On the Competing Roles of Heat and Fresh Water Flux in Forcing Thermohaline Oscillations

David W. Pierce,[1] Tim P. Barnett,[1]
and Uwe Mikolajewicz[2]

[1] *Climate Research Division, Scripps Institution of Oceanography, La Jolla, CA*

[2] *Max-Planck-Institut fuer Meteorologie, Hamburg, Germany*

Revised Edition, October 1994

Abstract

We describe the physical mechanisms causing century-scale Southern Ocean thermohaline oscillations in the Hamburg Large Scale Geostrophic model, a primitive equation oceanic general circulation model. The oscillations have been shown to occur on a 320-year time scale when random fluctuations are added to the fresh water flux field which forces the model; we extend this result to show that they occur in a variety of situations, including ones without added noise. The oscillations involve movement between two model states: one characterized by strong convection and an active thermohaline circulation, and the other with a halocline around Antarctica capping off the water column, preventing convection. The physical mechanism which forces the model from the quiescent state to an actively convecting one is subsurface (300 m) heating around Antarctica, which destabilizes the water column; the ultimate source of this heat is advected North Atlantic Deep Water. This leads to a teleconnection between forcing conditions in the North Atlantic and the thermohaline structure of the Southern Ocean. The mechanism which shuts off convection is surface freshening, primarily by precipitation, in the region poleward of the Antarctic Circumpolar Current. The oscillations are analyzed in terms of a simple "flip-flop" model which indicates that non-linearities in sea water's equation of state are necessary for the oscillations to occur. The spatial pattern of convection around Antarctica affects the time evolution of the Southern Ocean's thermohaline overturning, and the way in which different surface forcings cause the model to oscillate.

1. Introduction

Increasing attention is being focused on the role the oceans play in decade to century scale climate variability, as only the oceans seem to have the required thermal inertia on such scales. The atmosphere's smaller thermal capacity limits its memory to months, while large-scale changes in ice sheets can take a thousand years or more. Changes in solar radiation or internal forcings (such as volcanos) are either not yet well-characterized, or are currently thought to be too weak to have more than local effects. This leaves the oceans as the primary agent of variability on decadal time scales.

Stommel (1961) first suggested that the ocean's thermohaline circulation might have several stable states. Bryan (1986) confirmed this using an ocean general circulation model (OGCM), demonstrating the existence of two stable states. Recent work (see Weaver et al. (1993) for a comprehensive review) has increased understanding of how and when transitions between these states might occur; the ocean's sensitivity to surface fresh water flux seems to be particularly important. When transitions between the states are self-sustaining, oscillations in the thermohaline circulation arise. Winton and Sarachik (1993) demonstrated this process in a OGCM with simple geometry; they termed the oscillations "deep-decoupled" because during periods with no convection, the deep waters are "decoupled" from the atmosphere.

To date, the majority of work on thermohaline oscillations has been done with simplified ocean models, without realistic bottom topography or forcing fields. Mikolajewicz and Maier-Reimer (1990) (MMR hereafter) extended such work using a full 3-dimensional OGCM, the Hamburg Large Scale Geostrophic (LSG) model, which has realistic bottom topography and forcing fields. They found that the LSG model reached a steady state when driven with smooth, seasonally-varying forcing fields, but developed large thermohaline oscillations when noise was added to the surface fresh water flux. In subsequent work (Mikolajewicz and Maier-Reimer, 1991) they analyzed the fluctuations in terms of an advected salinity anomaly dipole similar to the loop model of Welander (1986), and observed that fluctuations were not generated when using a linear equation of state. Barnett et al. (1994) showed that a variety of realistically scaled, stochastic fresh water flux forcing models would make the LSG model oscillate in this same mode, particularly if the forcing were confined to the high latitudes of the Southern hemisphere. Stocker and Mysak (1992) found significant spectral peaks with about the same period in paleoclimate data, which suggests that the model results may have some counterpart in observations.

In this work we analyze the physical mechanisms driving the fluctuations found in MMR. We demonstrate a wide range of circumstances under which the fluctuations occur, including many without any noise added to the forcing fields. Our objectives are: 1) to describe and quantify the model processes which cause the fluctuations; 2) to determine to what degree the same physical processes are applicable to the real oceans.

2. Description of the Model

We use the Hamburg LSG OGCM, in a configuration identical to MMR's and that described in Maier-Reimer et al. (1993). Those references show the details of the model construction, and demonstrate that it does a reasonable job of reproducing the large-scale thermohaline structure and circulation. The model includes realistic geometry and topography; it is constructed on an E-grid with 72 longitudinal and 72 latitudinal points, and 11 vertical levels. It is designed to study climatic phenomenon whose spatial scale is much larger than the internal Rossby radius of deformation, and as such, neglects the nonlinear advection of momentum in the Navier-Stokes equations. This and the fully implicit time integration scheme allow a 30 day timestep. The model is hydrostatic, uses a full equation of state, and employs a convective adjustment scheme to eliminate static instabilities.

Seasonally varying wind forcing from Hellerman and Rosenstein (1983) is imposed at the surface. Mixed boundary conditions are used for surface heat and fresh water forcing; the surface temperature is relaxed to COADS air temperature (Woodruff et al., 1987) using a coupling coefficient of 40 w m^{-2} C^{-1} over a surface layer of 50 m, yielding a characteristic time of 60 days. Net precipitation minus evaporation (P-E) is applied as a flux to the surface of the model. Note that the Hamburg LSG model directly applies a surface fresh water flux, rather then using an "implied salt flux" forcing such as used by, for example, the GFDL OGCM (Cox, 1989). This difference should be kept in mind when comparing the forcing scenarios used here to other models.

In the standard case the fresh water flux is diagnosed from a 4000-year model run with surface salinities relaxed to observed values (Levitus, 1982). In various test cases this fresh water flux is modified as described in the following sections; typically this is done by applying a uniform scale factor to the entire field, or, for comparison with MMR, by adding white noise of a given standard deviation to the flux field.

A thermodynamic sea ice model is included, with the heat flux through the ice cover inversely proportional to the ice thickness. The ice is advected by the ocean's surface currents.

For the results shown here, the model was previously spun-up over 3500 years with both temperature and salinity boundary conditions relaxed to observed values, then integrated for an additional 500 years with the mixed boundary conditions. At this point the model is stable, and exhibited no oscillations or drift when integrated to year 10,000. Thus, the state of the model in year 4000 serves as an adequate initial condition for the experiments shown here. Only yearly averaged values are used in the analysis described below.

