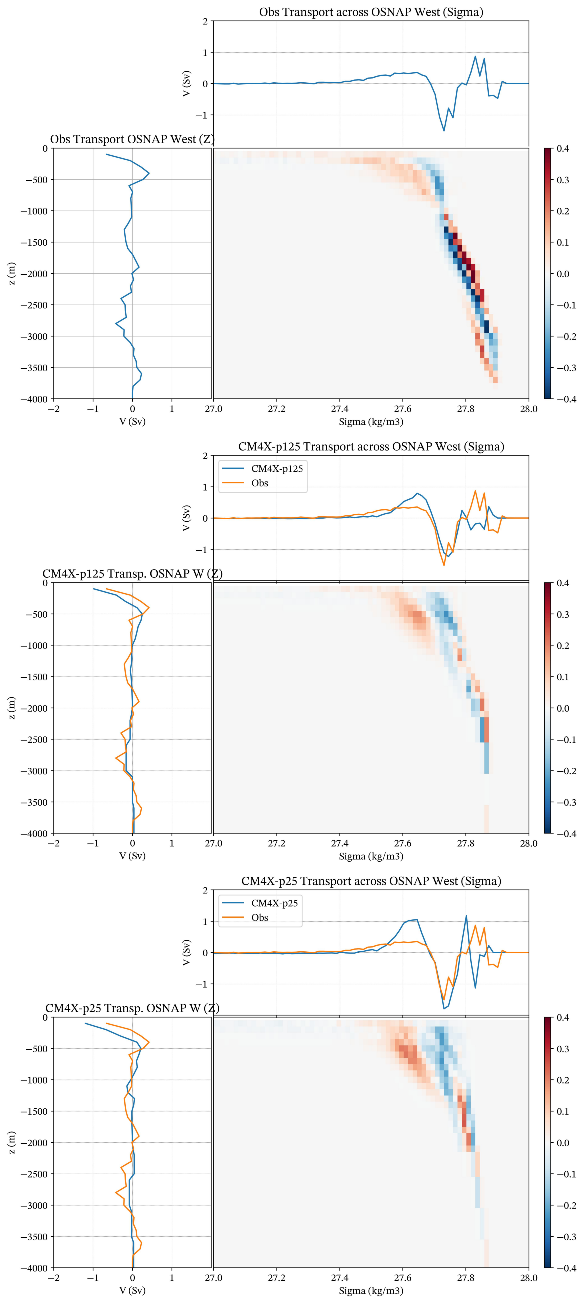

Figure 31. The Sigma-z diagram of volume transport across OSNAP West for OSNAP observations (Fu et al., 2023) (top), CM4X-p125 (middle), and CM4X-p25 (bottom), all computed over years 2014-2020. The color shading in each panel shows the integrated volume transport (Sv) across each subsection over each potential density (σ_0) bin (x-axis) and depth (z) bin (y-axis); the integrated transport across each subsection over each potential density bin summed over the entire depth range is shown in the blue curve above; the integrated transport across each subsection at each depth bin summed over the entire potential density range is shown in the blue curve on the left. The line plots for the CM4X results have the OSNAP measurements plotted for easier comparison.

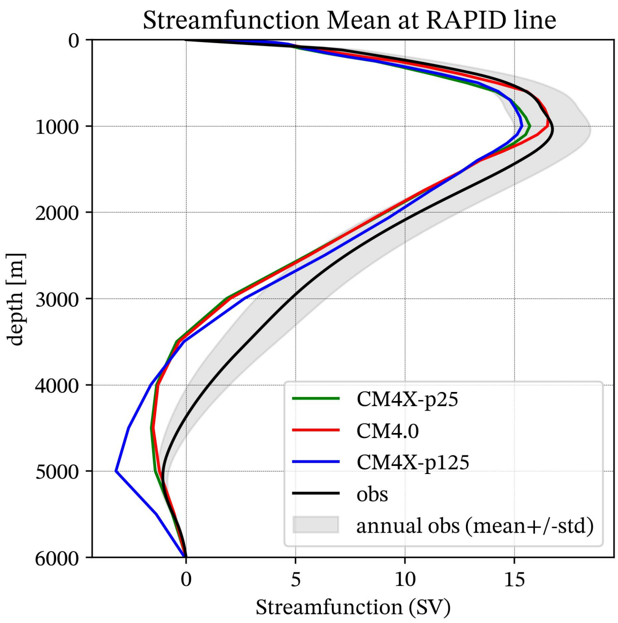

Figure 32. Climatological annual mean AMOC streamfunction strength across the RAPID section (26.5°N) for the models computed over years 2004-2022 using CM4.0, CM4X-p25, and CM4X-p125, in comparison with the RAPID observations over the period of 2004-2022.

tributions from different sections along the regions' boundaries and $\mathcal{G} \equiv \sum \mathcal{G}^{(\mathrm{process})}$ to distinguish contributions from boundary fluxes $\mathcal{G}^{(\mathrm{BF})}$ (primarily due to air-sea fluxes of heat, salt, and freshwater), parameterized mixing processes $\mathcal{G}^{(\mathrm{mix})}$, and spurious numerical mixing $\mathcal{G}^{(\mathrm{Spu})}$. All terms are diagnosed directly from mass, heat, and salt tendencies, except for the spurious numerical mixing $\mathcal{G}^{(\mathrm{Spu})}$, which we identify as the budget residual (see Drake et al. (2024) for details). Because we use monthly-mean budget tendencies that are binned online into depth coordinates, however, some error may be introduced by the omission of sub-monthly correlations between the tendencies and σ_2.

In the Labrador Sea (Figure 35a,b), there is a nearly perfect time-mean balance between surface-forced water mass transformations (peaking at $-7\,\mathrm{Sv}$ at $\sigma_2 = 36.6\,\mathrm{kg/m^3}$) and export across the OSNAP-West section. Contributions from other terms, such as the inflow of dense water through Davis Strait and local transformation by mixing processes, are negligible. In the Irminger Sea and Iceland Basin (Figure 35c,d), by contrast, the peak surface-forced water mass transformation is weaker ($10\,\mathrm{Sv}$) and occurs at a lighter density ($36.2\,\mathrm{kg/m^3}$) than what is exported southward across the OSNAP-East section, which peaks at $15\,\mathrm{Sv}$ for waters denser than $\sigma_2 = 36.4\,\mathrm{kg/m^3}$. The difference between the transformation and overturning transport is due to 1) relatively dense Nordic Sea waters overflowing across the Greenland Scotland Ridge (peaking at $36.6\,\mathrm{kg/m^3}$) and 2) spurious numerical entrainment of the lighter surface-forced waters (as opposed to parameterized entrainment, which is relatively weak). These model results are qualitatively consistent with the observation-based analysis of (Evans et al., 2023), in that the total surface-forced water mass transformation peaks at a lighter density than the overturning across OSNAP and that interior mixing processes play a non-negligible role in this transformation. In both cases, the differences between CM4Xp25 and CM4Xp125 are relatively modest: due to slightly less spurious numerical mixing, CM4Xp125 appears to export less deep water across OSNAP-East at the peak density but more at the highest densities.

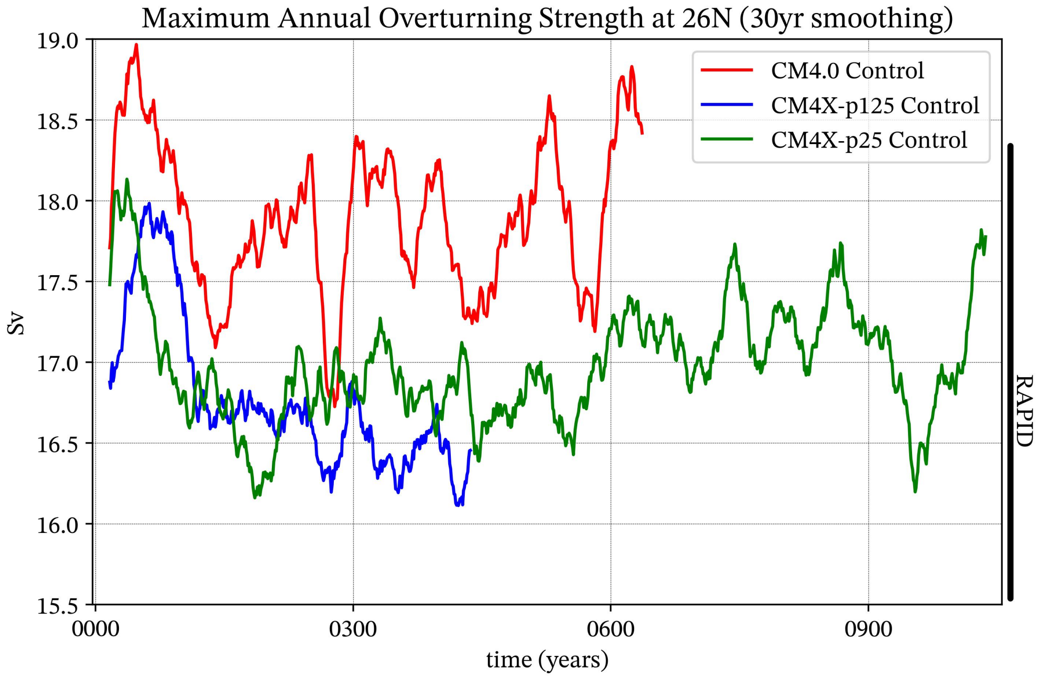

Figure 33. Time series of the maximum annual mean AMOC strength across the RAPID section (26.5°N) for the preindustrial control simulations using CM4.0, CM4X-25, and CM4X-p125. The time series are smoothed with a 30-year running mean. The dark line on the right depicts the range (mean +/- one standard deviation of annual mean) of the observed maximum annual mean AMOC strength across the RAPID section over the period of 2004-2022.

