Evaluation of MITgcm-based ocean ~~reanalysis~~ reanalyses for the Southern Ocean

Yoshihiro Nakayama[1], Alena Malyarenko[2,3], Hong Zhang[4], Ou Wang[4], Matthis Auger[5,6], Yafei Nie[7], Ian Fenty[4], Matthew Mazloff[8], Köhl Armin[9], and Dimitris Menemenlis[4]

[1]Institute of Low Temperature Science, Hokkaido University, Sapporo, Hokkaido, Japan.
[2]Te Kura Aronukurangi | School of Earth and Environment, Te Whare Wānanga o Waitaha | University of Canterbury, Ōtautahi | Christchurch, Aotearoa | New Zealand
[3]Te Puna Pātiotio | Antarctic Research Centre, Te Herenga-Waka Victoria University of Wellington, Aotearoa | New Zealand
[4]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA.
[5]Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Australia
[6]The Australian Centre for Excellence in Antarctic Science, University of Tasmania, Hobart, Australia
[7]School of Atmospheric Sciences, Sun Yat-Sen University, and Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China
[8]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA
[9]Center for Earth System Research and Sustainability (CEN), Universität Hamburg, Hamburg, Germany

Correspondence: Y. Nakayama (Yoshihiro.Nakayama@lowtem.hokudai.ac.jp)

Abstract. Global and basin-scale ocean reanalyses are becoming easily accessible and are utilized widely to study the Southern Ocean. Yet, such ocean reanalyses are optimized to achieve the best model-data agreement for their entire model domains and their ability to simulate the Southern Ocean requires investigations. Here, we compare several Massachusetts Institute of Technology general circulation model (MITgcm)-based ocean reanalyses (ECCOv4r5, ECCO LLC270, B-SOSE, and GECCO3) for the Southern Ocean. For the open ocean, the simulated time-mean hydrography and ocean circulation are ~~consistent with~~ similar to observations. The MITgcm-based ocean reanalyses show Antarctic Circumpolar Currents (ACC) measuring approximately 149±11Sv. The simulated 2 °C isotherms are located in positions similar to the ACC and roughly represent the southern extent of the current. Simulated Weddell and Ross Gyre strengths are 51±11Sv and 25±8Sv, respectively, consistent with observation-based estimates. However, our evaluation finds that the time evolution of the Southern Ocean is not well simulated in these ocean reanalyses. While observations showed little change in open-ocean properties in the Weddell and Ross Gyres, all simulations showed larger trends, most of which are excessive warming. For the continental shelf region, all reanalyses are unable to reproduce observed hydrographic features, suggesting that simulated physics determining on-shelf hydrography and circulation is ~~wrong~~not well represented. Nevertheless, ocean reanalyses are valuable resources and can be used for generating ocean lateral boundary conditions for regional high-resolution simulations. We recommend that future users of these ocean reanalyses pay extra attention if their studies target open-ocean Southern Ocean temporal changes or on-shelf processes.

1 Introduction

The Southern Ocean is a critical component of the Earth's climate system (e.g., Gille, 2002; Hellmer et al., 2012; Rintoul, 2018). The Antarctic Circumpolar Current (ACC), Ross Gyre (RG), and Weddell Gyre (WG) circulations are involved in the advection of warm ocean waters towards the Antarctic ice shelves (e.g., Orsi et al., 1999; Ryan et al., 2016; Nakayama et al., 2018). The formation of sea ice close to the Antarctic continent leads to the production of high salinity shelf water that sinks to the ocean's bottom, forming the Antarctic Bottom Water (AABW) — the coldest and densest water mass in the world's oceans (e.g., Jacobs et al., 1970; Foldvik et al., 2004; Williams et al., 2008; Ohshima et al., 2013). The Southern Ocean also serves as a significant sink for anthropogenic carbon, playing a crucial role in carbon cycling (e.g., Takahashi et al., 2012; Sallée et al., 2023; Williams et al., 2023). However, our knowledge of the Southern Ocean is still limited due to a lack of observations in the challenging polar conditions.

One approach to address the lack of observations is to create global, circum-Antarctic, or regional ocean models. In recent decades, several researchers have developed coupled sea ice, ice shelf, and ocean models, allowing us to study the interplay between large-scale ocean circulation, ice shelves, and sea ice on a global or circum-Antarctic scale (e.g., Timmermann et al., 2009; Nakayama et al., 2013; Kusahara and Hasumi, 2014; Dinniman et al., 2015; Schodlok et al., 2015; Mathiot et al., 2017; Kiss et al., 2020). To study processes over the Antarctic continental shelves, other researchers develop regional simulations (e.g., Dinniman et al., 2011; Schodlok et al., 2012; Gwyther et al., 2014; Nakayama et al., 2014a; Jourdain et al., 2017; Nakayama et al., 2017; Mack et al., 2019; Naughten et al., 2019; Nakayama et al., 2019, 2021a; Naughten et al., 2022). Regional models can yield a superior agreement between model and observations ~~and can~~ because (1) regional simulations can achieve high spatial and vertical resolutions and (2) model parameters can be tuned for the study regions. Thus, regional simulations can help better interpret observed on-shelf processes. For example, several studies have investigated on-shelf intrusions of warm Circumpolar Deep Water (CDW) or the formation of AABW (Assmann et al., 2013; Nakayama et al., 2018; Mensah et al., 2021). Such models can also be used to identify ideal locations for ocean observations (e.g., deployment locations of Argo floats). Nonetheless, constructing, evaluating, and improving these regional ocean models requires a significant amount of time and effort.

Another approach is to utilize existing ocean reanalyses, which are convenient and becoming an increasingly trusted tool. Over the past decades, following extensive developmental ~~effort~~efforts, several ocean reanalyses have become available. Previously Uotila et al. (2019) have analyzed the mean states of ten ocean reanalysis products, including GECCO2 (a prior version) for the Arctic and Antarctic focusing on sea ice and upper-ocean hydrography. Here, we will evaluate some recent ocean reanalyses based on the Massachusetts Institute of Technology general circulation model (MITgcm) and made available by the Estimating the Circulation and Climate of the Ocean consortium, including ECCO version 4 release 5 (ECCOv4r5; Forget et al., 2015), ECCO Latitude-Longtitude-polar Cap 270 (ECCO-LLC270; Zhang et al., 2018), the German contribution of the ECCO project version 3 (GECCO3; Köhl, 2020), and the Biogeochemical-Sea ice-Ocean State Estimate iterations 105 and 139 (B-SOSE; Mazloff et al., 2010; Verdy et al., 2017). As these are ocean reanalyses, researchers in various fields, especially non-ocean ~~modellers~~ modelers assume that these model outputs replicate past changes and variability in the Southern Ocean

to some extent. It is convenient for researchers because they can download one of these ocean reanalyses and use the same model output repeatedly for their different research. For example, they can use the same code to read the model data, conduct model-data comparisons, or test their hypothesis to strengthen their scientific claims.

Ocean reanalyses are generated by minimizing a cost function (weighted sum of model-data difference squared) by optimizing model parameters, including atmospheric forcing, ocean mixing parameters, and initial conditions (Forget et al., 2015; Zhang et al., 2018). All these MITgcm-ECCO assimilations (ECCOv4r5, ECCO-LLC270, B-SOSE, and GECCO) are unique in that they use the model as a hard constraint over many years. Long assimilation windows are chosen to yield the advantage of having closed budgets. This means that ECCO products assume the model is perfect and chaos does not impact the large scale over long time scales. In contrast, other groups nudge the model on daily to weekly time scales, keeping their models better on track with the observations. Challenges arise when applying this framework to improve the Southern Ocean. First, the observations used to constrain ocean models are primarily concentrated in lower-latitude, ice-free regions, and in summertime (Newman et al., 2019). Even if the cost is reduced for the global ocean, it is possible that the ocean state in the Southern Ocean remains suboptimal. Second, ~~with the exception of~~ except for ECCOv4r5, most ocean reanalyses do not explicitly model ice-shelf cavities (Forget et al., 2015; Zhang et al., 2018; Mazloff et al., 2010; Verdy et al., 2017; Köhl, 2020). We also note that ocean state estimates development takes a long time and model bathymetry can be outdated for the Antarctic continental shelf regions. Third, our optimization framework uses long assimilation windows. This means that we do not have controllability of the mesoscale dynamics, which may modulate the hydrography and circulation of the Southern Ocean (Ito and Marshall, 2008; Ruan et al., 2017). Fourth, polar scientists care about tiny changes in temperature or thermocline depth. Such changes may not be a priority for ocean reanalysis developers. Finally, while developers evaluate ocean reanalyses, there has been no cross-comparison of these MITgcm-based ocean reanalysis products and observations focusing on the entire Southern Ocean such as plotting sections and comparing time series against observations despite some regional attempts for reanalyses intercomparisons (e.g., Chen et al., 2023; Bailey et al., 2023; Cerovečki and Haumann, 2023).

In this study, we evaluate Southern Ocean circulation in the four, aforementioned ECCO reanalyses. We compare model output from these reanalyses to existing observations in the open ocean and over the continental shelves, examining how well they simulate the mean state, variability, and trends. We aim to provide guidelines for researchers on what they can and cannot do with these ocean reanalyses. We also would like to (1) make recommendations on how much these ocean reanalyses can be trusted for the purposes of creating initial conditions and lateral forcing for regional simulations and (2) suggest how we can improve the model-data agreement in the Southern Ocean for the existing ocean reanalyses.

2 Ocean reanalyses and datasets

The four ECCO reanalyses that will be evaluated in this study are listed in Table 1 and briefly described below.

2.1 ECCOv4r5

ECCO Version 4 release 5 (v4r5) is ECCO's latest global ocean state estimate covering the period of 1992–2019 (Table 1). This product is an updated version of the solution described in Forget et al. (2015). ECCOv4r5 uses the LLC90 grid, which has a nominal horizontal grid spacing of 1° between 70°S and 57°N. South of 70°S, the LLC90 configuration uses a bipolar grid. Horizontal grid spacing is approximately 40 km along the Antarctic Coast. Details of the previous releases can be found in Forget et al. (2015). ECCOv4r5 uses 6 hourly ERA-interim for the initial guess of atmospheric forcing.

We also highlight two important improvements related to polar regions compared to the previous versions (ECCOv4r4 and ECCO-LLC270). First, release 5 includes static ice shelf cavities. Ice shelf melt rates are calculated using a three-equation model (Hellmer and Olbers, 1989; Jenkins, 1991; Holland and Jenkins, 1999), with exchange coefficients of heat and salt at the ocean-ice shelf boundary being adjusted through optimization. Model bathymetry around Antarctica is from The International Bathymetric Chart of the Southern Ocean (IBCSO) Version 1.0. (Arndt et al., 2013). Ice shelf thickness is based on Bedmap-2 (Fretwell et al., 2013). Together, IBCSO bathymetry and Bedmap-2 ice-shelf thickness are used to update geometry around Antarctica. Second, starting from version 4 release 5, sea-ice thermodynamics are included in adjoint-model sensitivity computations achieving higher model ~~deta~~ data agreement for sea ice extent and variability as discussed below.

2.2 ECCO-LLC270

ECCO-LLC270 is a higher-resolution, global ECCO ocean state estimate covering 1992–2017 (Table 1). This reanalysis is an updated version of the solution described in Zhang et al. (2018). ECCO-LLC270 uses the LLC270 grid, which has a nominal horizontal grid spacing of 1/3° between 70°S and 57°N. South of 70°S, the LLC270 configuration also uses a bipolar grid with a horizontal grid spacing of approximately 15 km along the Antarctic Coast. ECCO-LLC270 does not include ice shelf cavities. Model bathymetry is based on the Global Sea Floor Topography from Satellite Altimetry and Ship Depth Soundings ~~Smith and Sandwell (1997)~~(Smith and Sandwell, 1997), which is problematic in some regions around Antarctica, e.g., in the Amundsen Sea. In this study, we evaluate Iteration 50 of ECCO-LLC270, which is similar to iteration 42 used in ECCO-Darwin (Carroll et al., 2020). ECCO-LLC270 uses 6 hourly ERA-interim for the initial guess of atmospheric forcing.

2.3 B-SOSE

B-SOSE is a circum-Antarctic regional ocean state estimate developed by the Scripps Institution of Oceanography (Table 1). The iteration 139 we used covers the period 2013-2021. The analysis model domain extends from 30°S to 78°S. The zonal grid spacing is 1/6°. B-SOSE does not include ice shelf cavities. Model bathymetry is based on ETOPO1 ~~(Amante and Eakins, 2009; Verdy et al.~~ (Amante and Eakins, 2009). Details can be found in Mazloff et al. (2010) and Verdy et al. (2017). B-SOSE includes biogeochemistry, but we do not evaluate this component in this manuscript. We also analyze the iteration 105, which covers 2008-2012. B-SOSE iteration 105 has a zonal grid spacing of 1/3° and has been used widely for Southern Ocean studies. B-SOSE uses hourly ERA5 for the initial guess of atmospheric forcing (Verdy et al., 2017).

2.4 GECCO3

GECCO3 (Köhl, 2020) is a global ocean state estimate developed by Universität Hamburg, Germany (Table 1). This simulation covers the period 1948-2018. GECCO3 uses the bathymetry and grid of the MPI-ESM with a quasi-uniform resolution of 0.4°.

115 GECCO3 does not include ice shelf cavities. Model bathymetry is from the MPI-HR configuration of the MPI-ESM climate model which is based on ETOPO5 (National Geophysical Data Center 1988) (Jungclaus et al., 2013). In this comparison, we use the results of GECCO3S6m with surface salinity relaxation with a relaxation time scale of 6 months. The detail can be found in Köhl (2020). GECCO3 uses 6 hourly NCEP/NCAR Reanalysis 1 for the initial guess of atmospheric forcing.

2.5 ~~Ocean~~ Sea ice observations used for ~~ocean~~ reanalysis evaluation~~We use~~

120 We use the monthly sea ice concentration observational data obtained from the National Oceanic and Atmospheric Administration/National Snow and Ice Data Center (NOAA/NSIDC) Climate Data Record of Passive Microwave Sea Ice Concentration, version 4 (CDR) (Meier et al., 2021). We use two observational sea ice thickness references. The first one, released by the European Space Agency Sea Ice Climate Change Initiative (SICCI), provides sea ice thickness records from June 2002 to April 2017 and has a spatial resolution of 50 km. The uncertainty of this dataset is considerably large (Wang et al., 2022; Hou et al., 2024). The

125 second sea ice thickness data is provided by the Laboratoire d'Etudes en Géophysique et Océanographie Spatiales (LEGOS) (Garnier et al., 2021, 2022) with a 12.5 km resolution. The LEGOS sea ice thicknesses benefit from improved snow depth measurements on sea ice, which are obtained by altimeters on board the satellites SARAL (Ka-band) and CS2 (Ku-band). To avoid the effect of snowpack property variabilities (e.g., density and grain size) on retrieval accuracy, the LEGOS sea ice thickness data is only available in winter months (from May to October) and covers from May 2013 to October 2018.

130 ## 2.6 Ocean observations used for reanalysis evaluation

We use multiple datasets for the ocean reanalysis evaluation. We use the World Ocean Circulation Experiment (WOCE) Hydrographic Sections S4P and SR04 for the Ross and Weddell Sea, respectively. The S4P section was observed in 1992 (WOCE), 2011, and 2018 (Talley, 2007; Purkey and Johnson, 2013; Purkey et al., 2019). The SR04 Section was observed in 1989, 1996 (WOCE Section 23), 2005, and 2010 (Fahrbach et al., 1991, 2004; Talley, 2007; Rücker van Caspel, 2016). We also use

135 World Ocean Atlas 2018 as a reference dataset for the entire Southern Ocean (Locarnini et al., 2018; Zweng et al., 2019). A Mean Dynamic Topography (MDT) obtained from satellite altimetry observations that include ice-covered regions (Armitage et al., 2018) is displayed to qualitatively compare with reanalysis stream functions. This MDT is referenced to the GOCO05c combined gravity field model ~~(Fecher et al. 2017).~~ (Fecher et al., 2017).

3 Results

3.1 Ocean bathymetry and ice shelves

The ocean reanalyses use ocean bathymetry derived from three different products (Fig. 1). We find obvious differences over the continental shelves. ECCO-LLC270 and GECCO3 use the older bathymetry product, which results in the absence of deep submarine glacial troughs (approximately 500 meters) connecting the open water and ice shelf cavities in the Amundsen Sea, Bellingshausen Sea, and western Ross Sea (Fig. 1). B-SOSE uses newer ETOPO1 bathymetry (Amante and Eakins, 2009), which is substantially improved compared to ECCO-LLC270 and GECCO3. However, there are still small differences between ECCOv4r5, including the absence of detailed structures of submarine glacial troughs in the Amundsen and Bellingshausen Seas. B-SOSE also cuts out the southernmost part of the Ross Sea continental shelf due to its horizontal gridding configuration. ECCOv4r5, the newest ocean reanalyses, utilizes the best available bathymetric data and has ice shelf cavities. For the open ocean region, the ocean bathymetries of these analyses are similar.

3.2 Sea ice

We first evaluate the sea ice area (SIA), which is the sum of grid cell areas multiplied by their respective sea ice concentrations, provided the concentration is at least 15%. For ECCOv4r5, ECCO-LLC270, B-SOSE (iter105), and GECCO3, ~~sea ice extent is highly consistent with~~ the monthly SIA time series are similar to observations (Fig. 2a). ~~For climatology, sea-ice concentration time series show their~~ All the reanalyses well capture the seasonal cycles of SIA, with minimum and maximum in February and September, respectively ~~, consistent with observations~~ (Fig. 2d). To further focus on ~~sea ice extent~~ SIA in winter and summer, we plot ~~September and March averages following previous modeling studies. Maximum values are mostly consistent with the satellite-derived estimates~~ detrended monthly anomalies in September and February (Figs. 2b-c)~~, except for B-SOSE(iter 139) with much higher sea ice extent by $\sim 5 \times 10^{12}$ (m^2)~~. All reanalyses show ~~simulated model extent within an error of $\sim 10^{12}$ (m^2). All reanalyses capture interannual variability of maximum sea ice extent slightly increasing from 1992-2016 and sudden drop after that consistent with observations.~~ near-realistic anomaly evolutions in September, but only ECCOv4r5~~shows the best agreement~~, ECCO-LLC270, and GECCO3 performed well in February, while B-SOSE has significantly larger variation than the observational data. The standard deviation of February and September anomalies of each reanalysis are quantitatively compared to the observations (Table 2). The ECCOv4r5 and ECCO LLC270 show the most realistic interannual variations in February and September respectively, with the former also having the highest correlation with the observation in both months. Overall, the ECCOv4r5 performs the best in the comparison, which likely indicates the effectiveness of sea-ice thermodynamics adjoint-model sensitivity computations. ~~Minimum values are also consistent with the satellite-derived estimates. All reanalyses show interannual changes consistent with observations (Figs. 2b-c).~~ Comparing reanalyses with long integration (ECCOv4r5, ECCO-LLC270, and GECCO), ECCOv4r5 and ECCO-LLC270 show good agreement at a similar level, suggesting that sea-ice thermodynamics adjoint sensitivity computation is not improving summer sea ice extent. ~~Amongst all reanalysis products, B-SOSE (iter 105) shows the largest error.~~

We also conduct the same analysis for sea ice volume (SIV) (Fig. 3). For the seasonality, ~~ECCOv4r5 and GECCO3 show minimum and maximum volumes in February and October, while~~ only ECCO LLC270 ~~and B-SOSE show~~ shows minimum and maximum in February and September ~~, respectively. Sea ice volumes of ECCOv4r5, ECCO LLC270~~respectively, ~~B-SOSE (iter105), and GECCO3 are similar, while the sea ice volume of~~ inconsistent with the observations (Fig. 3d). The maximum SIV month of other reanalyses all delayed to October, the same issue was found in the Global Ice-Ocean Modeling and Assimilation System (Liao et al., 2022), the ECMWF Ocean ReAnalysis System 5 (Nie et al., 2022) and many Coupled Model Intercomparison Project Phase 6 climate model results (Hou et al., 2024). All reanalyses underestimate the SIV throughout the year, but the B-SOSE (iter139) is ~~larger by ~60%~~the closest to the observations. For the interannual variability of sea ice volumes in September and ~~March~~February, all state estimates show similar temporal variability, and their variabilities are similar to that of sea ice extent (Figs. 2b-c and 3b-c).

