

**Feedbacks driving interdecadal variability in Southern Ocean convection in climate
models: A coupled oscillator mechanism**

Anand Gnanadesikan¹, Cassidy M. Speller¹, Grace Ringlein², John San Soucie^{2,3,4}, Jordan
Thomas^{1,5} and Marie-Aude Pradal¹

1: Department of Earth and Planetary Sciences, Johns Hopkins University

2: Department of Physics and Astronomy, University of Pennsylvania

3: Department of Earth and Environmental Sciences, University of Pennsylvania

4: MIT/WHOI Joint Program in Applied Ocean Physics and Engineering

5: Innovative Decisions, Inc.

Manuscript revised for Journal of Physical Oceanography

Wednesday, April 14, 2021

Corresponding Author: Anand Gnanadesikan, 327 Olin Hall, Johns Hopkins University, 3400 N.
Charles St., Baltimore, MD, USA. gnanades@jhu.edu

Abstract

Numerous climate models display large-amplitude, long-period variability associated with quasiperiodic convection in the Southern Ocean, but the mechanisms responsible for producing such oscillatory convection are poorly understood. In this paper we identify three feedbacks that help generate such oscillations within an Earth System Model with a particularly regular oscillation. The first feedback involves increased (decreased) upward mixing of warm interior water to the surface, resulting in more (less) evaporation and loss of heat to the atmosphere which produces more (less) mixing. This positive feedback helps explain why temperature anomalies are not damped out by surface forcing. A second key mechanism involves convective (nonconvective) events in the Weddell Sea causing a relaxation (intensification) of westerly winds, which at some later time results in a pattern of currents that reduces (increases) the advection of freshwater out of the Weddell Sea. This allows for the surface to become lighter (denser) which in turn can dampen (trigger) convection- so that the overall feedback is a negative one with a delay-helping to produce a multidecadal oscillation timescale. The decrease (increase) in winds associated with convective (nonconvective) states also results in a decrease (increase) in the upward mixing of salt in the Eastern Weddell Sea, creating a negative (positive) salinity anomaly that propagates into the Western Weddell Sea and dampens (triggers) convection-again producing a negative feedback with a delay. A principal oscillatory pattern analysis yields a reasonable prediction for the period of oscillation. Strengths of the feedbacks are sensitive to parameterization of mesoscale eddies.

1. Introduction

The Weddell Sea represents a major challenge in simulating the circulation of the global ocean within climate models. Within this basin, warmer and saltier waters influenced by North Atlantic Deep Waters intrude below cold, fresh Antarctic Surface Waters. Such intrusions result in conditional instability as deep waters brought near enough to the surface can become lighter than the surface waters. As a result, mixing between the deep and surface waters will produce cabelling. In many models that were part of the CMIP5 climate suite open-ocean deep convection within the Weddell Sea determines the temperature, salinity and density of the Antarctic Bottom Waters (Heuzé et al., 2013). Cessation of such convection in CMIP5 models was found by de Lavergne et al. (2014) to be associated with a more realistic lag in the warming of the Southern Ocean over the past century.

A number of current climate models exhibit regular, long-period variability in this convection. Models with regular periods range from ~40 years in the GFDL ESM2G model (Zanowski et al. 2015; Zanowski and Hallberg, 2017) to around 50 years in the GFDL ESM2Mc model (Galbraith et al., 2011, Bernardello et al. 2014) to multicentennial variability in the Kiel KCM1.2 model (Latif et al., 2013). Reintges et al. (2017) analyzed convective behavior in 15 such models and found an inverse relationship between the fraction of years exhibiting convection and the magnitude of stratification, with more stratified models exhibiting longer-period oscillations. Within individual models, such variability has been found to drive interdecadal variations in the formation of Antarctic Bottom Waters (Zanowski et al., 2015), Drake Passage transport and Weddell gyre strength (Behrens et al., 2016) ocean heat and carbon content (Bernardello et al. 2014, Thomas et al. 2018) and hemispheric atmospheric circulation (Cabr   et al., 2017). Zhang et al. (2017) argued that convective variability allows for

predictability of local climate with the GFDL ESM2M model. However, the mechanisms underlying such convective variability remain poorly understood.

A number of mechanisms have been proposed which might drive periodic convection in the ocean. The simplest of these relates to the fact that convection moves heat upwards in the water column, providing the primary balance to the downward transport of heat due to large-scale advection (Gnanadesikan et al., 2005). As a result, shutting off convection tends to trap heat at depth-which in turn over time creates instability that restarts convection. Martin et al. (2013) assert that this process is responsible for triggering convection in the Kiel Climate Model. The timescale of the variability is thus determined by the time required to build a sufficient warm anomaly at depth to counter the surface freshwater stratification.

An order-of-magnitude estimate of such a time scale in the real world can be taken from the amount of Atlantic, Indian and Pacific Deep Waters transformed within the Southern Ocean. Observationally-based estimates of this flux range from about 24 Sv (the inverse model of Talley, 2008) to 29 Sv (the ECCO data-assimilation solution of Forget, 2010) to 47 Sv (the inverse model of Sloyan and Rintoul, 2001). The volume of deep waters south of 50°S with a potential density $\sigma_0 > 27.7$ is $1.5 \times 10^{17} \text{ m}^3$, implying a residence time of order 200 years for the lower flux and 100 years for the higher one. Limiting the volume to waters between 300 and 2000m (the volume used in Reintges et al., 2017) causes the expected timescale for convection to shrink by about 2/3. On the other hand, if it is the volume of the entire deep ocean that needs to change (as in Loving and Vallis, 2005) the time scale rises to thousands of years.

An alternative possibility for periodic convection driven by surface forcing was proposed by Welander (1982). In this system there are two equilibrium states, one corresponding to low mixing and the other to high mixing. If the equilibrium density of the low mixing state is higher

than that of the high mixing state, and if the mixing is a positive function of density with a transition in between the two equilibrium states, the result is to produce an oscillation. If the mean buoyancy flux in the Southern Ocean is positive due to freshwater addition, as it is in the real world, the dominant driver of instability is expected to be cooling of the surface. However, it is unclear how such a mechanism could produce multidecadal convection, given that the relaxation time scale for surface temperature anomalies in the Southern Ocean is 3-6 months (Hausmann et al., 2016), as is the timescale for deep convection to replace the water in the surface layer (as shown by the annual resetting of surface nutrients to their deep values).

One possibility is that a decadal time scale could be related to slow vertical mixing across the pycnocline. A number of investigators (Timmermann and Beckmann, 2004; Heuzé et al., 2015b) have shown that the representation of background vertical mixing can affect whether a polynya occurs within the Southern Ocean in short-term simulations. However, both papers demonstrate that stronger wind-driven mixing suppresses convection- a result that would seem to conflict with the Reintges et al. (2017) observation that weaker stratification is associated with more frequent convection in long-term coupled simulations.

Another class of theories invokes slow integration of anomalies in freshwater forcing to drive convection, with the idea that salinity anomalies build up until they reach some critical value, at which point they trigger convection. Martinson et al., (1982) proposed that such preconditioning of the surface was important for driving the Weddell Polynya, with advection potentially helping to set a timescale. Weisse et al. (1994) found that stochastic freshwater forcing in an ocean-only GCM was able to generate decadal variability in North Atlantic overturning without invoking atmospheric coupling.

A final class of theories for periodic convection proposed by a number of investigators for the North Atlantic (Latif et al., 1996, Marshall et al., 2001) allows for longer timescales of convection by invoking a delayed oscillator mechanism involving atmospheric coupling. Larger SST gradients between the tropics and high latitudes produce a poleward shift of the westerly wind jet (positive North Atlantic Oscillation). The wind changes extract more heat from the subpolar gyre, causing a larger density gradient which in turn spins up the meridional overturning circulation. The stronger wind also spins up the subpolar gyre producing more heat transport. The changes in ocean circulation thus damp the SST gradient, but with a delay determined by the speed of Rossby wave propagation and the strength of air-sea coupling. Note, however, that both of these papers neglect the impact of salinity. Behrens et al. (2016) suggest that a similar oscillator might play a role in the generation of deep Southern Ocean convection and an important role for wind forcing in Southern Ocean convection was also proposed by Gordon et al. (2007) and Cheon et al. (2014). Cabré et al. (2017) demonstrate that wind stresses do vary in phase with Southern Ocean convection in one version of the GFDL ESM2Mc model. However, none of these papers clearly demonstrated how changes in winds would drive convection and what time scales might be involved.

A final complication is the range of physical processes that have been shown to affect convection in models. In addition to winds and vertical mixing, Thomas et al. (2018) show that increasing the eddy mixing parameter A_{Redi} (which governs stirring along density surfaces and is poorly constrained in models) reduces both the stratification and amplitude of convective variability. However increasing the interfacial diffusion coefficient A_{GM} (which governs how strongly eddies flatten strongly sloping isopycnals) increases the stratification, but also reduces the amplitude of convectively driven variability. Eddy-permitting simulations of the Southern

Ocean (Dufour et al., 2017) have been found to produce convective events but the length of the simulations is too short to draw strong conclusions about the inherent period of these oscillations. Kjellson et al. (2015) showed that in addition to vertical mixing, differing initializations of sea ice thickness or representations of ice sheet melting could also trigger deep convection in short historical simulations.

The primary focus of this manuscript is to describe the mechanisms that produce periodic convection in a version of the coarse-resolution GFDL ESM2Mc model, in which the changes in stratification leading to convection are dominated by changes in surface salinity. We choose this model because it has a relatively skillful simulation of wintertime temperatures and salinities within the Weddell Sea, winds over the Southern Ocean as a whole and the most regularly periodic variation in convection amongst the versions of ESM2Mc which we have run at Johns Hopkins. Previous work with ESM2Mc (Galbraith et al., 2011) suggested that convective initiation and shutoff within the Weddell Sea could be predicted from salinity anomalies originating near the prime meridian, but did not describe how such anomalies could be generated, how such generation might be linked to convective variability in the Weddell Sea, and why such a mechanism would produce periodic multidecadal variability.

The structure of this paper is as follows. In Section 2 we review the model to be used in this paper and outline the Principal Oscillation Pattern analysis that we use to identify critical regions for SST and SSS variability as well as the coupling between these regions. We also describe the term balance analysis that we use to elucidate the mechanisms involved. In Section 3 we show that the Principal Oscillation Pattern analysis reveals a coupling between EOFs of SST and SSS with centers of action in the Western Weddell and Eastern Weddell seas. We then examine the term balances of SSS and SST in each region- highlighting how variation in the

advective salt flux drives SSS variability in the Western Weddell Sea and the joint roles of variability in currents and surface salinity anomalies in producing these changes. We then examine how convection in turn results in changes in winds and mixing that produce anomalies in both surface velocities and sea surface salinity. In Section 4, we examine the realism of the physical processes involved in convection as well as the robustness of our story to changes in the model physics.

2. Methods

a.) The model

The model used here is the Geophysical Fluid Dynamics Laboratory (GFDL) fully coupled CM2Mc (Coupled Model 2 with Modular Ocean Model, coarse grid), comprised of separate atmosphere, ocean, sea ice, and land components that are linked through a flux coupler. The atmosphere model has 24 vertical levels, a 3.75° longitudinal resolution and a 3° latitudinal resolution and has modern parameterizations of radiation, clouds, and boundary layers. The ocean model has 28 vertical levels using a z^* coordinate and a nominal 3×1.5 degree resolution, with a latitudinal resolution which varies with latitude. The model was initialized with modern observations and then run out for 1500 years with preindustrial levels of greenhouse gasses and aerosols. In this paper, we consider three simulations branched from this point with three different mesoscale eddy parameterizations, but with atmospheric radiative forcing remaining at preindustrial levels.

In the control state of the model (described in Galbraith et al. 2011) eddy parameterizations were tuned to minimize both maximum and RMS SST errors, produce a

relatively realistic Southern Ocean circulation, a relatively realistic El Nino and a stable and realistic overturning circulation. As in most current climate models, the interfacial diffusion coefficient A_{GM} follows the theory of Visbeck et al. (1998), in which it is proportional to the slope of the isopycnal surface S_ρ and the stratification N :

$$A_{GM} = \alpha N S_\rho L^2, \quad (1)$$

where α is a scaling coefficient and L is generally chosen as the width of the baroclinic zone. Equation 1 works well in capturing the difference between subtropical gyres (where S_ρ is small) and boundary currents (where it is large). However, it leaves A_{GM} poorly defined in regions with weak stratification where the isopycnal slope goes to infinity and the buoyancy frequency goes to zero. To avoid instability A_{GM} is usually given a floor at some minimum value and a ceiling at some maximum value. In our baseline simulation, the minimum value of A_{GM} was set to 200 m^2/s and the maximum value was set to 1400 m^2/s . We also ran a second set of simulations where the minimum was raised to 600 m^2/s , closer to the value used in the higher resolution GFDL ESM2M model (Dunne et al., 2012). We refer to this simulation as AGMmin600.

In recent years, our group at JHU has examined the impact of changing the diffusion coefficient A_{Redi} that governs the stirring of passive tracers along isopycnals. This parameter is poorly constrained in both observations and models (Gnanadesikan et al., 2015) but has important impacts on convection (Pradal and Gnanadesikan, 2014) and biogeochemical cycling (Bahl et al., 2019). After our initial spinup, a number of additional simulations were generated in which the parameter was set to a lower value of 400 m^2/s and higher values of 1200 and 2400 m^2/s (a range similar to that seen in CMIP5 models) with all four cases being run out for an additional 1000 years. In Thomas et al. (2018), we noted that the version of the model with the

lowest mixing coefficient (denoted as AREDI400) had the most regular convection while the model with the highest mixing coefficient (denoted as AREDI2400) had much less variability in convection within the Weddell Sea. There is an initial adjustment period of about 100 years over which surface temperatures come to equilibrium. In order to avoid trends due to this adjustment, we analyze output from the last 600 years of each model simulation.

In this manuscript we will focus on the AREDI400 simulation, but will briefly compare the mechanisms found in this simulation to the AREDI2400 and AGMmin600 simulations at the end of the manuscript. Various diagnostics for the three models are compared with observations and the range of CMIP5 models in Table 1. Despite its coarse resolution, AREDI400 produces a relatively realistic mean state of the Southern Ocean compared to the CMIP5 suite of coupled climate models. The ACC transport through Drake Passage is 169 Sv-very close to the modern value of 170 Sv from Donahue et al. (2016) used as a metric of Southern Ocean circulation by Beadling et al. (2019). The mean density gradient across the ACC of 0.26 kg/m^3 is also close to the observed value. The peak wind stress of 0.163 Pa lies between the 0.14 Pa estimated by the NCEP-NCAR reanalysis between the years of 1949 and 1970 and more modern values of 0.18-0.19 Pa, which include significant impacts from the ozone hole (Thompson and Solomon, 2002). Maximum sea ice extent is 17.9 Mkm^2 in this model in the mean compared with 18.7 Mkm^2 from observations (Turner et al., 2012). Density stratification between 300m and the surface within the region of deep convection when convection is not present (a metric used by Reintges et al., 2017 to look at the relationship between stratification and convective period) is 0.16 kg/m^3 . All of these metrics are close to the middle of the suite of CMIP5 models. The model does have bottom temperatures and densities that are too warm and too light, but the wintertime density stratification within the Weddell Sea is close to modern observations (though this

involves compensating errors from temperature and salinity). The AREDI2400 model has less stratification, cooler deep waters, weaker winds and ACC transport and less sea ice. The AGMmin600 model is more stratified, with warmer deep waters and more sea ice, but surprisingly also has weaker winds. With an ACC transport at Drake Passage of 135 Sv the AGMmin600 is closer to the old estimate of 137 Sv of Cunningham et al. (2002) but still well within the CMIP5 model range.