Figure 35e shows the spatial distribution of surface-forced water mass transformations at the isopycnal of peak export across the OSNAP arrays in each region. In the Irminger Sea and Iceland Basin, transformations are concentrated immediately downstream of the Denmark Strait, further downstream along the East Greenland Current, along the Reykjanes Ridge, and on the southern edge of the Iceland Faroe Ridge, qualitatively consistent with observation-based estimates (Petit et al., 2020). In the Labrador Sea, transformations are concentrated along the northwestern continental slope, also qualitatively consistent with observation-based estimates (Zou et al., 2024).

As discussed in Section 6.3, the deep water export across OSNAP-West is unrealistically strong in both CM4X-p25 and CM4X-p125. Our water mass analysis suggests this result is a direct result of too-high surface-forced water mass transformations, supporting the previous section's hypothesis that this is related to a bias in dense layer outcrop areas or mixed-layer depths. Furthermore, our analysis suggests that the light and shallow biases of the CM4X AMOC is in part attributable to spurious numerical mixing, which transforms a few Sv of inflowing dense Nordic waters (with $36.8\,\mathrm{kg/m^3} \leq \sigma_2 \leq 37.2\,\mathrm{kg/m^3}$) towards lighter density classes.

6.5 Conclusions regarding the North Atlantic simulations

While representation of the deep western boundary current is expected to play a role in altering Gulf Stream vertical structure, Ezer (2016), Schoonover et al. (2017) and Debreu et al. (2022) stress the importance of realistically resolving a smooth continental slope so to allow for realistic flow-topography interactions. Resolving flow-topography

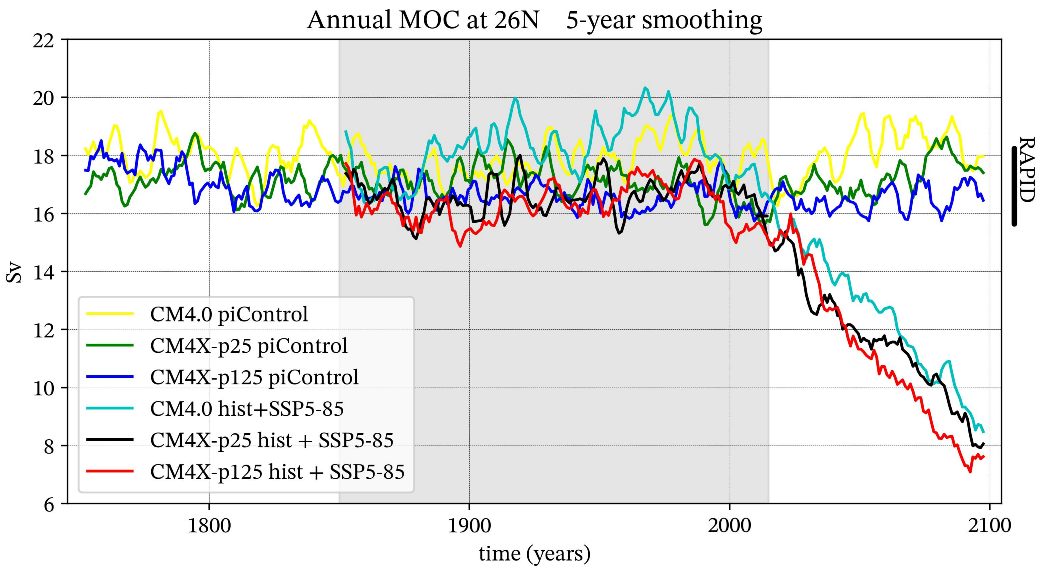

Figure 34. Time series from CM4.0, CM4X-p25, and CM4X-p125 of the maximum annual mean AMOC strength across the RAPID section (26.5°N) for the piControl, historical over years 1850-2014 (shaded region), and the SSP5-8.5 scenario experiment from 2014-2100. The time series are smoothed with a 5-year running mean. The dark line on the right depicts the range (mean +/- one standard deviation of annual mean) of the observed maximum annual mean AMOC strength across the RAPID section over the period of 2004-2022 for comparison.

interactions is a function of model horizontal and vertical grid spacing. Likely the result of improvements on many fronts, Chassignet and Xu (2017) conclude that a transition from unrealistic to realistic representation of the Gulf Stream and offshore mesoscale turbulence should be expected once reaching 1/12° and finer. Hence, we withhold expectation that the Gulf Stream in CM4X should match that in Todd (2021), with the slight improvements in CM4X-p125 relative to CM4X-p25 suggesting a trend in the right direction.

When compared to CM4.0, the CM4X simulations generally have a better representation of the maximum AMOC strength across both the subpolar and the subtropical sections, and multidecadal AMOC changes in the CM4X historical simulations are also improved. Comparing CM4X-p125 and CM4X-p25, the refined grid spacing in CM4X-p125 leads to a slightly better AMOC representation across the OSNAP section, especially across the OSNAP West subsection enclosing the Labrador Sea with improved (reduced) density contrast between the Labrador Sea boundary outflow and inflow. A linear relationship between subpolar gyre strength and the maximum subpolar overturning strength is generally observed in many climate models (Meccia et al., 2021). Thus, the relatively weaker subpolar gyre in CM4X-p125 compared to CM4X-p25 is consistent with the reduction in the maximum AMOC strength across the OSNAP section.

Many North Atlantic modeling biases in CM4.0 also exist in CM4X, such as the overestimation of winter deep convection in the Labrador Sea, the missing of winter deep convection in the Greenland Sea, the underestimation of the strength of the dense Nordic Sea overflows across the OSNAP section, the shallow AMOC depth, and the weak low-frequency AMOC variability in the control simulation. These long-standing modeling biases in the North Atlantic, which often appear in coupled climate models, are key areas for future improvements and more experiments are needed to address the processes