3.3 Large-scale ocean state

3.3.1 Large-scale hydrographic structure

We select 550 m to compare simulated outputs and the World Ocean Atlas (WOA) dataset, as this is a typical depth for CDW to travel in the open ocean before crossing the continental shelf break towards the Antarctic ice shelves (Nakayama et al., 2018). In the WOA dataset (Locarnini et al., 2018; Zweng et al., 2019), the annual mean 1°C isotherm and 34.6 isohaline encompass the Southern Ocean, separating warm and cold water masses (Figs. 4e,j). Similarly, the 1°C isotherms and 34.6 isohalines of all ocean reanalyses are located at similar positions, indicating that all four reanalyses generally simulate similar hydrographic structures (Figs. 4 and S1). Inside the 1°C isotherms, WOA shows two local temperature minima in the Weddell Sea and Ross Sea (Figs. 4e,j), which are also present in all four ocean reanalyses. Inside the 34.6 isohaline, observations show a local maximum in the Ross Sea (Figs. 4e,j), and such structures are also simulated in ECCOv4r5, ECCO-LLC270, and B-SOSE (iterations 105 and 139).

~~Because the WOA dataset is biased towards summer observations, we~~ We compare the February WOA mean to the February mean state of the reanalyses (Figs. 5 and S2) because the WOA dataset is biased towards summer observations. All reanalyses represent the surface temperature and salinity values well. The position of isolines and isohalines agrees between the reanalyses ~~,~~and matches the WOA for the Southern Ocean (Fig. 5). Close to the continent, the Weddell Sea matches the temperature and salinity values particularly well except for B-SOSE: the cold and relatively salty surface water mass is a feature of ~~all ocean state estimates~~ECCOv4r5, ECCO-LLC270, and GECCO3. However, the sector between the Antarctic Peninsula and the Ross Sea has a ~~surface fresh~~ fresh surface bias within the coastal current, confined to the top layer. Among these model outputs, B-SOSE (iter 139) shows the strongest freshening (Fig. 5h). We hypothesize that the fresh bias is related, in part, to the choice of freshwater flux from the Antarctic continent~~(e. g., see arrows in Fig. 5).~~. In ECCOv4r5, ice sheet melt is computed, and iceberg calving is prescribed (Hammond and Jones, 2017). In ECCO-LLC270, both melt and calving are prescribed (see Feng et al., 2021, Fig. S1b). In B-SOSE, melt and calving ~~is~~ are prescribed, also based on Hammond and Jones (2017)~~.~~, and the fresh bias is partially exacerbated by too strong a sea ice seasonal cycle (i.e., too much formation in winter). GECCO3 does

not prescribe any freshwater discharge around Antarctica. Local sea ice melting likely also contributes to the fresh cap creation during spring, which is not vertically mixed in summer. It is likely that the warm SST bias in this region is also caused by the fresh bias: stratified water column does not promote vertical mixing of the summer warming.

We also compare mixed layer depths (MLD) because ocean reanalyses are widely used for physical oceanography and biogeochemistry studies (Fig. 6). MLD is defined as the depth level where the density difference from the surface exceeds 0.03 kg m^{-1}. We select September and February, the same time as sea ice concentration and volume. These months coincide with or close to the MLD maximum and minimum. In September, observed MLD is located deep (about 500 m) at lower latitude regions ($\sim$50°S) in the Pacific sector. Similar patterns are simulated for all ocean reanalyses. We also find several spots with deep MLD along the Antarctic coasts, which are likely representing formation regions of AABW in the WOA. However, none of the ocean reanalyses show such deep MLD along the Antarctic coast. In February, MLDs become shallower for all locations but maintain deep MLD (about 100 m) regions at lower latitudes ($\sim$50°S). Such patterns are mostly present in all reanalyses but simulated MLDs are biased shallow in all ocean reanalyses presented in this study.

We show time series of MLD at the center of the Weddell and the Ross Gyres as ocean reanalyses tend to show better agreement for the open water regions (Fig. 7). All ocean reanalyses show similar seasonal cycles with maximum and minimum in December/January and September/October, respectively (Figs, 7c-d). Simulated MLDs show small variability, and the standard deviations of simulated MLDs are about or less than 30 m. Standard deviations become large between October and December during the shallowing of MLDs. The time series of the simulated MLD at the center of the Weddell Gyre shows larger interannual variability, while the time series of simulated MLD at the center of the Ross Gyre shows smaller interannual variability (Figs. 7a-b).

We also calculate the Root Mean Square Error (RMSE) of potential temperature and salinity at depth levels of 500 m, 2000 m, and 4000 m (Fig. 8). RMSE is calculated by interpolating each ocean state estimate output onto a common horizontal grid with a 1° resolution. At 500 m, most ocean reanalyses show stable RMSE indicating that ocean reanalyses have nearly constant error throughout the model simulations. ECCOv4r5 and ECCO-LLC270 perform better, maintaining similar RMSE during model integration over approximately 30 years. In contrast, GECCO shows a rapid increase in RMSE within the first 10 years, after which RMSE stabilizes or slightly decreases (Figs. 8a,d). B-SOSE exhibits the steepest increase in RMSE among the ocean analyses presented here. At the deeper levels (2000 m and 4000 m), all ocean reanalyses show increases in RMSE (except GECCO3 salinity) at similar rates (Figs. 8b,c,e,f). This consistent increase in RMSE may reflect the fact that all ocean reanalyses struggle to accurately simulate Antarctic Bottom Water (AABW) production, thereby misrepresenting key ocean processes leading to similar responses in all ocean reanalyses' RMSE (discussed further below).

3.3.2 Large-scale Ocean Circulation

Recent observations estimated the ACC strength to be 140-170 Sv using advanced technologies (Donohue et al., 2016; De Verdière and Ollitrault, 2016; Xu et al., 2020). Other canonical estimates are 130-140 Sv (Whitworth III et al., 1982; Whitworth, 1983; Whitworth and Peterson, 1985; Koenig et al., 2014). For ocean reanalysis, we also calculate the time mean ACC strengths by determining the flux through the Drake Passage for each simulated period. The mean ACC strengths for the

ECCOv4r5, ECCO-LLC270, B-SOSE (iter139), and GECCO3 models are 154 Sv, 140 Sv, 162 Sv, and 141 Sv, respectively (Table ~~2~~3). For simulated periods for each reanalysis, the ECCOv4r5, ECCO-LLC270, and GECCO3 ACC strengths do not

240 show unrealistic trends showing rather stable values, consistent with observations (Hogg et al., 2015; Donohue et al., 2016). We note, however, that the B-SOSE model's ACC (iter105) increases by ∼30 Sv within 5 years with a maximum strength of 200 Sv (Fig. 9), so B-SOSE (iter105)'s unrealistic ACC strengthening is fixed in the ~~most recent releases~~iteration 139. As ACC transport is a good measure for evaluating global simulations, we recommend checking ACC transport when using newer versions of ocean reanalysis products.

245 The estimates of WG and RG strengths are uncertain due to intense sea-ice cover, lack of observations, and large seasonal and internannual variability. The estimated WG strength has a wide range of values between 30 and 100 Sv based on previous observation-based and modeling studies (e.g., Fahrbach et al., 1991; Park and Gambéroni, 1995; Beckmann et al., 1999; Wang and Meredith, 2008; Mazloff et al., 2010; Cisewski et al., 2011; Reeve et al., 2019; Neme et al., 2023). The estimated RG strengths are between 10-30 Sv (Chu and Fan, 2007; Mazloff et al., 2010; Nakayama et al., 2014b) and the most recent study

250 using altimetry data estimated the RG strength to be 23±8Sv (Dotto et al., 2018). Following previous studies, we calculate the strengths of the WG and RG by extracting the minimum stream function values in the Weddell Sea and Ross Sea, respectively. In the ocean reanalyses, simulated WG strengths are 56 Sv, 60 Sv, 54 Sv, and 35 Sv for ECCOv4r5, ECCO-LLC270, B-SOSE (iter139), and GECCO3 models, respectively (Table ~~2~~3). The simulated RG strengths are 21 Sv, 32 Sv, 32 Sv, and 15 Sv for ECCOv4r5, ECCO-LLC270, B-SOSE (iter139), and GECCO3 models, respectively (Table ~~2).~~ 3). We note that GECCO's

255 WG and RG strengths are weaker by about 30-50%, which may be explained by the fact that ECCOv4r5, ECCO-LLC270, and SOSE use ERA products, while GECCO uses NCEP products as atmospheric forcing as initial guesses. Due to the large uncertainty of observational estimates, mean WG and RG strengths of ocean reanalyses are consistent with observations (Fig. 9).

The ACC strength from the reanalyses shows seasonal and interannual variabilities with magnitudes of about 5 Sv, which is

260 small considering the total ACC strength of about 150 Sv (Figs. ~~6-8~~9-11). Seasonally, the ACC tends to be marginally stronger in winter than in summer, a pattern observed consistently across all reanalyses (Fig. ~~7~~11a). For the interannual variability, simulated small fluctuations are consistent with observations, as Gutierrez-Villanueva et al. (2023) find almost no interannual changes in ACC strengths using repeated oceanographic measurements from 2005 (Fig. ~~6~~10a).

The WG strength from the reanalyses shows a substantial seasonal and interannual variability (Figs. ~~6-8~~9-11). For the

265 interannual variability, time series of ~~gyre depth~~ negative sea level anomalies in the center of the gyre obtained from SSH observations between 2011 to 2019 suggest an acceleration of the gyre beginning in 2014 and ending in 2017 (Armitage et al., 2018; Auger et al., 2022b), which seems to be captured to some extent in these ocean reanalyses (Fig. ~~7~~10b). There is, however, a substantial disparity in the interannual variability among the reanalyses (Fig. ~~7b~~10b). For the seasonality, the gyre tends to strengthen during winter and weaken during summer (Fig. ~~6~~11b), which aligns with SSH observations from 2013 to 2019

270 (Auger et al., 2022b).

The RG strength demonstrates substantial interannual variability according to reanalyses, though it exhibits less marked seasonal changes ~~compared to~~ than the WG (Figs. ~~6-8~~9-11). Observations of the Ross Gyre indicate a slowdown starting

in 2015 and continuing into 2016 (Dotto et al., 2018; Armitage et al., 2018). This is replicated across all reanalyses with varying amplitude (Fig. 710c). Notably, observations of the RG display distinct seasonal variability, weakening in summer and intensifying bi-annually in autumn and spring (Dotto et al., 2018; Auger et al., 2022b). While the winter slowdown is consistently observed in all reanalyses, there are differences in how they represent the intensity and timing of the bi-annual intensification (Fig. 611c).

The spatial structures of the ACC, WG, and RG are also similar across all ocean reanalyses (FigFigs. 4k-n). For example, the simulated 2°C isotherms, which roughly represent the southern extent of the Antarctic Circumpolar Current (ACC), are consistently located in all reanalyses. Additionally, the stream functions within the ACC exhibit similar patterns across the reanalyses. To assess the agreement between the reanalyses and observations, we employ the Mean Dynamic Topography (MDT) as a proxy. Although the MDT only represents the circulation of the surface geostrophic currents, it enables a qualitative comparison of the spatial structure of the Southern Ocean circulation. The MDT derived from the regional dataset (Fig. 4o) including ice-covered measurements from Armitage et al. (2018) demonstrates good agreement with the stream functions obtained from the reanalyses within the ACC. The WG structures also exhibit similarity across all ocean reanalyses and observations, as indicated by the zero stream function contours extending from 60°W to 120°E. The simulated RG size (defined here using zero stream function contour) is larger in ECCO-LLC270, extending from 160°E to 100°W, while ECCOv4r5 and GECCO3 show smaller RG extending from 160°E to 140°W. While the MDT derived from satellite altimetry confirms a larger RG, it does not account for deeper currents, water column height, or ageostrophic movement.

3.4 Weddell and Ross Gyre hydrography

We find qualitatively good agreement between ocean reanalyses and observations for the Southern Ocean. However, polar oceanographers and glaciologists care about modest changes in potential temperature as on-shelf warming by 0.5 or 1°C can substantially enhance ice shelf melting and impact abyssal overturning circulation (e.g., Dutrieux et al., 2014; Li et al., 2023). Here, we further evaluate MITgcm-based ocean reanalyses by focusing on detailed structures and time evolution in the open-water region and comparing them against in-situ observations.

3.4.1 Weddell Gyre hydrography: mean state

We use the repeatedly measured oceanographic sections from Joinville Island toward Kapp Norvegia (Section SR04 in Fig. 1a). Based on observations, we find four main water masses: Antarctic Surface Water (ASW), Warm Deep Water (WDW), Weddell Sea Deep Water (WSDW), and Weddell Sea Bottom Water (WSBW) as shown in Figs. 12i-j and 13i-j (e.g., Vernet et al., 2019). For example, in the central part of the WG, the Antarctic Surface Water is found at the top 50-100 m, characterized by a cold (colder than 0°C) and fresh (fresher than 34) layer. The WDW is located above the 0°C isotherms, WSDW is located between 0°C and −0.7°C, and WSBW is located below −0.7°C (Figs. 12-13). At the southwestern edge of the section over the continental shelf, we find a thickened layer of ASW. Over the slope, we find a warm core of WDW (or mCDW) flowing towards the Filchner-Ronne Ice Shelf (see red circles in Figs. 12i-j). At the northwestern edge of this section, the thickening of the ASW over the continental shelf is also observed. Over the continental slope, we find a layer of cold (< −1°C) and

fresh (<34.65) water over the continental slope (blue ellipses in Figs. 12i-j and 13i-j, S3), which is the newly formed Antarctic Bottom Water originating from the Filchner-Ronne Ice Shelf.

For the ocean reanalyses along the same section, all reanalyses show similar hydrographic characteristics. They all show (a) thick ASW over the continental shelf and (b) cores of WDW at the southwestern side of the section (Figs. 12-13). None of the simulations, however, show a layer of cold and freshwater over the continental slope at the northwestern side of this section, because none of the ocean reanalyses resolve the downslope descent of dense shelf water and formation of AABW from the Weddell Sea (Figs. 12-13, S3).

3.4.2 Weddell Gyre's hydrography time evolution

For most global ocean simulations without hundreds of years of spin-up, the simulated changes in the deep ocean are mostly caused by model drifts, a phenomenon well-recognized among ocean modelers (Rahmstorf, 1995). As the model's deep ocean circulation, mixing, diffusion, etc., are different from the real ocean, the model's state gradually shifts from the observed quasi-steady state to the model's steady state. Thus, we need to pay extra attention when analyzing time evolution at depths. When users want to discuss deep ocean changes, it is imperative to verify that (1) the magnitude of simulated changes exceeds that of model drift and (2) the identified changes stem from other external forcing mechanisms. In this section, we demonstrate the influence of model drifts for the Weddell Gyre section by comparing it with repeated observations.

Strass et al. (2020) studied the observed time evolution of deep water characteristics in the Weddell Gyre (WG) and found warming trends at most sampling locations. The mean warming rate is about ∼0.002°C per year below ~~2000~~ 3000 m, which exceeds that of the global ocean by a factor of about 5. We find a similar level of warming by comparing observations in 1989 and 2010 (Table ~~3~~ 4). Salinity also increases at most sites below 700 m. This salinity increase is not strong enough to fully compensate for the warming effect on seawater density and deep water masses show a general density decrease as shown in Strass et al. (2020). Despite bottom water property changes attracting scientific attention, these observed changes in temperature and salinity in the deep Weddell Gyre are small and water mass characteristics are highly stable. For example, in comparing observed sections for 1989, 1995, 2005, and 2010, it is difficult to spot differences for both temperature and salinity (Figs. 12i, 13i). When comparing temperature and salinity for 2010 and 1989, the observed temperature and salinity trends at the center of the WG below 2000 m are about ∼0.0019°C per year and ∼0.0001 per year, consistent with Strass et al. (2020).

The simulated WG sections show all observed water masses (ASW, WDW, WSDW, and WSBW) for the entire simulated periods for all ocean reanalyses (Figs. 12, 13). Year 1 shows excellent agreement with ocean reanalyses and observations since all the reanalyses are initialized with interpolated ocean observations such as the World Ocean Atlas (Locarnini et al., 2018; Zweng et al., 2019) with some adjustments by adjoint methods. However, all the reanalyses simulate larger changes than the observations. Such changes in deep watermasses can also be seen in the entire deep Southern Ocean as shown in Fig. 8.

For simulated potential temperature, ECCOv4r5 and GECCO3 show warming of WDW (located roughly at 1500-2000 m) and WSBW (located roughly below 3000 m) as shown in Fig. 12 and Table ~~3.~~ 4. The simulated warming of WDW in these two ocean reanalyses exceeds 0.0084°C per year (three times the observed value), while the observed mean temperature increase between ~~1000-2000~~ 1500-2000 m is 0.0028 °C per year (Fig. 12 and Table ~~3~~ 4). ECCO-LLC270 shows moderate warming but

340 the cores of WDW inflow along the continental slope show a maximum warming rate of 0.01 °C per year. B-SOSE (iterations 105 and 139) exhibits complex patterns with both warming and cooling and their simulated maximum and minimum trends are much larger than observations (Figs. 12e-f and ~~S3~~S4). Thus, in all reanalyses, simulated water mass changes for both WDW and WSBW are much larger than observations.

The simulated ECCOv4r5 and ECCO-LLC270 salinity increases below the halocline. These rates for both WDW and WSBW
345 are over 10 times larger than observations (Fig. 12 and Table ~~3~~4). B-SOSE demonstrates slightly different behaviors with some patches of freshening close to the coast but deep water shows a large salinity increase similar to other ocean reanalyses (Figs. 13, ~~S4, Table 3~~S5, Table 4). GECCO presents the best agreement with the least ~~amount of~~ salinity changes. Since observations show that the temperature, salinity, and thus density of the deep WS water masses remain relatively stable, ocean reanalyses fail to reproduce realistic changes. They all show excessive temperature and/or salinity changes (Table ~~3~~4).

350 The Hovmöller diagrams (Fig. 14) show simulated changes in water mass characteristics for the middle of the WG (see blue dot in Fig. 1 for the location). All ocean reanalyses considered in this study show gradual change, which is the model adjustment to the model steady state from the initial condition. Among all the ocean reanalyses considered in this study, ECCO-LLC270 shows the smallest temperature and salinity changes, which are still larger than observations (Table ~~3~~4). For ECCO-LLC270 temperature, we find simulated cooling above 1500 m and warming deeper than 1500 m. ECCO-LLC270 salinity increases
355 for all depths below the near-surface halocline by approximately 0.02 after 26 years. Other reanalyses show larger trends than that of ECCO-LLC270 (Table ~~3~~4). In ECCOv4r5, two warming peaks exist at depths of approximately 1000 m and 3000 m, representing the warming of WDW and AABW. B-SOSE simulates warming for all depths, with maximum warming below the mixed layer. For GECCO3, excessive warming can be found at most depths, and the maximum warming of 0.63°C is simulated at a depth of ~1700 m. The ECCOv4r5, B-SOSE, and GECCO3 salinity show freshening for depths shallower than
360 the halocline, and salinity increases below the halocline.

~~This section demonstrates that we need to pay extra attention when analyzing time evolution at depths. The simulated deep ocean changes are mostly caused by model drifts, a phenomenon well-recognized among ocean modelers. As the model's deep ocean circulation, mixing, diffusion, etc., are different from the real ocean, the model's state gradually shifts from the observed quasi-steady state to the model's steady states. In discussions concerning the drivers of changes in deeper ocean regions, it is~~
365 ~~imperative to verify that (1) the magnitude of simulated changes exceeds that of model drift and (2) the identified changes stem from other external forcing mechanisms.~~

3.4.3 Ross Gyre's hydrography mean state

We ~~utilized~~ utilize the S4P oceanographic section (Fig. 1) that traverses the Ross Gyre and the Pacific sector of the Southern Ocean. We identify three main water masses (Fig. 15): Antarctic Surface Water (ASW), CDW, and AABW. In the western
370 part of the section (corresponding to the Ross Gyre), we observe CDW or modified CDW at shallower depths and AABW at deeper depths. On the other hand, in the eastern part of the section, a warmer core of CDW is observed for the top 2000~~m~~ m, and we can find AABW below CDW (Fig. 15i). We compare the hydrographic characteristics of the ocean reanalyses along the same section and found that all simulations exhibited similar features (Figs. 15, 16). Similar to the Weddell Gyre section

(Figs. 12,13), none of the simulations show a layer of cold and freshwater over the continental slope at the western side of this section, because all MITgcm-based ocean reanalyses fail to simulate the descent of dense shelf water and the formation of AABW.