3. Character of the Model Variability

There are a number of ways to generate thermohaline oscillations in the model. Barnett et al. (1994) show how noise with a variety of spectral characteristics (but zero mean) added to the the surface fresh water flux can cause oscillations. However, the model can oscillate even with no imposed noise; simply reducing the magnitude of the monthly fresh water flux by as little as 7.5% or increasing it 10% is enough to initiate fluctuations in the thermohaline transport. For many of the results shown here, the variability will be excited by uniformly reducing the P-E flux to .8 of the standard value; this test case will be denoted by $\mathcal{F}(.8)$. This is an important case to analyze because it shows how the oscillations arise from internal model processes; added noise is not required. In later sections we will examine a variety of other cases, including one in which white noise with a standard deviation of 2 mm day^{-1} has been added to the P-E field, denoted by $\mathcal{F}(\textit{noise})$; and one where the P-E field has been increased to 1.1 times the standard value, denoted by $\mathcal{F}(1.1)$.

The oscillations are primarily a Southern Ocean phenomena. There is a North Atlantic response, but its amplitude (as measured by changes in the meridional streamfunction) is almost an order of magnitude smaller. Fig. 1 shows the maximum value of the meridional overturning streamfunction as a function of time for the Southern Ocean (top panel) and North Atlantic (bottom panel); the two are plotted with the same scale to facilitate direct comparison. Note that in this and all other plots of the Southern Ocean meridional streamfunction, the absolute value is used. The $\mathcal{F}(.8)$ case is shown; the $\mathcal{F}(noise)$ and $\mathcal{F}(1.1)$ cases are similar. Fluctuations in the Southern Ocean reach 50 Sv peak-to-peak, while the North Atlantic varies by about 5 Sv. Since most of the response is in the Southern Ocean, we will concentrate our analysis in that region.

The analyses presented here of the $\mathcal{F}(.8)$ case will focus on a selected subinterval of the model run: years 5100 to 5510 (marked "A" on Fig. 1). This segment includes one complete cycle of the thermohaline oscillation. We will also make use of two specific points in the model's time series. The Southern Ocean thermohaline circulation reaches a peak at year 5235; we will refer to this time as "THC-ON". At time "THC-OFF", year 5490, the Southern Ocean thermohaline circulation is almost completely shut off. These points are shown in Fig. 1.

The period of the oscillations is approximately 320 years. Although not shown here, this oscillation continues with only minor variations for at least 8000 years (about 25 cycles).

a. Transport through the Drake Passage

We will typically show fluctuations in the thermohaline circulation by using a proxy, the volume transport through the Drake Passage. This is a convenient and accurate index of the mass overturning streamfunction in the Southern Ocean, which is where the model's response is concentrated. The Drake Passage transport has the advantage that it has been the target of field observations which provide a basis for comparison (for a summary, see Nowlin and Klinck (1986)).

Fig. 2 shows the Drake Passage transport for the $\mathcal{F}(.8)$ case. The correspondence between the transport (Fig. 2) and the maximum of the Southern Ocean meridional overturning streamfunction (top panel of Fig. 1) is clear. The correlation coefficient between the two curves is .949. There are theoretical reasons to expect a good correlation between the Drake Passage transport and the strength of convection in the Southern Ocean. The convection is instrumental in determining the density gradient across the Antarctic Circumpolar Current (ACC); through the JEBAR effect (Holland, 1973), this will determine the barotropic Drake Passage transport to a large degree. This effect is illustrated for the ACC by Cox (1975) and Cai (1994).

b. Meridional Circulation

The changes in meridional overturning accompanying this oscillation are striking. Fig. 3 shows the zonally integrated meridional circulation at THC-ON, the peak of Southern Ocean thermohaline activity. There is an intense, deep circulation cell in the southern ocean, centered at about 45°S; the peak transport in this cell is over 45 Sverdrups. There is also a substantial amount (about 15 Sv) of deep water being formed in the northern hemisphere, and a strong cell at the extreme southern edge of the domain, producing about 15 Sv of deep water. The southern cell centered at 65°S is analogous to a Weddell/Ross sea overturning cell; the location is appropriate for sinking along the Antarctic continental margin, and the water production rate is in general agreement with estimates.

In the North Atlantic, there is about 20 Sv of water being renewed to a depth of 3000 meters, values consistent with observations of tracers in the region (Broecker, 1979). The water flows southward, and at about 15°N rises to 2000 meters and overrides the Antarctic Bottom Water (AABW), behavior which is consistent with the observations.

Overturning in the southern Pacific (bottom panel of Fig. 3) is enhanced during the time of maximum thermohaline circulation, as the higher AABW production rate forces stronger northward flow and upwelling.

The situation at THC-OFF, the nadir of the Southern Ocean thermohaline circulation, is shown in Fig. 4. The global picture shows that there is almost no sinking south of 30°S; compared to the 45 Sv of production during the peak of activity, this is quite a change. There is a corresponding decrease in upwelling in the Pacific, while the Atlantic, again, looks basically unchanged. Clearly, the 320 year oscillation is primarily a Southern Ocean phenomena, although the indirect effects of the cessation of the deep flow can be seen in the shifting of the circulation patterns in the Atlantic.

c. Salinity Field

Intimately associated with the thermohaline oscillations are changes in the salinity field. Fig. 5 shows the yearly-averaged surface salinity fields at times THC-ON and THC-OFF, and the difference between them (OFF - ON). A tremendous pool of fresh water accumulates around Antarctica, generally between the Weddell and Ross Seas, as the thermohaline circulation drops from its peak activity to the minimum. Again, there is little evidence of large-scale changes in the Atlantic or Pacific north of about 30°S. The fresh water pool caps off the water column and plays an important role in maintaining the oscillations, as will be described in the following section.

The first EOF of the global surface salinity field, calculated over years 5000 to 6000, is shown in the top panel of Fig. 6. This mode accounts for 70.2% of the total variance. The corresponding eigenamplitude is shown in the bottom panel. The sense of these fields is such that at times of enhanced Southern Ocean thermohaline circulation, the eigenamplitude is positive.