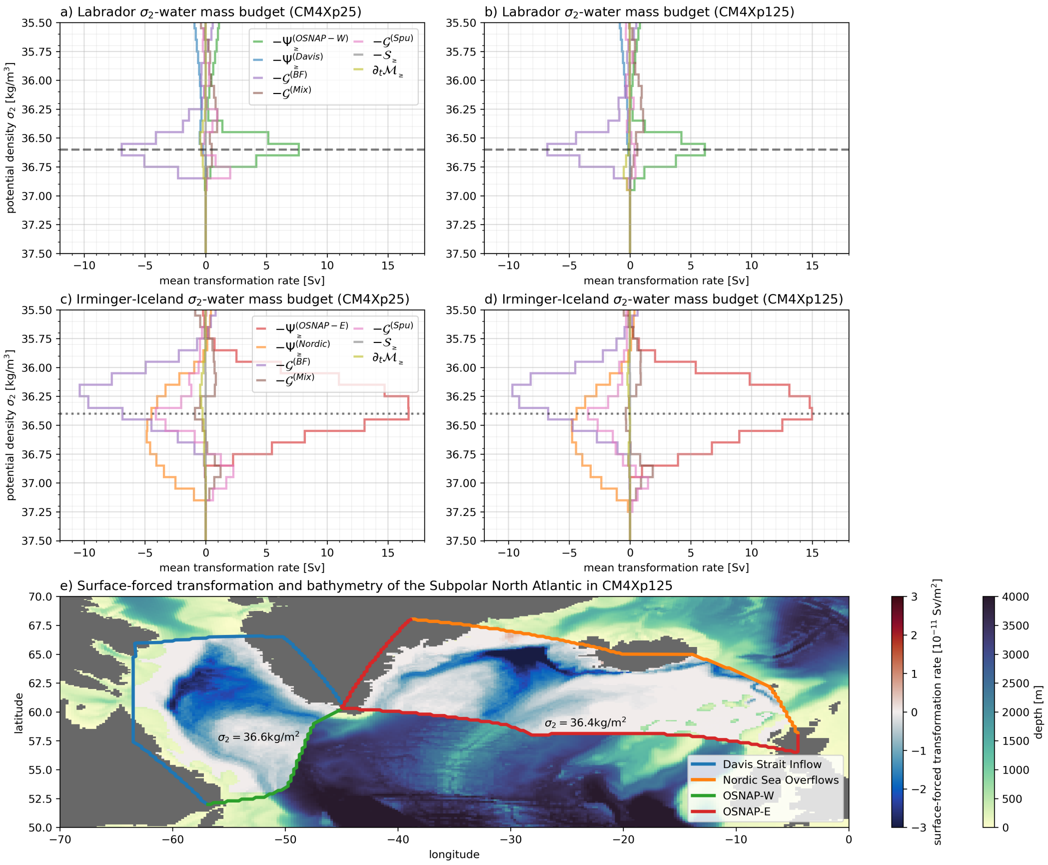

Figure 35. Water mass budgets in potential density (σ_2) coordinates in in CM4Xp25 (a,c) and CM4Xp125 (b,d,e) for a Labrador Sea region (a,b) and a Irminger Sea and Iceland Basin region (c,d). In panel (e), the Labrador region is bounded to the northwest by a Davis Strait section (blue) and to the southeast by the OSNAP-West section (green; see Fu et al. (2023)) while the Irminger-Iceland region is bounded to the northeast by the Greenland-Scotland Ridge (orange) and the to south by the OSNAP-East section (red). Within the two regions, colors represent the rates of water mass transformation per unit area across the isopycnal of peak transport across the OSNAP array (grey lines in a-d) due to boundary fluxes (i.e. the integrand of $\mathcal{G}^{(\mathrm{BF})}$, which is dominated by air-sea heat and freshwater fluxes)Outside of the two regions, colors represent seafloor depth. All terms in the water mass budget are computed following Drake et al. (2024) and represent the 2010-2024 time-mean of monthly-mean transformation rates diagnosed from the forced (historical + SSP5-8.5) scenarios. [Overturning transport values differ from those reported in Section 6.3 because we use different density coordinates, averaging intervals, and forcing scenarios.]

involved in reducing these biases. In particular, future improvements in the strength of the dense Nordic Sea overflows (including both the Denmark Strait overflow and the Iceland-Scotland overflow) across the subpolar section will enhance the vertical stratification in the Labrador Sea and thus reduce the Labrador Sea winter deep convection strength, as well as deepen the downstream AMOC. Future improvements in the Arctic dense water formation might also contribute to the improvement of the dense Nordic Sea overflows and the downstream AMOC.

7 Strategies for ocean climate model development

CM4X exemplifies the value of a hierarchy of coupled climate models where the only difference is the grid spacing. Such hierarchies provide a means to expose the importance, or lack thereof, for the enhanced representation of dynamical processes. Having two or more model configurations among a hierarchy provokes questions that go unasked with a single model. We propose that the expansion of model phase space to include carefully built hierarchies, such as the CM4X whose members represent a vigorous ocean mesoscale, is a useful, if not necessary, step to furthering the science going into climate models and the science emerging from simulations.

Although CM4X-p125 reaches thermal equilibrium in a remarkably short time (see the Part I analysis in Griffies et al. (2024)), the analysis in this paper revealed that there remain many familiar biases in need of being addressed in future development projects. Particular biases revealed in our analysis include: weak interannual variability in the tropics, as revealed by the skewness of the sea level (Figure 3 in Section 2.2); a poor representation of the eastern boundary upwelling zones, in part due to under-resolved ocean and atmosphere dynamics as well as the representation of low level clouds (Section 3); biases in the sea ice seasonal cycle, reflecting a number of possible process biases (Section 4); overly strong ventilation properties of mode and intermediate waters of the Southern Ocean, likely related to under-resolved mesoscale eddy processes at the high latitudes (Figure 9 of Part I); an overly diffuse Gulf Stream as it leaves the American coast, possibly due to over-dissipation (Sections 2.2 and 6.1); overly shallow overflows in the North Atlantic (Figure 27) and overly deep mixed layers in the Labrador Sea (Figure 28), likely due to under-resolved mesoscale processes and too much entrainment in the deep overflows. Even with these remaining issues, each of the biases are generally reduced in CM4X-p125 relative to CM4X-p25. From an oceanographic perspective, the key advantage of rapid thermal equilibration is that it allows one to examine ocean properties within a thermally equilibrated climate model, with those properties not having drastically drifted outside their physically sensible range. So although CM4X-p125 has many biases in need of further reduction, we propose that it provides a powerful venue for studying climate dynamics.

There are compelling arguments that, for purposes of centennial climate studies and climate change projections, the community should prioritize advances in numerical and computational methods to facilitate the direct simulation of the ocean mesoscale (Silvestri et al., 2024). The order of magnitude reduction in thermal equilibration time found in the CM4X-p125 simulation, relative to CM4X-p25, bolsters that argument. That is, the finer ocean grid used in CM4X-p125 has four times the number of grid points and uses half the time step; however, this factor of eight added expense for CM4X-p125 is offset by its factor of ten shorter thermal equilibration time. Advances signaled by CM4X-p125 were supported by advances in the numerical methods related to the MOM6 vertical Lagrangian dynamical core (Griffies et al., 2020). In particular, the quasi-isopycnal vertical coordinate used in the ocean interior greatly reduces spurious diapycnal mixing relative to a geopotential-like coordinate (Adcroft et al., 2019). Spurious numerical mixing is very difficult to minimize using standard Eulerian based numerical methods in the presence of a strong downscale cascade of tracer variance enabled by mesoscale eddies (Griffies et al., 2000). Developments leading to the MOM6 dynamical core were partly motivated to address this difficulty, particularly when confronted with the hundreds to thousands of mesoscale eddy turnover timescales accessed by climate simulations, thus allowing for seemingly small numerical errors to accumulate to have nontrivial detrimental impacts on stratification. Coupling advances in numerical methods to advances in computational hardware and software could render CM4X-p125 the coarsest, not the finest, member of a future climate model hierarchy.

8 Open Research

Software comprising the model as well as the software used for creating the figures will be placed on Zenodo at the revision stage of this work.

Observation-based datasets used in this paper are cited locally. We are indebted to the many efforts of the various programs providing observational-based data used to help evaluate these simulations, including the following.

- The Argo program provides data that were collected and made freely available by the International Argo Program and the national programs that contribute to it, with access available from

  http://www.argo.ucsd.edu and http://argo.jcommops.org

  The Argo Program is part of the Global Ocean Observing System.
- OSNAP data were collected and made freely available by the OSNAP (Overturning in the Subpolar North Atlantic Program) project and all the national programs that contribute to it (www.o-snap.org). The DOI for this data set is

  https://doi.org/10.35090/gatech/70342

Acknowledgments

This project started in May 2020, during the early stages of the Covid-19 pandemic shutdown. We are grateful to the GFDL computer operations team for keeping the computational resources reliable during the shutdown. We thank the GFDL management for providing the computer resources needed for the development and analysis documented here. We thank John Dunne and Matthew Harrison for very helpful comments on early drafts of this manuscript. Krista Dunne, Sergey Malyshev and Chris Milly kindly provided expertise in helping to update the rivers and lakes for use with the C192 atmosphere coupled to the p125 ocean. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce.

- The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce.
- A.A. was supported by Award NA18OAR4320123 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce.
- R.B. was supported under NSF Division of Polar Programs Grant NSF2319828.
- C.Y.C. was supported by Award NA19OAR4310365 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce.
- H.F.D. was supported by the NOAA Climate and Global Change Postdoctoral Fellowship Program, administered by UCAR's Cooperative Programs for the Advancement of Earth System Science (CPAESS) under Award NA18NWS4620043B.
- H.K. acknowledges the support from Natural Environment Research Council grants NE/T013494/1 and NE/W001543/1.
- M.L. was supported by award NA18OAR4320123 and NA23OAR4320198 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce.
- G.A.M was supported by NSF (PLR-1425989) and UKRI (MR/W013835/1).
- A.S. was supported by Schmidt Sciences, LLC under the M^{2}LInES project.
- K.E.T acknowledges support from the Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) Project under NSF Awards PLR-1425989 and OPP-1936222 and 2332379.
- L.Z. was supported by Schmidt Sciences, LLC under the M^{2}LInES project, NSF grant OCE 1912357 and NOAA CVP NA19OAR4310364.