3.4.4 Ross Gyre hydrography time evolution

In this section, we present an additional example where temporal evolutions of the deep ocean are not in good agreement with observations. Specifically, observations from the Ross Sea exhibit more significant warming and freshening than the Weddell Sea. Thus, our example in this section highlights how both model drifts and external forcing influence deep ocean circulation and hydrography. It appears that model drifts predominantly drive the simulated changes, as most simulations consistently show changes that are substantially larger than observed changes.

According to Purkey and Johnson (2013); Purkey et al. (2019), observations of the S4P section suggest the Ross Sea has warmed and freshened below 2000 m at a depth mean rate of 0.002–0.004 °C year^{-1} and ~0.001 year^{-1}, respectively. Along S4P we find smaller changes between 1989 and 2018 (Figs. 15, 16 and Table 34) compared to observed changes (Purkey et al., 2019) because we also include the eastern part of the S4P section for our analyses where water masses presents smaller changes.

Here, we plot simulated RG sections for different years (1, 5, 10, and the final year) for each ocean reanalysis (Figs. 12-1315-16). Hovmöller diagrams (Fig. 1417) also show simulated changes in water mass characteristics for the middle of the RG (see blue dot Fig. 1 for the location). Similar to the Weddell Sea, all ocean reanalysis show gradual change to the model's stable state. The simulated sections display three water masses (ASW, CDW, and AABW), consistent with observations. Again, Year 1 demonstrated excellent agreement with observations similar to the Weddell Sea. For the modified CDW temperature changes (between 1500-2000 m), all most ocean reanalysis show larger changes than observations (Fig. 12 and Table 315 and Table 4). Potential temperature changes for B-SOSE (iter139) seem to show the best agreement with observations (Table 34), but this is an artifact of spatial averaging. B-SOSE (iter139) shows strong warming and cooling and spatial patterns do not agree with observations. For the AABW (below 3000 m), most ocean reanalyses show too strong warming. The simulated changes exceed observed changes except for B-SOSE (iter 105). Again, potential temperature changes for B-SOSE (iter105) seem to show the best agreement with observations (Table 34), but this is an artifact of spatial averaging with strong cooling and warming on the western and eastern sides, respectively (not shown). The reason for the simulated excessive warming is likely the model drift and the lack of AABW formation in these ocean reanalyses. For the modified CDW salinity changes, all simulation shows larger changes in magnitude compared to observations. For depths between 1500 and 2000 m, ECCOv4r5, ECCO-LLC270, and B-SOSE (iter139) simulations show a salinity increase, while B-SOSE (iter 105) and GECCO3 show freshening (Fig. 13 and Table 316 and Table 4). Deeper than 3000m m, most ocean reanalyses show increases in salinity, while observations show freshening.

3.5 Continental shelf

For the continental shelf regions of Antarctica, a few regions have been attracting attention in the past decades for the changes observed in Antarctic ice sheet loss. There are permanent features, such as mCDW intrusions into the West Antarctic ice shelves and the formation of ~~DSW~~ Dense Shelf Water (DSW) over the Ross and Weddell Sea, that never changed during the period of observations. Thus, we compare last-year slices of simulated on-shelf vertical sections of each ocean reanalysis in the Amundsen Sea, Weddell Sea, Ross Sea, and off the Totten Ice Shelf at the East Antarctic coast (Fig. 1) to check in the first order ~~,~~ if models can simulate such permanent features based on observations.

For the Amundsen Sea, we plot repeatedly measured vertical sections connecting open water to the Pine Island Ice Shelf since the first measurements in 1994 (e.g., Hellmer et al., 1998; Jacobs et al., 2011; Nakayama et al., 2013; Dutrieux et al., 2014). For the hydrography in the Amundsen Sea, we find mCDW and WW and past observations show that these two water masses are always present with little change in water mass properties (~~Fig. 18~~Figs. 18, 19). Change of the thermocline depth controls the melting of the eastern Amundsen Sea ice shelves Dutrieux et al. (2014); De Rydt et al. (2014). For the ocean reanalyses, ECCOLLC270 and GECCO3 show very shallow thermocline depth likely caused by outdated bathymetric datasets used for these simulations (also seen in Fig. 1), choices of atmospheric forcing, lack of ice shelf cavities, etc. ECCOv4r5 and B-SOSE show better agreement with observations for the Pine Island Trough but still show shallower thermocline depths and warmer mCDW compared to observations.

For the Ross Sea, we compare the vertical section roughly following the Ross Ice Shelf front. This section plotted is a composite of the observations in 1984, 1994, and 2007 (Jacobs and Giulivi, 2010). The observed sections were all made during the summer months. Under warm AASW, the western Ross Sea is filled with high salinity shelf water (HSSW), and the central section is characterized by the presence of MCDW inflow and ice shelf water (ISW) outflow. The modeled sections all suffer from flooding by CDW (note the change in color for temperature sections). The lack of HSSW in all datasets is likely attributed to the poorly simulated sea ice production rates, lack of ice shelf cavities for LLC270, SOSE, and GECCO, and over-estimated onshore flux of mCDW due to too large GM diffusivity over the continental slope (Dettling et al., 2023). Lastly, even though ECCOV4r5 has an open cavity, the ISW with observed values cannot be produced, as the appropriate source water mass is missing. None of the reanalysis reflects the freshening trend (e.g., Fig 5 in Jacobs and Giulivi, 2010).

For the Weddell Sea, we compare the vertical section roughly following the Filchner-Ronne ice shelf front. This section has been repeatedly observed many times and we select 1989, 1996, 2005, and 2010 for model-data comparison. The modeled sections are much closer to observations than in the Ross Sea, in temperature, salinity, and bathymetry, with the exception of B-SOSE. ECCO-LLC270 and GECCO3 show a modified Warm Deep Water (mWDW) core in the Western part of the section but do not have HSSW or ISW in the Filchner trough. ECCOv4r5 shows the opposite section structure: it is able to produce the colder water mass in the trough (ISW in case of ECCOv4r5 in the presence of open cavity) but does not show mWDW inflow (Janout et al., 2021). BSOSE ~~has shown~~ shows a thick fresh layer at the top of the water column ~~,~~ and a warm intrusion below it, similar to the Ross Sea.

For East Antarctica, we select a recently observed section off the Totten Ice Shelf. The first observational campaign was conducted in 2015 (Rintoul et al., 2016) and repeated a few times after that (Hirano et al., 2023; Nakayama et al., 2023). The section is constructed in the following way: left-side panels are along-front and right-side panels are away toward the continental shelf break (~~Fig~~Figs. 18, 19, fourth column). Based on observations (~~Fig~~Figs. 18t,19t), we find thick winter water (WW) and a thin layer of mCDW close to the bottom. In contrast to the Amundsen Sea, thick WW regulates both the volume and temperature of mCDW flowing into the Totten Ice Shelf cavity (Hirano et al., 2023). All reanalyses show too thick mCDW, and mCDW temperature is more than 1°C warmer than observations at the Totten Ice Shelf front. ECCO-LLC270 and GECCO3 show mCDW filling the continental shelf up to 200m.

We note that BSOSE iteration 139 differs significantly from iteration 105 close to the coast (Fig. ~~S3~~S6). While the later iteration removed a cold and fresh water column immediately adjacent to the front of PIG, the significant freshening of the top 200 m in the Ross and Weddell Seas has likely stopped DSW and HSSW production in both regions and led to the flooding of the continental shelves by CDW.

4 Discussion

4.1 Downscaling regional simulation for the Southern Ocean

In recent regional simulations of the Southern Ocean, ocean reanalyses have been used to construct lateral boundary conditions. For instance, Nakayama et al. (2017, 2018, 2021a) utilized ECCO-LLC270 to simulate West Antarctic and East Antarctic configurations of the regional model. Other researchers (e.g., Ito, 2022) used B-SOSE to construct their model boundary condition and study carbon ~~cycle~~ cycles in the Central Pacific Sector of the Southern Ocean.

As discussed in the previous section, all MITgcm-based ocean reanalyses simulate large-scale Southern Ocean hydrography and circulation patterns similar to observations in the open ocean, especially in the Weddell Sea. None of the models, however, accurately capture the time evolution of the Weddell and Ross Gyre hydrography, which ~~is~~ contains the most stable (not changing) and repeatedly measured ~~water mass~~ watermasses in the Southern Ocean. All ocean reanalyses show excessive warming trends. Moreover, all reanalyses fail to successfully simulate on-shelf hydrographic structures and ~~circulation~~circulations. Simulated on-shelf ~~condition is~~ conditions are different from observations (e.g., different water ~~mass~~ masses found on the shelf, missing water masses, etc).

Therefore, when using ocean reanalyses to force regional ocean simulation, we recommend that (1) northern regional model boundaries should be located away from the continental shelf in the area with enough observations and the east-west extent of the model should be large enough to minimize the effect of boundary conditions on the study region and (2) intensive model-data comparison should be repeated to evaluate the new model performance. Sometimes, we can locate the east and west model boundaries at locations where we find little inflow into the model domain (e.g., Amundsen-Bellingshausen domain (Nakayama et al., 2017, 2018). Then, we can minimize the influence of on-shelf model boundary conditions (ocean reanalyses on-shelf hydrographic properties). For example, by doing so, the Amundsen Sea and East Antarctic regional simulations (Nakayama et al., 2018) successfully capture interannual changes of on-shelf ocean heat intrusions. We caution that the choice of boundary

15

condition is crucial for studying near-surface processes due to the large spread of surface properties amongst ocean reanalyses (Fig. 5).

4.2 Common problems and possible solutions

There are several common problems in the simulation of Southern Ocean states that may be addressed through various solutions.

One major issue is that all ocean reanalyses struggle to accurately simulate the formation of Antarctic Bottom Water (AABW) and its precursor Dense Shelf Water and exhibit excessive warming of deep water (Figs. 9,11 12,14). One potential solution is to implement a bottom boundary layer scheme that can control the transport of dense shelf water to deep layers. Studies using coarse-resolution NOAA/GFDL simulations have shown that the implementation of implementing such a scheme can significantly improve the representation of AABW formation and deep water masses (Snow et al., 2015). Increasing horizontal and vertical resolution can also help address this problem (Mensah et al., 2021), but this may only be practical for regional simulations due to the massive CPU time required for global simulations with fine grid spacing (e.g., Stewart et al., 2019). Conducting data assimilation requires an additional 5-10 times more CPU hours/computational time. Modifying the weight for cost calculation is another possible solution, as the current version of ocean reanalyses has too few observations in the Southern Ocean and is weakly constrained. Applying larger weights for stable water masses based on existing observations or separating the cost into climatological mean and anomaly components can help put different weights on the mean state and time evolution, which has been implemented for ECCOv4r5 and for the last 11 iterations for GECCO3. An alternative solution could be to conduct data assimilation for shorter time windows (e.g., a few months) as is done for the California Current ocean state estimates (Zaba et al., 2018). This allows us to increase model controllability while still enforcing governing physics and capturing the continuous evolution of large-scale dynamics over weeks to months. However, it also means that the state estimates can not provide closed budgets and there exists discontinuity over longer timescales as discontinuities will exist between each assimilation window.

Another common problem is that seasonal and interannual variabilities of large-scale ocean circulation, such as Weddell Gyre (WG), Ross Gyre (RG), and Antarctic Circumpolar Current (ACC) strengths, do not agree among ocean reanalyses, despite using sea surface height data to constrain their models (Figs. 6-8 9-11). This issue can be addressed by implementing newly emerging datasets for the polar region, such as surface height anomaly datasets that include sea ice-covered regions (Armitage et al., 2018; Dotto et al., 2018; Auger et al., 2022a). These datasets can help optimize seasonal and interannual changes in sea surface height and thus barotropic ocean circulation changes.

Next, fresh bias in the SSS Sea Surface Salinity (SSS) in the Amundsen, Bellingshausen and Ross Seas in reanalyses needs to be taken into account when creating a regional model boundary condition that crosses the continental shelf (Fig. 5). The fresh water is confined to the Coastal Current circling the continent, thus the eastern boundary of the regional domain will import the fresh water into the domain. The regional domain may experience a warm SST bias as a result of due to the shallow mixed layer.

16

Lastly, none of the ocean reanalyses successfully simulate on-shelf hydrography close to observations (Figs. ~~15-16~~18-19).
Nakayama et al. (2021b) demonstrated that adjoint optimization can improve the representation of on-shelf hydrography if there are enough ~~data~~ observations to constrain the model. Oceanographic observations are accumulating, particularly over the Antarctic shelf region close to rapidly melting ice shelves, and can help constrain ocean models. Mooring observations at critical locations (e.g., Webber et al., 2017; Darelius et al., 2023) and massive datasets from ship-based and seal-tag CTD measurements can be particularly useful. More deployment of under-ice capability Argo floats are underway and new satellite-based datasets are emerging, such as the temporally varying Antarctic ice shelf melt rate estimates Adusumilli et al. (2020); Paolo et al. (2022). These newly emerging datasets can further help constrain the ocean reanalyses for polar oceans.

5 Conclusions

We assess the performance of MITgcm-based ocean reanalyses for the Southern Ocean. For the mean states, the reanalyses agree well with sea ice, observed hydrographic structures ~~and~~, ocean circulation in the open ocean ~~, including the~~ (e.g., Antarctic Circumpolar Current and ~~the~~ Weddell and Ross ~~Gyres~~Gyre circulations), and spatial distributions of MLD.

For the time evolution, MITgcm-based ocean reanalyses show good agreement with observations for sea ice concentration for both seasonal cycles and interannual variability. Sea ice volume is also consistent with observations despite the large errors of satellite-based estimates. MITgcm-based ocean reanalyses do not always agree with observations. We need to pay extra attention and we recommend that each user conduct their own evaluation to support your research goals. For example, most reanalyses successfully reproduce sea ice seasonal and interannual variability. However, there is no coherent time evolution of the ACC, WG, and RG strength, indicating a need for improvements in the representation of sea surface height and barotropic stream function for the Southern Ocean region. We also find excessive trends in the Weddell and Ross Gyres' deep water hydrography, which are concerning given their observed temperature and salinity do not change significantly over time.

For the continental shelf region, we do not recommend using global MITgcm-based ocean state estimates for both mean states and time evolution. All reanalyses fail to reproduce the observed state. For most regions, we find substantial differences (e.g., different water masses found on the shelf or missing water masses), which suggests that different physical processes govern ocean circulation and determine ~~hydrpgrahic~~ hydrographic structures. This is a concern for recent studies (Walker and Gardner, 2017; Bronselaer et al., 2018; Rignot et al., 2019; Brancato et al., 2020; Millan et al., 2020; Kim et al., 2021) that utilize ocean reanalyses for studying ocean melting ice shelves, as they may not yet have the skill to produce meaningful output. Development of downscaling simulation of MITgcm-based ocean reanalyses is a solution for continental shelf region and we can achieve much model-data agreement including both mean states and time evolution (Nakayama et al., 2017; Taewook et al., 2024).

Overall, while the MITgcm-based reanalyses generally perform well globally, there is room for improvement in capturing variability in the open ocean and mean states over the Antarctic continental shelves.

Author contributions. Y.N. conceived the study. Y.N. and A.M. conducted the model and data analyses and wrote the initial draft of the paper.
Y.N., A.M., M.A., H.Z., O.W., Y.N., I.F., M. M., A. K., and D.M. discussed the results and implications and commented on the manuscript
at all stages.

Competing interests. The authors declare no competing interests.

6 Acknowledgments

References

560 Adusumilli, S., Fricker, H. A., Medley, B., Padman, L., and Siegfried, M. R.: Interannual variations in meltwater input to the Southern Ocean from Antarctic ice shelves, Nature Geoscience, 13, 616–620, 2020.

Amante, C. and Eakins, B. W.: ETOPO1 arc-minute global relief model: procedures, data sources and analysis, 2009.

Armitage, T. W., Kwok, R., Thompson, A. F., and Cunningham, G.: Dynamic topography and sea level anomalies of the Southern Ocean: Variability and teleconnections, Journal of Geophysical Research: Oceans, 123, 613–630, 2018.

565 Arndt, J. E., Schenke, H. W., Jakobsson, M., Nitsche, F. O., Buys, G., Goleby, B., Rebesco, M., Bohoyo, F., Hong, J., Black, J., et al.: The International Bathymetric Chart of the Southern Ocean (IBCSO) Version 1.0 A new bathymetric compilation covering circum-Antarctic waters, Geophysical Research Letters, 40, 3111–3117, 2013.

Assmann, K., Jenkins, A., Shoosmith, D., Walker, D., Jacobs, S., and Nicholls, K.: Variability of circumpolar deep water transport onto the Amundsen Sea continental shelf through a shelf break trough, Journal of Geophysical Research, 118, 6603–6620, 2013.

570 Auger, M., Prandi, P., and Sallée, J.-B.: Southern ocean sea level anomaly in the sea ice-covered sector from multimission satellite observations, Scientific Data, 9, 70, https://doi.org/10.1038/s41597-022-01166-z, number: 1 Publisher: Nature Publishing Group, 2022a.

Auger, M., Sallée, J.-B., Prandi, P., and Naveira Garabato, A. C.: Subpolar Southern Ocean Seasonal Variability of the Geostrophic Circulation From Multi-Mission Satellite Altimetry, Journal of Geophysical Research: Oceans, 127, e2021JC018 096, 2022b.

Bailey, S. T., Jones, C. S., Abernathey, R. P., Gordon, A. L., and Yuan, X.: Water mass transformation variability in the Weddell Sea in ocean
575 reanalyses, Ocean Science, 19, 381–402, 2023.

Beckmann, A., Hellmer, H. H., and Timmermann, R.: A numerical model of the Weddell Sea: Large-scale circulation and water mass distribution, Journal of Geophysical Research: Oceans, 104, 23 375–23 391, 1999.

Brancato, V., Rignot, E., Milillo, P., Morlighem, M., Mouginot, J., An, L., Scheuchl, B., Jeong, S., Rizzoli, P., Bueso Bello, J. L., et al.: Grounding line retreat of Denman Glacier, East Antarctica, measured with COSMO-SkyMed radar interferometry data, Geophysical
580 Research Letters, 47, e2019GL086 291, 2020.

Bronselaer, B., Winton, M., Griffies, S. M., Hurlin, W. J., Rodgers, K. B., Sergienko, O. V., Stouffer, R. J., and Russell, J. L.: Change in future climate due to Antarctic meltwater, Nature, 564, 53–58, 2018.

Carroll, D., Menemenlis, D., Adkins, J., Bowman, K., Brix, H., Dutkiewicz, S., Fenty, I., Gierach, M., Hill, C., Jahn, O., et al.: The ECCO-Darwin data-assimilative global ocean biogeochemistry model: Estimates of seasonal to multidecadal surface ocean pCO2 and air-sea
585 CO2 flux, Journal of Advances in Modeling Earth Systems, 12, e2019MS001 888, 2020.

Cerovečki, I. and Haumann, F. A.: Decadal reorganization of Subantarctic Mode Water, Geophysical Research Letters, 50, e2022GL102 148, 2023.

Chen, Y., Zhang, Z., Wang, X., Liu, X., and Zhou, M.: Interannual variations of heat budget in the lower layer of the eastern Ross Sea shelf and the forcing mechanisms in the Southern Ocean State Estimate, International Journal of Climatology, 2023.

590 Chu, P. C. and Fan, C.: An inverse model for calculation of global volume transport from wind and hydrographic data, Journal of Marine Systems, 65, 376–399, 2007.

Cisewski, B., Strass, V. H., and Leach, H.: Circulation and transport of water masses in the Lazarev Sea, Antarctica, during summer and winter 2006, Deep Sea Research Part I: Oceanographic Research Papers, 58, 186–199, 2011.

Darelius, E., Daae, K., Dundas, V., Fer, I., Hellmer, H. H., Janout, M., Nicholls, K. W., Sallée, J.-B., and Østerhus, S.: Observational evidence
595 for on-shelf heat transport driven by dense water export in the Weddell Sea, Nature Communications, 14, 1022, 2023.

De Rydt, J., Holland, P., Dutrieux, P., and Jenkins, A.: Geometric and oceanographic controls on melting beneath Pine Island Glacier, Journal of Geophysical Research: Oceans, 119, 2420–2438, 2014.

De Verdière, A. C. and Ollitrault, M.: A direct determination of the World Ocean barotropic circulation, Journal of Physical Oceanography, 46, 255–273, 2016.

600   Dettling, N., Losch, M., Pollmann, F., and Kanzow, T.: Towards parameterizing eddy-mediated transport of Warm Deep Water across the Weddell Sea continental slope, 2023.