The variance is concentrated in the Ross sea, and corresponds to the fresh water cap which can be seen in Fig. 5. Comparison of the eigenamplitude as a function of time with the transport through the Drake Passage (Fig. 2) shows that this first mode is associated with the periodic cessation of the thermohaline circulation, but *not* the gradual decline in thermohaline circulation which occurs after the peak in thermohaline activity is reached.

The second mode, which accounts for another 17.7% of the variability, is shown in Fig. 7. Unlike the first EOF, this one shows a strong signal in the Weddell and Enderby Seas. The eigenamplitude shows that the second mode generally represents the gradual diminution of the thermohaline circulation after the peak transport is reached.

Comparing the eigenamplitudes of the two EOFs, it can be seen that the troughs and peaks in EOF 2 lead those in EOF1 by 65 years. A rough estimate of the timescale for advection at 60°S halfway around Antarctica (from the center of EOF 2 in the Enderby Sea to the center of EOF 1 in the Ross Sea) by an average current of .5 cm sec^{-1} yields 63 years. The close agreement is somewhat coincidental given the difficulty of choosing a characteristic velocity, but nevertheless hints that an advective process may help carry information from the initial convection sites around the Antarctic continent.

d. Sea Ice

There are corresponding changes in the distribution and thickness of sea ice, particularly in the Ross and Weddell Seas. Fig. 8 shows sea ice thickness in the Southern Ocean at years THC-ON and THC-OFF. The distributions in the northern hemisphere are, by comparison, quite steady over time. At the peak of thermohaline activity, there is about a third to a half meter of ice covering the Ross Sea, and half a meter or less covering the eastern part of the Weddell. At the minimum in thermohaline overturning—year 5490—the area covered by ice has increased somewhat, and the thickness increased greatly. The Ross Sea is now covered by 2 to 10 meters of ice, and the Weddell Sea by 2 to 15 meters.

4. Origins of the Model Variability

The physical causes of the thermohaline variability in the LSG model are intimately associated with the movement between two extreme states. The time THC-ON, the Southern Ocean is overturning rapidly, and has small vertical salinity and temperature gradients and a minimum of sea ice cover. At time THC-OFF, there is weak or non-existent Southern Ocean thermohaline circulation, an extensive fresh water surface pool capping off convection in the Ross and Weddell Seas, and heavy sea ice cover. What needs to be understood is why the LSG model produces *transitions* between these states once they are established. By contrast, Bryan (1986) found no such transitions using the Princeton GFDL OGCM; here, we shall see that the competing effects of a stabilizing fresh water flux and destabilizing heat flux are necessary (but not sufficient) to produce oscillations.

To initiate convection in a water column capped with a layer of fresh water, there must be a mechanism which reduces its static stability. A trend of decreasing precipitation is not possible, since the P-E flux does not change from year to year in the $\mathcal{F}(.8)$ case. Greater surface cooling is unlikely, since (as shown above) there is a thick, insulating sea ice cover when the column is most stable. This suggests examining the components of the density balance in the *interior* of the water column. Fig. 9 shows time versus depth contours of the temperature and salinity in the Weddell Sea for test case $\mathcal{F}(.8)$, as well as the Convective Adjustment Index (CAI), defined as the fraction of time which a model gridbox is convected into from above.

Moving from the time of greatest Southern Ocean thermohaline circulation (indicated by "THC-ON") to the minimum of activity ("THC-OFF"), there are distinct changes in the temperature and salinity structure, and the convective activity, in the Weddell Sea. The biggest changes in the temperature field (top panel of Fig. 9) appear at middepth, between 150 and 700 meters. There is strong net heating throughout these depths *during the time of decreased thermohaline circulation only.* Comparing the temperature field to the CAI (bottom panel of Fig. 9), we can see that the temperature does not start increasing until the CAI drops below about .4; once it does, the subsurface waters start to warm. The temperature increase is gradual at first, .5 C in the first 80 years; in the last 50 years, however, the temperature increases 1.5 C. When convection is more active the subsurface heat is vented to the atmosphere, and does not accumulate locally.

The evolution of salinity is complementary to that of the temperature, in the sense that the stabilizing fresh water cap gets thicker and fresher as convection decreases. The minimum surface salinity, and hence the strongest stabilizing effect, occurs at time THC-OFF.

The subsurface waters eventually become warm and light enough to overcome the fresh water cap and overturn the water column. Note how the convective activity rapidly increases once the intermediate waters reach 3.5 C; the subsequent deep convection vents the heat to the atmosphere, reducing the intermediate temperatures and distributing the accumulated fresh water throughout the water column. It is the destabilizing subsurface heating which is the cause of the model's transition from the "off" to the "on" state. A similar effect was found in a OGCM with simplified topography (Winton and Sarachik, 1993).

The effect which temperature and salinity have on column stability, N^2, is shown in Fig. 10. The water column reaches its maximum stability just slightly before THC-OFF, the time of least thermohaline circulation. As can be seen from the bottom panel, this maximum is produced by simultaneous peak in the salinity contribution, at a depth of 70 m. Shortly thereafter, the *destabilizing* contribution of the temperature field reaches a maximum at a depth of 150 m, stimulating convection from below, rather than from above. After the convection has proceeded for a while, venting the heat accumulated below the surface, the depth at which the temperature is most destabilizing moves to the surface. The convection once again becomes the more familiar case of convection from above.

The presence of subsurface convection can be verified by examining the depth at which convection first occurs in the model after year THC-OFF. This is shown for the Southern Ocean in Fig. 11. The great majority of the convecting domain first experiences convection at the surface; but there is an extensive region along the Antarctic coast around 10°E where the first convection occurs at a depth > 700 meters.

It is clear that *deep heating in the Southern Ocean* is the mechanism which initiates the model's convective overturning. The ultimate source of this heat is NADW, which has carried heat by advection and diffusion into the circumpolar waters.

a. Antarctic Zonation

The results shown in the previous section are for the Weddell Sea region of the Southern Ocean. An important feature of the oscillations can be seen by examining the behavior of other Antarctic sectors. For the purposes of this analysis, the Southern Ocean will be divided into five zones: Enderby (0 to 97.5°E), Wilkes (97.5 to 172.5°E, Ross (172.5°E to 137.5°W), Amundsen (137.5 to 72.5°W), and Weddell (72.5°W to 0). These are shown in Fig. 12. The northern boundary of each zone is taken to be 60°S.