- W.Z. was supported by the National Science Foundation under Grant Number F1240-01(NSF OCE 1912357). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Adcroft, A., Anderson, W., Blanton, C., Bushuk, M., Dufour, C. O., Dunne, J. P., ... Zhang, R. (2019). The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *Journal of Advances in Modeling the Earth System, JAMES*. doi: 10.1029/2019MS001726

Albert, A., Echevin, V., Lévy, M., & Aumont, O. (2010). Impact of nearshore wind stress curl on coastal circulation and primary productivity in the Peru upwelling system. *Journal of Geophysical Research: Oceans*, *115*(C12). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JC006569` doi: https://doi.org/10.1029/2010JC006569

Andrews, T., Gregory, J. M., Paynter, D., Silvers, L. G., Zhou, C., Mauritsen, T., ... Titchner, H. (2018). Accounting for changing temperature patterns increases historical estimates of climate sensitivity. *Geophysical Research Letters*, *45*, 8490–8499. doi: 10.1029/2018GL078887

Argo. (2023). *Argo float data and metadata from Global Data Assembly Centre (Argo GDAC)* [dataset]. SEANOE. Retrieved from `https://www.seanoe.org/data/00311/42182/` doi: 10.17882/42182

Armour, K. C., Bitz, C. M., & Roe, G. H. (2013). Time-varying climate sensitivity from regional feedbacks. *Journal of Climate*, *26*, 4518–4534. doi: 10.1175/JCLI-D-12-00544.1

Bakun, A., Black, B. A., Bograd, S. J., Garcia-Reyes, M., Miller, A. J., Rykaczewski, R. R., & Sydeman, W. J. (2015). Anticipated effects of climate change on coastal upwelling ecosystems. *Current Climate Change Reports*, *1*, 85–93.

Beadling, R. L. (2023). Global consequences of regional connectivity along the Antarctic margin. *Journal of Geophysical Research: Oceans*, *128*. doi: 10.1029/2023JC019908

Beadling, R. L., Krasting, J. P., Griffies, S. M., Hurlin, W. J., Bronselear, B., Russell, J. L., ... Winton, M. (2022). Importance of the Antarctic Slope Current in the Southern Ocean response to ice sheet melt and wind stress change. *Journal of Geophysical Research - Oceans*, *127*. doi: 10.1029/2021jc017608

Beadling, R. L., Russell, J. L., Stouffer, R. J., Mazloff, M., Talley, L. D., Goodman, P. J., ... Pandde, A. (2020). Representation of Southern Ocean properties across coupled model intercomparison project generations: CMIP3 to CMIP6. *Journal of Climate*, *33*, 6555–6581. doi: 10.1175/JCLI-D-19-0970.1

Bograd, S. J., Jacox, M. G., Hazen, E., Lovecchio, E., Montes, I., Buil, P. M., ... Rykaczewski, R. R. (2023). Climate change impacts on eastern boundary upwelling systems. *Annual Review of Marine Science*, *15*(1), 303-328. Retrieved from `https://doi.org/10.1146/annurev-marine-032122-021945` (PMID: 35850490) doi: 10.1146/annurev-marine-032122-021945

Bronselaer, B., Winton, M., Griffies, S. M., Hurlin, W. J., Rodgers, K. B., Sergienko, O. V., ... Russell, J. L. (2018). Change in future climate due to Antarctic meltwater. *Nature*, *564*(7734), 53.

Bushuk, M., Zhang, Y., Winton, M., Hurlin, B., Delworth, T., Lu, F., ... others (2022). Mechanisms of regional Arctic sea ice predictability in two dynamical seasonal forecast systems. *J. Climate*, *35*(13), 4207–4231.

Ceppi, P., Myers, T. A., Nowack, P., Wall, C. J., & Zelinka, M. D. (2024). Implications of a pervasive climate model bias for low-cloud feedback. *Geophysical Research Letters*, *51*. doi: 10.1029/2024GL110525

Chassignet, E. P., & Marshall, D. P. (2008). Gulf Stream separation in numerical ocean models. *Geophysical Monograph Series*, *177*. doi: https://doi.org/10.1029/177GM05

Chassignet, E. P., & Xu, X. (2017). Impact of horizontal resolution (1/12° to 1/50°) on Gulf Stream separation, penetration, and variability. *Journal of Physical Oceanography*, *47*(10), 1999-2021. doi: 10.1175/JPO-D-17-0031.1

Cheng, L., Trenberth, K. E., Fasullo, J., Boyer, T., Abraham, J., & Zhu, J. (2017). Improved estimates of ocean heat content from 1960 to 2015. *Science Advances*, *3*(3), e1601545. doi: 10.1126/sciadv.1601545

Cunningham, S. A., Kanzow, T., Rayner, D., Baringer, M. O., Johns, W. E., Marotzke, J., … Bryden, H. L. (2007). Temporal variability of the Atlantic meridional overturning circulation at 26.5° n. *Science*, *317*. doi: 10.1126/science.1141304

Danabasoglu, G., Large, W., & Briegleb, B. (2010). Climate impacts of parameterized Nordic Sea overflows. *Journal of Geophysical Research*, *115, C11005*. doi: 10.1029/2010JC006243

Danabasoglu, G., Yeager, S. G., Bailey, D., Behrens, E., Bentsen, M., Bi, D., … Wang, Q. (2014). North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part I: Mean states. *Ocean Modelling*, *73*(0), 76–107. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1463500313001868` doi: 10.1016/j.ocemod.2013.10.005

Dangendorf, S., Hay, C., Calafat, F. M., Marcos, M., Piecuch, C. G., Berk, K., & Jensen, J. (2019). Steric sea level rise in GFDL CM4 and ESM4: Insights into model drift and water mass representation. *Nature Climate Change*, *9*, 705–710. doi: 10.1038/s41558-019-0531-8

Debreu, L., Kevlahan, N., & Marchesiello, P. (2022). Improved Gulf Stream separation through Brinkman penalization. *Ocean Modeling*, *179*. doi: https://doi.org/10.1016/j.ocemod.2022.102121

Delworth, T. L., Broccoli, A. J., Rosati, A., Stouffer, R. J., Balaji, V., Beesley, J. A., … Zhang, R. (2006). GFDL's CM2 global coupled climate models - Part 1: Formulation and simulation characteristics. *Journal of Climate*, *19*, 643–674. doi: 10.1175/JCLI3629.1

DeRepentigny, P., Jahn, A., Holland, M. M., & Smith, A. (2020). Arctic sea ice in two configurations of the CESM2 during the 20th and 21st centuries. *J. Geophys. Res.: Oceans*, *125*(9), e2020JC016133.

DeVries, T., Le Quéré, C., Berthet, S., Hauck, J., Ilyina, T., Landschützer, P., … Séférian, R. (2019). Decadal trends in the ocean carbon sink. *Proceedings of the National Academy of Sciences*, *116*, 11646–11651. doi: 10.1073/pnas.1900371116

Donner, L. J., O'Brien, T. A., Rieger, D., Vogel, B., & Cooke, W. F. (2016). Are atmospheric updrafts a key to unlocking climate forcing and sensitivity? *Atmospheric Chemistry and Physics*, *16*. doi: 10.5194/acp-16-12983-2016

Donohue, K. A., Tracey, K. L., Watts, D. R., Chidichimo, M. P., & Chereskin, T. K. (2016). Mean Antarctic Circumpolar Current transport measured in Drake Passage. *Geophysical Research Letters*, *43*, 11,760–11,767. doi: 10.1002/2016GL070319

Drake, H. F., Bailey, S., Dussin, R., Griffies, S. M., Krasting, J. P., MacGilchrist, G. A., … Zika, J. D. (2024, April). *Water mass transformation budgets in finite-volume generalized vertical coordinate ocean models.* Retrieved 2024-05-31, from `https://essopenarchive.org/users/242115/articles/788549-water-mass-transformation-budgets-in-finite-volume-generalized-vertical-coordinate-ocean-models?commit=e6178873ac99d67c98963b54ba60699d57b7703f` doi: 10.22541/essoar.171284935.57181910/v1

Dunne, J. P., Horowitz, L. W., Adcroft, A., Ginoux, P., Held, I. M., John, J. G., … Zhao, M. (2020). The GFDL Earth system model version 4.1 (GFDL-ESM4.1):

Model description and simulation characteristics. *Journal of Advances in Modeling Earth Systems*, *12*. doi: 10.1029/2019MS002015

Dunne, J. P., John, J. G., Hallberg, R. W., Griffies, S. M., Shevliakova, E. N., Stouffer, R. J., . . . Zadeh, N. (2012). GFDL's ESM2 global coupled climate-carbon Earth System Models Part I: Physical formulation and baseline simulation characteristics. *Journal of Climate*, *25*, 6646-–6665. doi: 10.1175/JCLI-D-11-00560.1