Dinniman, M. S., Klinck, J. M., and Smith Jr, W. O.: A model study of Circumpolar Deep Water on the West Antarctic Peninsula and Ross Sea continental shelves, Deep Sea Res. II, 58, 1508–1523, 2011.

Dinniman, M. S., Klinck, J. M., Bai, L.-S., Bromwich, D. H., Hines, K. M., and Holland, D. M.: The effect of atmospheric forcing resolution
605   on delivery of ocean heat to the Antarctic floating ice shelves, Journal of Climate, 28, 6067–6085, 2015.

Donohue, K., Tracey, K., Watts, D., Chidichimo, M. P., and Chereskin, T.: Mean antarctic circumpolar current transport measured in Drake Passage, Geophysical Research Letters, 43, 11–760, 2016.

Dotto, T. S., Naveira Garabato, A., Bacon, S., Tsamados, M., Holland, P. R., Hooley, J., Frajka-Williams, E., Ridout, A., and Meredith, M. P.: Variability of the Ross Gyre, Southern Ocean: Drivers and responses revealed by satellite altimetry, Geophysical Research Letters, 45,
610   6195–6204, 2018.

Dutrieux, P., De Rydt, J., Jenkins, A., Holland, P. R., Ha, H. K., Lee, S. H., Steig, E. J., Ding, Q., Abrahamsen, E. P., and Schröder, M.: Strong sensitivity of Pine Island ice-shelf melting to climatic variability, Science, 343, 174–178, 2014.

Fahrbach, E., Knoche, M., and Rohardt, G.: An estimate of water mass transformation in the southern Weddell Sea, Marine Chemistry, 35, 25–44, 1991.

615   Fahrbach, E., Hoppema, M., Rohardt, G., Schröder, M., and Wisotzki, A.: Decadal-scale variations of water mass properties in the deep Weddell Sea, Ocean Dynamics, 54, 77–91, 2004.

Fecher, T., Pail, R., Gruber, T., GOCO Consortium University of Bonn, I. o. G., Schall, G. W. S. J. K. J. B. I. L. S. M. A. E. J., Graz University of Technology, I. o. G. T. M.-G. A. K. B. K. D. R. N. Z., Austrian Academy of Sciences, S. R. I. O. B. E. H. S. K., and University of Bern, A. I. A. J. U. M. L. P. A. M.: GOCO05c: a new combined gravity field model based on full normal equations and
620   regionally varying weighting, Surveys in geophysics, 38, 571–590, 2017.

Feng, Y., Menemenlis, D., Xue, H., Zhang, H., Carroll, D., Du, Y., and Wu, H.: Improved representation of river runoff in Estimating the Circulation and Climate of the Ocean Version 4 (ECCOv4) simulations: implementation, evaluation, and impacts to coastal plume regions, Geoscientific Model Development, 14, 1801–1819, 2021.

Foldvik, A., Gammelsrød, T., Østerhus, S., Fahrbach, E., Rohardt, G., Schröder, M., Nicholls, K. W., Padman, L., and Woodgate, R.: Ice
625   shelf water overflow and bottom water formation in the southern Weddell Sea, Journal of Geophysical Research: Oceans, 109, 2004.

Forget, G., Campin, J.-M., Heimbach, P., Hill, C. N., Ponte, R. M., and Wunsch, C.: ECCO version 4: An integrated framework for non-linear inverse modeling and global ocean state estimation, 2015.

Fretwell, P., Pritchard, H. D., Vaughan, D. G., Bamber, J. L., Barrand, N. E., Bell, R., Bianchi, C., Bingham, R. G., Blankenship, D. D., Casassa, G., Catania, G., Callens, D., Conway, H., Cook, A. J., Corr, H. F. J., Damaske, D., Damm, V., Ferraccioli, F., Forsberg, R., Fujita,
630   S., Gim, Y., Gogineni, P., Griggs, J. A., Hindmarsh, R. C. A., Holmlund, P., Holt, J. W., Jacobel, R. W., Jenkins, A., Jokat, W., Jordan, T., King, E. C., Kohler, J., Krabill, W., Riger-Kusk, M., Langley, K. A., Leitchenkov, G., Leuschen, C., Luyendyk, B. P., Matsuoka, K., Mouginot, J., Nitsche, F. O., Nogi, Y., Nost, O. A., Popov, S. V., Rignot, E., Rippin, D. M., Rivera, A., Roberts, J., Ross, N., Siegert, M. J., Smith, A. M., Steinhage, D., Studinger, M., Sun, B., Tinto, B. K., Welch, B. C., Wilson, D., Young, D. A., Xiangbin, C., and Zirizzotti,

A.: Bedmap2: improved ice bed, surface and thickness datasets for Antarctica, The Cryosphere, 7, 375–393, https://doi.org/10.5194/tc-7-375-2013, 2013.

Garnier, F., Fleury, S., Garric, G., Bouffard, J., Tsamados, M., Laforge, A., Bocquet, M., Fredensborg Hansen, R. M., and Remy, F.: Advances in altimetric snow depth estimates using bi-frequency SARAL and CryoSat-2 Ka–Ku measurements, The Cryosphere, 15, 5483–5512, 2021.

Garnier, F., Bocquet, M., Fleury, S., Bouffard, J., Tsamados, M., Remy, F., Garric, G., and Chenal, A.: Latest Altimetry-Based Sea Ice Freeboard and Volume Inter-Annual Variability in the Antarctic over 2003–2020, Remote Sensing, 14, 4741, 2022.

Gille, S. T.: Warming of the Southern Ocean since the 1950s, Science, 295, 1275–1277, 2002.

Gutierrez-Villanueva, M. O., Chereskin, T. K., and Sprintall, J.: Compensating transport trends in the Drake Passage frontal regions yield no acceleration in net transport, Nature Communications, 14, 7792, 2023.

Gwyther, D., Galton-Fenzi, B., Hunter, J., and Roberts, J.: Simulated melt rates for the Totten and Dalton ice shelves, Ocean Science, 10, 267–279, 2014.

Hammond, M. D. and Jones, D. C.: Freshwater flux from ice sheet melting and iceberg calving in the Southern Ocean, Geoscience Data Journal, 3, 60–62, https://doi.org/10.1002/gdj3.43, 2017.

Hellmer, H. and Olbers, D.: A two-dimensional model for the thermohaline circulation under an ice shelf, Antarctic Science, 1, 325–336, 1989.

Hellmer, H. H., Jacobs, S. S., and Jenkins, A.: Oceanic erosion of a floating Antarctic glacier in the Amundsen Sea, Ocean, Ice, and Atmosphere: Interactions at the Antarctic continental margin (S Jacobs, R Weiss, eds) Antarctic Research Series, AGU, Washington DC, USA, 75, 319–339, 1998.

Hellmer, H. H., Kauker, F., Timmermann, R., Determann, J., and Rae, J.: Twenty-first-century warming of a large Antarctic ice-shelf cavity by a redirected coastal current, Nature, 485, 225–228, 2012.

Hirano, D., Tamura, T., Kusahara, K., Fujii, M., Yamazaki, K., Nakayama, Y., Ono, K., Itaki, T., Aoyama, Y., Simizu, D., et al.: On-shelf circulation of warm water toward the Totten Ice Shelf in East Antarctica, Nature Communications, 14, 4955, 2023.

Hogg, A. M., Meredith, M. P., Chambers, D. P., Abrahamsen, E. P., Hughes, C. W., and Morrison, A. K.: Recent trends in the S outhern O cean eddy field, Journal of Geophysical Research: Oceans, 120, 257–267, 2015.

Holland, D. M. and Jenkins, A.: Modeling thermodynamic ice-ocean interactions at the base of an ice shelf, J. Phys. Oceanogr., 29, 1787–1800, 1999.

Hou, Y., Nie, Y., Min, C., Shu, Q., Luo, H., Liu, J., and Yang, Q.: Evaluation of Antarctic sea ice thickness and volume during 2003–2014 in CMIP6 using Envisat and CryoSat-2 observations, Environmental Research Letters, 19, 014 067, 2024.

Ito, T.: Development of the Regional Carbon Cycle Model in the Central Pacific Sector of the Southern Ocean, Journal of Advances in Modeling Earth Systems, 14, e2021MS002 757, 2022.

Ito, T. and Marshall, J.: Control of lower-limb overturning circulation in the Southern Ocean by diapycnal mixing and mesoscale eddy transfer, Journal of Physical Oceanography, 38, 2832–2845, 2008.

Jacobs, S. S. and Giulivi, C. F.: Large multidecadal salinity trends near the Pacific-Antarctic continental margin, J. Clim., 23, 4508–4524, 2010.

Jacobs, S. S., Amos, A. F., and Bruchhausen, P. M.: Ross Sea oceanography and Antarctic bottom water formation, in: Deep Sea Research and Oceanographic Abstracts, vol. 17, pp. 935–962, Elsevier, 1970.

Jacobs, S. S., Jenkins, A., Giulivi, C. F., and Dutrieux, P.: Stronger ocean circulation and increased melting under Pine Island Glacier ice shelf, Nature Geoscience, 4, 519–523, 2011.

Janout, M. A., Hellmer, H. H., Hattermann, T., Huhn, O., Sültenfuss, J., Østerhus, S., Stulic, L., Ryan, S., Schröder, M., and Kanzow, T.: FRIS Revisited in 2018: On the Circulation and Water Masses at the Filchner and Ronne Ice Shelves in the Southern Weddell Sea,

675   Journal of Geophysical Research: Oceans, 126, e2021JC017 269, https://doi.org/https://doi.org/10.1029/2021JC017269, e2021JC017269 2021JC017269, 2021.

Jenkins, A.: A one-dimensional model of ice shelf-ocean interaction, Journal of Geophysical Research: Oceans, 96, 20 671–20 677, 1991.

Jourdain, N. C., Mathiot, P., Merino, N., Durand, G., Le Sommer, J., Spence, P., Dutrieux, P., and Madec, G.: Ocean circulation and sea-ice thinning induced by melting ice shelves in the Amundsen Sea, Journal of Geophysical Research: Oceans, 122, 2550–2573, 2017.

680 Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and Von Storch, J.: Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model, Journal of Advances in Modeling Earth Systems, 5, 422–446, 2013.

Kim, T.-W., Yang, H. W., Dutrieux, P., Wåhlin, A. K., Jenkins, A., Kim, Y. G., Ha, H. K., Kim, C.-S., Cho, K.-H., Park, T., et al.: Interannual variation of modified circumpolar deep water in the Dotson-Getz Trough, West Antarctica, Journal of Geophysical Research: Oceans, 126,

685   e2021JC017 491, 2021.

Kiss, A. E., Hogg, A. M., Hannah, N., Boeira Dias, F., Brassington, G. B., Chamberlain, M. A., Chapman, C., Dobrohotoff, P., Domingues, C. M., Duran, E. R., et al.: ACCESS-OM2 v1. 0: a global ocean–sea ice model at three resolutions, Geoscientific Model Development, 13, 401–442, 2020.

Koenig, Z., Provost, C., Ferrari, R., Sennéchael, N., and Rio, M.-H.: Volume transport of the A ntarctic C ircumpolar C urrent: Production

690   and validation of a 20 year long time series obtained from in situ and satellite observations, Journal of Geophysical Research: Oceans, 119, 5407–5433, 2014.

Köhl, A.: Evaluating the GECCO3 1948–2018 ocean synthesis–a configuration for initializing the MPI-ESM climate model, Quarterly Journal of the Royal Meteorological Society, 146, 2250–2273, 2020.

Köhl, A.: GECCO3 Ocean Synthesis (41S6m) (Version 41S6m) [Data set], Quarterly J Royal Met Soc.,

695   https://doi.org/10.25592/uhhfdm.14187, 2024.

Kusahara, K. and Hasumi, H.: Pathways of basal meltwater from Antarctic ice shelves: A model study, J. Geophys. Res. Lett., 2014.

Li, Q., England, M. H., Hogg, A. M., Rintoul, S. R., and Morrison, A. K.: Abyssal ocean overturning slowdown and warming driven by Antarctic meltwater, Nature, 615, 841–847, 2023.

Liao, S., Luo, H., Wang, J., Shi, Q., Zhang, J., and Yang, Q.: An Evaluation of Antarctic Sea Ice Thickness from the Global Ice-Ocean

700   Modeling and Assimilation System based on In-situ and Satellite Observations, The Cryosphere, 16, 1807–1819,, 2022.

Locarnini, M., Mishonov, A., Baranova, O., Boyer, T., Zweng, M., Garcia, H., Seidov, D., Weathers, K., Paver, C., Smolyar, I., et al.: World ocean atlas 2018, volume 1: Temperature, 2018.

Mack, S. L., Dinniman, M. S., Klinck, J. M., McGillicuddy Jr, D. J., and Padman, L.: Modeling ocean eddies on Antarctica's cold water continental shelves and their effects on ice shelf basal melting, Journal of Geophysical Research: Oceans, 124, 5067–5084, 2019.

705 Mathiot, P., Jenkins, A., Harris, C., and Madec, G.: Explicit representation and parametrised impacts of under ice shelf seas in the z_* coordinate ocean model NEMO 3.6, Geoscientific Model Development, 10, 2849–2874, 2017.

Mazloff, M. R., Heimbach, P., and Wunsch, C.: An eddy-permitting Southern Ocean state estimate, Journal of Physical Oceanography, 40, 880–899, 2010.

Meier, W. N., Fetterer, F., Windnagel, A. K., and Stewart, J. S.: NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 4, https://doi.org/10.7265/efmz-2t65, 2021.

Mensah, V., Nakayama, Y., Fujii, M., Nogi, Y., and Ohshima, K. I.: Dense water downslope flow and AABW production in a numerical model: Sensitivity to horizontal and vertical resolution in the region off Cape Darnley polynya, Ocean Modelling, 165, 101 843, 2021.

Millan, R., St-Laurent, P., Rignot, E., Morlighem, M., Mouginot, J., and Scheuchl, B.: Constraining an ocean model under Getz Ice Shelf, Antarctica, using a gravity-derived bathymetry, Geophysical Research Letters, 47, e2019GL086 522, 2020.

Nakayama, Y.: ECCOv4r5 output analyzed in "Evaluation of MITgcm-based ocean reanalysis for the Southern Ocean", Zenodo, https://doi.org/10.5281/zenodo.10930853, 2024a.

Nakayama, Y.: ECCO-LLC270 output analyzed in "Evaluation of MITgcm-based ocean reanalysis for the Southern Ocean" (Part 1), Zenodo, https://doi.org/10.5281/zenodo.10934678, 2024b.

Nakayama, Y.: ECCO-LLC270 output analyzed in "Evaluation of MITgcm-based ocean reanalysis for the Southern Ocean" (Part 2), Zenodo, https://doi.org/10.5281/zenodo.10935131, 2024c.

Nakayama, Y.: SOSE output analyzed in "Evaluation of MITgcm-based ocean reanalysis for the Southern Ocean" (Part 1), Zenodo, https://doi.org/10.5281/zenodo.10935396, 2024d.

Nakayama, Y.: SOSE output analyzed in "Evaluation of MITgcm-based ocean reanalysis for the Southern Ocean" (Part 2-1), Zenodo, https://doi.org/10.5281/zenodo.13117970, 2024e.

Nakayama, Y. and Malyarenko, A.: Analysis scripts developed for "Evaluation of MITgcm-based ocean reanalysis for the Southern Ocean", Zenodo, https://doi.org/10.5281/zenodo.10958934, 2024.

Nakayama, Y., Schröder, M., and Hellmer, H. H.: From circumpolar deep water to the glacial meltwater plume on the eastern Amundsen Shelf, Deep Sea Research Part I: Oceanographic Research Papers, 77, 50–62, 2013.

Nakayama, Y., Ohshima, K. I., Matsumura, Y., Fukamachi, Y., and Hasumi, H.: A numerical investigation of formation and variability of Antarctic bottom water off Cape Darnley, East Antarctica, Journal of Physical Oceanography, 44, 2921–2937, 2014a.

Nakayama, Y., Timmermann, R., Rodehacke, C. B., Schröder, M., and Hellmer, H. H.: Modeling the spreading of glacial meltwater from the Amundsen and Bellingshausen Seas, Geophysical Research Letters, 41, 7942–7949, 2014b.

Nakayama, Y., Menemenlis, D., Schodlok, M., and Rignot, E.: Amundsen and B ellingshausen S eas simulation with optimized ocean, sea ice, and thermodynamic ice shelf model parameters, Journal of Geophysical Research: Oceans, 122, 6180–6195, 2017.

Nakayama, Y., Menemenlis, D., Zhang, H., Schodlok, M., and Rignot, E.: Origin of Circumpolar Deep Water intruding onto the Amundsen and Bellingshausen Sea continental shelves, Nature communications, 9, 1–9, 2018.

Nakayama, Y., Manucharayan, G., Kelin, P., Torres, H. G., Schodlok, M., Rignot, E., Dutrieux, P., and Menemenlis, D.: Pathway of Circumpolar Deep Water into Pine Island and Thwaites ice shelf cavities and to their grounding lines, Scientific Reports, 2019.

Nakayama, Y., Greene, C. A., Paolo, F. S., Mensah, V., Zhang, H., Kashiwase, H., Simizu, D., Greenbaum, J. S., Blankenship, D. D., Abe-Ouchi, A., et al.: Antarctic slope current modulates ocean heat intrusions towards Totten glacier, Geophysical Research Letters, 48, e2021GL094 149, 2021a.

Nakayama, Y., Menemenlis, D., Wang, O., Zhang, H., Fenty, I., and Nguyen, A. T.: Development of adjoint-based ocean state estimation for the Amundsen and Bellingshausen seas and ice shelf cavities using MITgcm–ECCO (66j), Geoscientific Model Development, 14, 4909–4924, 2021b.

745 Nakayama, Y., Wongpan, P., Greenbaum, J. S., Yamazaki, K., Noguchi, T., Simizu, D., Kashiwase, H., Blankenship, D. D., Tamura, T., and Aoki, S.: Helicopter-Based Ocean Observations Capture Broad Ocean Heat Intrusions Toward the Totten Ice Shelf, Geophysical Research Letters, 50, e2022GL097 864, 2023.

Naughten, K. A., Jenkins, A., Holland, P. R., Mugford, R. I., Nicholls, K. W., and Munday, D. R.: Modeling the influence of the Weddell polynya on the Filchner–Ronne ice shelf cavity, Journal of Climate, 32, 5289–5303, 2019.

750 Naughten, K. A., Holland, P. R., Dutrieux, P., Kimura, S., Bett, D. T., and Jenkins, A.: Simulated Twentieth-Century Ocean Warming in the Amundsen Sea, West Antarctica, Geophysical Research Letters, 49, e2021GL094 566, 2022.

Neme, J., England, M. H., Hogg, A. M., Khatri, H., and Griffies, S. M.: The Role of Bottom Friction in Mediating the Response of the Weddell Gyre Circulation to Changes in Surface Stress and Buoyancy Fluxes, Journal of Physical Oceanography, 54, 217–236, https://doi.org/10.1175/JPO-D-23-0165.1, publisher: American Meteorological Society Section: Journal of Physical Oceanography, 2023.

755 Newman, L., Heil, P., Trebilco, R., Katsumata, K., Constable, A., Van Wijk, E., Assmann, K., Beja, J., Bricher, P., Coleman, R., et al.: Delivering sustained, coordinated, and integrated observations of the Southern Ocean for global impact, Frontiers in Marine Science, 6, 433, 2019.

Nie, Y., Uotila, P., Cheng, B., Massonnet, F., Kimura, N., Cipollone, A., and Lv, X.: Southern Ocean sea ice concentration budgets of five ocean-sea ice reanalyses, Climate Dynamics, 59, 3265–3285, 2022.

760 Ohshima, K. I., Fukamachi, Y., Williams, G. D., Nihashi, S., Roquet, F., Kitade, Y., Tamura, T., Hirano, D., Herraiz-Borreguero, L., Field, I., et al.: Antarctic Bottom Water production by intense sea-ice formation in the Cape Darnley polynya, Nature Geoscience, 6, 235–240, 2013.

Orsi, A. H., Johnson, G. C., and Bullister, J. L.: Circulation, mixing, and production of Antarctic Bottom Water, Progress in Oceanography, 43, 55–109, 1999.

765 Paolo, F., Gardner, A., Greene, C., Nilsson, J., Schodlok, M., Schlegel, N., and Fricker, H.: Widespread slowdown in thinning rates of West Antarctic Ice Shelves, EGUsphere, pp. 1–45, 2022.

Park, Y.-H. and Gambéroni, L.: Large-scale circulation and its variability in the south Indian Ocean from TOPEX/POSEIDON altimetry, Journal of Geophysical Research: Oceans, 100, 24 911–24 929, 1995.