The average water column temperature in each of these sectors as a function of time is shown in Fig. 13

(the time period covered is the same as in Fig 9). The bottom panel shows the maximum of the Southern Ocean meridional overturning streamfunction, replotted from Fig. 2. It is evident that the sectors can be grouped. The Weddell and Enderby sectors generally have temperatures less than 1.6 C, and have sharp peaks near year THC-OFF. The Amundsen and Ross sectors always have temperatures greater than 1.6 C, have no peak near THC-OFF, but show sudden drops just before time THC-ON. The Wilkes sector is intermediate between these two cases, with a broad peak near THC-OFF and a decline before THC-ON.

Comparing the two panels of Fig. 13, it can be seen that the distinctive shape of the curve for maximum overturning streamfunction as a function of time is coming from different regions convecting at different times. Convection vents subsurface heat to the atmosphere, resulting in a sharp temperature drop coincident with the increase in convective activity. In particular, the increase in transport shortly after the overall minimum near year THC-OFF is associated with increases in convection in the Enderby and Weddell sectors. This corresponds to the structure seen in EOF mode 2 (Fig. 7). The Enderby sector starts convecting 15 years before the Weddell; this is in agreement with Fig. 11, which shows that the initial convection sites lie in the Enderby sector.

The peak in meridional streamfunction at time THC-ON, on the other hand, arises due to sudden increases in convection in the Amundsen, Ross, and Wilkes sectors. This corresponds to the structure seen in EOF mode 1 (Fig. 6). Activity in the Amundsen sector appears to precede that in the Ross sector by about 10 years, but the peaks are too broad to determine the time lapse precisely.

The spread of convection in the Southern Ocean from its initial locations can be seen in Fig. 14, which shows the CAI at a depth of 250 m as a function of longitude and time. The values have been meridionally averaged from 75°S to 60°S. Starting at the bottom of the Figure, we see that convection first starts on the border between the Enderby and Wilkes sectors (marked 'A' on Fig. 14). Once convection starts at point A, it spreads outward, filling the Enderby and Wilkes sectors. Fifteen years later, convection starts at point 'B' in the Weddell sea. Convection spreads eastwards from B, and joins with the advancing front from A to initiate convection in the entire Weddell sector. There is a delay of about 50 years before the advancing front manages to pass westward through the Drake Passage, after which it initiates convection in the Amundsen sector, and finally reaches the Ross sector. It has been shown that the mechanism for this propagating "seed" of convection is the advection of salty upper-layer waters; once subsurface heating initiates convection, salty water is mixed into the surface layers, is advected over adjoining marginally-stable water columns, and initiates convection in them (T. Osborn, *pers. comm.*).

The subsurface heat released by convection also forces changes in the sea ice, as can be seen in Fig. 15. Sea ice in the Weddell melts back precipitously after time THC-OFF, as convective activity in the Weddell sector picks up. Sea ice in the Ross sector, on the other hand, does not show any large changes around THC-OFF; rather, it decreases suddenly just before time THC-ON, as the convective activity in that sector increases. These results suggest that the model's sea ice cover responds to changes in the ocean's thermohaline circulation, but does not drive them.

b. Atlantic Teleconnnection

It is shown above that subsurface heating in the Weddell and Enderby sectors is the key process in maintaining the large fluctuations in thermohaline circulation. This heat ultimately is supplied by NADW, which flows south from its formation regions in a deep western boundary current, mixes with the circumpolar waters in the ACC, and spreads out at shallow depths around the Antarctic continent. It has been conjectured that NADW production halted, or at least was seriously reduced in magnitude, during the last glacial maximum (Broecker et al., 1985). Such a cessation might prevent the model from undergoing the large scale thermohaline oscillations it normally exhibits, as there would be no source of subsurface heat.

An experiment to test this hypothesis was carried out by reducing, in the North Atlantic only, the coupling coefficient between the observed temperature and the ocean surface temperature. This prevented the cooling and consequent sinking which forms NADW. Normally, the coupling coefficient has a constant value of γ_o = 40 w m^{-2} K^{-1}, but in this test case it dropped quadratically from γ_o at 32.5°N to .17 w m^{-2} K^{-1} at 65°N. The change was implemented at model year 4600.

Fig. 16 shows the transport through the Drake Passage for this test case. The transport is the same as the $\mathcal{F}(.8)$ case until year 4650; since the change was implemented at year 4600, this is a lag of about 50 years, shorter than the advective timescale through the deep western boundary current (about 300 years). Rather, this is the spin-down time of the North Atlantic's thermohaline circulation.

After the cessation of NADW, the thermohaline circulation continues to oscillate, but in a distinctly different way. The time scale is now about 600 years; the amplitude of the oscillation decreases from 90 to

40 Sv; and there is superposed on the 600 year oscillation a smaller oscillation with a period of 50 years. The heat advected by the NADW to the high latitudes of the southern hemisphere clearly is a key element determining the structure of the thermohaline oscillations.

A final point worth noting is that both the temperature and formation rate of NADW set an upper limit on how much heat NADW can supply to the southern ocean. Either a reduced formation rate, as already discussed, or a shift in NADW temperature and salinity towards colder, fresher conditions (while maintaining the same density) will reduce the destabilizing subsurface heat flux in the southern ocean.

5. Other Forcing Scenarios

The results presented in Section 4 focussed on experiments based on the $\mathcal{F}(.8)$ case. The work in MMR and Barnett et al. (1994), however, demonstrated the existence of thermohaline fluctuations using a different forcing scheme: the addition of noise to the fresh water flux field. In this section we will show that the mechanism described above plays a crucial role in generating oscillations in the noise-forced case, as well as in a case where the fresh water flux field is increased 10%.

a. The Noise-Forced Case

MMR first demonstrated that the LSG model develops fluctuations in the strength of the Drake Passage transport when noise is added to the fresh water flux field. Barnett et al. (1994) extended this result to show the effect of the standard deviation, spectral characteristics, and geographical application of this noise on the character of the variability.

The results in this section are from a case where normally distributed random numbers with zero mean and a standard deviation of 2 mm/day were added to the fresh water forcing field at each ocean gridpoint. A value of 2 mm/day is roughly equal to the amplitude of the seasonal cycle found by Roads et al. (1992) in an analysis of National Meteorological Center data; Barnett et al. (1994) examine the effect of the amplitude of this noise on the model's behavior. There was no spatial or temporal correlation to the values added to the flux field; thus, this case is identical to the "white-white" case of Barnett et al. (1994). This case will be referred to as $\mathcal{F}(noise)$.