Enriquez, A. G., & Friehe, C. A. (1995). Effects of wind stress and wind stress curl variability on coastal upwelling. *Journal of Physical Oceanography*, *25*(7), 1651 - 1671. Retrieved from `https://journals.ametsoc.org/view/journals/phoc/25/7/1520-0485_1995_025_1651_eowsaw_2_0_co_2.xml` doi: https://doi.org/10.1175/1520-0485(1995)025⟨1651:EOWSAW⟩2.0.CO;2

Evans, D. G., Holliday, N. P., Bacon, S., & Le Bras, I. (2023, June). Mixing and air–sea buoyancy fluxes set the time-mean overturning circulation in the subpolar North Atlantic and Nordic Seas. *Ocean Science*, *19*(3), 745–768. Retrieved 2023-09-18, from `https://os.copernicus.org/articles/19/745/2023/` (Publisher: Copernicus GmbH) doi: 10.5194/os-19-745-2023

Eyring, V., Bony, S., Meehl, J. A., Senior, C., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organisation. *Geoscientific Model Development*, *9*, 1937–1958. Retrieved from `http://www.geosci-model-dev.net/9/1937/2016/` doi: 10.5194/gmd-9-1937-2016

Ezer, T. (2016). Revisiting the problem of the Gulf Stream separation: on the represenation of topography in ocean models with different types of vertical grids. *Ocean Modelling*, *104*. doi: https://doi.org/10.1016/j.ocemod.2016.05.008

Feltham, D. L. (2008). Sea ice rheology. *Annu. Rev. Fluid Mech.*, *40*, 91–112.

FRAM Group. (1991). An eddy-resolving model of the Southern Ocean. *EOS Transactions of the American Geophysical Union*, *72*, 169–175. doi: 10.1029/90EO00128

Fraser, A., Wongpan, P., Langhorne, P., Klekociuk, A., Kusahara, K., Lannuzel, D., . . . others (2023). Antarctic landfast sea ice: A review of its physics, biogeochemistry and ecology. *Reviews of Geophysics*, *61*(2), e2022RG000770.

Frederikse, T., Landerer, F., Caron, L., Adhikari, S., Parkes, D., Humphrey, V. W., . . . Wu, Y.-H. (2020). The causes of sea-level rise since 1900. *Nature*, *584*(7821), 393–397. doi: 10.1038/s41586-020-2591-3

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., & Winton, M. (2015). Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *Journal of Climate*, *28*, 862-886. doi: 10.1175/JCLI-D-14-00117.1

Fu, Y., Lozier, M. S., Biló, T. C., Bower, A. S., Cunningham, S. A., Cyr, F., . . . Yashayaev, I. (2023). Seasonality of the meridional overturning circulation in the subpolar north atlantic [dataset]. *Communications Earth & Environment*, *4*(1), 181.

Gent, P. R., Yeager, S. G., Neale, R. B., Levis, S., & Bailey, D. A. (2010). Improvements in a half degree atmosphere/land version of the ccsm. *Climate Dynamics*, *34*, 819–833.

Goddard, P., Dufour, C., Yin, J., Griffies, S. M., & Winton, M. (2017). CO2-induced ocean warming of the Antarctic continental shelf in an eddying global climate model. *Journal of Geophysical Research - Oceans*, *122*. doi: 10.1002/2017JC012849

Gordon, C., Rosati, A., & Gudgel, R. (2000). Tropical sensitivity of a coupled model to specified isccp low clouds. *Journal of Climate*, *13*(13), 2239–2260.

Griffies, S. M., Adcroft, A., Alberty, M. S., Beadling, R. L., Bushuk, M., Chang, C.-Y., . . . Zika, J. D. (2024). The GFDL-CM4X climate model hierarchy, Part I: thermal properties and the mesoscale dominance hypothesis. *Journal of*

Advances in Modeling Earth Systems.
Griffies, S. M., Adcroft, A., & Hallberg, R. W. (2020). A primer on the vertical Lagrangian-remap method in ocean models based on finite volume generalized vertical coordinates. *Journal of Advances in Modeling Earth Systems*, *12*. doi: 10.1029/2019MS001954
Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Böning, C. W., ... Yeager, S. G. (2016). OMIP contribution to CMIP6: experimental and diagnostic protocol for the physical component of the ocean model intercomparison project. *Geoscientific Model Development*, *9*, 3231–3296. doi: 10.5194/gmd-9-3231-2016
Griffies, S. M., Pacanowski, R. C., & Hallberg, R. W. (2000). Spurious diapycnal mixing associated with advection in a *z*-coordinate ocean model. *Monthly Weather Review*, *128*, 538–564. doi: 10.1175/1520-0493(2000)128⟨0538:SDMAWA⟩2.0.CO;2
Griffies, S. M., Winton, M., Anderson, W. G., Benson, R., Delworth, T. L., Dufour, C., ... Zhang, R. (2015). Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *Journal of Climate*, *28*, 952–977. doi: 10.1175/JCLI-D-14-00353.1
Griffies, S. M., Winton, M., Donner, L. J., Horowitz, L. W., Downes, S. M., Farneti, R., ... others (2011). The GFDL CM3 coupled climate model: Characteristics of the ocean and sea ice simulations. *J. Climate*, *24*(13), 3520–3544.
Hallberg, R. W. (2013). Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects. *Ocean Modelling*, *72*, 92-103. doi: 10.1016/j.ocemod.2013.08.007
Hallberg, R. W., Adcroft, A. J., Dunne, J. P., Krasting, J. P., & Stouffer, R. (2013). Sensitivity of twenty-first-century global-mean steric sea level rise to ocean model formulation. *Journal of Climate*, *26*, 2947-–2956. doi: 10.1175/JCLI-D-12-00506.1
Hallberg, R. W., & Gnanadesikan, A. (2006). On the role of eddies in determining the structure and response of the wind-driven southern hemisphere overturning: Results from the Modeling Eddies in the Southern Ocean (MESO) project. *Journal of Physical Oceanography*, *36*, 2232–2252. doi: 10.1175/JPO2980.1
Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., ... Zadeh, N. (2019). Structure and performance of GFDL's CM4.0 climate model. *Journal of Advances in Modeling the Earth System*, *11*, 3691–3726. doi: 10.1029/2019MS001829
Heuzé, C. (2021). Antarctic Bottom Water and North Atlantic Deep Water in CMIP6 models. *Ocean Science*, *17*. doi: 10.5194/os-17-59-2021
Holland, W. R. (1967). On the wind-driven circulation in an ocean with bottom topography. *Tellus*, *19*(4), 582–600. doi: 10.3402/tellusa.v19i4.9825
Horvat, C. (2021). Marginal ice zone fraction benchmarks sea ice and climate model skill. *Nature Communications*, *12*(1), 2221.
Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., ... Zhang, H.-M. (2020). Improvements of the daily optimum interpolation sea surface temperature (DOISST) Version 2.1. *Journal of Climate*, *34*, 2923–2939. doi: 10.1175/JCLI-D-20-0166.1
Hughes, C. W., & De Cuevas, B. A. (2001). Why western boundary currents in realistic oceans are inviscid: A link between form stress and bottom pressure torques. *Journal of Physical Oceanography*, *31*(10), 2871–2885. doi: 10.1175/1520-0485(2001)031⟨2871:WWBCIR⟩2.0.CO;2
Hughes, C. W., Thompson, A. F., & Wilson, C. (2010). Identification of jets and mixing barriers from sea level and vorticity measurements using simple statistics. *Ocean Modelling*, *32*, 44–57. doi: 10.1016/j.ocemod.2009.10.004
Huneke, W. G. C., Morrison, A. K., & Hogg, A. M. (2022). Spatial and subannual

variability of the Antarctic Slope Current in an eddying ocean sea ice model. *Journal of Physical Oceanograph*, *52*, 347–361. doi: 10.1175/jpo-d-21-0143.1

Hurlburt, H. E., & Hogan, P. J. (2000). Impact of 1/8 to 1/64 resolution on Gulf Stream model-data comparisons in basin-scale subtropical atlantic ocean models. *Dynamics of Atmospheres and Oceans*, *32*. doi: https://doi.org/10.1016/S0377-0265(00)00050-6

Hutter, N., Losch, M., & Menemenlis, D. (2018). Scaling properties of Arctic sea ice deformation in a high-resolution viscous-plastic sea ice model and in satellite observations. *Journal of Geophysical Research: Oceans*, *123*(1), 672–687.