Purkey, S. G. and Johnson, G. C.: Antarctic Bottom Water warming and freshening: Contributions to sea level rise, ocean freshwater budgets,
770 and global heat gain, Journal of Climate, 26, 6105–6122, 2013.

Purkey, S. G., Johnson, G. C., Talley, L. D., Sloyan, B. M., Wijffels, S. E., Smethie, W., Mecking, S., and Katsumata, K.: Unabated bottom water warming and freshening in the South Pacific Ocean, Journal of Geophysical Research: Oceans, 124, 1778–1794, 2019.

Rahmstorf, S.: Climate drift in an ocean model coupled to a simple, perfectly matched atmosphere, Climate Dynamics, 11, 447–458, 1995.

Reeve, K. A., Boebel, O., Strass, V., Kanzow, T., and Gerdes, R.: Horizontal circulation and volume transports in the Weddell Gyre derived
775 from Argo float data, Progress in Oceanography, 175, 263–283, 2019.

Rignot, E., Mouginot, J., Scheuchl, B., van den Broeke, M., van Wessem, M. J., and Morlighem, M.: Four decades of Antarctic Ice Sheet mass balance from 1979–2017, Proceedings of the National Academy of Sciences, 116, 1095–1103, 2019.

Rintoul, S. R.: The global influence of localized dynamics in the Southern Ocean, Nature, 558, 209–218, 2018.

Rintoul, S. R., Silvano, A., Pena-Molino, B., van Wijk, E., Rosenberg, M., Greenbaum, J. S., and Blankenship, D. D.: Ocean heat drives
780 rapid basal melt of the Totten Ice Shelf, Science Advances, 2, e1601 610, 2016.

Ruan, X., Thompson, A. F., Flexas, M. M., and Sprintall, J.: Contribution of topographically generated submesoscale turbulence to Southern Ocean overturning, Nature Geoscience, 10, 840–845, 2017.

Rücker van Caspel, M.: The importance of the western Weddell Sea to Weddell Sea deep water formation, Ph.D. thesis, Dissertation, Bremen, Universität Bremen, 2016, 2016.

785 Ryan, S., Schröder, M., Huhn, O., and Timmermann, R.: On the warm inflow at the eastern boundary of the Weddell Gyre, Deep Sea Research Part I: Oceanographic Research Papers, 107, 70–81, 2016.

Sallée, J., Abrahamsen, E., Allaigre, C., Auger, M., Ayres, H., Badhe, R., Boutin, J., Brearley, J., de Lavergne, C., et al.: Southern ocean carbon and heat impact on climate, Philosophical Transactions of the Royal Society A, 381, 20220 056, 2023.

Schodlok, M., Menemenlis, D., and Rignot, E.: Ice shelf basal melt rates around Antarctica from simulations and observations, Journal of

790 Geophysical Research: Oceans, 2015.

Schodlok, M. P., Menemenlis, D., Rignot, E., and Studinger, M.: Sensitivity of the ice shelf ocean system to the sub-ice shelf cavity shape measured by NASA IceBridge in Pine Island Glacier, West Antarctica, Annals of Glaciology, 53, 156–162, 2012.

Smith, W. H. and Sandwell, D. T.: Global sea floor topography from satellite altimetry and ship depth soundings, Science, 277, 1956–1962, 1997.

795 Snow, K., Hogg, A. M., Downes, S. M., Sloyan, B. M., Bates, M. L., and Griffies, S. M.: Sensitivity of abyssal water masses to overflow parameterisations, Ocean Modelling, 89, 84–103, 2015.

Stewart, A. L., Klocker, A., and Menemenlis, D.: Acceleration and overturning of the Antarctic Slope Current by winds, eddies, and tides, Journal of Physical Oceanography, 49, 2043–2074, 2019.

Strass, V. H., Rohardt, G., Kanzow, T., Hoppema, M., and Boebel, O.: Multidecadal warming and density loss in the deep Weddell Sea,

800 Antarctica, Journal of Climate, 33, 9863–9881, 2020.

Taewook, P., Yoshihiro, N., and SungHyun, N.: Amundsen Sea Circulation Controls Bottom Upwelling and Antarctic Pine Island and Thwaites Ice Shelf Melting, Nature Communications, 15, 2024.

Takahashi, T., Sweeney, C., Hales, B., Chipman, D. W., Newberger, T., Goddard, J. G., Iannuzzi, R. A., and Sutherland, S. C.: The changing carbon cycle in the Southern Ocean, Oceanography, 25, 26–37, 2012.

805 Talley, L. D.: Hydrographic Atlas of the World Ocean Circulation Experiment (WOCE) Volume 2: Pacific Ocean, WOCE International Project Office Southampton, 2007.

Timmermann, R., Danilov, S., Schröter, J., Böning, C., Sidorenko, D., and Rollenhagen, K.: Ocean circulation and sea ice distribution in a finite element global sea ice–ocean model, Ocean Modelling, 27, 114–129, 2009.

Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthélemy, A., Bricaud, C., Carton, J., Fučkar, N., Garric, G., Iovino, D., et al.: An

810 assessment of ten ocean reanalyses in the polar regions, Climate Dynamics, 52, 1613–1650, 2019.

Verdy, A., Cornuelle, B., Mazloff, M. R., and Rudnick, D. L.: Estimation of the tropical Pacific Ocean state 2010–13, Journal of Atmospheric and Oceanic Technology, 34, 1501–1517, 2017.

Vernet, M., Geibert, W., Hoppema, M., Brown, P. J., Haas, C., Hellmer, H., Jokat, W., Jullion, L., Mazloff, M., Bakker, D., et al.: The Weddell Gyre, Southern Ocean: present knowledge and future challenges, Reviews of Geophysics, 57, 623–708, 2019.

815 Walker, C. and Gardner, A. S.: Rapid drawdown of Antarctica's Wordie Ice Shelf glaciers in response to ENSO/Southern Annular Mode-driven warming in the Southern Ocean, Earth and Planetary Science Letters, 476, 100–110, 2017.

Wang, J., Min, C., Ricker, R., Shi, Q., Han, B., Hendricks, S., Wu, R., and Yang, Q.: A comparison between Envisat and ICESat sea ice thickness in the Southern Ocean, The Cryosphere, 16, 4473–4490, 2022.

Wang, Z. and Meredith, M.: Density-driven Southern Hemisphere subpolar gyres in coupled climate models, Geophysical Research Letters,

820 35, 2008.

Webber, B. G., Heywood, K. J., Stevens, D. P., Dutrieux, P., Abrahamsen, E. P., Jenkins, A., Jacobs, S. S., Ha, H. K., Lee, S. H., and Kim, T. W.: Mechanisms driving variability in the ocean forcing of Pine Island Glacier, Nature communications, 8, 1–8, 2017.

Whitworth, T.: Monitoring the transport of the Antarctic circumpolar current at Drake Passage, Journal of Physical Oceanography, 13, 2045–2057, 1983.

Whitworth, T. and Peterson, R.: Volume transport of the Antarctic Circumpolar Current from bottom pressure measurements, Journal of Physical Oceanography, 15, 810–816, 1985.

Whitworth III, T., Nowlin Jr, W., and Worley, S.: The net transport of the Antarctic circumpolar current through Drake Passage, Journal of Physical Oceanography, 12, 960–971, 1982.

Williams, G., Bindoff, N., Marsland, S., and Rintoul, S.: Formation and export of dense shelf water from the Adélie Depression, East Antarctica, Journal of Geophysical Research: Oceans, 113, 2008.

Williams, R. G., Ceppi, P., Roussenov, V., Katavouta, A., and Meijers, A. J.: The role of the Southern Ocean in the global climate response to carbon emissions, Philosophical Transactions of the Royal Society A, 381, 20220 062, 2023.

Xu, X., Chassignet, E. P., Firing, Y. L., and Donohue, K.: 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, e2020JC016 365, 2020.

Zaba, K. D., Rudnick, D. L., Cornuelle, B. D., Gopalakrishnan, G., and Mazloff, M. R.: Annual and interannual variability in the California current system: Comparison of an ocean state estimate with a network of underwater gliders, Journal of Physical Oceanography, 48, 2965–2988, 2018.

Zhang, H., Menemenlis, D., and Fenty, I.: ECCO LLC270 Ocean-Ice State Estimate, 2018.

Zweng, M., Seidov, D., Boyer, T., Locarnini, M., Garcia, H., Mishonov, A., Baranova, O., Weathers, K., Paver, C., Smolyar, I., et al.: World ocean atlas 2018, volume 2: Salinity, 2019.
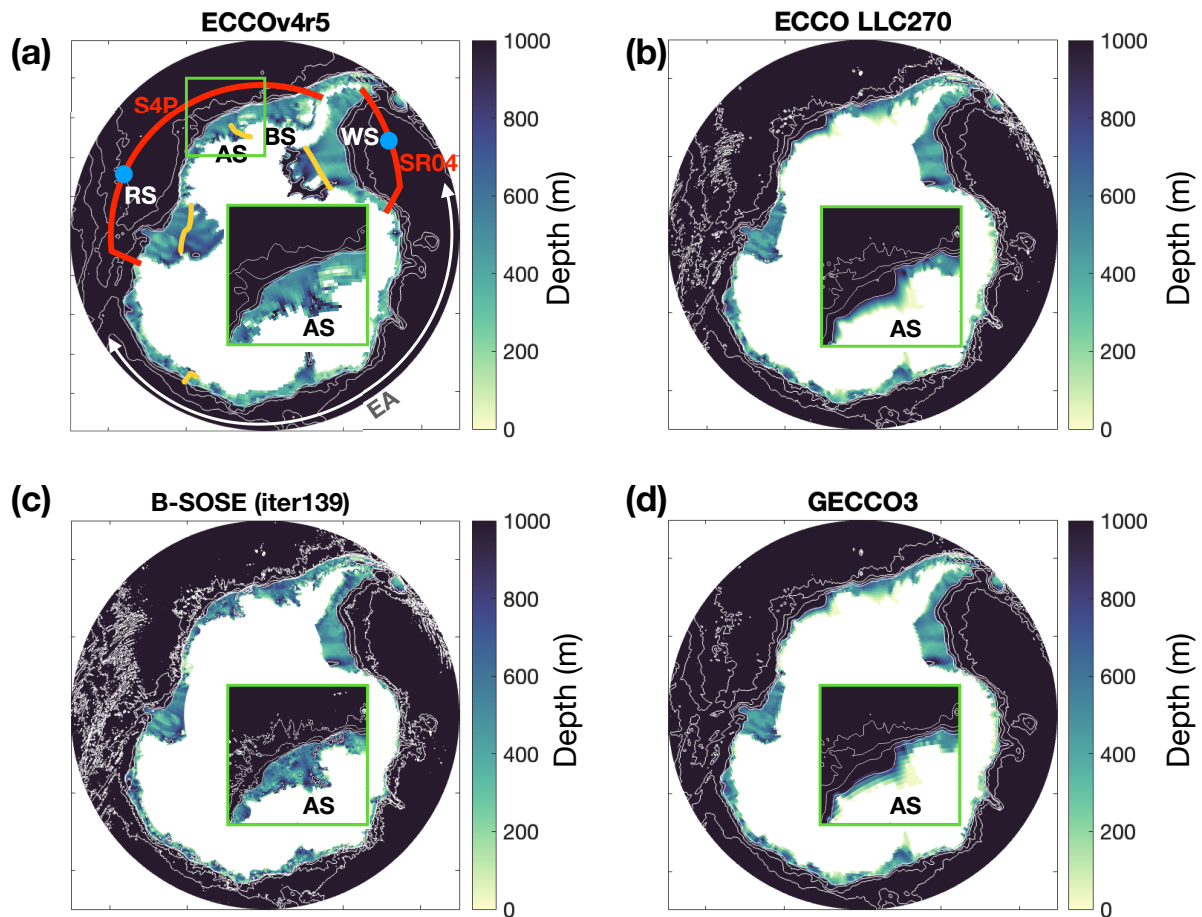
Figure 1. Model bathymetry (color) of (a) ECCOv4r5, (b) ECCO-LLC270, (c) B-SOSE, and (d) GECCO3 with contours of 1000 m, 2000 m, 3000 m, and 4000 m in white. Letters RS, AS, BS, WS, and EA denote the Ross Sea, Amundsen Sea, Bellingshausen Sea, Weddell Sea, and East Antarctica, respectively. For all panels, insets show close-ups of the Amundsen region enclosed by green boxes. Red lines show the locations of vertical sections SR04 and S4P. Orange lines show the location of vertical on-shelf sections shown in Figs. ~~15~~ 18 and ~~16.~~ 19. Blue dots indicate the location where Hovmöller diagrams are shown in Figs. ~~11~~ 14 and ~~14.~~ 17.
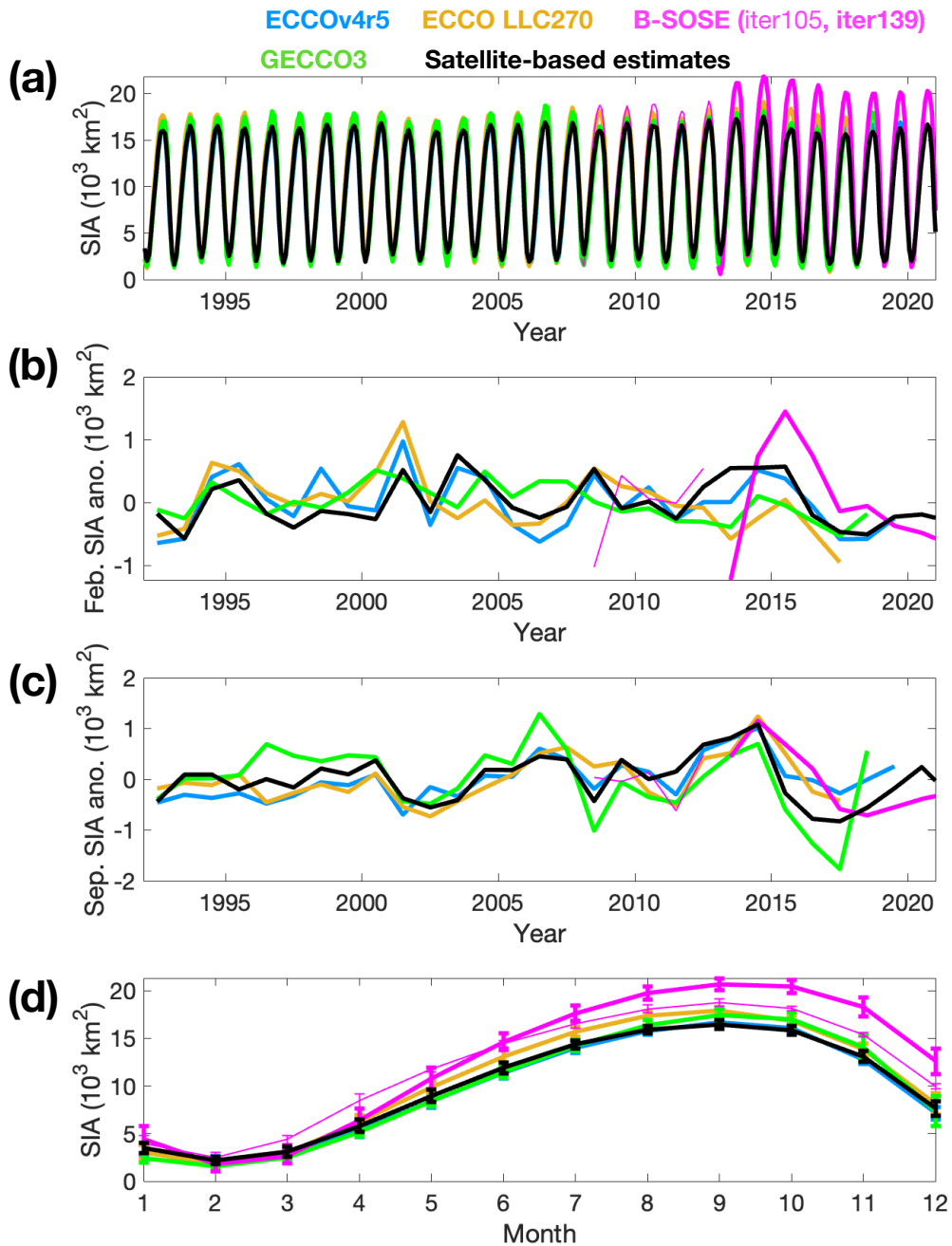
Figure 2. (a) Time series of ~~(a)~~ sea ice ~~concentration, area (b~~SIA)~~September sea ice extent~~, (b, c) ~~March sea ice extent~~SIA anomalies for February and September, and (d) ~~sea ice concentration~~ climatological SIA seasonal ~~climatology~~cycle. ECCOv4r5, ECCO-LLC270, B-SOSE (iter105), B-SOSE (iter139) GECCO3, and satellite-based estimates are shown in blue, ~~red~~orange, magenta (thin), magenta, green, and black, respectively. For (d), error bars represent one standard deviation of variability. ~~For~~ The 5-day SOSE outputs ~~, only 5-day means~~ are ~~available.~~ ~~Thus, 5-day means are plotted and a 30-day moving average is applied~~ averaged monthly for ~~all plots for B-SOSE outputs. For (b) and (c),~~ **28** ~~we plot 30-day means that nearly match~~ comparison with ~~September and March means, respectively~~other datasets.
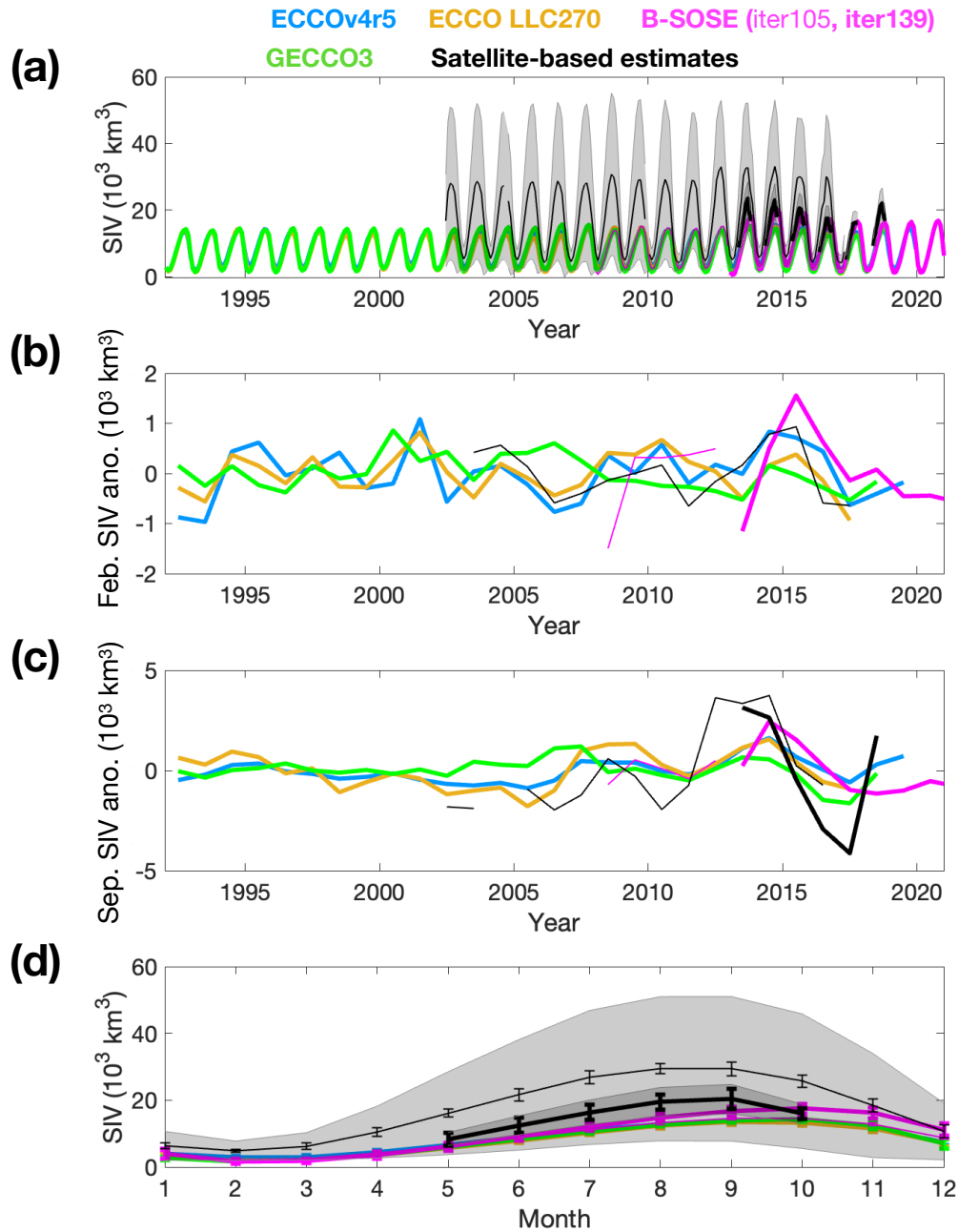
Figure 3. ~~(a) Time series of (a)~~ Same as Figure. 2 but for sea ice volume ~~, (b~~SIV)~~September sea ice volume, (c) March sea ice volume, and (d) sea ice volume seasonal climatology~~. ~~ECCOv4r5, ECCO-LLC270, B-SOSE, and GECCO3~~ Two observational data are ~~shown in blue~~plotted, ~~red, magenta,~~ thin and ~~green~~thick black lines indicate products from SICCI and LEGOS, respectively. ~~For~~ The shadings in ~~(c),~~ ~~error bars represent one standard deviation of variability. For B-SOSE outputs, only 5-day means are available. Thus, 5-day means are plotted and a~~30-day moving average is applied for all plots for B-SOSE outputs. For (b) and ~~(~~ed) ~~, we plot 30-day means that nearly match with September and March means, respectively~~represent the uncertainties of the observations.