Fig. 17 shows the transport through the Drake Passage for the $\mathcal{F}(noise)$ case. The transport undergoes substantial variations, with a peak to peak amplitude of 90 Sv. For this analysis we will examine, in detail, the model's behavior from year 5350 to 5665. This is the segment marked "A" in Fig. 17. The maximum of the meridional overturning stream function in the Southern Ocean during this time interval is shown in panel D of Fig. 18; the correlation with the Drake Passage transport is again quite good, .887.

The differences in activity by Antarctic zone, which are described for the $\mathcal{F}(.8)$ case in Section , are equally important for the $\mathcal{F}(noise)$ case. This is demonstrated in Fig. 18. Panel A shows the average CAI for the Weddell and Enderby sectors, normalized by the number of active gridpoints in the region. Panel B shows the same data for the Amundsen, Ross, and Wilkes sectors. Within each of these groupings by sector, differences in CAI were minor.

The timing of convection in these two groups is quite different. The Weddell and Enderby have a sharp peak near year 5375, then a slowly decaying tail. The Amundsen, Ross, and Wilkes, by contrast, are relatively quiet during most of this time period, and only show signs of strong convection around year 5530. The Weddell and Enderby sectors show little activity during that year. The CAI of all five sectors summed together (panel C) is quite similar to the maximum of the meridional overturning streamfunction in the Southern Ocean (panel D). In particular, note that the sharp peak in meridional overturning near year 5530 is driven by convection in the Amundsen, Ross, and Wilkes sectors, while the initial turn-on and gradual decay are due to activity in the Weddell and Enderby sectors. Thus we see that regional details in convective activity are reflected in the meridional overturning streamfunction and, therefore, the Drake Passage transport. This suggests that long-term measurements of flow through the Drake Passage might be a convenient way of inferring such local changes, which would otherwise be difficult to measure directly.

Fig. 19 shows the CAI at a depth of 250 m as a function of longitude and time. As in the $\mathcal{F}(.8)$ case, convection starts first in the Enderby sector (although a bit farther west than in the $\mathcal{F}(.8)$ case), then, about 15 years later, begins in the Weddell. Convection spreads outward from both locations; after a delay of 70 years, the front passes eastwards through the Drake Passage and triggers convection in the Amundsen sector. Comparing this to the $\mathcal{F}(.8)$ case (Fig. 14), we can see that the major difference between the two is that the Ross sector convects sooner in the $\mathcal{F}(.8)$ case. This can be understood by noting that the $\mathcal{F}(.8)$ case has less net precipitation in the Ross sector, implying less stability. This difference can be seen quite directly in the Drake Passage transport versus time curves; the sharp peak due to convection in the Ross sector comes distinctly later in the cycle for the $\mathcal{F}(noise)$ case (Fig. 17) than for the $\mathcal{F}(.8)$ case (Fig. 2).

Fig. 20 shows the average temperature, salinity, and Frequency of Convection (CAI) as a function of time and depth in the Weddell sector. The same cycle as demonstrated above for the $\mathcal{F}(.8)$ case can be seen. The initial strong, fresh cap on the water column prevents convection; heat accumulates at depths of 150 to 400 meters, destabilizing the water column, until overturning abruptly starts at year 5360. Subsurface heat then vents rapidly to the atmosphere, reducing the destabilizing density contribution from temperature and slowing the rate of convection. This allows fresh water to accumulate on the surface once more, and the cycle repeats.

The evolution of the water column in the Ross sector is similar, but with different timing (Fig. 21). A subsurface temperature of 6 C is not sufficient to initiate catastrophic overturning given the strength of the surface fresh water cap; convection does not begin in earnest until year 5500, when it can be seen from Fig. 19 that the propagating signal enters the Ross sector from the east. Then there is a very quick, short-lived overturning, resulting in the sharp peak in Drake Passage transport at year 5530 which can be seen in Fig. 17.

b. The $\mathcal{F}$(1.1) Case

The model also develops fluctuations in the strength of the Southern Ocean thermohaline circulation when driven by a P-E field which has been multiplied by 1.1 at every point. In this case, however, they are more irregular in nature, occurring in bursts of activity separated by relatively quiescent intervals. Fig. 22 shows the Drake Passage transport for this case. An examination of the cycle which occurs around year

4950 shows the same physics as described previously for the $\mathcal{F}(.8)$ and $\mathcal{F}(noise)$ cases, but with a much stronger fresh water cap in the Southern Ocean.

Multiplying the P-E field by 1.5 or larger values did not lead to a repeating series of fluctuations; rather, the thermohaline circulation completely collapsed, in a halocline catastrophe similar to that reported by (Bryan, 1986).

c. The Non-Oscillating Case

Considering the degree to which the model is prone to develop thermohaline oscillations, as shown in the previous sections, it is worth addressing the question of why the model does *not* oscillate in the "standard" configuration as described in, for example, the control case of Mikolajewicz and Maier-Reimer (1990).

It can be seen in Figs. 14 and 19 that the oscillation cycle consistently starts with convection in the Enderby sector. In the standard configuration, there is convection here every winter; this constant venting prevents the catastrophic buildup of heat which is characteristic of the thermohaline oscillation cycle. At the same time, it mixes salty mid-depth water to the surface, encouraging convection in adjoining sectors via the mechanism described in Section 4a. In the oscillating cases, by contrast, convection in the Enderby periodically shuts off, leading to the subsurface warming and overturning.

What, then, encourages steady convection in the Enderby sector in the standard configuration? Detailed examination of the model's forcing fields shows strong local contrasts in the P-E field in this region. For example, a gridpoint with an applied P-E flux of 5.6×10^{-8} m s^{-1} might be next to a gridpoint with -1.1×10^{-8} m s^{-1}. It is reasonable to suppose that such strong local variations might encourage convection at points which have a small or negative P-E, while still maintaining a realistic value of P-E averaged over the region.

To test this conjecture, we ran a case where the standard forcing of monthly P-E values was modified only in the eastern half of the Enderby sector, between 60°E and 100°E, and south of 60°S. All P-E values in this region were replaced with the monthly area-weighted average value of the original field in the region. The result is a P-E field which has the same net P-E input to the region and has the same monthly variation as before, but no spatial variation in the affected region.