The strongest interdecadal sea surface temperature (SST) variability amongst the three simulations is found in the AREDI400 model. A particularly important region of variability is found in the Western Weddell Sea, between 50-20°W and 70-55°S (Fig. 1a, red box). Periodic increases in annual mean SST of up to 1.2°C above the average are associated with extraction of heat from the water column (Fig. 1b), higher salinity at the surface (Fig. 1c) and higher densities throughout the water column, but particularly at the surface (Fig. 1d). Decreases in SST are associated with a buildup of heat, lower surface salinity, and lower density throughout the water column, but again particularly at the surface. This pattern is consistent with a periodic extraction of heat by convection, but also with that convection being driven primarily from the surface. The RMS variation in surface density (0.11 kg/m^3) is about four times the variation at the depth where we find the largest buildup of heat (0.03 kg/m^3). The fact that high surface density is associated with higher SSTs (which would lighten the water column) means it is the higher SSSs that must be responsible for the increase in density. There is some hint of density anomalies propagating downwards from the surface. The period is about 50 years. Similar dynamics are found for salinity in the ESM2G model (Zanowski et al., 2015) which has a very similar atmosphere but different ocean dynamics. This suggests that the simplest explanation for convection (resetting of an unstable warm deep reservoir by large-scale circulation as in Loving

and Vallis, 2005) is not a tenable explanation for periodicity of that convection and that we need to look at how the surface anomalies are generated and sustained over multiple decades.

The cycles in sea surface temperature are also linked to variations in sea ice cover. Fig. 2 shows composites of the 15% sea ice extent contour during September taken relative to the time at which the minimum ice cover occurs in the Western Weddell region. The variation is largely expressed in terms of variation of the sea ice edge, rather than by the formation of a polynya within the ice pack. The ice edge reaches its maximum extent about 25 years before a minimum and drifts polewards over a period of about 15 years across the Weddell Sea (Fig. 2a). From about 10 years to 4 years before the minimum, the sea ice retreats faster in the Eastern Weddell Sea along the Prime Meridian, but from that time onward the retreat is much faster in the Western Weddell Sea, where the wintertime ice edge retreats by over 1200 km (Fig. 2a). After the minimum in ice cover associated with the maximum in convection, the ice cover advances rapidly in the Western Weddell Sea for about 6 years, advances broadly across the Weddell Sea for about 4 years (Fig. 2b), and then primarily advances in the Eastern Weddell Sea for the remainder of the cycle (Fig. 2b). As we will see asymmetric behavior between the Eastern and Western Weddell sectors (shown by the yellow and cyan boxes within the plot and denoted in this paper by the acronyms EW and WW respectively) plays an important role in explaining this oscillation.

b. Principal Oscillation Pattern Analysis

In order to isolate dynamics that can result in oscillations we need to identify centers of action where variability in temperature and salinity is large and varies coherently. Such centers

may be separated in space and peak at different times. We isolate both the key locations and the interactions between them using a method similar to the Principal Oscillation Pattern methodology developed by Hasselmann (1988) and applied by Xu and von Storch (1990) to analyze El Nino. While Weisse et al. (1994) applied this methodology to the North Atlantic to our knowledge it has not been applied to look at Southern Ocean dynamics.

We apply the following steps to generate the analysis (a more complete mathematical description is provided in the Supplemental material and the Matlab routines are provided via the data archive). 1. In order to focus on decadal variability, we smooth the surface temperature and salinity time series with a ten-year boxcar. 2. In order to avoid drifts, we limit our analysis to the last 600 years of this smoothed time series. 3. We are interested in covariation of temperature and salinity, but the range of temperature is much larger than that of salinity. We therefore normalize the anomalies by the standard deviation of the decadal variability in each field averaged over the Southern Ocean. This ensures that any patterns of variability are not dominated by temperature. 4. In order to identify the principal centers of action we perform an empirical orthogonal function on the normalized SST and SSS fields together, performing a singular vector decomposition on a data matrix containing both fields. We then extract the first ten modes, which will capture the spatial patterns associated with greatest joint variability in temperature and salinity. 5. Standard EOF analysis allows statistically insignificant correlations between centers of action to cause them to become unphysically coupled, producing modes that represent either a sum across the leading centers of action or of differences between these centers. However, the leading modes can be rotated to a new set of modes that are not only orthogonal but are also strongly peaked in space (though they are no longer strictly orthogonal in time). This rotation thus isolates primary centers of action (see Lian and Chen, 2012 for an in-

depth discussion of this). In our case we rotate the first ten modes using the varimax criterion (Kaiser, 1958), which ensures that each rotated mode is as “peaked” in space as possible. 6. The rotated modes are projected onto the anomaly fields to generate the associated time varying amplitudes χ_{1-4} which can be lumped together as a time-varying vector $\vec{\chi}$ (note that in the notation of Hasselmann, 1988, these correspond to the amplitudes of what he terms Principal Prediction Patterns). 7. In order to identify relationships between the centers of action defined by the rotated EOF patterns we use multiple linear regression to define a matrix equation

$$\frac{\partial \vec{\chi}}{\partial t} = R * \vec{\chi} \quad (2)$$

This equation isolates how anomalies in one convective region can produce changes in other regions and how these changes may feed back onto the convective region. In general we are looking for equation sets of the form

$$\frac{\partial \chi_{conv}}{\partial t} = -\lambda_{conv} \chi_{conv} + \sigma_{cr} \chi_{resp} \quad (3a)$$

$$\frac{\partial \chi_{resp}}{\partial t} = \sigma_{rc} \chi_{conv} - \lambda_{resp} \chi_{resp} \quad (3b)$$

Where χ_{conv} represents the amplitude of a pattern of anomalous SSTs and SSSs that is centered in a region with deep convection, χ_{resp} represents the amplitude of a similar pattern that responds to this deep convection, the λ terms represent the rate at which these anomaly patterns are damped and the σ terms represent couplings between regions, where anomalies in one region cause a change in another region. We are particularly interested in equation sets where $\sigma_{rc} * \sigma_{cr} < 0$, since such sets of equations would correspond to a harmonic oscillator. 8. In order to find a prediction for the natural period of oscillation we extract the eigenvectors and eigenvalues of the R matrix in equation (2). Eigenvectors with imaginary eigenvalues are rotary solutions and

come in pairs. The least damped oscillatory state will then be one that can amplify that part of random atmospheric forcing that has the appropriate range of frequencies (as in Xu and Von Storch, 1990 for ENSO and Griffies and Tziperman, 1995 for the North Atlantic Overturning). Such a state may also be sustained by coupling if one of the damping coefficients λ is negative.

c.) Term balances

While the principal oscillation pattern analysis will show linkages between temperature in one location and salinity in another, it cannot by itself tell us why such linkages exist or why they have the sizes they do- i.e. what determines the size and sign of the λ and σ coefficients in equation 3. In order to better understand such linkages we examine the term balances that lead to changes. The MOM4 model breaks the time rate of change of any variable into several terms. Physically, this can be described as

$$\text{Time rate of change of tracer } X = \text{Advection} + \text{Vertical diffusion from below} + \text{Nonlocal}_{KPP} + \text{Sources} + \text{Mesoscale eddy} + \text{Submesoscale eddy} + \text{Surface fluxes} \quad (4)$$

In (4), “Advection” refers to the three-dimensional advection from the resolved flow, “Vertical diffusion from below” is that component of the subgridscale diffusive flux associated with downgradient diffusive transport by small-scale turbulent diffusion and “Nonlocal_{KPP}” refers to the transport due to large eddies within a mixed layer following Large et al., (1994). “Source” terms are zero for salinity but may include the impact of vertically penetrating shortwave radiation for temperature. “Mesoscale eddy” includes both the diffusive Redi and advective GM terms. “Submesoscale eddy” transport is due to mixed layer restratification following Fox-Kemper et al. (2011). As in previous work (Seiour et al., 2017) we sum these tendencies over

some set of layers to define flux terms, Q_X^{tag} where X denotes the relevant tracer and the tag denotes the process. We combine vertical diffusion and Nonlocal_{KPP} terms (which only depend on boundary fluxes and the vertical structure of tracers and momentum) with any source terms (shortwave heating in particular) to form a vertical mixing term Q_X^{vmix} . We also combine the mesoscale and submesoscale eddy terms (which depend on lateral as well as vertical gradients) to form an eddy term, Q_X^{eddy} . Our term balance equation then becomes

$$Q_X^{tend} = Q_X^{adv} + Q_X^{vmix} + Q_X^{eddy} + Q_X^{surf} \quad (5)$$

As we will see below, some of these terms can be in phase with peaks in temperature and salinity corresponding to convective events. However, others (particularly the tendency term) will be out of phase with these peaks. By looking for terms on the right-hand side of the equation that are similarly out of phase with the temperature and salinity peaks we can gain insight into how we can generate anomalies that persist over many decades.

3. Results

a.) EOF and POP analysis: AREDI400 model

We begin by analyzing the AREDI400 model, which as previously stated has the most regular variation of any model we have run. The first ten unrotated EOFs in this model account for 85% of the overall variance in temperature and salinity with the first mode accounting for 39.4% and the first four modes accounting for about 70%. The SST and SSS patterns associated with the first rotated EOF are shown in Figure 3a and b. The EOFs are normalized so that the

corresponding modal amplitude time series $\chi(t)$ have standard deviation of 1 (the impact of a mode on a particular point is given by the value of the pattern at that point times the value of the modal amplitude over time). The first rotated EOF is associated with covariation of temperature and salinity centered in the Weddell Sea- exactly the area where we see the maximum temperature variability in Fig. 1a. The yellow boxes in Fig. 3 correspond to the WW region where the signature of deep convection is seen in Figs. 1b-d. Figures 3c and 3d show the third rotated EOF, which is the mode that exhibits the strongest two-way coupling with the first mode (modes 2 and 4 are shown in the Supplemental material). The third rotated EOF is associated with a negative salinity anomaly to the east and a positive temperature anomaly to the north of the center of action of the first rotated EOF. The cyan boxes in Fig.3 show that the EW region (20°W-20°E, 70-55°S) corresponds to the maximum salinity anomaly in Fig. 3d.

As shown in Fig. 4a, the time series of the modal amplitude for the third EOF χ_3 (red line) is out of phase with that for χ_1 (blue line) and when plotted against it (Fig. 4b) produces a tilted ellipse that corresponds to oscillatory behavior. It is thus unsurprising that the time tendencies of each mode are related to the amplitude of the other mode. A multiple linear regression produces the following set of relationships

$$\frac{\partial \chi_1}{\partial t} = 0.075 \text{yr}^{-1} * \chi_1 - 0.128 \text{yr}^{-1} * \chi_3 \quad (6a)$$

$$\frac{\partial \chi_3}{\partial t} = 0.132 \text{yr}^{-1} * \chi_1 - 0.063 \text{yr}^{-1} * \chi_3 \quad (6b)$$

which produce correlation coefficients of 0.90 for the prediction of χ_1 (blue line, Fig. 4c) and 0.92 for the prediction of χ_3 (red line, Fig. 4c).

The physical story told by this set of equations helps to explain how we can get a sustained oscillation over many decades. An initial positive perturbation in temperature and salinity over the Western Weddell Sea (positive χ_1) will grow, but also generate a negative salinity perturbation to the east and a positive temperature perturbation to the north (positive χ_3). But positive perturbations in χ_3 generate a negative trend in χ_1 —resulting in oscillatory behavior. Simply solving equation (6a,b) for an oscillatory solution produces a resonant period of 57 years.

Performing the full principal oscillation pattern analysis (using the full R matrix with the top 4 modes shown in Table S1) yields a complex eigenvector involving all four of these modes whose structure is shown in Fig. 3d. The length of each arrow shows the relative contribution of each mode to the pattern and demonstrates that the oscillation is dominated by modes 1 and 3. As the corresponding eigenvalue is $0.0016 + i*0.11 \text{ yr}^{-1}$, the vectors rotate clockwise with a period of 57 years (with a statistically insignificant growth). At any time we can get the actual physical expression of the mode by taking the projection of these vectors on the real axis and summing the resulting weighted spatial patterns. As suggested from the raw analysis of equation 6, positive values of mode 1 (corresponding to high temperatures and deep convection across the Weddell Sea) are followed by high values of mode 3 (corresponding to relative freshening in the Eastern Weddell Sea) with a lag of ~ 9.5 years. High mode 3 then results in a slow reduction of the convection and buildup of a salty anomaly in the Eastern Weddell Sea over ~ 19 years (corresponding to half of the cycle of 57 years). Extending this to more modes results in a drop of the time scale of the oscillation to 53 years when 10 modes are used in the regression and 51 years when 12 modes are used, but does not yield additional physical insight about how different

regions are coupled. As we are primarily using the POP analysis to identify key regions that are involved with the oscillation, we do not explore these results further here.

b.) Understanding the oscillatory dynamics using heat and salt budgets in the centers of action

While the modal analysis is helpful in identifying the centers of action that are connected with each other and in characterizing the timescales governing such connections, it does not by itself provide an explanation for the size of the relevant coefficients that set the timescales of variability. To get a better sense for this we turn to an analysis of the term balances in the Eastern Weddell (EW) and Western Weddell (WW) boxes shown in Figs. 2 and 3. As would be expected from Figs 3 a and b, the decadal smoothed SST and SSS in each of these regions correlate with χ_1 with a value of ~ 0.95 and the sum of the anomalies in the two regions has an even higher correlation with χ_1 of 0.96 for SST and 0.97 for SSS. We also consider the SSS difference between the Western Weddell Sea and Eastern Weddell Sea which as would be expected from Fig. 3d projects strongly onto mode 3 variability (with a correlation of 0.8). In the analysis that follows time mean heat and salt balances in each region are examined, followed by a discussion of which terms are responsible for temporal variation in heat and salt.

i. Mean balances in the Western and Eastern Weddell Seas

In the Western Weddell region (Table 2, left column) the net freshwater balance at the surface is positive, with about 0.46 m.yr^{-1} of freshwater being added. Export of this freshwater from the surface layer removes $-15.2 \text{ kg.m}^{-2}.\text{yr}^{-1}$ of salt, which must be supplied from other sources. The leading source of salt is vertical mixing, which supplies $14.1 \text{ kg.m}^{-2}.\text{yr}^{-1}$ while advection supplies $7.5 \text{ kg.m}^{-2}.\text{yr}^{-1}$. Because eddies tend to bring fresher surface waters into convective regions and export saltier denser waters, they actually remove salt at a rate of around $-6.4 \text{ kg.m}^{-2}.\text{yr}^{-1}$, closing the balance. The mean surface heat flux in the Western Weddell Sea is -8.7 Wm^{-2} , requiring an oceanic supply of heat. The two primary sources of heat to the upper 50m are vertical mixing ($+14.6 \text{ W.m}^{-2}$) and lateral eddy mixing ($+6.2 \text{ W.m}^{-2}$). The mean effect of advection is to cool this region (-12 W.m^{-2}).