29
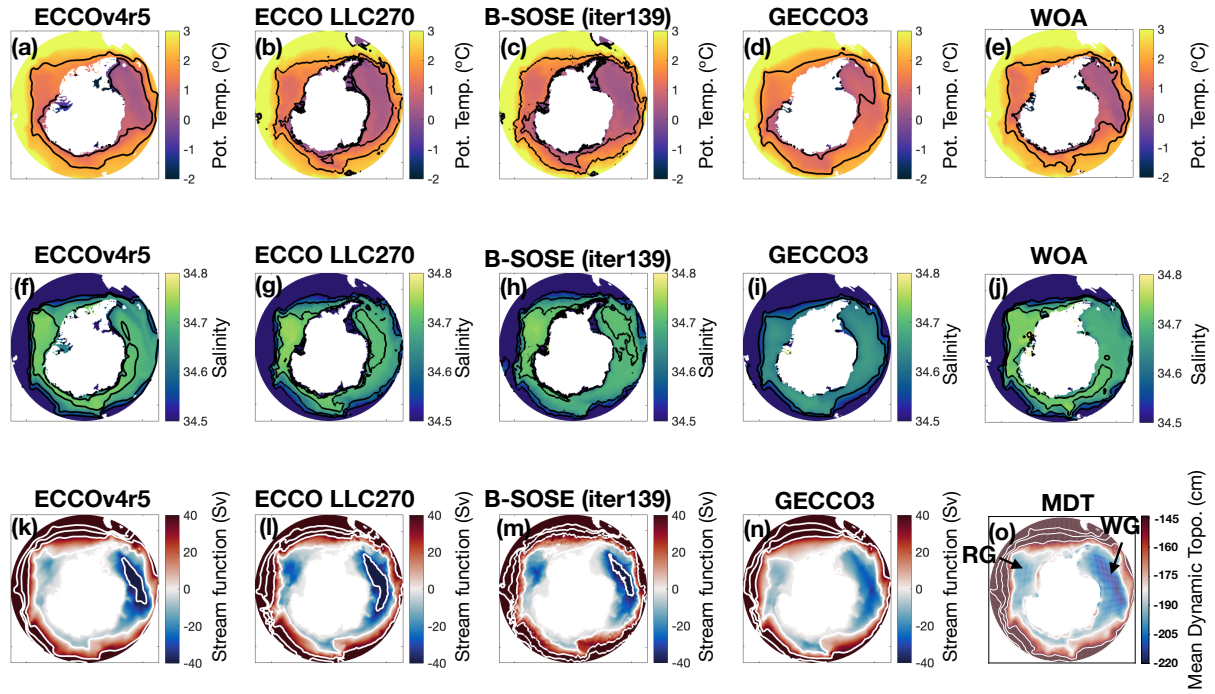
Figure 4. (a,~~g~~f,~~l~~k) ECCOv4r5, (b,~~h~~g,~~m~~l) ECCO-LLC270, (c,~~i~~h,~~n~~m) B-SOSE, (d,~~j~~i,~~o~~n) GECCO3 mean potential temperature and salinity, and stream function for the model simulated periods, respectively. (~~f~~e,~~k~~j) World Ocean Atlas climatology potential temperature and salinity, respectively. (~~p~~o) Mean dynamic topography from Armitage et al. (2018). WG and RG denote the Weddell Gyre and Ross Gyre, respectively. For (a)-(e), potential temperature contours of -1 °C, 0 °C, 1 °C, and 2 °C are shown. For (~~g~~f)-(j), salinity contours of 34.5, 34.6, 34.7, and 34.8 are shown. For (k)-(o), stream function contours of -40 Sv, 0 Sv, 40 Sv, 80 Sv, and 120 Sv are shown. As the vertical grids are slightly different, we extract potential temperature and salinity at 553 m, 553 m, 552 m, 560 m, and 550 m for ECCOv4r5, ECCO-LLC270, B-SOSE, GECCO3, and WOA, respectively.
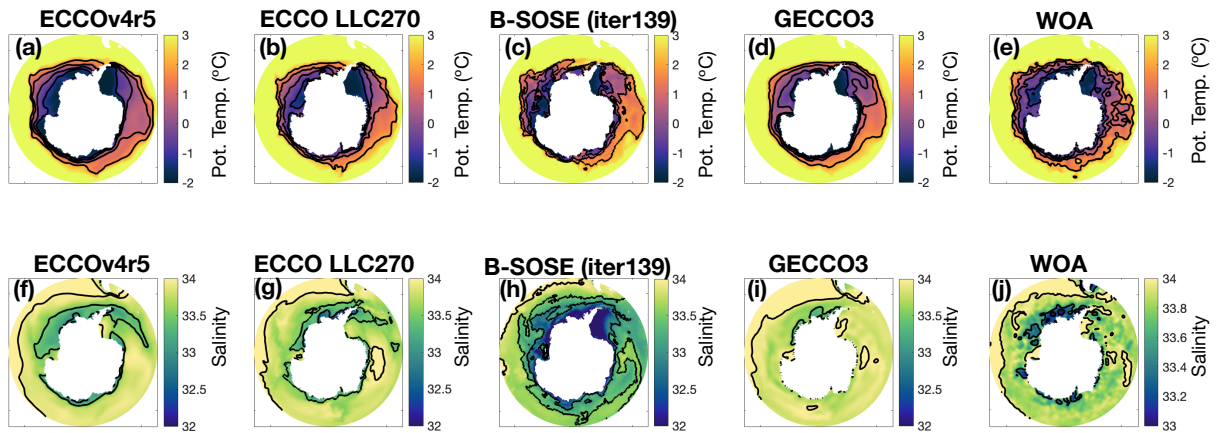
Figure 5. (a,f) ECCOv4r5, (b,g) ECCO-LLC270, (c,h) B-SOSE, (d,i) GECCO3 February mean surface potential temperature and salinity. (e,j) World Ocean Atlas climatology potential temperature and salinity, respectively. ~~Red arrows point to~~ Warm cores of WDW flowing toward the Filchner-Ronne Ice Shelf and layers of cold ($< -1°$C) and fresh ($< -1°$C) water over the continental slope are indicated by red and blue ellipses, respectively. For (a)-(e), potential temperature contours of -1 °C, 0 °C, 1 °C, and 2 °C are shown. For (f)-(j), salinity ~~bias~~ contours of 33.0, 33.5, and 34.0 are shown.
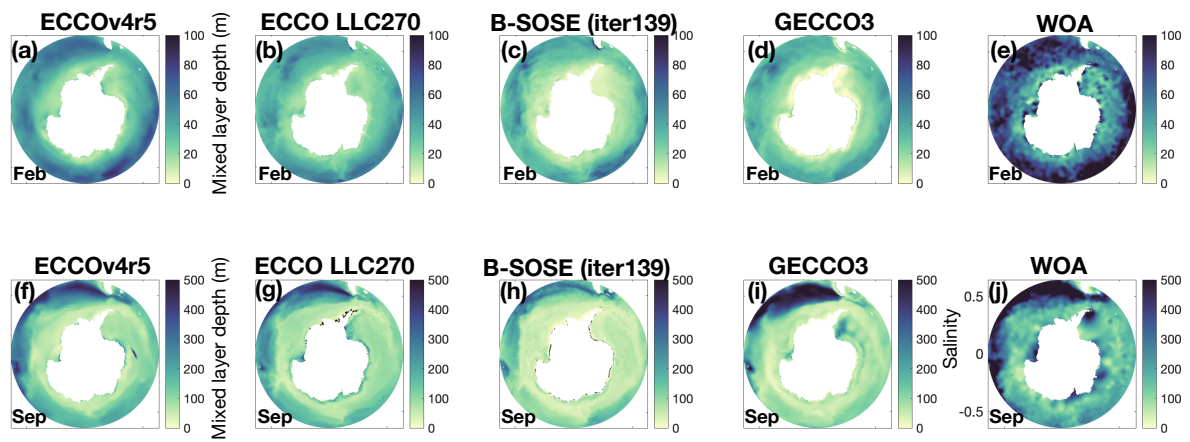
Figure 6. (a,f) ECCOv4r5, (b,g) ECCO-LLC270, (c,h) B-SOSE, (d,i) GECCO3 February and September mean MLD. MLD is defined as the depth level where the density difference from the surface exceeds 0.03 kg m^{-1}. (e,j) World Ocean Atlas monthly climatology of MLD for February and September, respectively.
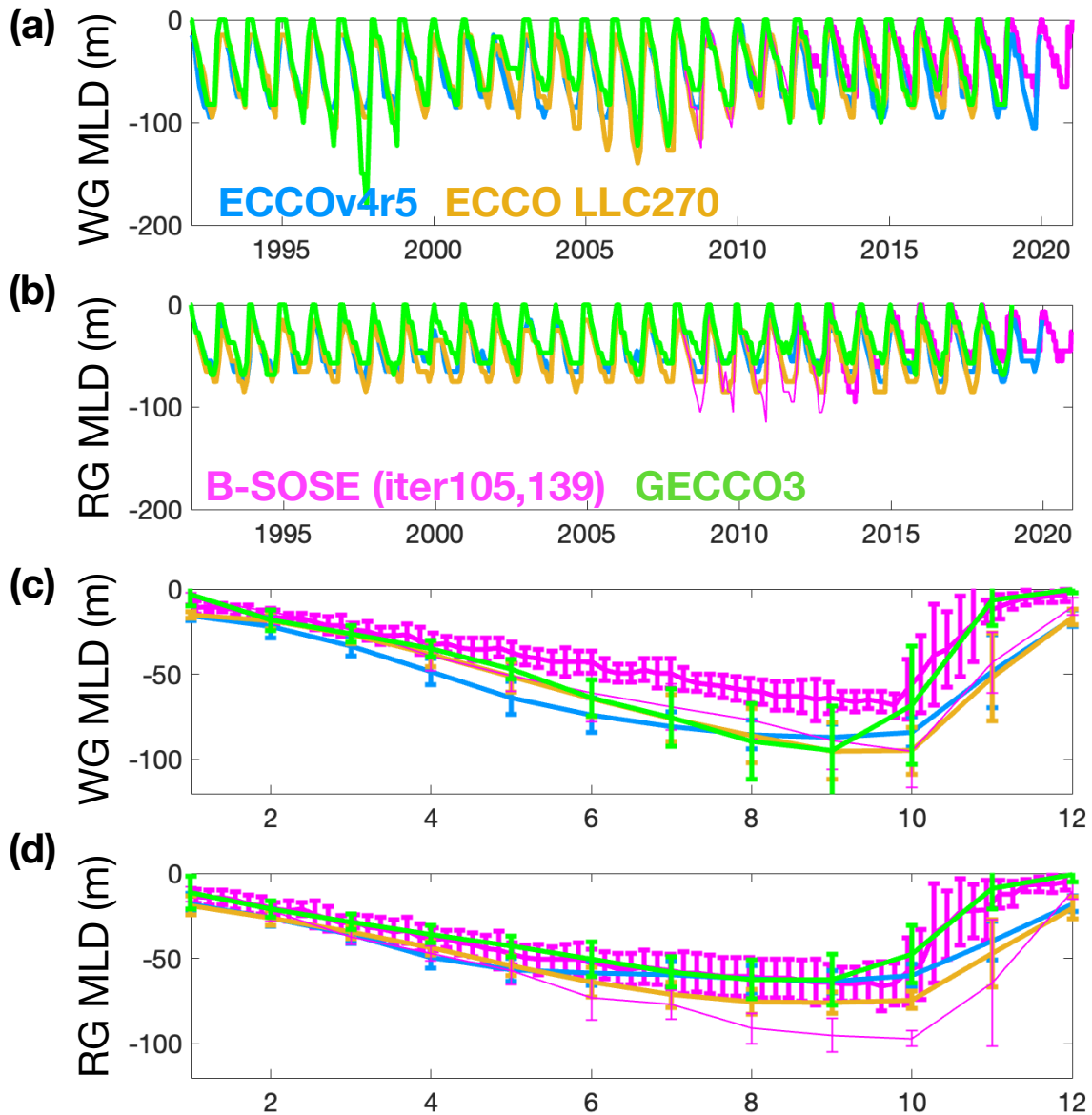
Figure 7. Time series of MLD for the ~~strength of~~ (a) ~~ACC,~~ Weddell Gyre and (b) ~~WG,~~ Ross Gyre for locations shown by blue dots in Figure 1. Climatological MLD seasonal cycle for (c) Weddell Gyre and (d) Ross Gyre. For (c,d)~~RG stream function~~, error bars represent one standard deviation of variability. ECCOv4r5, ECCO-LLC270, B-SOSE (iter105), B-SOSE (iter139), and GECCO3 are shown in blue, ~~red~~orange, magenta (thin), magenta and green, respectively.

~~Same as Fig. 9 but for ACC, WG, and RG anomalies.~~

Simulated SR04 vertical sections ~~RMSE~~ of (a~~)~~ ~~ECCOv4r5,(b)~~ ~~ECCO-LLC270,(c)~~ ~~B-SOSE,~~ potential temperature and (d~~)~~ ~~GECCO3~~ ~~potential temperature for year 1,year 5e,year 10f)~~ salinity at 500 m, 2000m, and ~~final years of~~ 4000 m depth levels, respectively. RMSE is calculated by interpolating each ~~reanalysis along SR04 section (Fig~~ocean state estimate output into a common horizontal grid with 1º resolutions. ~~1). B-SOSE only extends for 9 years and only years 1~~ECCOv4r5, ~~5~~ECCO-LLC270, ~~and 9 are plotted. B-SOSE~~ ~~(i~~iter105~~)Observed vertical section of potential temperature for 1989,~~ ~~1996~~B-SOSE (iter139), ~~2005, and~~ ~~2010 along the SR04 section.~~ ~~Potential temperature (b~~GECCO3 are shown in blue, ~~o~~orange, ~~f,h,j~~magenta ~~(thin)differences~~, magenta and ~~trends (defined as difference~~ green, respectively.

Simulated SR04 vertical sections ~~RMSE~~ of (a~~)~~ ~~ECCOv4r5,(b)~~ ~~ECCO-LLC270,(c)~~ ~~B-SOSE,~~ potential temperature and (d~~)~~ ~~GECCO3 potential temperature for year 1,year 5e,year 10f)~~ salinity at 500 m, 2000m, and ~~final years of~~ 4000 m depth levels, respectively. RMSE is calculated by interpolating each ~~reanalysis along SR04 section (Fig~~ocean state estimate output into a common horizontal grid with 1º resolutions. ~~1). B-SOSE only extends for 9 years and only years 1~~ECCOv4r5, ~~5~~ECCO-LLC270, ~~and 9 are plotted. B-SOSE~~ ~~(i~~iter105~~)Observed vertical section of potential temperature for 1989,~~ ~~1996~~B-SOSE (iter139), ~~2005, and~~ ~~2010 along the SR04 section. Potential temperature (b~~GECCO3 are shown in blue, ~~o~~orange, ~~f,h,j~~magenta ~~(thin)differences~~, magenta and ~~trends (defined as difference~~ green, respectively.