The results of this case are shown in Fig. 23. The model develops thermohaline oscillations of the kind previously seen in the $\mathcal{F}(.8)$, $\mathcal{F}(\textit{noise})$, and $\mathcal{F}(1.1)$ cases.

It is vital to note that the surface fresh water flux field, which ultimately served to suppress the thermohaline oscillations in the standard case, was in no way selected by the experimenters. Rather, the field was diagnosed from the fluxes evolved by the LSG model during the spin-up period, when it was relaxed to observed climatological conditions. In essence, relaxing the model to a steady state stimulated the model to evolve a fresh water flux field which *could support* a steady state.

There is no reason to think this is a particular characteristic of the LSG model. Diagnosing the fresh water forcing field from a run relaxed to steady conditions might generally underestimate the true variability. Certainly, in this case, this widely-used technique had the result of explicitly suppressing the model's thermohaline oscillations. On the other hand, until accurate observations of (P-E) are available for the Southern Ocean, the possibility that the fresh water flux which drives the model to a steady state is the more realistic case must be kept in mind.

Another process which helps keep the standard configuration free of thermohaline oscillations is the stabilizing effect of the seasonal sea ice cycle. Growing sea ice rejects brine, which is exported from the surface layer; but melting sea ice releases fresh water, which (being less dense than seawater) floats on the surface. This process stabilizes the water column, with greater stabilizations achieved for greater changes in sea ice thickness. The LSG model exports the brine rejected from sea ice into the top *two* gridboxes (which have thicknesses 50 and 62.5 m, respectively). Therefore vertical mixing, a somewhat slow process, is not necessary to achieve the unmixing effect in the model.

The effect which this unmixing has on the stability of the thermohaline circulation can easily be seen by removing brine rejection from the thermodynamic sea ice model. This was accomplished by setting the salinity of newly grown sea ice to 35 ppt; thus, no brine was rejected while the ice was forming, and no fresh water was released upon melting. Without this stabilizing factor, the model develops the same type of thermohaline oscillations; the Drake Passage transport varies by about 90 Sv, with a time period of about 200 years.

It should be kept in mind that the stability imparted by the sea ice cycle is not an anomalous effect of the model; the brine rejection effect occurs in nature as well.

6. A Simple Flip-Flop Model

The physics of the thermohaline oscillation, as described above, can be further elucidated by analyzing a simple "flip-flop" model in the spirit of Welander (1982), and the extension by Winton (1993). Unlike those models, the one presented here keeps a total budget of both heat and salt, and allows the temperature and salinity of both layers to vary.

A schematic of the model is shown in Fig. 24. The temperature of the surface box is relaxed to the atmospheric temperature; in a situation which is not experiencing a secular drift in climate, this lost heat must be reintroduced into the system. This is accomplished by an interior heat flux Q_i, which is the time-average of the heat lost through the surface. Similarly, the fresh water gained at the surface through precipitation and brine rejection, F_w, must be removed from the deep box at the same rate. This formulation emphasizes the important point that a surface heat loss demands subsurface heating, and net precipitation requires subsurface export of fresh water. It also allows both the box temperatures to vary, so convection can be initiated either by the traditional cooling from above, or by heating from below; the choice is determined by the system itself, rather than by the model formulation.

The model equations are

$$h_1\dot{T}_1 = K_T\,(T_a - T_1) + K\,(T_2 - T_1) \quad (1)$$

$$h_1\dot{S}_1 = -F_w + K(S_2 - S_1) \quad (2)$$

$$h_2\dot{T}_2 = Q_i + K\,(T_1 - T_2) \quad (3)$$

$$h_2\dot{S}_2 \;=\; F_w + K(S_1 - S_2) \tag{4}$$

where h_1 and h_2 are the thicknesses of the upper and lower layers, respectively; K_T/h_1 is the relaxation coefficient between the surface box and the atmosphere; K is the mixing coefficient between the two boxes; Q_i is $(\rho_o c_p)$ times the interior heat flux; and $F_w \equiv \overline{S}F_{P-E}$, where $\overline{S}$ is a characteristic surface salinity and F_{P-E} is the applied net freshwater flux. The interior mixing coefficient K is taken to be large when the system is statically unstable, small otherwise, and is always positive.

As in Winton (1993), a non-linear equation of state is used:

$$\rho(T,S) = \rho_o - \alpha(T - T_o) + \beta(S - S_o) + \gamma(T - T_o)^2 \tag{5}$$

where the reference state will be taken as (T_o, S_o) = (0°C, 35 ppt). This choice will be discussed further below.

In steady state, the density difference between the layers is given by

$$\Delta\rho \equiv \rho_1 - \rho_2 = K^{-1}\left\{\alpha Q_i - \beta F_w - 2Q_i\gamma\left(T_a + \frac{Q_i}{K_T} + \frac{Q_i}{2K}\right)\right\} \tag{6}$$

A negative value means the system is statically stable; a positive value indicates static instability and consequently the initiation of convection. The flip-flop model will spontaneously oscillate between convecting and non-convecting states if the steady state is statically stable *when the model is convecting*, and is statically unstable *when the model is not convecting*. The requirement for oscillations then becomes:

$$\Delta\rho \begin{cases} > 0 & \text{when K small} \\ < 0 & \text{when K large} \end{cases} \tag{7}$$

Examining Eq. 6, we see that the only term which can change the sign of $\Delta\rho$ as K varies is $-Q_i^2\gamma/K$. For oscillations to be present, this term must *increase* as K becomes small; this is only possible if $\gamma < 0$. It is therefore not possible for a simple system such as modeled here, with cold, fresh water on the surface, to oscillate unless a non-linear equation of state is used and $\gamma < 0$. The requirement of a non-linear equation of state was also found to be true in the three-box model of Winton (1993). By contrast, the oscillations found in Welander (1982), where a linear equation of state was used, can only exist when warm, salty water convects downwards from the surface. The case of interest here is with cold, fresh surface water, and so that result is not directly applicable. For seawater at 0°C and 35 ppt, the International Equation of State of Sea Water, 1980 (UNESCO, 1981) can be expanded to show $\gamma = -6.8 \times 10^{-3}$ kg m^{-3} °C^{-2}, and oscillations such as described here are possible.