Ishii, M., Fukuda, Y., Hirahara, S., Yasui, S., Suzuki, T., & Sato, K. (2017). Accuracy of global upper ocean heat content estimation expected from present observational data sets. *SOLA*, *13*(0), 163–167. doi: 10.2151/sola.2017-030

Jacobs, S. S. (2004). Bottom water production and its links with the thermohaline circulation. *Antarctic Science*, *16*(4), 427–437. doi: 10.1017/S095410200400224X

Jacox, M. G., & Edwards, C. A. (2012). Upwelling source depth in the presence of nearshore wind stress curl. *Journal of Geophysical Research: Oceans*, *117*(C5). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JC007856` doi: https://doi.org/10.1029/2011JC007856

Jahn, A., Holland, M. M., & Kay, J. E. (2024). Projections of an ice-free Arctic Ocean. *Nature Reviews Earth & Environment*, 1–13.

Jahn, A., Kay, J. E., Holland, M. M., & Hall, D. M. (2016). How predictable is the timing of a summer ice-free Arctic? *Geophys. Res. Lett.*, *43*(17), 9113–9120.

Jevrejeva, S., H, P., & Jackson, L. P. (2021). Global mean thermosteric sea level projections by 2100 in CMIP6 climate models. *Environmental Research Letters*, *16*(1), 705–710. doi: 10.1088/1748-9326/abceea

Junker, T., Schmidt, M., & Mohrholz, V. (2015). The relation of wind stress curl and meridional transport in the Benguela upwelling system. *Journal of Marine Systems*, *143*, 1-6. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0924796314002425` doi: https://doi.org/10.1016/j.jmarsys.2014.10.006

Kaspar, F., Hollmann, R., Lockhoff, M., Karlsson, K.-G., Dybbroe, A., Fuchs, P., ... Schulz, J. (2009). Operational generation of AVHRR-based cloud products for Europe and the Arctic at EUMETSAT's Satellite Application Facility on Climate Monitoring (CM-SAF). *Advances in Science and Research*, *3*(1), 45–51.

Kay, J. E., Wall, C., Yettella, V., Medeiros, B., Hannay, C., Caldwell, P., & Bitz, C. (2016). Global climate impacts of fixing the Southern Ocean shortwave radiation bias in the Community Earth System Model (CESM). *Journal of Climate*, *29*(12), 4617–4636.

Keen, A., Blockley, E., Bailey, D., Debernard, J. B., Bushuk, M., Delhaye, S., ... Wyser, K. (2021). An inter-comparison of the mass budget of the Arctic sea ice in CMIP6 models. *The Cryosphere*, *15*(2), 951–982.

Khatri, H., Griffies, S. M., Storer, B. A., Buzzicotti, M., Aluie, H., Sonnewald, M., ... Shao, A. (2024). A scale-dependent analysis of the barotropic vorticity budget in a global ocean simulation. *Journal of Advances in Modeling Earth Systems*, *16*(6), e2023MS003813. doi: 10.1029/2023MS003813

Krasting, J. P., Griffies, S. M., Tesdal, J.-E., MacGilchrist, G., Beadling, R. L., & Little, C. M. (2024). Steric sea level rise and relationships with model drift and water mass representation in GFDL CM4 and ESM4. *Journal of Climate*, *accepted*.

LaCasce, J. (2017). The prevalence of oceanic surface modes. *Geophysical Research Letters*, *44*, 11,097–11,105. doi: 10.1002/2017GL075430

Landy, J. C., Dawson, G. J., Tsamados, M., Bushuk, M., Stroeve, J. C., Howell, S. E., ... others (2022). A year-round satellite sea-ice thickness record from

CryoSat-2. *Nature*, *609*(7927), 517–522.

Lavergne, T., Eastwood, S., Teffah, Z., Schyberg, H., & Breivik, L.-A. (2010). Sea ice motion from low-resolution satellite sensors: An alternative method and its validation in the Arctic. *J. Geophys. Res.: Oceans*, *115*(C10).

Lellouche, J.-M., Greiner, E., Bourdallé-Badie, R., Garric, G., Melet, A., Drévillon, M., … Le Traon, P.-Y. (2021). The Copernicus global 1/12° oceanic and sea ice GLORYS12 reanalysis. *Frontiers in Earth Science*, *9*. doi: 10.3389/feart.2021.698876

Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., Locarnini, R. A., … Zweng, M. M. (2012, 05). World ocean heat content and thermosteric sea level change (0-2000 m), 1955-2010. *Geophysical Research Letters*, *39*(10), n/a – n/a. doi: 10.1029/2012gl051106

Li, F., Lozier, M. S., Bacon, S., Bower, A. S., Cunningham, S. A., de Jong, M. F., … Zhou, C. (2021). Subpolar North Atlantic western boundary density anomalies and the meridional overturning circulation. *Nature Communications*, *12*. doi: 10.1038/s41467-021-23350-2

Lim, F., Lozier, M. S., Danagasoglu, G., Holliday, N. P., Kwon, Y.-O., Romanou, A., … Zhang, R. (2019). Local and downstream relationships between Labrador Sea Water volume and North Atlantic meridional overturning circulation variability. *Journal of Climate*, *32*, 3883-3898. doi: 10.1175/JCLI-D-18-0735.1

Lockwood, J. W., Griffies, C. O. D. S. M., & Winton, M. (2021). On the role of the Antarctic Slope Front on the occurrence of the Weddell Sea polynya under climate change. *Journal of Climate*, *34*, 2529–2548. doi: 10.1175/JCLI-D-20-0069.1

Loose, N., Abernathey, R., Grooms, I., Busecke, J., Guillaumin, A., Yankovsky, E., … others (2022). GCM-filters: A Python package for diffusion-based spatial filtering of gridded data. *Journal of Open Source Software*, *7*(70).

Lozier, M. S., Li, F., Bacon, S., Bahr, F., Bower, A. S., Cunningham, S. A., … Zhao, J. (2019). A sea change in our view of overturning in the subpolar North Atlantic. *Science*, *363*(6426), 516-521. Retrieved from `https://www.science.org/doi/abs/10.1126/science.aau6592` doi: 10.1126/science.aau6592

McCarthy, G. D., Smeed, D. A., Johns, W. E., Frajka-Williams, E., Moat, B. I., Rayner, D., … Bryden, H. L. (2015). Measuring the Atlantic meridional overturning circulation at 26°N. *Progress in Oceanography*, *130*, 91-111. doi: 10.1016/j.pocean.2014.10.006

Meccia, V. L., Iovino, D., & Bellucci, A. (2021). North Atlantic gyre circulation in PRIMAVERA models. *Climate Dynamics*, *56*(11), 4075–4090. doi: 10.1007/s00382-021-05686-z

Meehl, G. A., Arblaster, J. M., Bitz, C. M., Chung, C. T., & Teng, H. (2016). Antarctic sea-ice expansion between 2000 and 2014 driven by tropical Pacific decadal climate variability. *Nat. Geosci.*, *9*(8), 590–595.

Meier, W. N., Fetterer, F., Windnagel, A. K., & Stewart, S. (2021). *NOAA/NSIDC climate data record of passive microwave sea ice concentration, version 4.* National Snow and Ice Data Center. Retrieved from `https://nsidc.org/data/g02202` doi: 10.7265/efmz-2t65

Menary, M. B., Robson, J., Allan, R. P., Booth, B. B. B., Cassou, C., Gastineau, G., … Zhang, R. (2020). Aerosol-forced AMOC changes in CMIP6 historical simulations. *Geophysical Research Letters*, *47*. doi: 10.1029/2020GL088166

Moorman, R., Morrison, A. K., & Hogg, A. M. C. (2020). Thermal responses to Antarctic ice shelf melt in an eddy-rich global ocean-sea ice model. *Journal of Climate*, *33*. doi: 0.1175/jcli-d-19-0846.1

Morrison, A. K., Waugh, D. W., Hogg, A. M., Jones, D. C., & Abernathey, R. P. (2022). Ventilation of the Southern Ocean pycnocline. *Annual Review of Marine Science*, *14*, 15.1–15.26. doi: 10.1146/annurev-marine-010419-011012

Paolo, F. S., Fricker, H. A., & Padman, L. (2015). Volume loss from Antarctic ice shelves is accelerating. *Science*, *348*, 327–331. doi: 10.1126/science.aaa0940

Parsons, A. (2006). A two-layer model of Gulf Stream separation. *Journal of Fluid Mechanics*, *39*. doi: https://doi.org/10.1017/S0022112069002308

Petit, T., Lozier, M. S., Josey, S. A., & Cunningham, S. A. (2020). Atlantic Deep Water Formation Occurs Primarily in the Iceland Basin and Irminger Sea by Local Buoyancy Forcing. *Geophysical Research Letters*, *47*(22), e2020GL091028. Retrieved 2024-07-03, from `https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL091028` (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL091028) doi: 10.1029/2020GL091028

Pickett, M. H., & Paduan, J. D. (2003). Ekman transport and pumping in the California Current based on the U.S. Navy's high-resolution atmospheric model (COAMPS). *Journal of Geophysical Research: Oceans*, *108*(C10). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003JC001902` doi: https://doi.org/10.1029/2003JC001902

Purich, A., Cai, W., England, M. H., & Cowan, T. (2016). Evidence for link between modelled trends in Antarctic sea ice and underestimated westerly wind changes. *Nat. Comms.*, *7*, 10409.