Figure 8. ~~Same as Fig. 9 but for ACC, WG, and RG seasonality. Error bars represent one standard deviation~~ Time series of ~~variability.~~ ~~Simulated SR04 vertical sections~~ RMSE of (a~~)~~ ~~ECCOv4r5,(b)~~ ~~ECCO-LLC270,(c)~~ ~~B-SOSE,~~ potential temperature and (d~~)~~ ~~GECCO3 potential~~ ~~temperature for year 1,year 5e,year 10f)~~ salinity at 500 m, 2000m, and ~~final years of~~ 4000 m depth levels, respectively. RMSE is calculated by interpolating each ~~reanalysis along SR04 section (Fig~~ocean state estimate output into a common horizontal grid with 1º resolutions. ~~1).~~ ~~B-SOSE only extends for 9 years and only years 1~~ECCOv4r5, ~~5~~ECCO-LLC270, ~~and 9 are plotted.~~ B-SOSE ~~(i~~iter105~~)Observed vertical~~ ~~section of potential temperature for 1989,~~ ~~1996~~B-SOSE (iter139), ~~2005,~~ and ~~2010 along the SR04 section. Potential temperature (b~~GECCO3 are shown in blue, ~~o~~orange, ~~f,h,j~~magenta ~~(thin)differences~~, magenta and ~~trends (defined as difference~~ green, respectively.
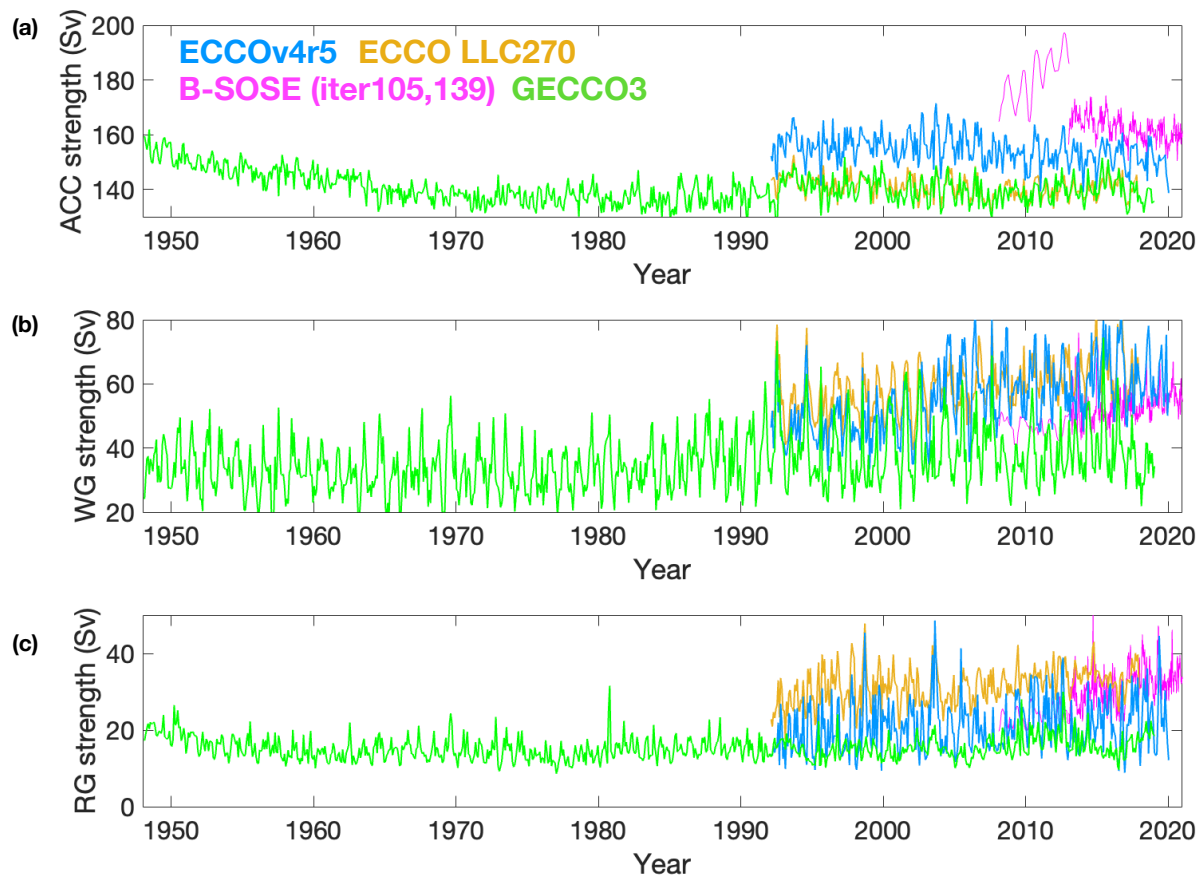
Figure 9. ~~Time series of the~~ first and last years divided by total model period). The potential temperature contours ~~strength~~ of ~~-0.7 °C and 0 °C are also in black for panels~~ (a) ACC, (b) WG, and (c) RG stream function. ECCOv4r5, ECCO-LLC270, B-SOSE (~~e~~iter105), B-SOSE (iter139), and GECCO3 are shown in blue, orange, magenta (~~g~~thin), magenta and green, respectively.
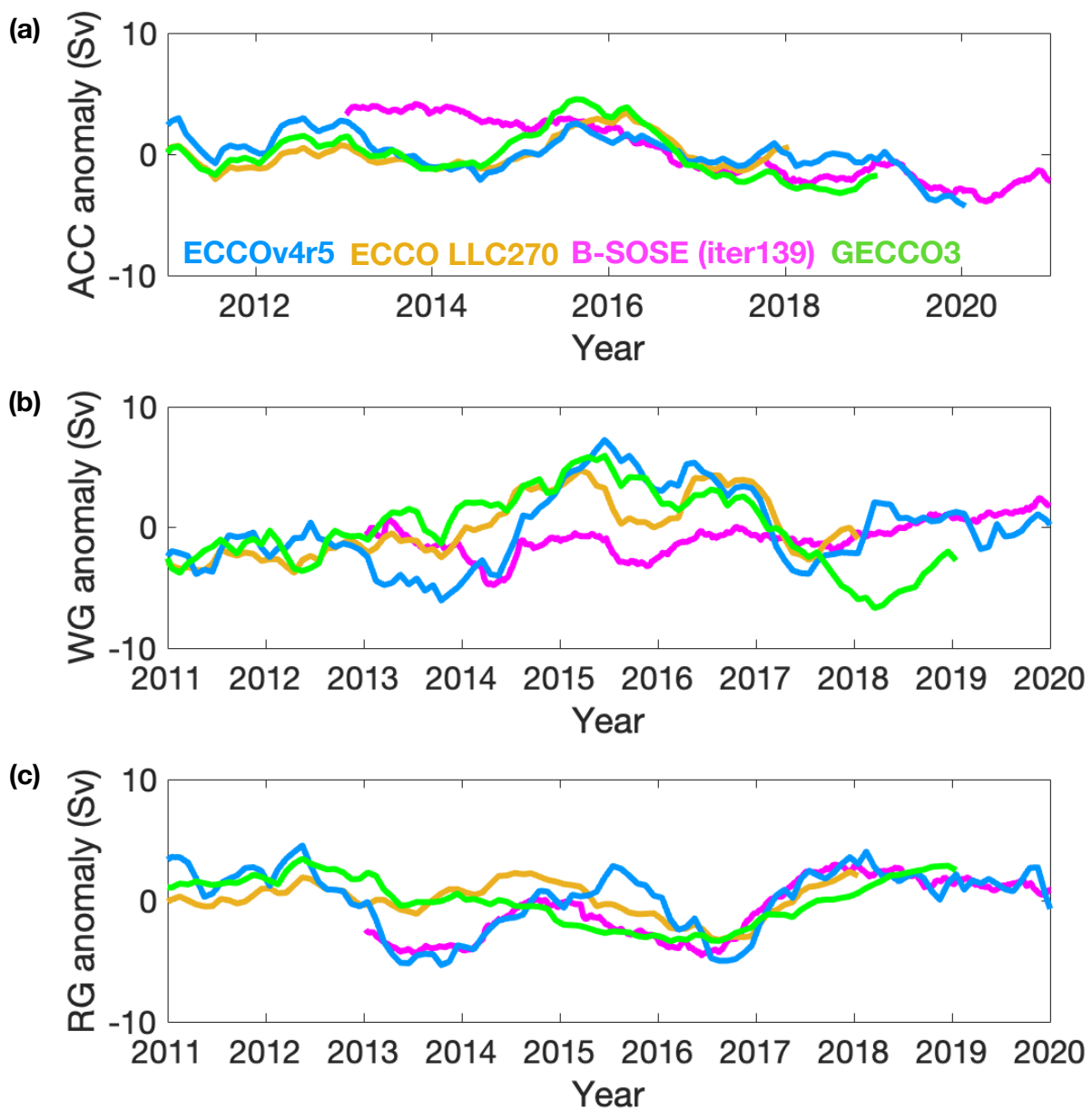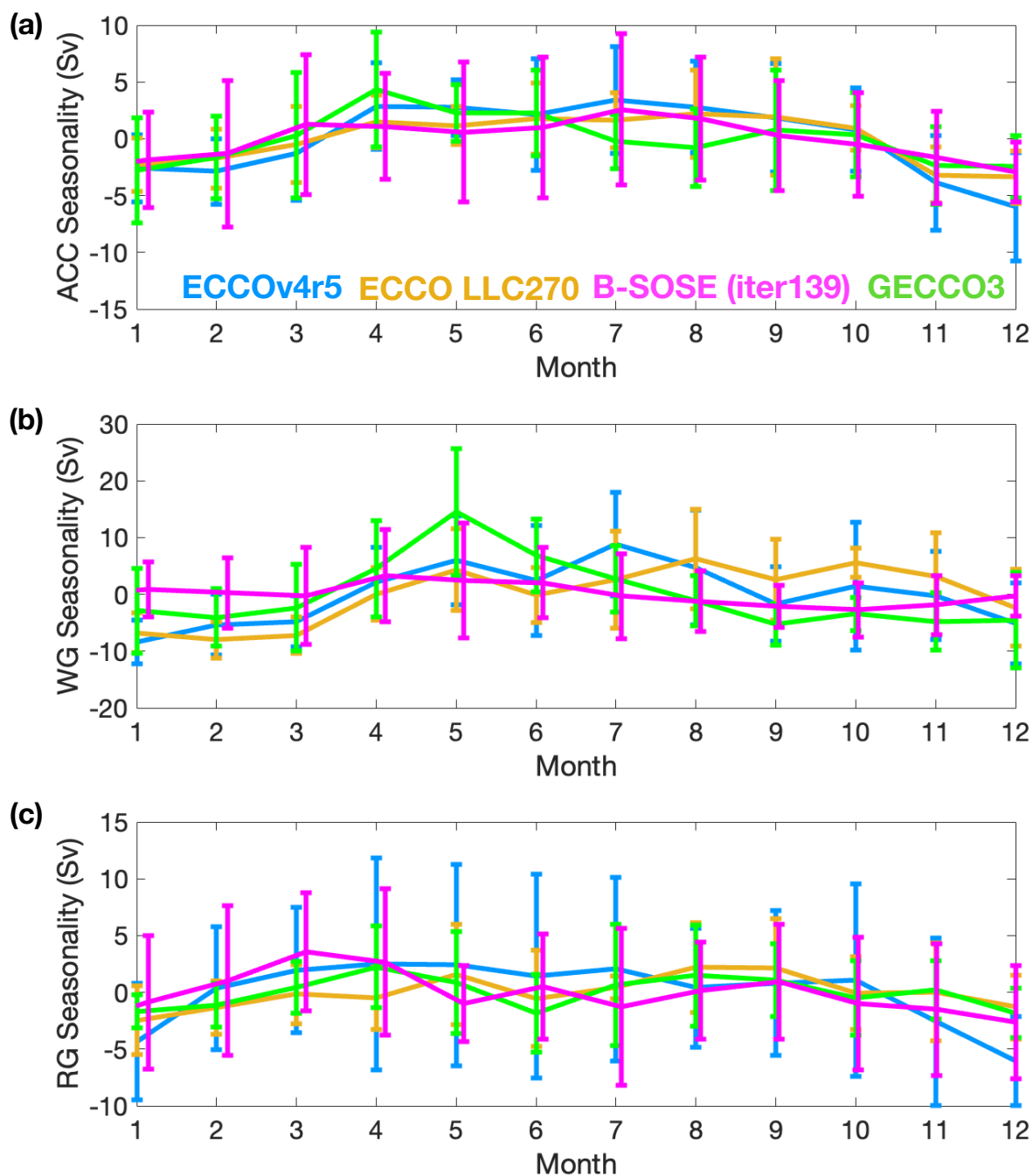
Figure 10. Same as Fig. ~~12~~ 9 but for ~~salinity~~ACC, WG, and RG anomalies.

Figure 11. ~~Times series of Weddell Gyre (a,e,i,m) temperature and (c,g,k,o) salinity (blue dots in~~ Same as Fig. ~~1) 9 but for~~ ~~ECCO-v4r5~~ACC, ~~ECCO-LLC270~~WG, ~~B-SOSE,~~ and ~~GECCO3~~RG seasonality. ~~Time series~~ Error bars represent one standard deviation of ~~(b,f,j,n)temperature and (d,h,l,p)salinity difference to year 1 mean for ECCO-v4r5, ECCO-LLC270, B-SOSE, and GECCO3~~variability.