The physics which *require* a non-linear equation of state for oscillations to be possible with cold, fresh surface waters are straightforward. The stabilizing buoyancy flux due to the fresh water input, $(\partial\rho/\partial S)F_w$, competes with the destabilizing buoyancy flux due to heat, $(\partial\rho/\partial T)Q_i$, to determine the overall stability of the water column. Oscillations in a "flip-flop" model occur when alternately one, then the other, of these terms dominates. Given fixed values for F_w and Q_i, a linear equation of state means that $(\partial\rho/\partial S)F_w$ and $(\partial\rho/\partial T)Q_i$ are also fixed, the buoyancy fluxes do not change, and no oscillations are possible. Oscillations were observed in Welander (1982) even with a linear equation of state because Q_i was not fixed. As described there, however, that formulation had the characteristic that only situations with warm, salty surface waters could show oscillations.

With a *non*-linear equation of state, the buoyancy flux due to heat is not fixed, since it then depends on the average temperature of the water column; when $\gamma < 0$, as in sea water, a fixed heat flux gives a progressively larger buoyancy flux as the temperature is increased. Since heat is the destabilizing factor, the water column gets less statically stable as the average temperature increases. When convection ceases, the average temperature in the water column increases to a point where the destabilizing heat flux becomes greater than the stabilizing fresh water flux, and convection starts. Once that happens, the average temperature falls, the destabilizing buoyancy flux decreases, and convection shuts off. Thus the cycle maintains itself.

It is worth noting that the same type of oscillation can be caused by a different non-linearity: the changing average salinity in the surface layer. This arises when the characteristic surface salinity ($\overline{S}$) is so modified by the precipitation that it alters the stabilizing buoyancy flux due to fresh water input, $(\partial\rho/\partial S)\overline{S}F_{P-E}$. The physical idea in this case is that constant rainfall on a confined patch of seawater results in a progressively smaller buoyancy flux as the seawater gets fresher. The size of this effect in the LSG model can be estimated from Fig. 5, which shows that the surface salinity varies by about 2 ppt, giving a change in buoyancy flux of about 6%. By contrast, the increase in temperature from 1.5 C to 4.0 C which can be seen in Fig. 20 gives a change in buoyancy flux of about 42%. The sense of these two effects is such that the oscillations are encouraged; i.e., when convection shuts off, the subsurface waters heat up, increasing the destabilizing buoyancy flux due to heat; at the same time, the surface waters freshen, decreasing the stabilizing buoyancy flux due to precipitation.

The period of the oscillations in the flip-flop model is given by the temperature rise ΔT needed in layer 2 to overcome the stabilizing salinity gradient, divided by the net heating rate of the lower box, $[Q_i - K(T_2 - T_1)]/h_2$. Estimating from Fig. 9 that $\Delta T \sim 2$°C, and taking h_2 = 500 m, $K = 6 \times 10^{-7}$ m s^{-1}, and the interior heat flux to be 10 w m^{-2} yields a period of about 30 years. This discrepancy with the period observed in the model (320 years) is directly attributable to the simplistic vertical heat transfer formulation in the flip-flop model. The net deep heating predicted by this simple scheme would amount to .7 °C per decade, far more than the heating seen in the model (Fig. 9). Clearly, even a very few convecting gridboxes in the numerical model are sufficient to vent the subsurface heat far more rapidly than would be possible by vertical mixing alone; this lengthens the period of oscillations considerably.

a. Application of the Flip-Flop Model to the full LSG Model

Mikolajewicz and Maier-Reimer (1991) noted that the thermohaline oscillations which they observed in the full LSG model, forced with noise added to the fresh water flux, did not occur when an equation of state linearized about T_o=3.0°C, S_o=35 ppt, and p_o=0 db was used. Since the flip-flop model predicts that no oscillations should be seen with a linearized equation of state, we ran a number of additional cases to verify that this finding holds under a wider range of conditions.

For our runs, we tried all combinations of linearizing about three center points, (T_o, S_o, p_o) = (3.0, 35.0, 500.0), (1.0, 35.0, 1000.0), (-1.0, 35.0, 1000.0), and using three forcing scenarios: $\mathcal{F}(.8)$, $\mathcal{F}(1.1)$, and $\mathcal{F}(noise)$. Each of the nine cases was run for at least 3000 model years. The results are shown in Fig. 25. None of the cases showed oscillations approaching the magnitude of those seen when the full equation of state is used.

A run with a *quadratic* equation of state expanded around 0°C and 35 ppt, however, oscillated readily when forced with .8 of the standard fresh water flux. The transport through the Drake Passage in this case is shown in Fig. 26.

Thermohaline oscillations in the full LSG model, then, are highly sensitive to the presence of a non-linear term in the equation of state. This provides strong evidence that the physics embodied in the simple flip-flop model are essential to driving the oscillations in the full LSG model.

7. Discussion

Solar insolation falling upon the Earth is such that there is net heating in the tropical regions and net cooling in the polar regions. In the present climate regime, there also tends to be positive P-E in the polar regions, which in turn must be fed by negative P-E over the remaining part of the globe. This combination of circumstances mandates that polar regions are subjected to the conflicting tendencies of a destabilizing surface heat loss and a stabilizing net fresh water flux.

There are only three general static stability regimes possible under such circumstances. The water column might be always stable, if strong precipitation is coupled with weak cooling; it might be always unstable (and therefore convecting) if weak precipitation is coupled with strong cooling; or it might vary between these two regimes. Such variation can occur over time, such that the large-scale thermohaline circulation oscillates; this is the physical essence of the "flip-flop" model discussed in Section 5. Alternatively, the variation can be over space, such that some locales are continually convecting and others always capped off; this is the situation with the "standard" version of the LSG model, which exhibits no thermohaline oscillations over time but has a strong spatial pattern of convective activity. The test case where the fresh water flux applied to the Southern Ocean was smoothed mitigated these regional differences, and caused the model to change from the spatially-varying convection case to the temporally-varying case.

The fluctuations can form self-sustaining oscillations if non-linearities in sea water's equation of state are taken into account. The sense of the non-linearities are such that the destabilizing buoyancy flux due to heat increases in magnitude with increasing temperature. When convection in a region shuts off, the interior temperature increase until the destabilizing heat flux dominates, reinitiating convection. In our case, the destabilizing heat flux is provided by NADW.