In the Eastern Weddell region (Table 2, right-hand column) the surface fluxes also act to dilute salt and remove heat, but at a slightly lower rate than in the Western Weddell region. ($-12.8 \text{ kg.m}^{-2}.\text{yr}^{-1}$ and -10.8 W.m^{-2} respectively). Again vertical mixing ($+12.2 \text{ kg.m}^{-2}.\text{yr}^{-1}$) and advection ($+4.8 \text{ kg.m}^{-2}.\text{yr}^{-1}$) supply salt while lateral eddy fluxes ($-4.2 \text{ kg.m}^{-2}.\text{yr}^{-1}$) remove it. Similarly vertical mixing ($+12 \text{ W.m}^{-2}$) and lateral eddy mixing ($+4.3 \text{ W.m}^{-2}$) supply heat while advection closes the balance by removing heat (-5.7 W.m^{-2}).

ii. Variation of term balances in the Western Weddell Sea-identifying the role of salinity advection

We computed composites of the term balances in the AREDI400 simulation by picking 11 peaks where χ_1 was maximal and where we had at least 50 years of decadal smoothed term

balances before and after the peak. As shown in Fig. 5 the composited SST and SSS from the Western Weddell (WW) region show peaks that are in phase with the χ_1 time series, and which have a periodicity of about 50 years -very close to what we get from the linear POP analysis. The composite SSS variation from trough to peak is 0.54 PSU (which would produce a density increase of 0.43 kg.m^{-3} at the mean SST of 1.9°C) while the SST variation from trough to peak is 1.76°C (which would produce a density decrease of -0.14 kg.m^{-3} at the mean salinity of 33.9 PSU). This supports the previous assertion of Galbraith et al. (2011) that it is the surface salinity cycle in the Weddell Sea that holds the key to understanding why ESM2Mc shows such regular convection.

Examining the composite salinity balance over the top 60m (the region where we see the large changes in density in Fig. 1) shows that Q_{salt}^{tend} is the result of several terms acting together. We begin by looking at the evolution of Q_{salt}^{tend} (black line, Fig. 5c) over the 50 year convective cycle starting at a lag of zero with respect to the peak in χ_1 . At this point in the cycle Q_{salt}^{tend} is near zero (unsurprising given that such times correspond to a maxima in sea surface salinity). The following ~20 years see a period of freshening, which is strongest about 7 years after the maximum in χ_1 with a value of $-2 \text{ kg.m}^{-2}.\text{yr}^{-1}$ (corresponding to a salinity decline of 0.04 PSU yr^{-1}). Following this minimum, the remaining 30 years of the cycle see salinification in this region, with an initial 15-20 year period characterized by relatively weak salinification of about $0.06\text{-}0.1 \text{ kg.m}^{-2}.\text{yr}^{-1}$ (corresponding to a change of about $0.012\text{-}0.02 \text{ PSU.yr}^{-1}$). At this point the salinification accelerates to a peak value of around $2 \text{ kg.m}^{-2}.\text{yr}^{-1}$ about 7 years before the next maximum before dropping back to zero at the end of the 50 year cycle. Examination of the individual cycles (not shown) shows that this pattern is seen in most of the cases where we have a convective event.

The term balances for salinity project onto this cycle in a number of informative ways. The dilution flux driven by net surface freshwater balance (Q_{salt}^{surf} cyan line, Fig. 5c) tends to reinforce the convective cycle, as warmer waters and an ice-free surface during the winter enhance evaporation during convective periods. The net impact of the changing water flux across the surface is of the same order of magnitude as the vertical mixing flux (Q_{salt}^{vmix} dark blue line, Fig. 5c) so that the higher densities at the surface during convection are actually due to both processes. Eddy fluxes (Q_{salt}^{eddy} magenta line, Fig. 5c), by contrast, represent a strong negative feedback which is strongly anticorrelated with the surface salinity anomaly. When summed together, the eddy, vertical mixing and surface dilution fluxes cancel over most of the cycle and seem to be most important during convective initiation. None of them, however, is strongly enough out of phase with convection to produce the regular oscillatory dynamics we see. The key to explaining the 50 year time scale that emerges for convection must thus be tied to the advective flux (Q_{salt}^{adv} red line), which has a positive peak of about $2 \text{ kg.m}^{-2}.\text{yr}^{-1}$ about 15 years before the peak (or 10 years after the previous minimum) in χ_1 and a negative trough of about the same size about 10 years after the peak.

It is interesting to contrast the salt balance with the heat balance (Fig. 5d). In the WW region Q_{temp}^{tend} is much smaller than the Q_{temp}^{vmix} and Q_{temp}^{surf} , and shows very little trend during the period about 15-30 years before a convective event when the salinity is slowly increasing. To first order, Q_{temp}^{vmix} is in phase with the surface temperature anomaly, while Q_{temp}^{surf} , Q_{temp}^{eddy} , and Q_{temp}^{adv} damp this anomaly. Taken together, Figs. 5c and d show that the primary driver of changes in the surface density in the Western Weddell Sea that are out of phase with the convective cycle (and thus correlate with $\partial\chi_1/\partial t$) is the advection of salinity.

As would be expected from Fig. 5d, a parallel analysis of the heat budget in WW region from 60-1000m (Fig. S2a) shows that upward mixing of heat through the top of this box is the primary means by which heat is exported from intermediate depths. This export $Q_{temp}^{vmix}(60m)$ is largely balanced by eddy fluxes of heat moving warm deep waters into convective regions. However, the temperature signal at depth (Fig. S2a) is clearly out of phase with that at the surface, though the amplitude is far smaller. The net storage of heat at depth is driven by variations in the advective flux of heat (black and red lines, Fig. S2c).

Repeating this analysis in the Eastern Weddell region (Fig. 6) shows a broadly similar picture for SSS and SST variability (Figures 6a and b) although the peak to trough differences in SST (1.35°C) and salinity (0.48 PSU) are slightly smaller than in the Western Weddell region. The intermediate-depth temperatures (Fig. S2b) vary much less, reflecting the fact that we do not get deep convection in the EW region in this model. The term balances for salinity (Fig. 7c) show some interesting differences with those in the WW. Peaks in Q_{salt}^{tend} (black line) have a noticeably different shape are shifted slightly earlier, and so when χ_1 peaks, EW salinity is already dropping. Q_{salt}^{adv} (red line, Fig. 6c) has a relatively similar out-of-phase cycle to χ_1 , adding salt before the peak and removing it after the peak. The surface dilution flux (Q_{salt}^{surf} cyan line, Fig. 6c) is correlated with the temperature and the lateral eddy flux (Q_{salt}^{eddy} magenta line, Fig. 6c) is anticorrelated with χ_1 and SSS as in the Western Weddell. Both fluxes, however, have a much reduced amplitude relative to the Western Weddell region. The biggest difference is the vertical mixing flux Q_{salt}^{vmix} (dark blue line, Fig. 6c). Instead of being in phase with χ_1 , mixing of salt in the Eastern Weddell is out of phase with it, with the highest signal seen before the peak in temperature.

iii. Explaining the drivers of variation in near-surface salinity advection in the Western Weddell Sea.

To first order, the anomalous advective salt tendency $Q_{salt}^{adv'}$ can be written as the anomalous advection of the mean salinity plus the mean advection of anomalous salinity

$$Q_{salt}^{adv'} = \vec{u}' \nabla \langle S \rangle + \langle \vec{u} \rangle \nabla S' \quad (7)$$

where the brackets denote the mean field and the primes denote anomalies. Analysis of $Q_{salt}^{adv'}$ in the Western Weddell Sea shows that both terms in equation (7) contribute to the total tendency. Figure 7a shows the regression between Q_{salt}^{adv} (red line in Fig. 5c) and anomalous sea surface salinity S' with vectors of the mean horizontal transport overlaid. We see that a +1 kg.m⁻².yr⁻¹ tendency within the Western Weddell Sea corresponds to higher salinity in the Eastern Weddell Sea, with a spatial pattern that maps closely onto the Mode 3 salinity anomaly field seen in Fig. 3b. Regressing Q_{salt}^{adv} onto the salinity difference between the Eastern Weddell and Western Weddell boxes gives a correlation coefficient of 0.64 and a regression coefficient of 0.034 PSU/(kg.m⁻².yr⁻¹). As seen in Fig. 7a, the eastern boundary of the Western Weddell Sea box has an inflow on its Southern flank (with a transport of 0.8 Sv) and an outflow on its northern flank (with a transport of about 2 Sv). Multiplying these flows by the anomalous gradient between the boxes and dividing by the area of the Western Weddell box (2.3×10^{12} m²) gives a range of 0.37-1.0 kg.m⁻².yr⁻¹ so that this mechanism is indeed big enough to play an important role in explaining the variation in Q_{salt}^{adv} seen in Fig. 5c. Combined with the fact that the salinity difference between the Western Weddell and Eastern Weddell Seas correlates with the

χ_3 coefficient with a correlation of 0.8, this analysis helps explain why negative values of χ_3 lead to positive trends in χ_1 .

Moving on to the effect of anomalous velocities, Figure 7b shows the regression between advective salt tendencies Q_{salt}^{adv} and transport anomalies with the mean salinity overlaid. The perturbation that emerges has a complex structure with two counter-rotating gyres to the north of the Western Weddell box which produce a jet leaving the box across the northern boundary. Examining the relationship of this pattern of anomalous currents to the mean currents shown in Fig. 7a, we see that the anomaly corresponds to a diversion of some of the transport within the Circumpolar Current into the Western Weddell Sea with stronger transport removing more fresh water across the northern face. The transport across this northern face is well-correlated with $Q_{salt}^{adv'}$ (0.73) and the anomalous transport associated with a $+1 \text{ kg.m}^{-2}.\text{yr}^{-1}$ increase in $Q_{salt}^{adv'}$ is 0.23 Sv. This implies an effective salinity difference associated with this flow of about 0.3 PSU, about what is seen between the northern face of the Western Weddell Sea box where there is anomalous outflow and the eastern and western faces where there is anomalous inflow.

iv. Understanding the evolution in spatial salinity fluxes between the Eastern and Western Weddell Seas-the role of mixing

We have shown that spatial salinity gradients are an important part of explaining the advective salinity tendencies, and thus in connecting χ_3 to $\partial\chi_1/\partial t$. But how does χ_1 connect to $\frac{\partial\chi_3}{\partial t}$? We explore this question by looking at the composite difference of the salinity between the EW and WW regions (Fig. 8a) and the individual physical terms that combine to produce the

time tendency of this difference. During the peak of convection, the salt tendency difference $Q_{salt}^{tend'}(EW) - Q_{salt}^{tend'}(WW)$ (black line Fig. 8b) is negative, so that the EW freshens relative to the WW. Both the surface dilution flux Q_{salt}^{surf} (light blue line) and vertical mixing flux Q_{salt}^{vmix} (dark blue line) are positively correlated with this change, while the advective (Q_{salt}^{adv} , red line) and eddy mixing flux (Q_{salt}^{eddy} , magenta line) act to oppose it.

Examining the composited cycle of Q_{salt}^{vmix} in Figs. 5, 6 and 8 we see that during deep convective events this flux acts to bring warm salty water up from below in the WW region, but not in the EW region. This then makes the Eastern Weddell “relatively fresh” and allows for advection to shut off the convection. The fact that this happens in the Western Weddell and not the Eastern Weddell is a function of the surface densities. Throughout the convective cycle, the Eastern Weddell is fresher and lighter than the Western Weddell (as shown in Fig. 8a the difference between these basins becomes small in magnitude before convection initiates, but it is still negative). We see the impact of this by examining two typical points in the Western Weddell region at 63°S, 40°W (located in the region of the highest climatological surface salinities seen in Fig. 7b and denoted by the yellow + in that figure) and one in the Eastern Weddell region at 60.8°S, 1.5°W. (located in the center of the region of highest anomalous salinity in Fig. 7a and denoted by the blue + in that figure). We examine individual points rather than averaging over the entire region because spatial averaging masks the sharp thresholds associated with spatially heterogeneous convection. We plot sea surface salinity (Fig. 9a) and sea surface density (Fig. 9b) during the month of September against the monthly mean thickness of the average boundary layer depth H_{blt} produced by the KPP mixing scheme at these two points. This depth represents the depth to which active mixing occurs and is generally shallower than the classic mixed layer depth computed using density stratification. Both locations have relatively similar mixed layer

depths (with an average ~60m and maxima around 100m) for salinities below about 34.4 PSU and densities below 27.7 kg m^{-3} . As salinities and densities rise above this threshold at the Western Weddell site (red +) monthly mean boundary layer depths deepen to thousands of meters. However, the salinities and densities at the Eastern Weddell site (blue o) never cross this threshold, and the boundary layer depths never break through the permanent halocline at this point.

Q_{salt}^{vmix} for the month of September shows a clear relationship with density and salinity at the Western Weddell Sea point (red + marks, Figs. 9c and d). The salt mixed up by deep convective events just during this month can be significant, particularly as the mixed layer is deepening. However, in the Eastern Weddell Sea the relationship between Q_{salt}^{vmix} and density actually drops as the density increases, with a correlation between the two time series of -0.3. As the salinities and temperatures at these two points are reasonably well correlated (0.75 and 0.70 respectively), this suggests that 15 years before the peak of convection in the Eastern Weddell $Q_{salt}^{vmix'} > 0$ (more salt than average is being mixed up) whereas in the Western Weddell $Q_{salt}^{vmix'} < 0$ (less salt than average is being mixed up). The spatial pattern of correlation between $Q_{salt}^{vmix'}$ and SST in the Weddell Sea is shown in Fig. 10a, and shows very large positive values in the Western Weddell Sea and along the Antarctic continent but strong anticorrelation centered along latitude 55°S between the prime meridian and about 100°E .

What accounts for such a relationship? If we examine the surface salinity and heat fluxes, in Fig. 6 it is not clear why we would get enhanced $Q_{salt}^{vmix'}$ during cold periods, as the anomalous surface heat flux is positive and surface salt flux is negative- resulting in an addition of buoyancy. $Q_{salt}^{vmix'}$ in the Eastern Weddell also shows relatively little correlation with surface density, salinity, or mixed layer depth. However, we do see a relationship with local winds. As

shown in Fig. 10a,b the regions which see the greatest reduction in $Q_{salt}^{vmix'}$ during convective events also see a reduction in wind stress (similar to that reported by Cabré et al., 2017). At any given month, we see reasonable correlation between the strength of the winds and the amount of mixing from below. As discussed in Ragen et al. (2020) this relationship arises because a warmer Southern Ocean produces less thermal wind shear in the atmosphere, which in turn results in a reduced eddy flux of momentum into the jet stream. The CM2Mc models have a reasonable representation of this process, with a marginal sensitivity to gradients in atmosphere geopotential height similar to that seen in the ECMWF reanalysis of Dee et al. (2011).