Purich, A., & England, M. H. (2021). Historical and future projected warming of Antarctic Shelf Bottom Water in CMIP6 models. *Geophysical Research Letters*, *48*. doi: 10.1029/2021GL092752

Rackow, T., Danilov, S., Goessling, H. F., Hellmer, H. H., Sein, D. V., Semmler, T., ... Jung, T. (2022). Delayed Antarctic sea-ice decline in high-resolution climate change simulations. *Nature communications*, *13*(1), 637.

Richter, I. (2015). Climate model biases in the eastern tropical oceans: Causes, impacts and ways forward. *Wiley Interdisciplinary Reviews: Climate Change*, *6*(3), 345–358.

Ricker, R., Hendricks, S., Helm, V., Skourup, H., & Davidson, M. (2014). Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation. *The Cryosphere*, *8*(4), 1607–1622.

Rintoul, S. R. (2018). The global influence of localized dynamics in the Southern Ocean. *Nature*, *558*, 209–218. doi: 10.1038/s41586-018-0182-3

Rintoul, S. R., Hughes, C. W., & Olbers, D. (2001). The Antarctic Circumpolar Current system. In G. Siedler, J. Gould, & J. Church (Eds.), *Ocean circulation and climate, 1st edition* (Vol. 77, pp. 271–301). Academic Press.

Rintoul, S. R., & Naveira Garabato, A. C. (2013). Dynamics of the Southern Ocean circulation. In G. Siedler, S. M. Griffies, J. Gould, & J. Church (Eds.), *Ocean circulation and climate, 2nd edition: A 21st century perspective* (Vol. 103, pp. 471–492). Academic Press. doi: 10.1016/B978-0-12-391851-2.00018-0

Risien, C. M., & Chelton, D. B. (2008). A global climatology of surface wind and wind stress fields from eight years of QuikSCAT scatterometer data. *Journal of Physical Oceanography*, *38*(11), 2379 - 2413. Retrieved from `https://journals.ametsoc.org/view/journals/phoc/38/11/2008jpo3881.1.xml` doi: https://doi.org/10.1175/2008JPO3881.1

Roach, L. A., Dörr, J., Holmes, C. R., Massonnet, F., Blockley, E. W., Notz, D., ... others (2020). Antarctic sea ice area in CMIP6. *Geophys. Res. Lett.*, *47*(9), e2019GL086729.

Robson, J., Menary, M. B., Sutton, R. T., Mecking, J., Gregory, J. M., Jones, C., ... Wilcox, L. J. (2022). The role of anthropogenic aerosol forcing in the 1850–1985 strengthening of the AMOC in CMIP6 historical simulations. *Journal of Climate*, *35*, 6843–6863. doi: 10.1175/JCLI-D-22-0124.1

Roemmich, D., Church, J., Gilson, J., Monselesan, D., Sutton, P., & Wijffels, S. (2015). Unabated planetary warming and its ocean structure since 2006. *Nature Climate Change*, *5*, 240–245. doi: 10.1038/nclimate2513

Ross, A. C., Stock, C. A., Adcroft, A., Curchitser, E., Hallberg, R., Harrison, M. J., . . . Simkins, J. (2023). A high-resolution physical-biogeochemical model for marine resource applications in the northwest atlantic (mom6-cobalt-nwa12 v1.0). *Geoscientific Model Development*, *16*. doi: https://doi.org/10.5194/gmd-16-6943-2023

Rykaczewski, R. R., Dunne, J. P., Sydeman, W. J., García-Reyes, M., Black, B. A., & Bograd, S. J. (2015). Poleward displacement of coastal upwelling-favorable winds in the ocean's eastern boundary currents through the 21st century. *Geophysical Research Letters*, *42*(15), 6424-6431. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL064694` doi: https://doi.org/10.1002/2015GL064694

Sato, Y., Goto, D., Michibata, T., Suzuki, K., Takemura, T., Tomita, H., & Nakajima, T. M. (2018). Aerosol effects on cloud water amounts were successfully simulated by a global cloud-system resolving model. *Nature Communications*, *9*. doi: 10.1038/s41467-018-03379-6

Schmidt, G. A., Romanou, A., Roach, L. A., Mankoff, K. D., Li, Q., Rye, C. D., . . . Busecke, J. J. (2023). Anomalous meltwater from ice sheets and ice shelves is a historical forcing. *Geophys. Res. Lett.*, *50*(24), e2023GL106530.

Schmidtko, S., Heywood, K. J., Thompson, A. F., & Aoki, S. (2014). Multidecadal warming of Antarctic waters. *Science*, *346*, 1227–1231. doi: 10.1126/science.1256117

Schoonover, J., Dewar, W. K., Wienders, N., & Deremble, B. (2017). Local sensitivities of the Gulf Stream separation. *Journal of Physical Oceanography*, *47*. doi: https://doi.org/10.1175/JPO-D-16-0195.1

Shu, Q., Song, Z., & Qiao, F. (2015). Assessment of sea ice simulations in the CMIP5 models. *The Cryosphere*, *9*(1), 399–409.

Silvano, A., Purkey, S., Gordon, A. L., Castagno, P., Stewart, A. L., & S. R. Rintoul, e. a. (2023). Observing Antarctic Bottom Water in the Southern Ocean. *Frontiers in Marine Science*, *10*. doi: 10.3389/fmars.2023.1221701

Silvestri, S., Wagner, G. L., Campin, J.-M., Constantinou, N. C., Hill, C. N., Souza, A., & Ferrari, R. (2024). A new WENO-based momentum advection scheme for simulations of ocean mesoscale turbulence. *Journal of Climate*, *16*. doi: 10.1029/2023MS004130

SIMIP Community, T. (2020). Arctic sea ice in CMIP6. *Geophys. Res. Lett.*, *47*(10), e2019GL086749.

Small, R. J., Curchitser, E., Hedstrom, K., Kauffman, B., & Large, W. G. (2015). The benguela upwelling system: Quantifying the sensitivity to resolution and coastal wind representation in a global climate model. *Journal of Climate*, *28*(23), 9409 - 9432. Retrieved from `https://journals.ametsoc.org/view/journals/clim/28/23/jcli-d-15-0192.1.xml` doi: https://doi.org/10.1175/JCLI-D-15-0192.1

Smeed, D. A., Josey, S. A., Beaulieu, C., Johns, W. E., & amd E. Frajka-Williams., B. I. M. (2018). The North Atlantic Ocean is in a state of reduced overturning. *Geophysical Research Letters*, *45*, 1527–1533. doi: 10.1002/2017GL076350

Smith, R. L. (1968). Upwelling. *Oceanography and Marine Biology: an Annual Review*, *6*, 11–46.

Solodoch, A., Stewart, A. L., Hogg, A. M., Morrison, A. K., Kiss, A. E., Thompson, A. F., . . . Cimoli, L. (2022). How does Antarctic Bottom Water cross the Southern Ocean? *Geophysical Research Letters*, *49*. doi: 10.1029/2021gl097211

Stewart, A., Klocker, A., & Menemenlis, D. (2018). Circum-Antarctic shoreward heat transport derived from an eddy- and tide-resolving simulation. *Geophysical Research Letters*, *45*. doi: 10.1002/2017GL075677

Stewart, A., Klocker, A., & Menemenlis, D. (2019). Acceleration and overturning of the Antarctic Slope Current by winds, eddies, and tides. *Journal of Physical*