37
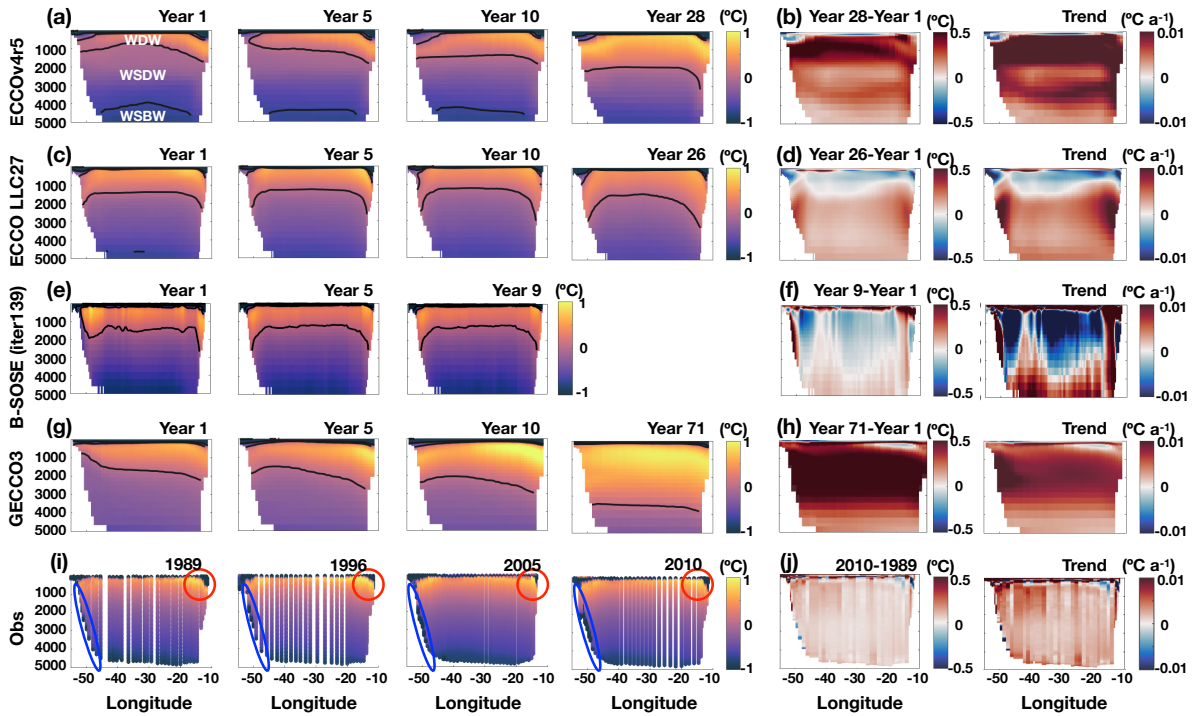
Figure 12. Simulated SR04 vertical sections of (a) ECCOv4r5, (b) ECCO-LLC270, (c) B-SOSE, and (d) GECCO3 potential temperature for year 1, year 5, year 10, and final years of each reanalysis along SR04 section (Fig. 1). B-SOSE only extends for 9 years and only years 1, 5, and 9 are plotted. (i) Observed vertical section of potential temperature for 1989, 1996, 2005, and 2010 along the SR04 section. Potential temperature (b,d,f,h,j) differences and trends (defined as the difference of the first and last years divided by the total model period). The potential temperature contours of -0.7 °C and 0 °C are also in black for panels (a), (c), (e), and (g). Close-ups of the shelf break regions enclosed by blue circles are also shown in Fig. S3.
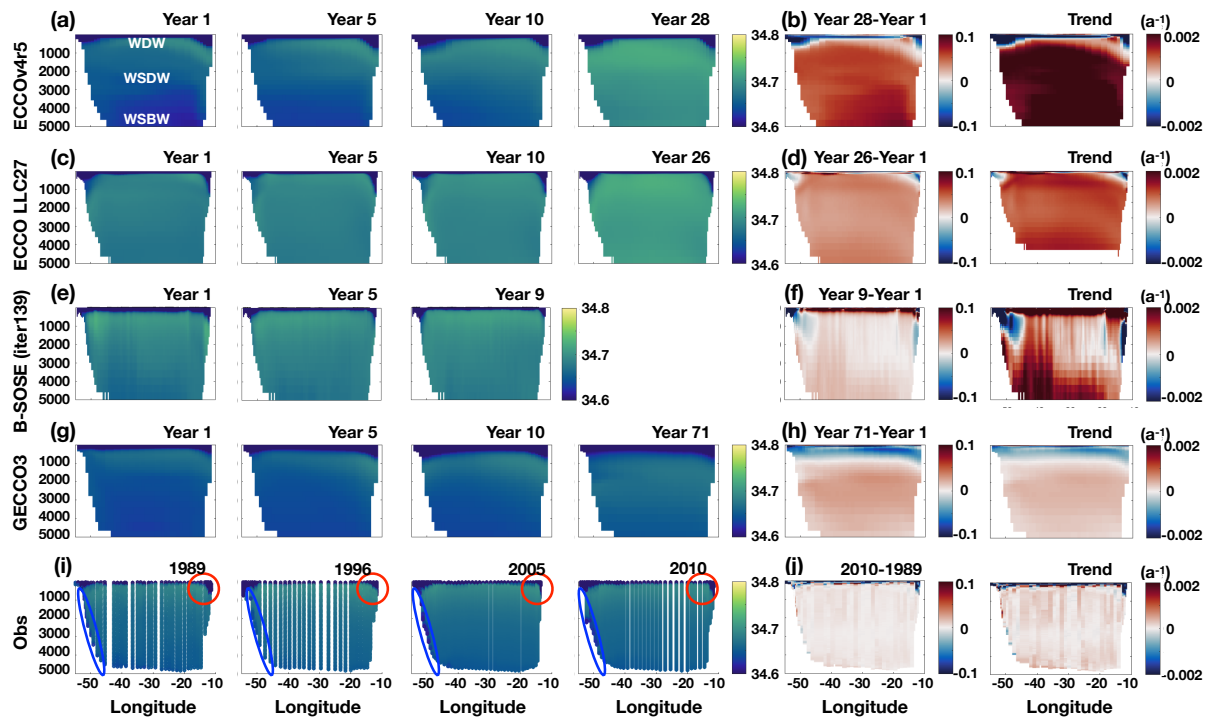
Figure 13. Same as Fig. 12 but for salinity. Close-ups of the shelf break regions enclosed by blue circles are also shown in Fig. S3.
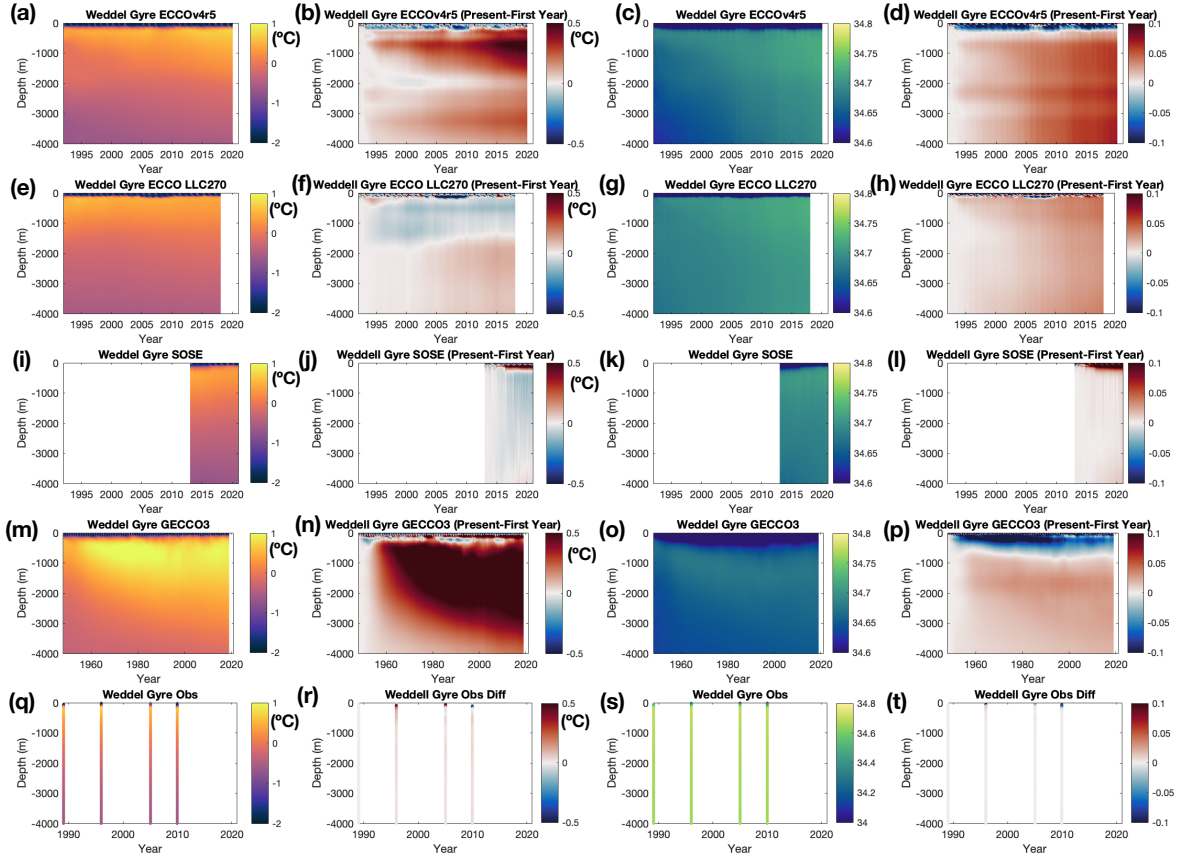
Figure 14. Times series of Weddell Gyre (a,e,i,m) temperature and (c,g,k,o) salinity (blue dot in Fig. 1) for ECCO-v4r5, ECCO-LLC270, B-SOSE, and GECCO3. Time series of (b,f,j,n)temperature and (d,h,l,p)salinity difference to year 1 mean for ECCO-v4r5, ECCO-LLC270, B-SOSE, and GECCO3. Vertical plots of observed (q) potential temperature, (r) potential temperature difference to 1989, (s) salinity, and (t) salinity difference to 1989 are also shown. These hydrographic properties are sampled from the nearest locations from $330.0°$E and $66.0°$S.
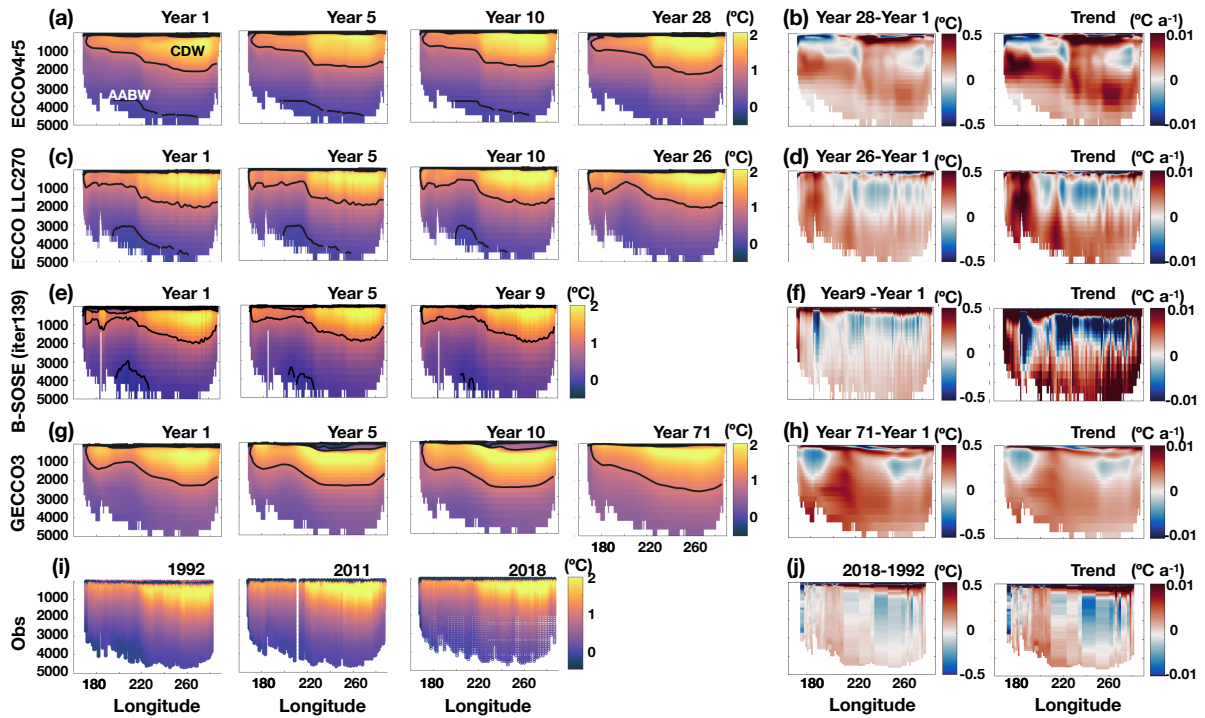
Figure 15. Same as Figure 12 but for the S4P section (Fig. 1). The potential temperature contours of 0 °C and 1 °C are also in black for panels (a), (c), (e), and (g).
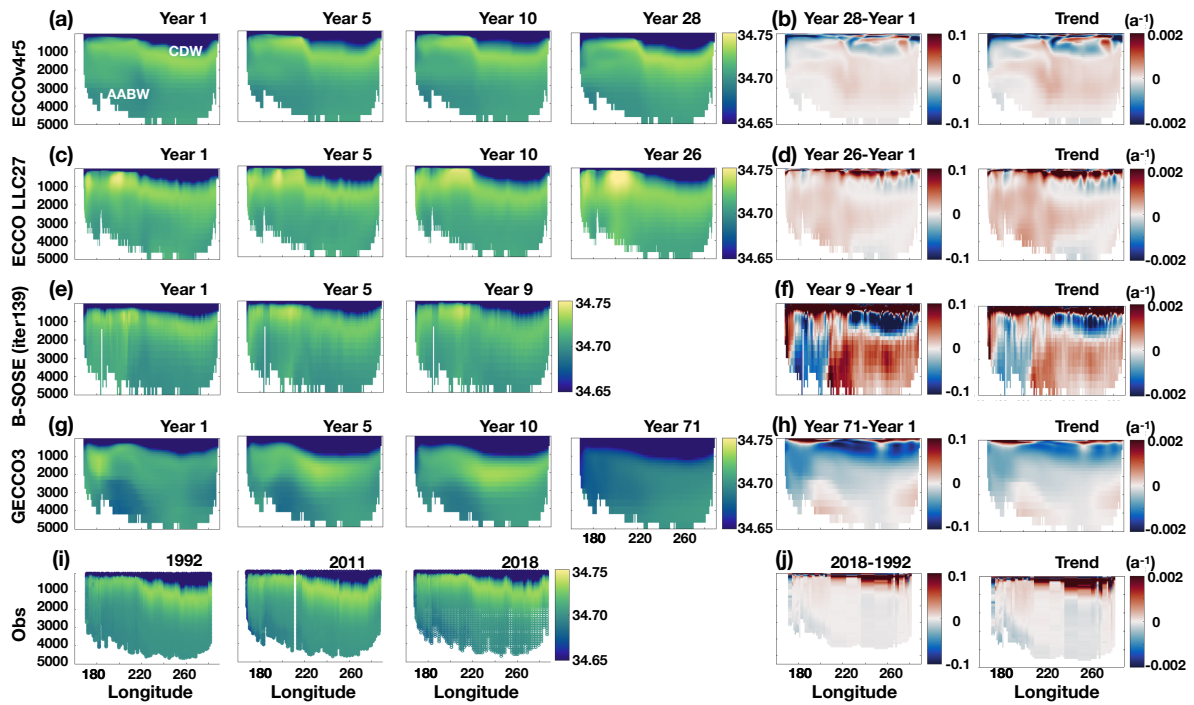
Figure 16. Same as Figure 12 but for salinity along the S4P section (Fig. 1).
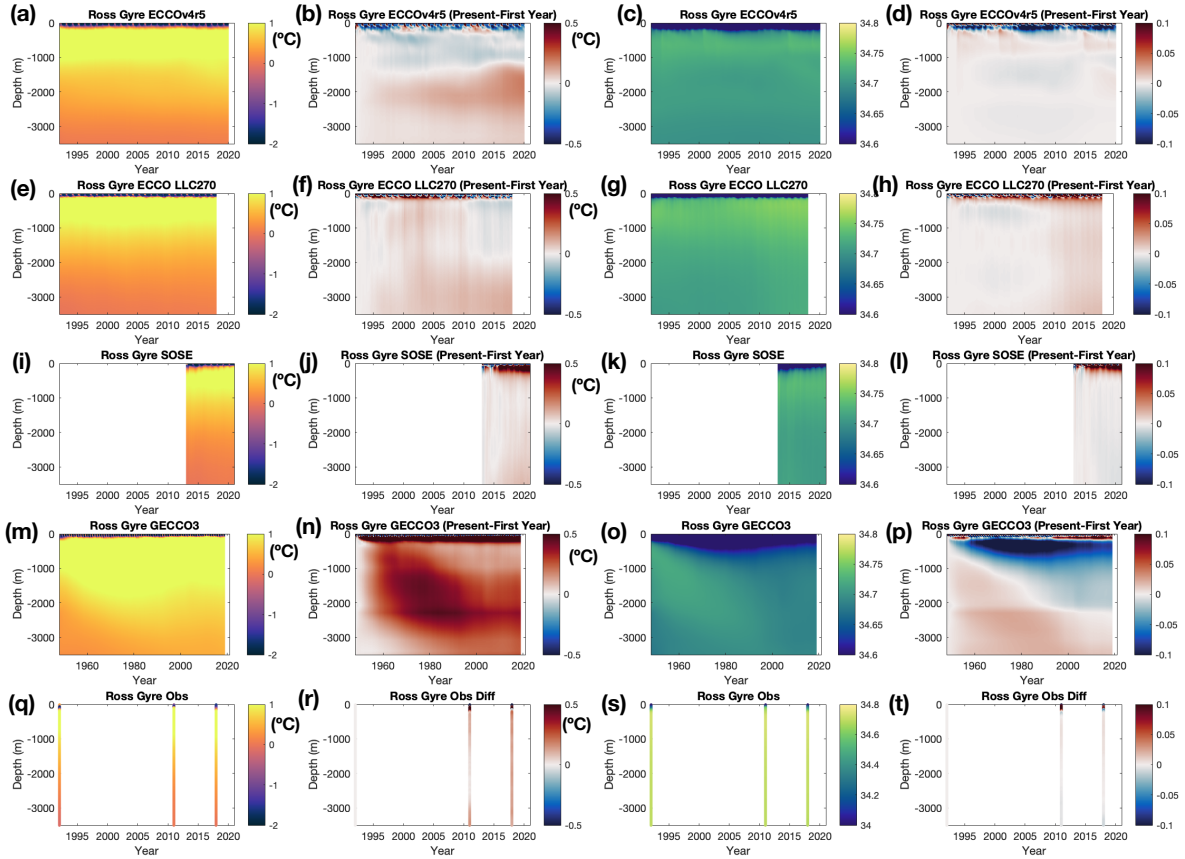
Figure 17. Times series of Ross Gyre (a,e,i,m) temperature and (c,g,k,o) salinity (cyan ~~dots~~ dot in Fig. 1) for ECCO-v4r5, ECCO-LLC270, B-SOSE, and GECCO3. Time series of (b,f,j,n) temperature and (d,h,l,p) salinity difference to year 1 mean for ECCO-v4r5, ECCO-LLC270, B-SOSE, and GECCO3. Vertical plots of observed (q) potential temperature, (r) potential temperature difference to 1992, (s) salinity, and (t) salinity difference to 1992 are also shown. These hydrographic properties are sampled from the nearest locations from 199.7°E and 67.8°S.
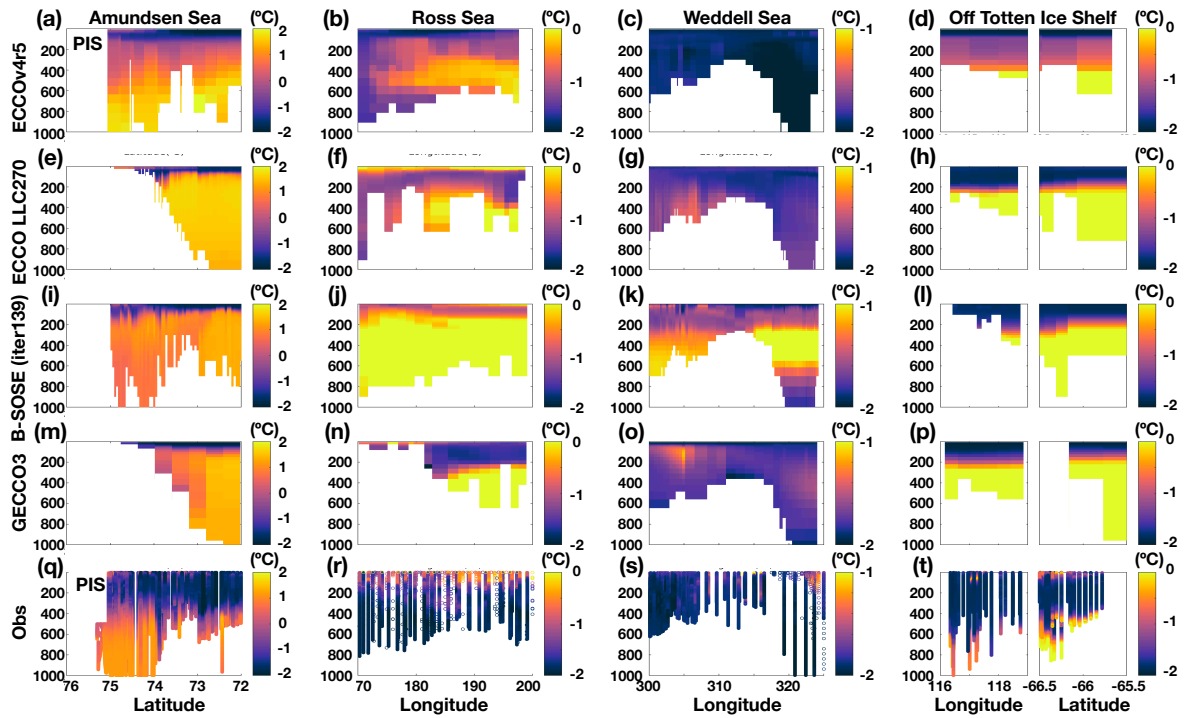
Figure 18. Simulated on-shelf vertical sections of (a-d) ECCOv4r5, (e-h) ECCO-LLC270, (j-l) B-SOSE, and (m-p) GECCO3 potential temperature of the final model year for the Amundsen Sea, Ross Sea, Weddell Sea, and Off the Totten Ice Shelf. (q-t) The same observed vertical sections of potential temperature compiled ~~available~~ all available years of observations.
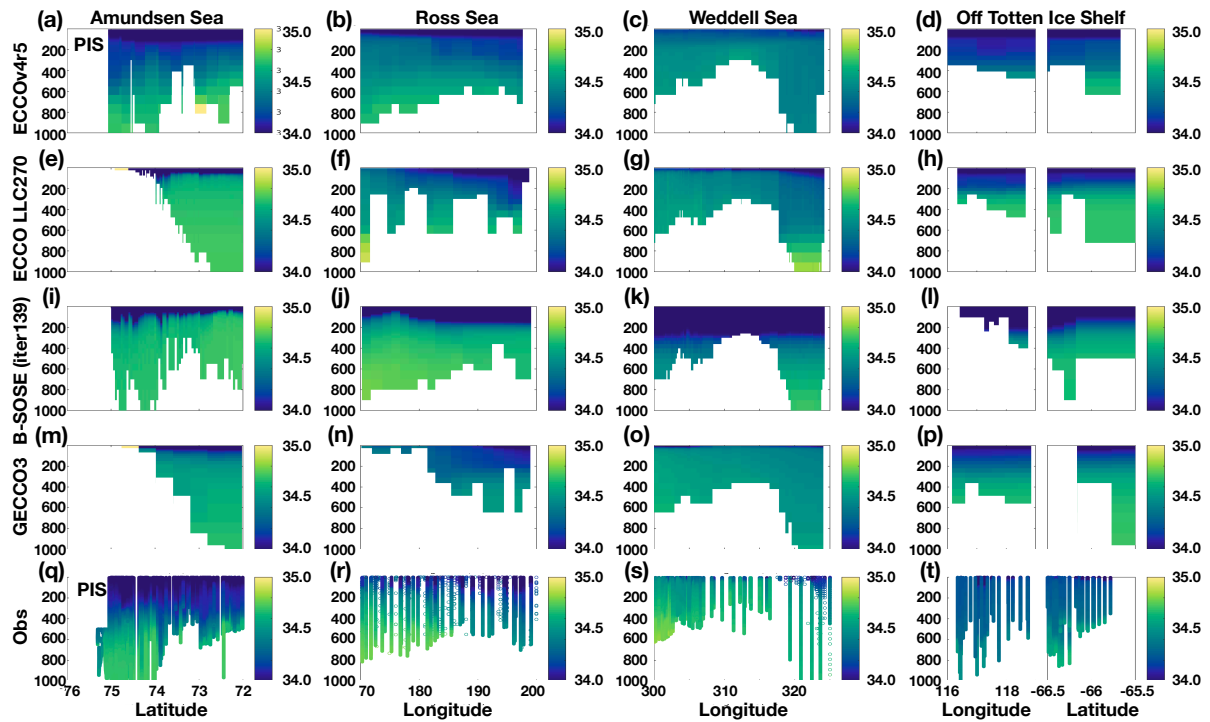
Figure 19. Same as Fig. 15 18 but for salinity.

Table 1. General information on ocean reanalyses evaluated in this manuscript.

Name	Year	Horizontal Grid	Reference	Initial atmospheric forcing guess	Bathymetry	Details
ECCOv4r5	1992-2019	~1°	Forget et al. (2015)	6 hourly ERA-interim	Arndt et al. (2013)	Ice shelf cavities
ECCO-LLC270	1992-2017	~1/3°	Zhang et al. (2018)	6 hourly ERA-interim	Smith and Sandwell (1997)	No ice shelf cavities
B-SOSE (iter139)	2013-2021	1/6°	Verdy et al. (2017)	Hourly ERA5	Amante and Eakins (2009)	No ice shelf cavities
B-SOSE (iter105)	2008-2012	1/3°	Verdy et al. (2017)	Hourly ERA5	Amante and Eakins (2009)	No ice shelf cavities
GECCO3	1948-2018	~0.4°	Köhl (2020)	6 hourly NCEP/NCAR Reanalysis	Jungclaus et al. (2013)	No ice shelf cavities and surface salinity restoring

Table 2. ~~Simulated strength~~ The standard deviation (SD) of ~~the Antarctic Circumpolar Current~~ reanalyzed and observed (in parenthesis) SIA anomalies for February, ~~Weddell Gyre~~ September, and ~~Ross Gyre~~ all months. Correlation coefficients followed by ** indicate that they pass the 95% significance test. The last column represents the root mean square error (RMSE) of the monthly climatology for the 12 months of reanalysis and observation.

	February		September		All months	
	$SD_{SIA}(10^3 \text{ km}^2)$	Correlation	$SD_{SIA}(10^3 \text{ km}^2)$	Correlation	$SD_{SIA}(10^3 \text{ km}^2)$	RMSE of the seasonal cycle (10^3 km^2)
ECCOv4r5	0.44 (0.38)	0.80 **	0.40 (0.47)	0.72 **	0.55 (0.55)	0.36
ECCO-LLC270	0,46 (0.37)	0.34	0.45 (0.47)	0.70 **	0.65 (0.53)	0.90
B-SOSE (iter139)	0.83 (0.45)	0.29	0.65 (0.67)	0.63	0.99 (0.78)	3.37
B-SOSE (iter105)	0.62 (0.31)	-0.59	0.40 (0.42)	0.31	0.37 (0.40)	1.87
GECCO3	0.27 (0.38)	0.20	0.67 (0.47)	0.68 **	0.76 (0.54)	0.74

Table 3. ~~Observed~~ Simulated mean strengths and ~~simulated trends in~~ standard deviations of the Antarctic Circumpolar Current, Weddell ~~Sea and Ross Sea at 500m and 1000 m. Trends that are three times larger/smaller than the observed values are in bold. Trends in opposite directions are also in bold. Note that simulated trends are calculated for the entire simulated years~~ Gyre, ~~which are different for each reanalysis product. Observed trends are calculated by using the temperature~~ and ~~salinity difference for 1989-2010 and 1992-2018 for the Weddell and~~ Ross ~~Sea, respectively~~ Gyre.

	ACC	WG	RG
ECCOv4r5	154Sv ± 5.6	56Sv ± 10.6	21Sv ± 7.0
ECCO-LLC270	140Sv ± 3.9	60Sv ± 8.1	32Sv ± 4.8
B-SOSE (iter139)	162Sv ± 5.3	54Sv ± 6.6	32Sv ± 6.0
B-SOSE (iter105)	181Sv ± 9.5	48Sv ± 4.1	24Sv ± 3.5
GECCO3	141Sv ± 5.7	35Sv ± 8.9	15Sv ± 3.0

Table 4. Observed and simulated trends in the Weddell Sea and Ross Sea for depth ranges between 1500-2000m and 3000m-bottom. Trends that are three times larger/smaller than the observed values are in bold. Trends in opposite directions are also in bold. Note that simulated trends are calculated for the entire simulated years, which are different for each reanalysis product. Observed trends are calculated by using the temperature and salinity difference for 1989-2010 and 1992-2018 for the Weddell and Ross Sea, respectively.

	ECCOv4r5	ECCO-LLC270	B-SOSE (iter139)	B-SOSE (iter105)	GECCO3	Observations
WS Pot. Temp. 1500-2000m ($°C\ year^{-1}$)	**0.0097**	0.0033	**-0.0026**	**-0.0060**	**0.0089**	0.0028
WS Pot. Temp. 3000m-bottom ($°C\ year^{-1}$)	**0.0061**	0.0039	0.0040	**-0.0054**	0.0041	0.0021
WS Salinity 1500-2000m ($year^{-1}$)	**0.0020**	**0.0011**	3.6×10^{-4}	**0.0020**	3.4×10^{-4}	1.4×10^{-4}
WS Salinity 3000m-bottom ($year^{-1}$)	**0.0023**	**0.0012**	**0.0013**	0.0012	2.5×10^{-4}	1.2×10^{-4}
RS Pot. Temp. 1500-2000m ($°C\ year^{-1}$)	**0.0045**	0.0014	-1.3×10^{-4}	**-0.016**	**0.0019**	-7.1×10^{-4}
RS Pot. Temp. 3000m-bottom ($°C\ year^{-1}$)	0.0034	0.0046	**0.0073**	0.0014	0.0031	0.0017
RS Salinity 1500-2000m ($year^{-1}$)	$\mathbf{1.3\times10^{-4}}$	$\mathbf{1.8\times10^{-4}}$	$\mathbf{1.1\times10^{-4}}$	$\mathbf{-1.9\times10^{-4}}$	$\mathbf{-3.1\times10^{-4}}$	-2.2×10^{-5}
RS Salinity 3000m-bottom ($year^{-1}$)	$\mathbf{7.1\times10^{-5}}$	$\mathbf{1.4\times10^{-4}}$	$\mathbf{2.4\times10^{-4}}$	$\mathbf{2.0\times10^{-4}}$	-3.5×10^{-6}	-1.0×10^{-4}