As shown in Fig. 10c, the changes in wind stress resulting from a 1°C change in SST over the Weddell Sea are not small, with a relaxation of winds within the “Furious Fifties” of 0.012 Pa.K⁻¹, implying declines of ~20% during the peak of convective events. Such declines in wind suppress mixing in this region both by reducing mechanical stirring and by reducing the transport of denser water from the south within the Ekman layer. By contrast, closer to the continent higher surface densities do appear to be capable of driving more mixing from below.

v. Understanding the evolution of currents

The changes in wind stress seen in Fig. 10c are a big part of understanding the changes in advective salt flux driven by anomalous currents. If we examine anomalous currents on the northern side of the Western Weddell box, we find that they are only weakly anticorrelated with SST in this box at zero lag (-0.29) but are strongly anticorrelated with the SST when lagged several years (the peak of -0.84 is found at a lag of 9 years). That northward transport hits a minimum (maximum) about a decade after a warming (cooling) helps explain why $Q_{salt}^{adv'}$ has an

extremum with the opposite sign as each peak in χ_1 (Fig. 5c) with about the same lag. Such a lag is broadly consistent with wind stress changes generating a Rossby wave that requires some time to propagate westward. Mean westward barotropic velocities in the model are less than 0.005 m.s⁻¹ and with a Rossby radius in this region of around 15 km, the 1st baroclinic mode has a speed of 0.003 m.s⁻¹. With a total westward propagation speed of 0.005 m.s⁻¹ the time required to transit 20 degrees of longitude at 55°S is about 10 years.

4. Discussion and conclusions

a.) Summary of mechanisms involved in producing periodic convection

A summary of the mechanisms involved in generating multidecadal variability in our AREDI400 simulation is shown in Fig. 11a. Loop 1 is on the left, and involves convection bringing warmer water to the surface and melting ice, allowing for more loss of heat to the atmosphere and evaporation and thus increasing the buoyancy flux driving convection (cyan lines in Fig. 5c and d). This positive feedback on convection helps to explain how convection can persist over many decades and is reminiscent of the mechanisms suggested as long ago as Martinson et al., 1981. In terms of equations 6 a and b, the positive relationship between χ_1 and $\partial\chi_1/\partial t$ (salinity perturbations in Western Weddell grow with about a 12 year timescale) is able to overcome the negative relationship between χ_3 and $\partial\chi_3/\partial t$ (salinity differences between Eastern Weddell and Western Weddell are damped with about a 15 year timescale). The second loop in Fig. 11 involves increases in Western Weddell temperature producing a relaxation in winds (Fig. 10b,c) which we assert produces lagged changes in currents (Fig. 7b) which in turn cause a reduction in the supply of saltier water to the Western Weddell (red lines Fig. 6b). This

mechanism is similar to those of Latif et al. (1998) and Marshall et al. (2001) for the North Atlantic, but with the difference that the focus is on salinity rather than temperature.

The relaxation of winds is also involved in a third feedback loop, as it suppresses mixing in the Eastern Weddell (Fig. 10a), resulting in a stronger salinity gradient between the Eastern and Western Weddell. This helps to drive a greater import of fresher waters into the Western Weddell Sea (Fig. 7a). The resulting negative advective salinity tendency acts to shut off the convection, so that positive χ_3 leads to negative $\partial\chi_1/\partial t$. A fourth feedback loop involving a greater supply of salt to the surface through deep convection in the Western Weddell Sea (Fig. 5c) also increases the salinity gradient between the Eastern and Western Weddell Seas, but with less of a delay. In the “convection-off” phase, these processes work in reverse.

Another way of visualizing the evolution of the system is via a T-S diagram of wintertime properties (Fig. 11b). Convection is initiated when the Western Weddell becomes salty enough for heat and salt to start mixing up from below- a time during which the density and salinity difference between the Eastern and Western Weddell is at a minimum. Both the Eastern and Western Weddell then warm as heat and salt are mixed up in the Western Weddell (dark blue line Figs 5c and d) and advected into the Eastern Weddell (red lines Fig. 6c and d). As noted before, warmer surface waters lose more heat to the atmosphere and also evaporate more. This produces a negative buoyancy flux that enhances mixing, and we see that the wintertime densities in both regions increase. Simultaneously, however, the difference in salinity between the Eastern Weddell and Western Weddell is increasing, and eventually results in a shutoff of convection. Note that even at the height of convection, the average surface water density over the Western Weddell region almost never reaches the average density at 300m (shown by the red X in Fig. 11b)-implying that convection is still patchy. The system now moves to a state where the

Eastern Weddell is much fresher and lighter than the Western Weddell. At this point, the increase in winds results in an increase both in vertical mixing in the Eastern Weddell (dark blue line, Fig. 6c) and advective salt supply to/freshwater export out of the Western Weddell (red line Fig. 5c), with the Eastern Weddell salinity increasing more quickly than the Western Weddell salinity (Fig. 8a) so that the salinity difference drops back to its minimum value and convection is able to begin again.

b.) Evaluating model realism

A number of questions arise from this study. The first is whether there is any basis for either the oscillations or the proposed mechanisms in the real world. A huge challenge to doing this is the relatively short observational record. Although open-ocean convection in the Weddell Sea was observed in the late 1970s at the start of the satellite era (Zwally and Gloerson, 1977) and there is some evidence that the so-called Weddell Polynya was present even earlier than this (Meier et al., 2013), the original polynya disappeared after only a few years. While its reappearance in 2016 and 2017 (Campbell et al., 2019) is technically consistent with multidecadal variability of the sort seen in many models, the record of direct observations is too short to constrain a period of variability. Moreover, the modern observational period has occurred during a period of anthropogenic global warming with an enhanced Southern Hemisphere hydrological cycle (Durack and Wijffels, 2010) as well as shifts in winds due to the ozone hole (Thompson and Solomon, 2002) both of which would be expected to produce changes in Southern Ocean convection (de Lavergne et al., 2014; Ferreira et al., 2015; Seviour et

al., 2017). Lack of measurements is also a problem in reconstructing regional decadal hydrographic trends (Heuzé et al. 2015a), particularly for the wintertime salinities under sea ice that are important in initiating deep convection in our models (Fig. 9).

A number of investigators have used proxy records to make estimates of convection, vertical exchange and winds but these records are inconsistent. Latif et al. (2013) argued that a multicentennial signal could be found in tree-ring records from long-lived Huon pine in Tasmania. Estimates of radiocarbon reservoir ages have also shown multidecadal variability in Southern Ocean vertical exchange (Hua et al., 2015; Paterne et al., 2019), but because of radiocarbon's long equilibration time it is difficult to isolate which regions drive such changes. Proxy records of sedimentation in New Zealand suggest multidecadal variability in winds (Hinajosa et al., 2017)- however, in our model New Zealand is outside of the main area where winds driven by Weddell Sea convection would be expected to vary.

When it comes to individual mechanisms, however, we may be on more solid ground. Campbell et al., (2019) used float observations to look at the conditions that led to the most recent polynya and concluded that surface density changes driven by salinity were more important than deep density changes. The relationship between Southern Ocean winds and temperature gradients between the tropics and high Southern Ocean latitudes is attested in reanalysis with a marginal dependence similar to what is found in our models (Ragen et al., 2020). Sea surface salinity anomalies that were initiated by winds, propagated to the Western Weddell Sea and excited convection there were also seen in the higher-resolution regional ice-ocean model of Hellmer et al., (2009). Moreover, changes in Southern Ocean winds have been associated with a “two-timescale” response in which a persistent increase in winds produces an initial cooling of the Southern Ocean associated with advection of cold, fresh water from the

south (Ferreira et al., 2015) and consequent suppression of mixing (Seviour et al., 2017) followed by warming either driven by upwelling subsurface warm water (Ferreira et al., 2015) or export of freshwater from convective regions (Seviour et al., 2017). Although Doddridge et al. (2019) do not find a strong warming phase in an idealized eddy-resolving model, Kostov et al. (2017) argued that such a response was not only found across a range of climate models but could also be inferred from observational data. Mathematically, such a response could lead to a delayed oscillator with dynamics similar to that seen here, as the cooling stage would be expected to enhance winds while the warming stage would be expected to cause them to relax.

c. Evaluating model robustness

Finally, we briefly consider the robustness of these results to the parameterization of mesoscale eddies. In previous work (Thomas et al., 2018) we showed that the AREDI400 model produced regular, large-amplitude variability in convection, the AREDI2400 model produced much smaller-amplitude variability and the AGMmin600 produced very little convection at all. T-S plots of wintertime surface waters for these two simulations are shown in Fig. 12a and b. The decline in wintertime variability is clearly visible in both model simulations but has a very different expression. In AREDI2400 simulation (Fig. 12a) both regions sit on the “convection on” part of the cycle in Fig. 10b but oscillate back and forth along a relatively small range of densities. The differences between the Eastern and Western Weddell salinities are much smaller, implying that the ability to generate advective salinity tendency anomalies is also much less. Qualitatively we can explain this in terms of the negative feedback loops in Fig. 11 not being

activated in this model. Note also that the subsurface densities at 300m are actually slightly larger- making it more difficult for the surface densities to break through to the deep over a wide area. In the AGMmin600 simulation (Fig. 12b) the two regions sit on the “convection off” part of the cycle in Fig. 11b, with very little variation in temperature, and some variation in salinity. Note again that the maximum densities are always significantly lighter than the densities at 300m. For this case, it appears that the higher value of A_{GM} within convective regions prevents the initial positive feedback loop in Fig. 11a involving mixing, temperature and buoyancy loss. Given the strong sensitivity of the feedback loops to processes that vary significantly across climate models, it is unsurprising that such models do not agree about the periodicity of Southern Ocean convection.

The lack of cyclicity in watermasses in the Weddell Sea does not, however, mean that the AREDI400 and AREDI2400 models have no variability at all. As shown in Figs. 12c and d, the first rotated EOF mode for both simulations still shows strong centers of action for temperature variability-just not in the Weddell Sea. In AREDI2400, the first mode of temperature variability is shifted to the west of the Antarctic Peninsula, while in AGMmin600 the primary center of action moves even further to the west. This latter region is one where the sea ice edge intersects a region of low stratification and has been shown to be sensitive to changes in winds in Seviour et al. (2017). The SST structure shown in Fig. 12d for AGMmin600 is very similar to the SST structure in Mode 2 from AREDI400 (Fig. S1a) though the salinity structure (not shown) matches less well. Preliminary analysis indicates the variability of these centers of action is governed by dynamics similar to those found in AREDI400, with salinity gradients within the Ross Sea varying out of phase with SSTs. A full analysis will be provided in future work.

748 *d.) Summary of key result*

749 In conclusion, we show that a coarse-resolution climate model can generate periodic
750 multidecadal variability as a result of surface forcing. The long time period arises from:

751 1. The self-reinforcing nature of convection on buoyancy fluxes as heat mixed to the surface
752 results in buoyancy loss.

753 2. The long time scale required for salinity anomalies forced by small imbalances in mixing and
754 advection to build up in the mixed layer

755 3. A delay between the near-instantaneous response of Southern Ocean winds to the warming
756 caused by deep ocean convection and changes in the wind-driven currents in the convective
757 region.

758 We suggest that it would be profitable to examine other, higher resolution climate models as well
759 as observational data for these mechanisms.

760

761 **Data availability statement:** Monthly SST, SSS mixed layer depths, winds and regional term
762 balances as well as the Matlab code used to do the Principal Oscillation Pattern analysis has been
763 made available through the Johns Hopkins Data Archive at <https://doi.org/10.7281/T1/FIS2QT>

764 **Acknowledgments:** Simulations were performed on the Homewood High Performance
765 Computing Cluster supported by the JHU Institute for Data-Intensive Science and Engineering.
766 AG, CS and MAP acknowledge support under NSF Grant OCE-1756568. GR was supported
767 with a Vagelos Fellowship from the University of Pennsylvania. We thank Atousa Saberi, Ali

768 Siddiqui, Renske Gelderloos, Karen Heywood and three anonymous reviewers for comments on
769 earlier drafts of this manuscript.

References

- Bahl, A., A. Gnanadesikan, A., and M.A. Pradal, 2019: Variations in ocean deoxygenation across Earth System Models: Isolating the role of parametrized lateral mixing. *Global Biogeochemical Cycles*, **33**, 703-724, doi://10.1029/2018GB006121.
- Behrens, E., G. Rickard, O. Morgenstern, T. Martin, A. Osprey, and M. Joshi, 2016: Southern Ocean deep convection in global climate models: A driver for variability of subpolar gyres and Drake Passage transport on decadal timescales, *J. Geophys. Res.-Oceans*, **121**, 3905–3925, doi:10.1002/2015JC011286.
- Bernardello, R., I. Marinov, J. B. Palter, E. D. Galbraith, and J. L. Sarmiento , 2014: Impact of Weddell Sea deep convection on natural and anthropogenic carbon in a climate model. *Geophys. Res. Lett.*, **41**, 7262-7269.
- Beadling, R.L., J.L. Russell, R.J. Stouffer, P.J. Goodman, and M. Mazloff, 2019: Assessing the Quality of Southern Ocean Circulation in CMIP5 AOGCM and Earth System Model Simulations. *J. Climate*, **32**, 5915–5940, <https://doi.org/10.1175/JCLI-D-19-0263.1>
- Cabré, A., I. Marinov and A. Gnanadesikan, 2017: Global atmospheric teleconnections and multidecadal climate oscillations driven by Southern Ocean convection, *J. Climate*, **30**, 8107-8126.
- Campbell, E.C., E.A Wilson, G.W.K Moore, S.C. Riser, C.E. Brayton, M.R. Mazloff and L.D. Talley, 2019: Antarctic offshore polynyas linked to Southern Hemisphere climate anomalies. *Nature*, **570**(7761), 319-325. doi:10.1038/s41586-019-1294-0.

790 Cheon, W. G., Y.G. Park, J.R. Toggweiler, and S.K. Lee, 2014: The relationship of Weddell
791 Polynya and open-ocean deep convection to the Southern Hemisphere westerlies. *Journal of*
792 *Phys. Oceanogr.*, **44**, 694-713.

793 de Lavergne, C. and Coauthors, 2014: Cessation of deep convection in the open Southern Ocean
794 under anthropogenic climate change, *Nat. Climate Change*, **4**, 278-282.

795 Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of
796 the data assimilation system. *Q.J.R. Meteorol. Soc.*, **137**, 553-597, doi:10.1002/qj.828

797 Doddridge, E. W., J. Marshall, H. Song, J.-M. Campin, M. Kelley, & L.S. Nazarenko, (2019).
798 Eddy compensation dampens Southern Ocean sea surface temperature response to westerly
799 wind trends. *Geophysical Research Letters*, **46**, 4365– 4377.
800 <https://doi.org/10.1029/2019GL082758>

801 Dufour, C. O., A.K. Morrison, S. M. Griffies, I. Frenger, H. Zanowski, and M. Winton, 2017:
802 Preconditioning of the Weddell Sea polynya by the ocean mesoscale and dense water
803 overflows. *J. Climate*, **30**, 7719-7737.

804 Durack, P. J., S. E. Wijffels, 2010: Fifty-Year Trends in Global Ocean Salinities and Their
805 Relationship to Broad-Scale Warming. *J. Climate*, **23**, 4342–4362.