Oceanograph, *49*, 2043–2074. doi: 10.1175/JPO-D-18-0221.1
Stommel, H. (1948). The westward intensification of wind-driven ocean currents. *EOS, Transactions American Geophysical Union*, *29*(2), 202–206. doi: 10.1029/TR029i002p00202
Strub, P. T., Combes, V., Shillington, F. A., & Pizarro, O. (2013). Currents and processes along the eastern boundaries. In G. Siedler, S. M. Griffies, J. Gould, & J. Church (Eds.), *Ocean circulation and climate, 2nd edition: A 21st century perspective* (Vol. 103, pp. 339–384). Academic Press. doi: 10.1016/B978-0-12-391851-2.00014-3
Sylla, A., Sanchez Gomez, E., Mignot, J., & López-Parages, J. (2022). Impact of increased resolution on the representation of the Canary upwelling system in climate models. *Geoscientific Model Development*, *15*(22), 8245–8267. Retrieved from `https://gmd.copernicus.org/articles/15/8245/2022/` doi: 10.5194/gmd-15-8245-2022
Taboada, F. G., Stock, C. A., Griffies, S. M., Dunne, J., John, J. G., Small, R. J., & Tsujino, H. (2018). Surface winds from atmospheric reanalysis lead to contrasting oceanic forcing and coastal upwelling patterns. *Ocean Modelling*, *133*, 79–111. doi: 10.1016/j.ocemod.2018.11.003
Tagklis, F., Bracco, A., Ito, T., & Castelao, R. M. (2020). Submesoscale modulation of deep water formation in the Labrador Sea. *Scientific Reports*, *10*. doi: 10.1038/s41598-020-74345-w
Tesdal, J.-E., MacGilchrist, G. A., Beadling, R. L., Griffies, S. M., Krasting, J. P., & Durack, P. J. (2023). Revisiting interior water mass responses to surface forcing changes and the subsequent effects on overturning in the Southern Ocean. *Journal of Geophysical Research - Oceans*, *128*. doi: 10.1029/2022JC019105
Thomas, M. D., Tréguier, A. M., Blanke, B., Deshayes, J., & Voldoire, A. A. (2015). A Lagrangian method to isolate the impacts of mixed layer subduction on the meridional overturning circulation in a numerical model. *Journal of Climate*, *28*, 7503–7517. doi: 10.1175/JCLI-D-14-00631.1
Thompson, A. F., Stewart, A. L., Spence, P., & Heywood, K. J. (2018). The Antarctic Slope Current in a changing climate. *Reviews of Geophysics*, *56*, 741–770.
Thompson, K. R., & Demirov, E. (2006). Skewness of sea level variability of the world's oceans. *Journal of Geophysical Research: Oceans*, *111*. doi: 10.1029/2004JC002839
Todd, R. E. (2021). Gulf Stream mean and eddy kinetic energy: Three-dimensional estimates from underwater glider observations. *Geophysical Research Letters*, *48*. doi: https://doi.org/10.1029/2020GL090281
Varela, R., DeCastro, M., Rodriguez-Diaz, L., Dias, J. M., & Gómez-Gesteira, M. (2022). Examining the ability of CMIP6 models to reproduce the upwelling SST imprint in the Eastern Boundary Upwelling Systems. *Journal of Marine Science and Engineering*, *10*(12). Retrieved from `https://www.mdpi.com/2077-1312/10/12/1970` doi: 10.3390/jmse10121970
Verdy, A., & Mazloff, M. R. (2017). A data assimilating model for estimating Southern Ocean biogeochemistry. *Journal of Geophysical Research: Oceans*, *122*, 6968–6988. doi: 10.1002/2016JC012650
Waldman, R., & Giordani, H. (2023). Ocean barotropic vorticity balances: theory and application to numerical models. *Journal of Advances in Modeling Earth Systems*, *15*(4), e2022MS003276. doi: 10.1029/2022MS003276
Wang, C., Zhang, L., Lee, S.-K., Wu, L., & Mechoso, C. R. (2014). A global perspective on CMIP5 climate model biases. *Nature Climate Change*, *4*. doi: 10.1038/NCLIMATE2118
Wang, H., Legg, S. A., & Hallberg, R. W. (2015). Representations of the Nordic Seas overflows and their large scale climate impact in coupled models. *Ocean Modelling*, *86*, 76–92. doi: 10.1016/j.ocemod.2014.12.005
Wang, X., Key, J., Kwok, R., & Zhang, J. (2016). Comparison of Arctic sea ice

thickness from satellites, aircraft, and PIOMAS data. *Remote Sensing*, *8*(9), 713.

Winton, M., Adcroft, A., Dunne, J. P., Held, I. M., Shevliakova, E., Zhao, M., ... Zhang, R. (2020). Climate sensitivity of GFDL's CM4.0. *Journal of Advances in Modeling the Earth System*, *12*. doi: 10.1029/2019MS001838

Xu, X., Chassignet, E. P., Firing, Y. L., & Donohue, K. (2020). Antarctic Circumpolar Current transport through Drake Passage: What can we learn from comparing high-resolution model results to observations? *Journal of Geophysical Research: Oceans*, *125*. doi: 10.1029/2020jc016365

Xu, X., Hurlburt, H., Schmitz Jr, W., Zantopp, R., Fischer, J., & Hogan, P. (2013). On the currents and transports connected with the Atlantic meridional overturning circulation in the subpolar North Atlantic. *Journal of Geophysical Research: Oceans*, *118*(1), 502–516. doi: 10.1002/jgrc.20065

Yan, X., Zhang, R., & Knutson, T. R. (2018). Underestimated AMOC variability and implications for AMV and predictability in CMIP models. *Geophysical Research Letters*, *45*. doi: 10.1029/2018GL077378

Yan, X., Zhang, R., & Knutson, T. R. (2019). A multivariate AMV index and associated discrepancies between observed and CMIP5 externally forced amv. *Geophysical Research Letters*, *46*. doi: 10.1029/2019GL082787

Yeager, S. (2015). Topographic coupling of the Atlantic overturning and gyre circulations. *Journal of Physical Oceanography*, *45*(5), 1258–1284. doi: 10.1175/JPO-D-14-0100.1

Yeager, S., Castruccio, F., Chang, P., Danabasoglu, G., Maroon, E., Small, J., ... Zhang, S. (2021). An outsized role for the Labrador Sea in the multidecadal variability of the Atlantic overturning circulation. *Science Advances*, *7*. doi: 10.1126/sciadv.abh3592

Zanna, L., Khatiwala, S., Gregory, J., Ison, J., & Heimbach, P. (2019). Global reconstruction of historical ocean heat storage and transport. *Proceedings of the National Academy of Science*. doi: 10.1073/pnas.1808838115

Zhang, J., & Rothrock, D. (2003). Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates. *Mon. Wea. Rev.*, *131*(5), 845–861.

Zhang, L., Delworth, T. L., Cooke, W., Goosse, H., Bushuk, M., Morioka, Y., & Yang, X. (2021). The dependence of internal multidecadal variability in the Southern Ocean on the ocean background mean state. *Journal of Climate*, *34*(3), 1061–1080. doi: 0.1175/JCLI-D-20-0049.1

Zhang, L., Delworth, T. L., Cooke, W., & Yang, X. (2019). Natural variability of Southern Ocean convection as a driver of observed climate trends. *Nature Climate Change*, *9*(1), 59–65.

Zhang, R., Delworth, T. L., Rosati, A., Anderson, W. G., Dixon, K. W., Lee, H.-C., & Zeng, F. (2011). Sensitivity of the North Atlantic Ocean circulation to an abrupt change in the Nordic Sea overflow in a high resolution global coupled climate model. *Journal of Geophysical Research*, *116*. doi: 10.1029/2011JC007240

Zhang, R., & Thomas, M. (2021). Horizontal circulation across density surfaces contributes substantially to the long-term mean northern Atlantic meridional overturning circulation. *Communications Earth & Environment*, *2*. doi: 10.1038/s43247-021-00182-y

Zhang, R., & Vallis, G. K. (2007). The role of bottom vortex stretching on the path of the North Atlantic Western Boundary Current and on the Northern Recirculation Gyre. *Journal of Physical Oceanography*, *37*, 2053–2080. doi: 10.1175/JPO3102.1

Zhao, M. (2020). Simulations of atmospheric rivers, their variability, and response to global warming using GFDL's new high-resolution general circulation model. *Journal of Climate*, *33*, 10287–10303. doi: 10.1175/JCLI-D-20-0241.1

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., … Xiang, B. (2018a). The GFDL Global Atmosphere and Land Model AM4.0/LM4.0 - Part II: Model Description, Sensitivity Studies, and Tuning Strategies. *Journal of Advances in Modeling Earth Systems*, *10*. doi: 10.1002/2017MS001209

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., … Xiang, B. (2018b). The GFDL Global Atmosphere and Land Model AM4.0/LM4.0 - Part I: Simulation Characteristics with Prescribed SSTs. *Journal of Advances in Modeling Earth Systems*, *10*. doi: 10.1002/2017MS001208

Zou, S., Petit, T., Li, F., & Lozier, M. S. (2024, April). Observation-based estimates of water mass transformation and formation in the Labrador Sea. *Journal of Physical Oceanography*, *-1*(aop). Retrieved 2024-04-28, from `https://journals.ametsoc.org/view/journals/phoc/aop/JPO-D-23-0235.1/JPO-D-23-0235.1.xml` (Publisher: American Meteorological Society Section: Journal of Physical Oceanography) doi: 10.1175/JPO-D-23-0235.1