806 Ferreira, D., J. Marshall, C. M. Bitz, S. Solomon, and A. Plumb, 2015: Antarctic Ocean and Sea
807 Ice Response to Ozone Depletion: A Two-Time-Scale Problem. *J. Climate*, **28**, 1206–1226.
808 doi: <http://dx.doi.org/10.1175/JCLI-D-14-00313.1>

809 Forget, G., 2010: Mapping ocean observations in a dynamical framework: A 2004–06 ocean
810 atlas. *J. Phys. Oceanogr.*, **40**, 1201–1221.

811 Fox-Kemper, B. and Coauthors, 2011: Parameterization of mixed layer eddies. III:
812 Implementation and impact in global ocean climate simulations. *Ocean Modelling*, **39**, 61-78.

813 Galbraith, E.D. and Coauthors, 2011: Climate Variability and Radiocarbon in the CM2Mc Earth
814 System Model, *J. Climate*, **24**, 4230-4254.

815 Gent, P.R., and J.C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *Journal*
816 *of Phys. Oceanogr.* **20**, 150-155.

817 Gnanadesikan, A., R. D. Slater, P. S. Swathi, and G. K Vallis, 2005: The energetics of ocean heat
818 transport. *J. Climate*, **18**(14),2604-2616. doi:10.1175/JCLI3436.1.

819 Gnanadesikan, A., R. Abernathey and M.-A. Pradal, 2015: Exploring the isopycnal mixing and
820 helium-heat paradoxes in a suite of Earth System Models, *Ocean Sci.*, **11**,591-605

821 Gordon, A.L., M. Visbeck, and J.C. Comiso, 2007: [A Possible Link between the Weddell](#)
822 [Polynya and the Southern Annular Mode](#). *J. Climate*, **20**, 2558–2571,
823 <https://doi.org/10.1175/JCLI4046.1>

824 Griffies S.M. and E. Tziperman. 1995: A linear thermohaline oscillator driven by stochastic
825 atmospheric forcing. *J. Climate*, **8**(10):2440–2453.

826 Hasselmann, K., 1988: PIPs and POPs' The Reduction of Complex Dynamical Systems Using
827 Principal Interaction and Oscillation Patterns, *J. Geophys. Res.*, **93**(D9), 11,015-11,021.

828 Hausmann, U., A. Czaja, and J. Marshall, 2016: Estimates of Air–Sea Feedbacks on Sea Surface
829 Temperature Anomalies in the Southern Ocean. *J. Climate*, **29**, 439–454.

830 Hellmer, H. H., F. Kauker, and R. Timmermann, 2009: Weddell Sea anomalies: Excitation,
831 propagation, and possible consequences, *Geophys. Res. Lett.*, **36**, L12605,
832 doi:10.1029/2009GL038407.

833 Heuzé, C., K.J. Heywood, D.P. Stevens and J.K. Ridley, 2013: Southern Ocean Bottom Water
834 Characteristics in CMIP5 models, *Geophys. Res. Lett.*, **40**(7), pp 1409-1414,
835 doi:10.1002/grl.50287.

836 Heuzé, C., F. Vivier, J. Le Sommer, J.-M. Molines, and T. Penduff, 2015a: Can we map the
837 interannual variability of the whole upper Southern Ocean with the current database of
838 hydrographic observations?, *J. Geophys. Res. Oceans*, **120**, 7960– 7978,
839 doi:10.1002/2015JC011115.

840 Heuzé, C., J.K. Ridley, D. Calvert, D.P. Stevens and K.J. Heywood, 2015b: Increasing vertical
841 mixing to reduce Southern Ocean deep convection in NEMO3.4, *Geosci. Model Dev.*,
842 NEMO special issue, **8**, 3119-3130 doi:10.5194/gmd-8-3119-2015.

843 Hinajosa, J.L., C. M. Moy, C.S. Stirling, G.S. Wilson, T.I. Eglinton, 2017: A New Zealand
844 perspective on centennial-scale Southern Hemisphere westerly wind shifts during the last two
845 millennia, *Quatern. Sci. Rev.*, **172**, 32-43. <https://doi.org/10.1016/j.quascirev.2017.07.016>

846 Hua, Q., G.E. Webb, J.X. Zhao, L.D. Nothdurft, M. Lybolt, G.J. Price, and B.N. Opdyke, 2015:
847 Large variations in the Holocene marine radiocarbon reservoir effect reflect ocean circulation
848 and climatic changes. *Earth and Planetary Science Letters*, **422**, 33-44.

849 Kaiser, H. F., 1958: The Varimax criterion for analytic rotations in factor analysis.
850 *Psychometrika*, **23**, 187–200.

851 Kostov, Y., J. Marshall, U. Hausmann, K.C. Armour, D. Ferreira and M.M. Holland, 2017: Fast
852 and slow responses of Southern Ocean sea surface temperature to SAM in coupled climate
853 models. *Clim. Dyn.* **48**, 1595–1609. <https://doi.org/10.1007/s00382-016-3162-z>

854 Large, W. G., J.C. McWilliams, and S.C. Doney, 1994: Oceanic vertical mixing: A review and a
855 model with a nonlocal boundary layer parameterization. *Rev. Geophys.*, **32**(4), 363–403.

856 Latif M., A. Grötzner, M. Münnich, E. Maier-Reimer, S. Venzke and T.P. Barnett, 1996: A
857 Mechanism for Decadal Climate Variability. In: Anderson D.L.T., Willebrand J. (eds)
858 *Decadal Climate Variability*. NATO ASI Series (Series I: Global Environmental Change),
859 vol 44. Springer, Berlin, Heidelberg. doi://10.1007/F978-3-662-03291-6_6.

860 Latif, M., T. Martin, and W. Park, 2013: Southern Ocean sector centennial climate variability
861 and recent decadal trends. *J. Climate*, **26**, 7767–7782.

862 Lian, T. and D. Chen, 2012: [An Evaluation of Rotated EOF Analysis and Its Application to Tropical](https://doi.org/10.1175/JCLI-D-11-00663.1)
863 [Pacific SST Variability](https://doi.org/10.1175/JCLI-D-11-00663.1). *J. Climate*, **25**, 5361–5373, <https://doi.org/10.1175/JCLI-D-11-00663.1>

864 Loving, J. L., and G. K. Vallis, 2005: Mechanisms for climate variability during glacial and
865 interglacial periods. *Paleoceanography*, **20**, PA4024, doi:10.1029/2004PA00111

866 Marshall, J., H. Johnson, and J. Goodman, 2001: A study of the interaction of the North Atlantic
867 Oscillation with ocean circulation. *J. Climate*, **14**, 1399–1421.

868 Martin, T., W. Park, and M. Latif, 2013: Multi-centennial variability controlled by Southern
869 Ocean convection in the Kiel Climate Model. *Clim Dyn* **40**, 2005–2022, doi:10.1007/s00382-
870 012-1586-7

871 Martinson, D.G., P.D. Killworth, and A.L. Gordon, 1981: [A Convective Model for the Weddell](#)
872 [Polynya](#). *J. Phys. Oceanogr.*, **11**, 466–488, [https://doi.org/10.1175/1520-](https://doi.org/10.1175/1520-0485(1981)011<0466:ACMFTW>2.0.CO;2)
873 [0485\(1981\)011<0466:ACMFTW>2.0.CO;2](#)

874 Meier, W. N., D. Gallaher, G.G. Campbell, 2013: New estimates of Arctic and Antarctic sea ice
875 extent during September 1964 from recovered Nimbus I satellite imagery. *The Cryosphere*,
876 **7**(2), 699-705.

877 Paterne, M., E. Michel, V. Héros, 2019: Variability of marine ^{14}C reservoir ages in the Southern
878 Ocean highlighting circulation changes between 1910 and 1950. *Earth and Planetary Science*
879 *Letters*, **511**, 99-104.

880 Pradal, M.A. and A. Gnanadesikan, 2014: How does the Redi parameter for mesoscale mixing
881 impact global climate in an Earth System Model?, *J. Adv. Model. Earth Syst.*, **6**, 586– 601,
882 doi:[10.1002/2013MS000273](https://doi.org/10.1002/2013MS000273).

883 Ragen, S., M.-A. Pradal and A. Gnanadesikan, 2020: Impact of changing isopycnal mixing on
884 the Antarctic Circumpolar Current in a coupled climate model, *J. Phys. Oceanogr.*, **50**, 965-
885 982. <https://doi.org/10.1175/JPO-D-19-0249.1>.

886 Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.* **12**,
887 1154-1158.

888 Reintges, A., T. Martin, M. Latif, and W. Park, 2017: Physical controls of Southern Ocean
889 deep-convection variability in CMIP5 models and the Kiel Climate Model, *Geophys. Res.*
890 *Lett.*, **44**, 6951– 6958, doi:10.1002/2017GL074087.

891 Seviour, W., A. Gnanadesikan, D.W. Waugh and M.A. Pradal, 2017: Transient response of the
892 Southern Ocean to changing ozone: Regional responses and physical mechanisms, *J. Climate*
893 **30**, 2463-2480.

894 Sloyan, B. M., and S. R. Rintoul, 2001: The Southern Ocean limb of the global deep overturning
895 circulation. *J. Phys. Oceanogr.*, **31**, 143–173.

896 Talley, L. D., 2008: Freshwater transport estimates and the global overturning circulation:
897 Shallow, deep and throughflow components. *Prog. Oceanogr.*, **78**, 257–303.

898 Thomas, J., D.W. Waugh, and A. Gnanadesikan, 2018: Understanding interannual variability in
899 ocean heat and carbon content in a suite of coupled climate models, *J. Climate*, **31**, 1467-
900 1482.

901 Thompson D, S. Solomon, 2002: Interpretation of recent Southern Hemisphere climate change.
902 **296**(5569):895–899. doi: 10.1126/science.1069270

903 Timmermann, R. and A. Beckmann, 2004: Parameterization of vertical mixing in the Weddell
904 Sea , *Ocean Modelling*, **6** (1), pp. 83-100 . doi: 10.1016/S1463-5003(02)0061-6

905 Turner, J., T.J. Bracegirdle, T. Phillips, G.J. Marshall, and J.S. Hosking, 2013: An Initial
906 Assessment of Antarctic Sea Ice Extent in the CMIP5 Models. *J. Climate*, **26**, 1473–1484,
907 <https://doi.org/10.1175/JCLI-D-12-00068.1>

908 Visbeck, M., Marshall, J., Haine, T., & Spall, M. (1997). Specification of eddy transfer
909 coefficients in coarse-resolution ocean circulation models. *J. Phys. Oceanogr.*, **27**(3), 381-
910 402.

911 Welander, P., 1982: A simple heat-salt oscillator, *Dyn. Atmos. Oceans*, **6**, 233-242.

912 Weisse, R., U. Mikolajewicz, and E. Maier-Reimer, 1994: Decadal variability of the North
 913 Atlantic in an ocean general circulation model. *J. of Geophys. Res.-Oceans*, **99**, 12411-
 914 12421.

915 Xu, J. S. and H. Von Storch, 1990: Predicting the state of the Southern Oscillation using
 916 principal oscillation pattern analysis. *J. Climate*, **3**(12), 1316-1329.

917 Zanowski, H., R. Hallberg, and J.L. Sarmiento, 2015: Abyssal Ocean Warming and
 918 Salinification after Weddell Polynyas in the GFDL CM2G Coupled Climate Model. *J. Phys.*
 919 *Oceanogr.*, **45**, 2755–2772, <https://doi.org/10.1175/JPO-D-15-0109.1>

920 Zanowski, H. and R. Hallberg, 2017: [Weddell Polynya Transport Mechanisms in the Abyssal](https://doi.org/10.1175/JPO-D-17-0091.1)
 921 [Ocean](https://doi.org/10.1175/JPO-D-17-0091.1). *J. Phys. Oceanogr.*, **47**, 2907–2925, <https://doi.org/10.1175/JPO-D-17-0091.1>

922 Zhang, L., T.L. Delworth, and L. Jia, 2017: [Diagnosis of Decadal Predictability of Southern](https://doi.org/10.1175/JCLI-D-16-0537.1)
 923 [Ocean Sea Surface Temperature in the GFDL CM2.1 Model](https://doi.org/10.1175/JCLI-D-16-0537.1). *J. Climate*, **30**, 6309–6328,
 924 <https://doi.org/10.1175/JCLI-D-16-0537.1>

925 Zwally, H. J. and P. Gloersen, 1977: Passive microwave images of the polar regions and
 926 research applications. *Polar Rec.*, **18**, 431–450.

927

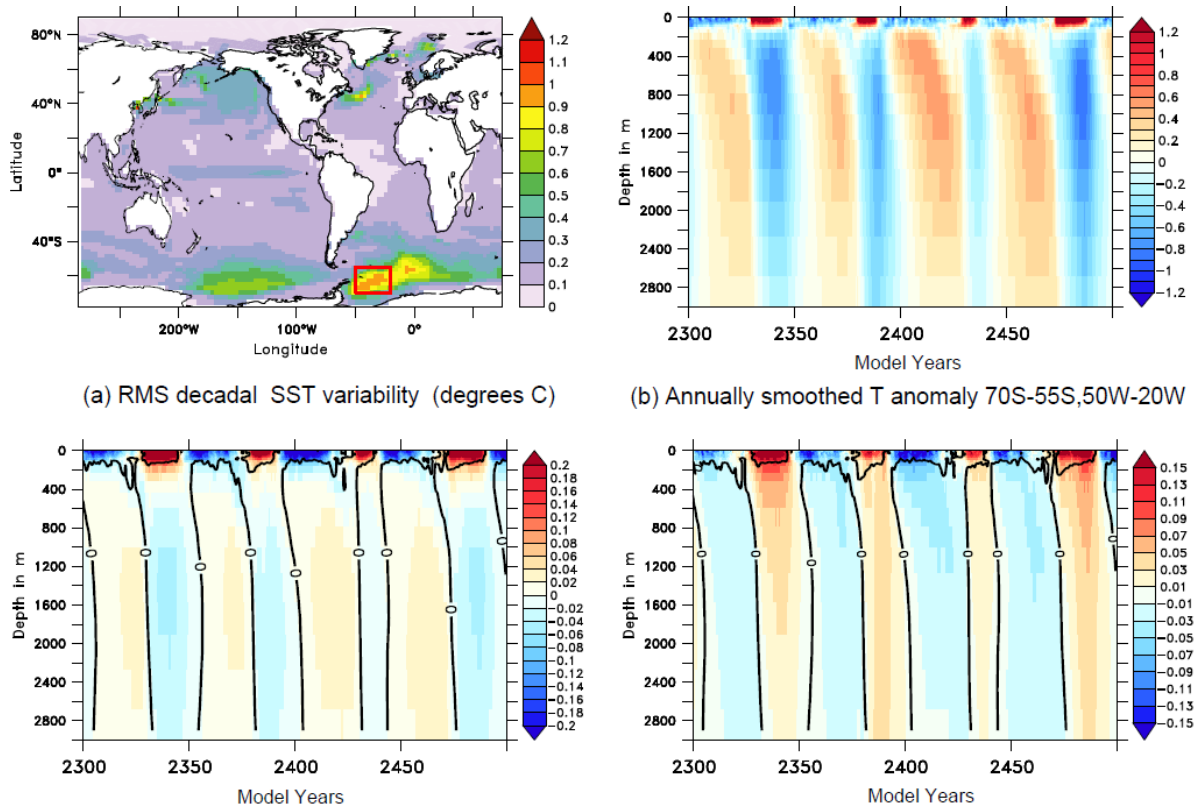
928 Table 1: Metrics of Southern Ocean Circulation and hydrography from three simulations considered in
929 this paper compared with observations and CMIP5 model range from recent publications.

	Observational range	CMIP5 range	ARED1400	ARED12400	AGMmin600
ACC transport in Drake Passage (Sv)	137±8 (Cunningham et al. 2003) 170±11 (Donahue et al., 2016)	88-242 (Beadling et al., 2019)	169	146	135
Maximum wind stress (Pa)	0.14 (NCEP pre-1972) 0.15 (Verdy and Mazloff, 2017) 0.19±.01 (ERA-Interim Dee et al., 2011)	0.14-0.21 (Beadling et al., 2019)	0.163	0.147	0.154
$\Delta\sigma$ depth averaged 65S-45S (kg/m ³)	0.27	0.15-0.47 (Beadling et al., 2019)	0.26	0.20	0.24
Error in mean bottom temperature <50S (°C)	0.0	-0.94-1.46 Heuzé et al. (2013)	1.01	0.33	1.61
Error in mean bottom σ_2 $z>1000\text{m}$, lat<50S, (kg/m ³)	0.0	-0.17-0.63 Heuzé et al. (2013)	-0.13	-0.09	-0.16
September sea ice extent (Mkm ²)	18.7 (Turner et al., 2012)	6.9-23.6 (Turner et al., 2012)	17.0	12.9	18.9
Nonconvective Sept. σ_θ difference (300m-surf, regions showing deep convection in Weddell Sea between 70S-50S)	~0.12 (Reintges et al., 2017)	0.05-0.3 (Reintges et al. 2017)	0.167	0.115	0.07
Sept. $\Delta\sigma_\theta$ 300m-surf, 70S-50S, 50W-20E	0.246 (WOA09)		0.249	0.182	0.280
September ΔT 300m-surf, 70S-50S, 50W-20E	1.99 (WOA09)		2.48	1.15	2.68

930

Table 2: Mean heat and salt balances in the upper 60 meters of the Western and Eastern Weddell Seas for a 100 year climatology. Q^{surf} refers to surface fluxes, either dilution due to net precipitation minus evaporation for salt or air-sea heat fluxes for temperature. Q^{vmix} is the flux due to the one-dimensional mixed layer entering the box across the bottom face. Q^{adv} is the total flux associated with three-dimensional advection at the grid scale and Q^{eddy} refers to the tendency associated with three-dimensional mixing and advection from meso- and submesoscale eddies.

Region/ term	Western Weddell Sea (70°W-50°W, 70°S-55°S)		Eastern Weddell Sea (20°W-20°E, 70°S-55°S)	
	Salt (kg/m ² /yr)	Temp (W/m ²)	Salt (kg/m ² /yr)	Temp (W/m ²)
Q^{surf}	-15.2	-8.7	-12.8	-10.8
Q^{vmix}	14.1	14.6	12.2	12.0
Q^{adv}	7.5	-12.0	4.8	-5.7
Q^{eddy}	-6.4	6.2	-4.2	4.3



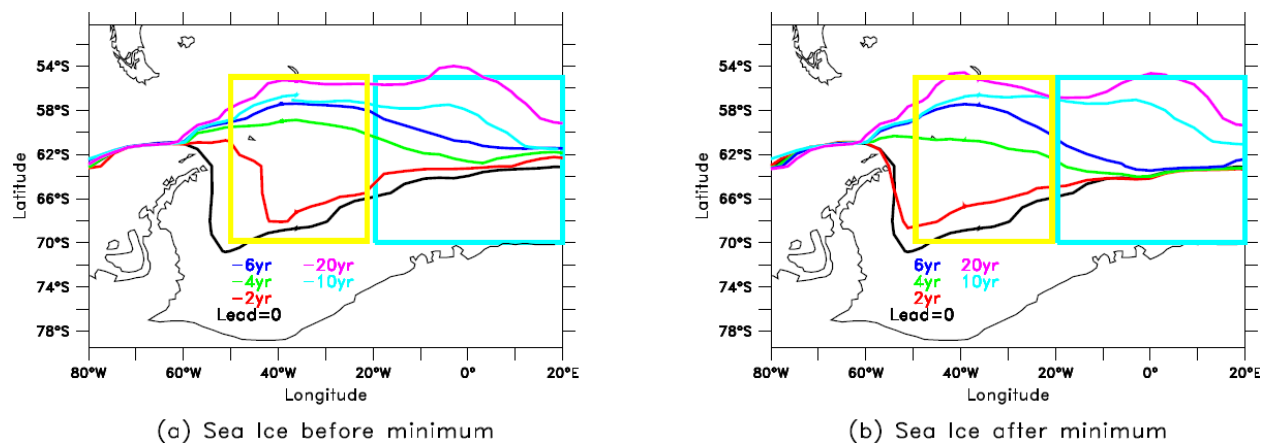
(c) Annually smoothed S anomaly 70S-55S,50W-20W

(d) Annually smoothed S anomaly 70S-55S,50W-20W

Figure 1: Convective variability in the AREDI400 simulation. (a) Root-mean square variability of decadally smoothed SST. Red box shows Western Weddell Sea box (70°S-55°S, 50°W-20°W) (b) Annually smoothed temperature anomaly (color) averaged over the Western Weddell Sea box showing that warming (cooling) at the surface is associated with cooling (warming) at depth. (c) Annually smoothed salinity anomaly Western Weddell Sea box, zero contour from (b) overlaid. (d) Annually smoothed density anomaly over Western Weddell Sea box, zero contour for (b) overlaid.

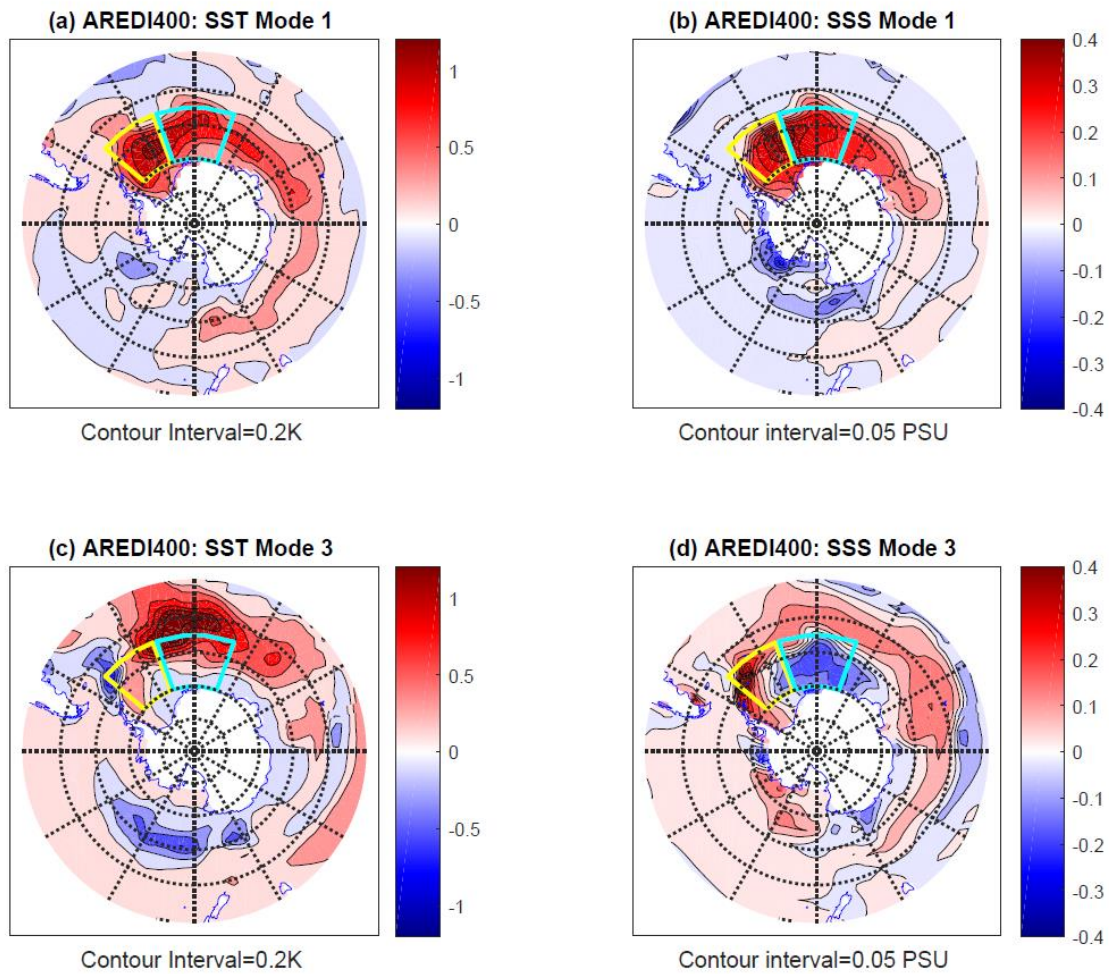
947

948



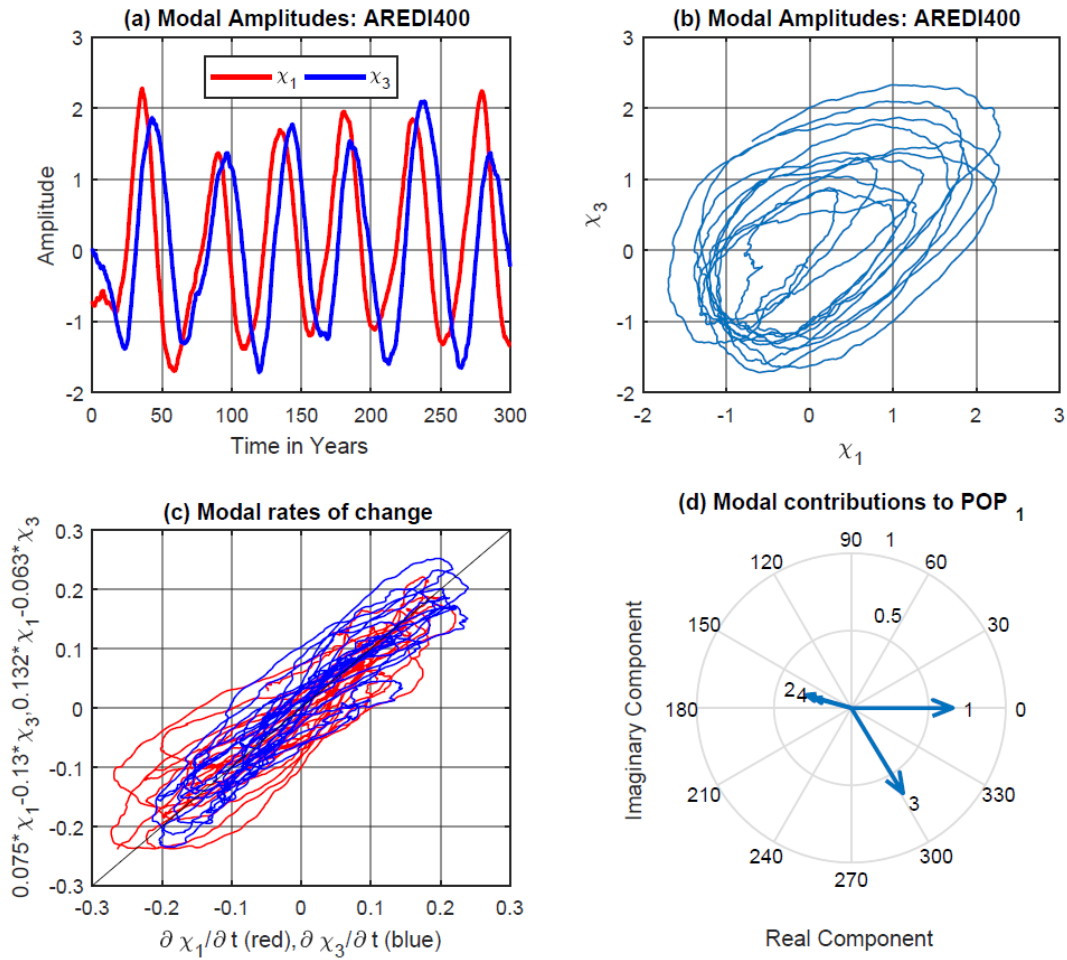
949

950 **Figure 2:** Evolution of September sea ice edge over the convective cycle in AREDI400 model. All lines
 951 show 15% sea ice extent contour. Black lines in all four plots show composite contour (based on 11
 952 events) for all times when the ice concentration in the Western Weddell Box (yellow rectangle)<5%,
 953 corresponding to the minimum in ice extent. Cyan box shows Eastern Weddell Sea region used for
 954 analysis later in this paper. Colored lines are shown in terms of time lagged from this, with negative
 955 numbers referring to the ice edge before the minimum and positive numbers referring to the ice edge after
 956 the minimum. (a) Composites at 0-20 years before the minimum. (b) Composites at 0-20yr after the
 957 minimum.



958

959 **Figure 3:** Structure of rotated EOFs in AREDI400 that dominate periodic convective dynamics. Modes
 960 are normalized so that the time series of the coefficients has a standard deviation of 1. Yellow and cyan
 961 boxes outline regions in the Western and Eastern Weddell Seas that are used for analysis in this paper. (a)
 962 SST Mode 1. (b) SSS Mode 1. (c) SST Mode 3 (d) SSS Mode 3.



964

965 **Figure 4:** Evolution of modal amplitudes, AREDI400. (a) Time series of χ_1 (red) χ_3 (blue). (b) χ_1 plotted

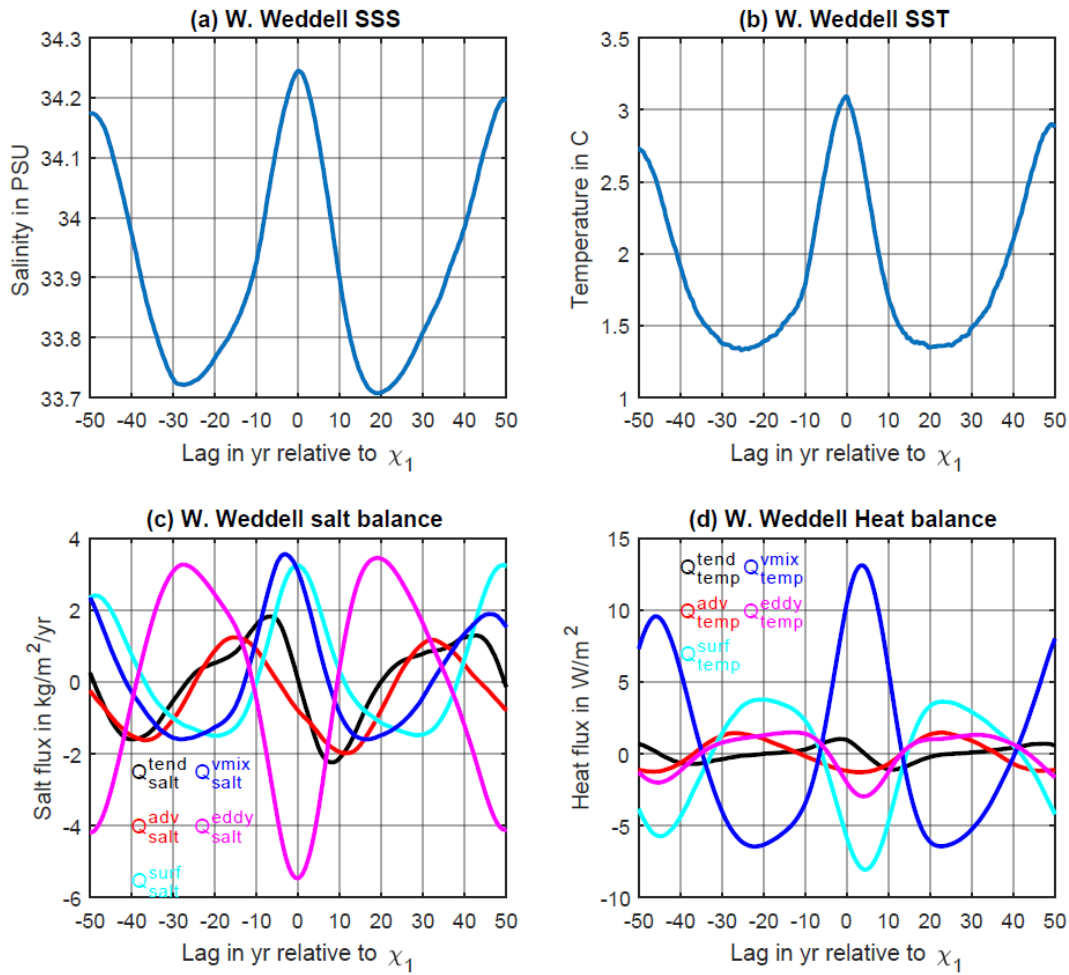
966 against χ_3 showing oscillatory behavior. (c) Evaluation of how equations (6a,b) predict the time series of

967 $\partial \chi_1 / \partial t$ (red) and $\partial \chi_3 / \partial t$ (blue). (d) Structure of POP in terms of the first 4 modes. Length of vector

968 shows contribution from each mode, angle shows relative phase. Italic numbers identify which rotated

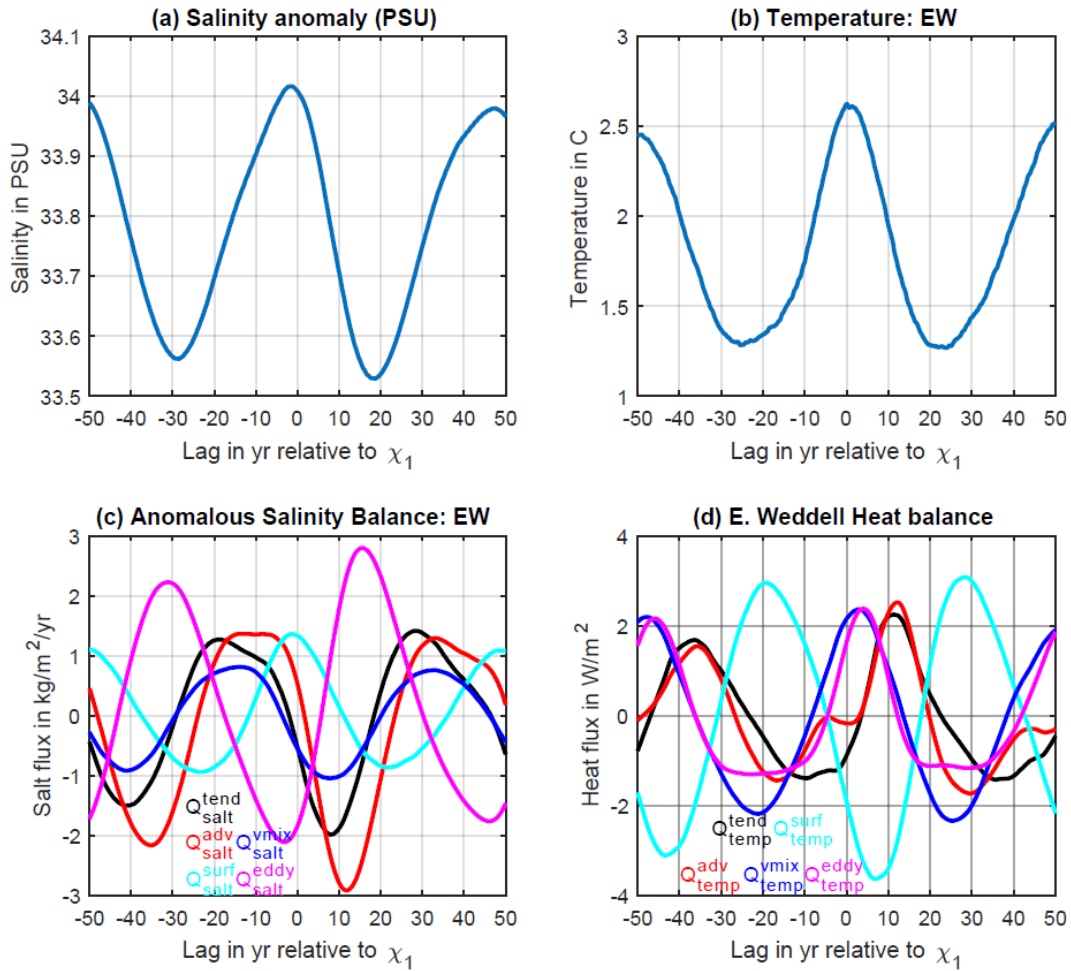
969 EOF corresponds to which vector.

970



971

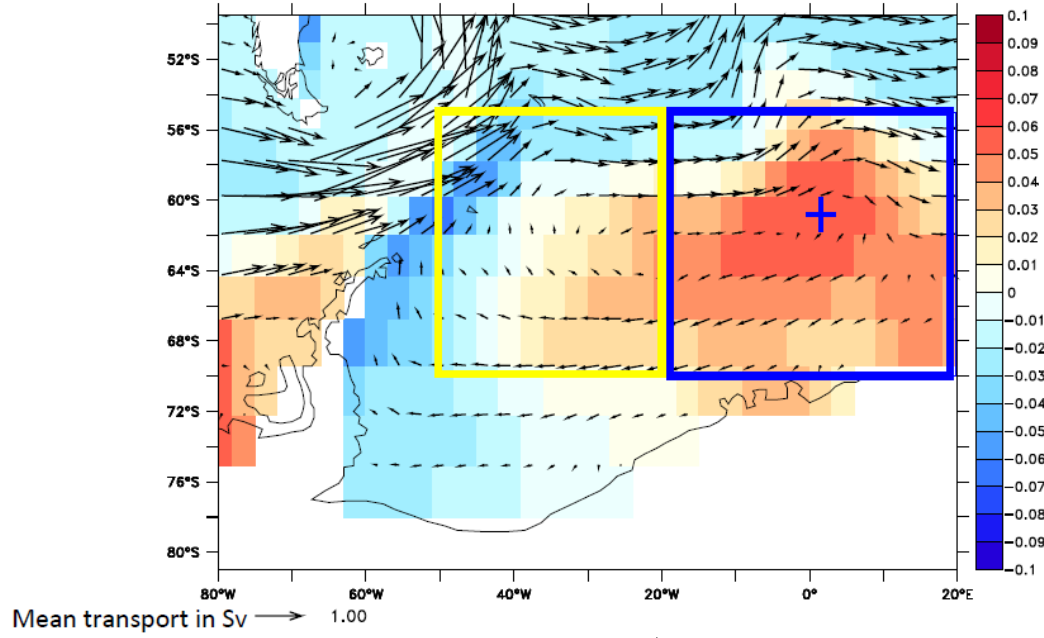
972 **Figure 5:** Composited time series in Western Weddell Sea region (50°W - 20°W , 70°S - 55°S), lagged
 973 relative to peaks in χ_1 , for AREDI400 simulation. (a) SSS in PSU. (b) SST in $^\circ\text{C}$. (c) Salt balance over
 974 top 60m. (d) Heat balance over top 60m. See text surrounding equation 5 for definitions of how the term
 975 balances are calculated.



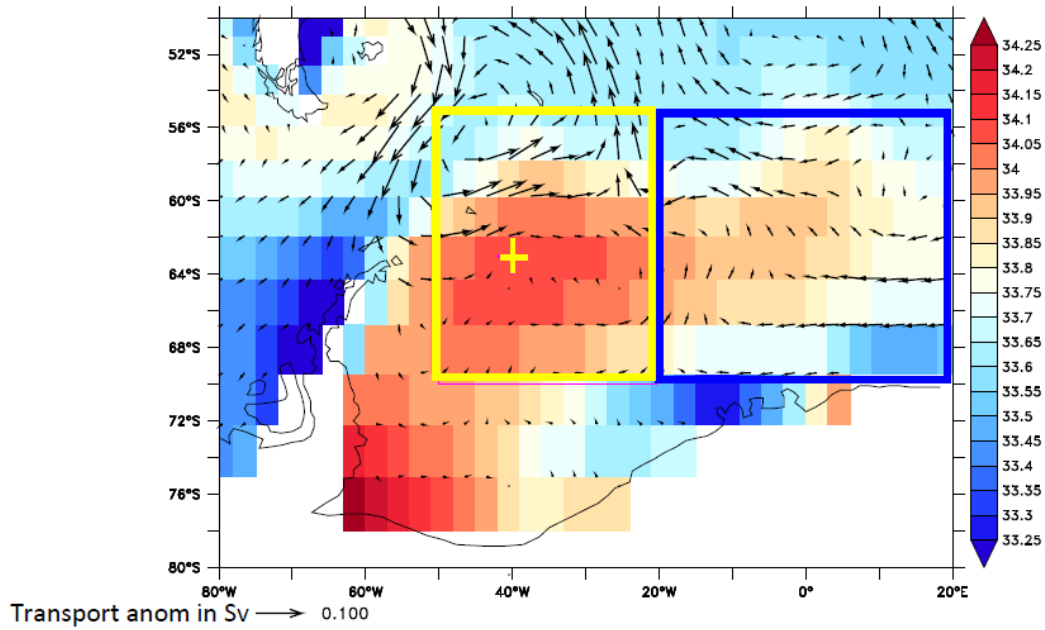
977

978 **Figure 6:** Composited time series in Eastern Weddell Sea region (20°W - 20°E , 70°S - 55°S), lagged
 979 relative to peaks in χ_1 , for AREDI400 simulation. (a) SSS in PSU. (b) SST in $^\circ\text{C}$. (c) Salt balance over
 980 top 60m. (d) Heat balance over top 60m. See text surrounding equation 5 for definitions of how the term
 981 balances are calculated.

982



(a) SSS anom vs. $Q_{\text{salt}}^{\text{adv}}(\text{WW})$ anom



(b) Trans anom vs. $Q_{\text{salt}}^{\text{adv}}(\text{WW})$ anom (mean SSS overlaid)

Figure 7: Explaining the time variation of $Q_{\text{salt}}^{\text{adv}}$ in Western Weddell box. Yellow box shows the WW box, Dark blue box shows the EW box. + marks are individual grid points which are used in Fig. 9 to illustrate the different mixing dynamics in the two regions. (a) Regression between $Q_{\text{salt}}^{\text{adv}}$ in Western Weddell box and SSS (color), mean 0-60m transport (vectors) overlaid. (b) Regression between $Q_{\text{salt}}^{\text{adv}}$ in Western Weddell box and 0-60m transport (vectors), mean SSS (colors) overlaid.

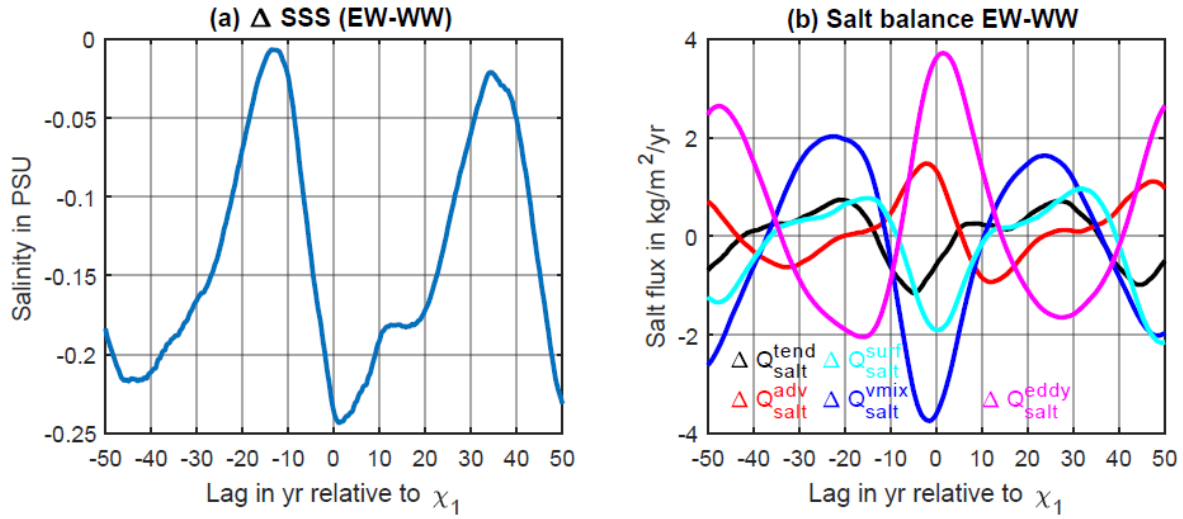
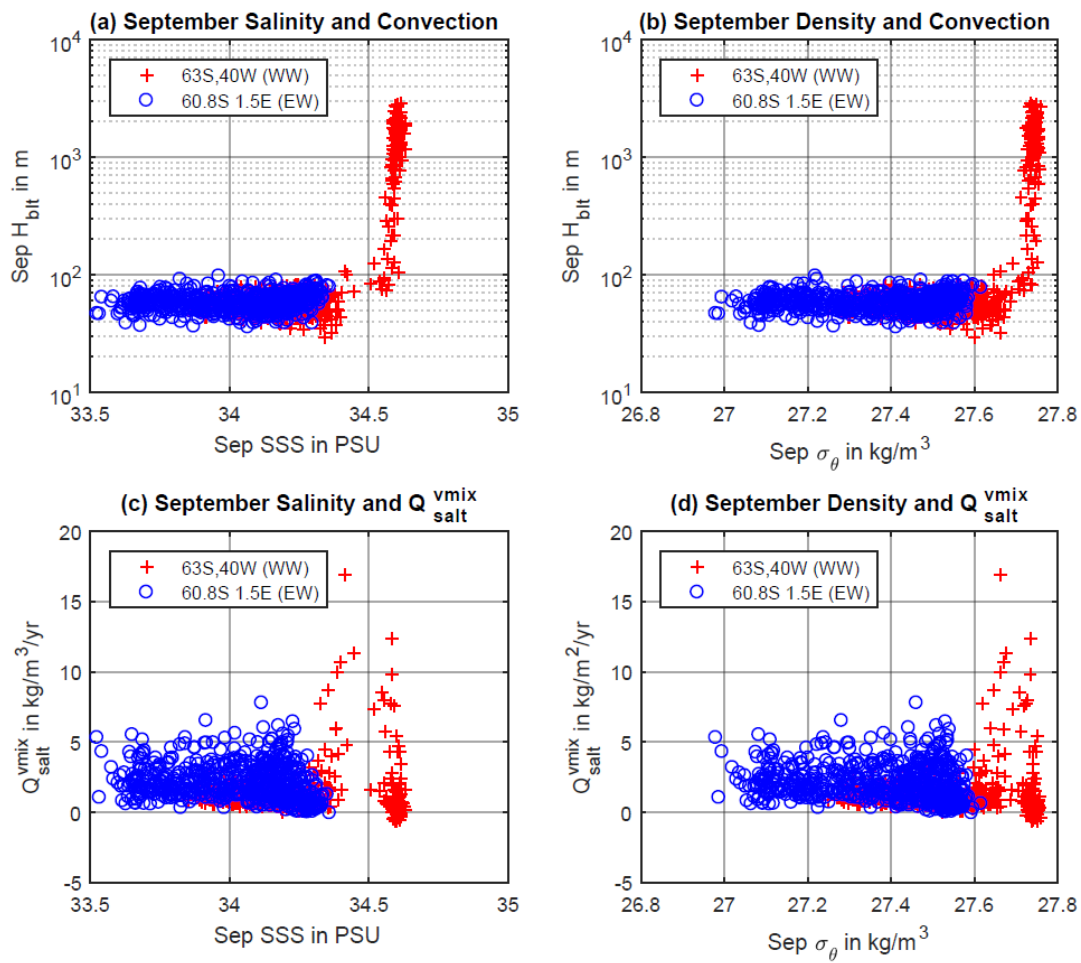


Figure 8: Evolution of difference in SSS between Eastern Weddell and Western Weddell (strongly correlated with EOF mode 3) in AREDI400 simulation. (a) Composed SSS difference lagged relative to peaks in χ_1 . (b) Difference in salinity term balances between the two regions. See text surrounding equation 5 for definitions of how the term balances are calculated.

994



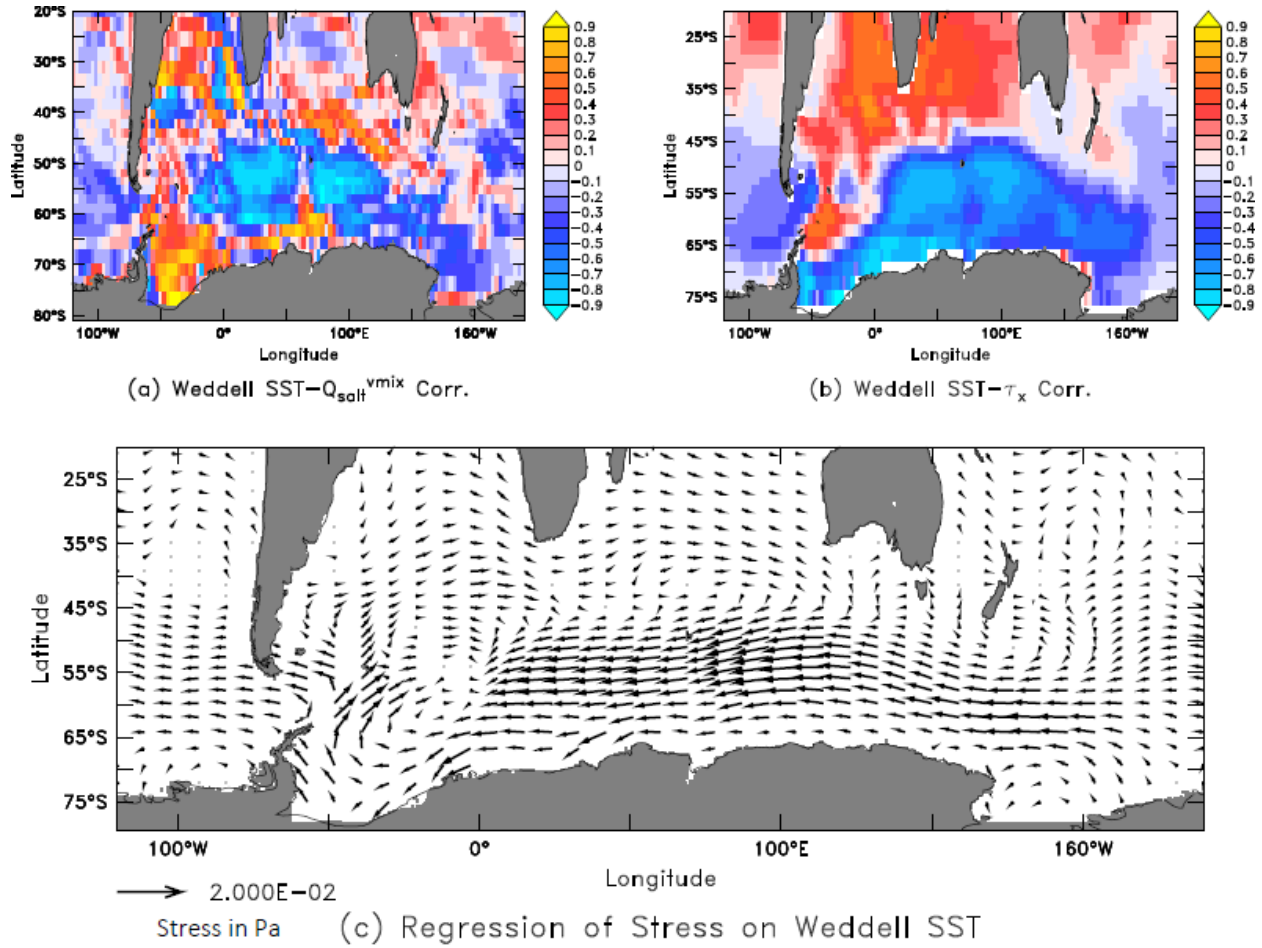
995

996 **Figure 9:** Relationship between surface salinity (left column) and density (right column) and
 997 monthly mean depth of surface boundary layer from KPP code (top row) and Q_{salt}^{vmix} bottom row,
 998 at two points- red is in the Western Weddell and shows deep mixing (yellow + in Fig. 7b), blue
 999 is in the Eastern Weddell (blue + in Fig. 7a).

1000

1001

1002



1003

1004

1005

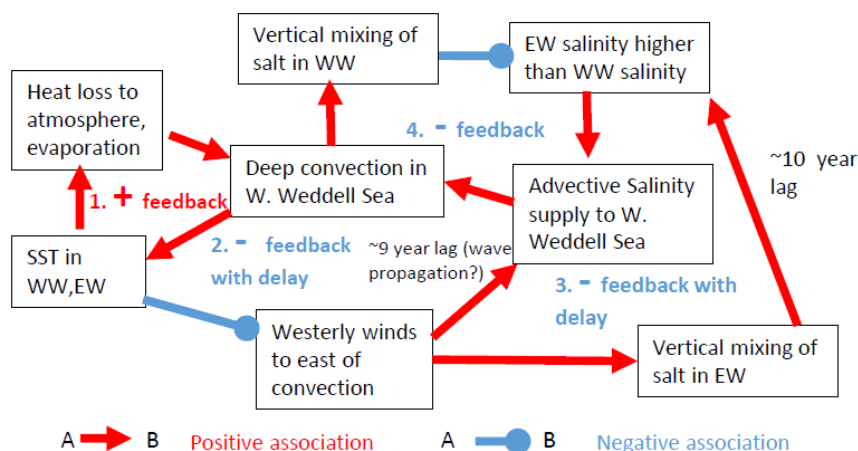
1006

1007

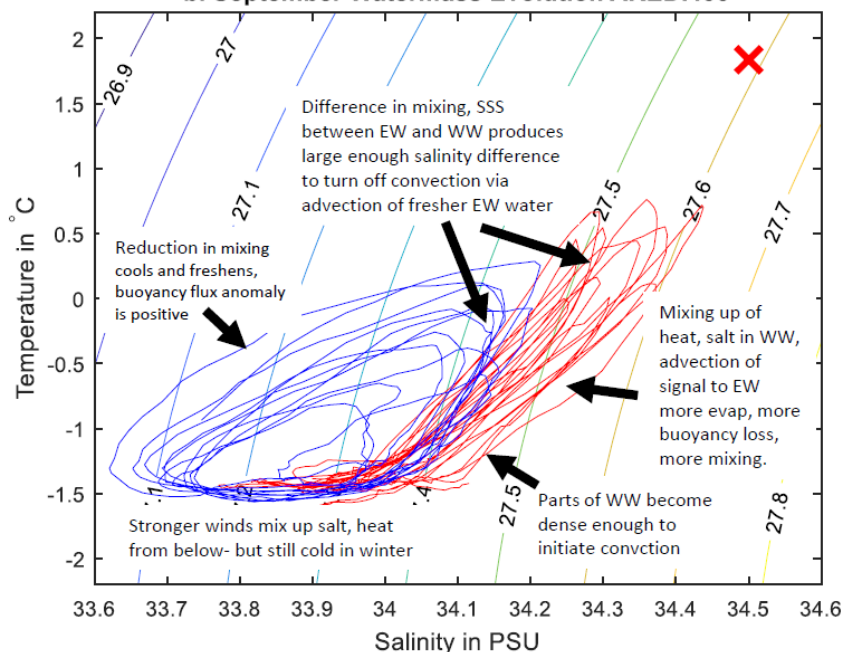
1008

Figure 10: Understanding drivers of Q_{salt}^{vmix} . (a) Correlation between decadally smoothed Q_{salt}^{vmix} and decadally smoothed SST in Western Weddell Sea (50°W-20°W, 70°S-55°S). (b) Correlation between decadally smoothed eastward wind stress and decadally smoothed SST in the Western Weddell Sea. (c) Regression of decadally smoothed eastward stress on decadally smoothed SST in the Western Weddell Sea.

a. Schematic of feedback loops involved in producing oscillatory convection



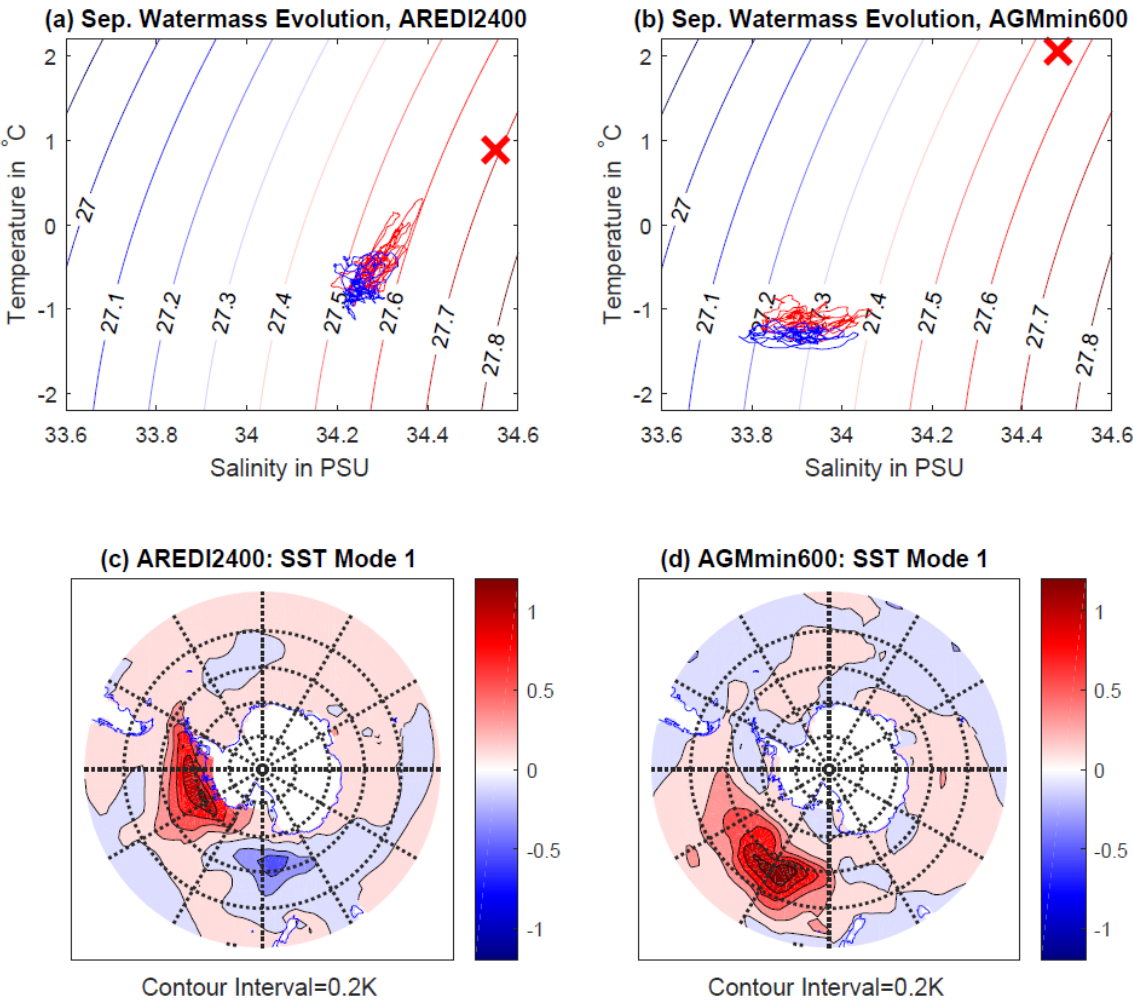
b. September Watermass Evolution AREDI400



1009

1010 **Figure 11:** Schematics illustrating mechanisms responsible for convective oscillations in
 1011 AREDI400. (a) Feedback diagram, red arrows with triangular heads show positive associations
 1012 (increase in A leads to increase in B), blue arrows with circular heads show negative association
 1013 (increase in A leads to decrease in B). Net feedback of each loop shown in colored labels. Note
 1014 that the negative relationship producing the negative feedback in loop 3 (which starts and ends
 1015 with the deep convection in the WW) is that between SST and winds. (b) Evolution in
 1016 watermass space showing only the month of September. Average surface watermasses in
 1017 Eastern (blue) and Western Weddell (red) cycle counterclockwise, coming closest together when
 1018 there is convective initiation. Red X shows average water mass properties in Western Weddell
 1019 box at 300m.

1020



1021

1022

1023

1024

1025

1026

Figure 12: Comparison of results with two other simulations with different representations of eddy mixing. (a) September water mass evolution (compare with Fig. 11b) but for AREDI2400 case. (b) September water mass evolution (compare with Fig. 11b) for AGMmin600 case. (c) SST structure of 1st rotated mode (compare with Fig. 3a) in AREDI2400 case. (d) SST structure of 1st rotated mode (compare with Fig. 3a) in AGMmin600 case.