It should be fully appreciated that this destabilizing heating is *not* a fortuitous happenstance, peculiar to the topographic details of a particular ocean basin or water column structure. It is simply the heat necessary to replace that lost through the surface, and so destabilizing subsurface heating is a feature of *every* polar ocean. The ubiquity of the subsurface heating in polar regions helps explain the presence of coherent "zones" of convective activity described above.

Subsurface heating is not the exact counterpart of destabilizing surface cooling, however; in the present climactic regime, polar ocean surface temperatures in the winter quickly reach the freezing point, and cannot get any colder or denser due to temperature. However, intermediate waters *can* always get lighter due to warming. This leads to the situation observed in the results described here, with convection beginning at depth rather than at the surface.

a. Application to the Real Oceans

The likelihood that this kind of oscillation occurs in the real oceans is a function not only of the physical mechanism described here, which is quite general, but of the particular heat and fresh water fluxes which drive the Earth's oceans.

Consider the buoyancy flux due to the net fresh water flux, $B_{FW} \equiv g\beta\overline{S}F_{FW}$ (where g is the gravitational constant, β the coefficient of haline contraction, $\overline{S}$ a characteristic surface salinity, and F_{FW} the net rate of P-E), and the buoyancy flux due to heat, $B_H \equiv g\alpha F_H/(\rho_o c_p)$ (where α is the coefficient of thermal expansion, F_H is the heat flux, and c_p is the heat capacity of seawater). As demonstrated in the flip-flop model, B_{FW} and B_H must be nearly equal in the polar regions for these oscillations to occur.

In the LSG model, the heat transport southward across 60 °S is about 1.5 PW; using $\alpha = .5 \times 10^{-4}$ C^{-1}, a value appropriate for sea water at 0 C, this yields an average destabilizing buoyancy flux of $B_H = 7.7 \times 10^{-9}$ m^2 s^{-3}. The model's average P-E south of 60°S is 2.6×10^{-8} m s^{-1}, which yields a stabilizing buoyancy flux due of $B_{FW} = 6.9 \times 10^{-9}$ m^2 s^{-3}. The closeness of B_H and B_{FW} in the LSG model is key to allowing the oscillations between convecting and non-convecting states. The range over which oscillations exist in the model can be roughly estimated by noting that oscillations do not occur if the P-E field is changed by more than about $\pm$ 30%. This is equivalent to a range of buoyancy fluxes from 4.8 to 9.0 $\times 10^{-9}$ m^2 s^{-3}. The range of heat fluxes across 60°S which would give an equivalent buoyancy flux is .9 to 1.7 PW. The 1.5 PW heat transport in the LSG model is within, but near the high end, of this range.

Unfortunately, there is a tremendous range of uncertainty in the actual values of heat flux across 60°S and P-E around Antarctica. Estimates of the heat flux, in particular, vary widely; Table 1 shows some estimates, from both observations and models, of this quantity. The range is too large to establish what the behavior of the real oceans should be in this respect. A value at the lower end of the estimates, .2 PW, would imply that the destabilizing buoyancy flux due to heat is only about 1/7 of the stabilizing fresh water flux; no large-scale oscillations proceeding via the mechanism described here would be expected in such a case. On the other hand, an assumed heat transport of 2 PW might be too large to allow the fresh water cap to ever become established. Observations which more tightly constrain the heat flux across 60°S are needed to resolve this question.

Finally, we consider the degree to which the LSG's numerics may be contributing to the oscillatory behavior. The LSG's 1.5 PW of poleward heat flux across 60°S is near the upper end of the estimates shown in Table 1. The LSG model has an upwind advection algorithm which is intrinsically quite diffusive, and which therefore, will tend to increase the heat flux across the ACC. However, it is conceivable that the lowest estimates shown in Table 1, which come from eddy-resolving OGCMs, may be too low due to underrepresentation of the eddy activity.

8. Conclusions

Mikolajewicz and Maier-Reimer (1990) showed that the Hamburg LSG OGCM could exhibit the sorts of thermohaline oscillations previously found in simplified models, when noise is added to the fresh water flux field. There and in Mikolajewicz and Maier-Reimer (1991) they analyzed the oscillation in terms of a advected salinity anomaly dipole of the kind described by Welander (1986), and noted that a case with a linearized equation of state did not oscillate.

The work here continues the exploration the physical causes of the thermohaline oscillations. We have demonstrated the following additional aspects:

1. *Added noise is not required to obtain the oscillations.* The LSG model develops thermohaline fluctuations without any added noise whatsoever, such as when the fresh water flux field is uniformly decreased by 7.5% to 30%; when the fresh water flux field is increased by 10%; when a quadratic equation of state is used; when the brine rejection effect is removed from the sea ice model; and when a smoothed precipitation field around Antarctica is used.

2. *The mechanism of the oscillations is destabilizing subsurface heating around Antarctica competing with stabilizing precipitation.* When a fresh water cap prevents convection, heat accumulates between 200 and 700 m; once this water is warm and light enough, it can overturn the water column and mix away the

fresh water cap. The heat then vents to the atmosphere, reducing the rate of convection; fresh water begins to accumulate on the surface once more, eventually shutting of convection. The cycle then repeats.

3. *The Flip-Flop model explains why the LSG model does not oscillate readily with a linearized equation of state*. The flip-flip model described in this work is a general extension of the work in Welander (1982) to polar regions, which have cold, fresh water atop warm, salty water. Additionally, it is suitable for considering multi-century timescales, as it conserves heat and salt. A straightforward analysis of the flip-flop model shows that it cannot oscillate unless a non-linear equation of state is used. This arises in the following way: with a non-linear equation of state, a constant *heat* flux becomes an ever-greater *buoyancy* flux as the average temperature increases. Therefore, a buoyancy flux which initially is too small to overturn a water column can increase until it is able to do so, if it heats up the water. Once overturning starts, the average temperature drops, and the buoyancy flux decreases again. To test this explanation, we ran nine test cases of the LSG model with the equation of state linearized around various points; none showed large-scale thermohaline fluctuations.

4. *The "standard" LSG model does not oscillate because of the particular pattern of P-E around Antarctica used to drive it.* Averaging the P-E field in the Enderby sector started the model oscillating. The original field has local P-E contrasts strong enough to keep some model gridpoints convecting consistently, preventing the buildup of heat associated with the thermohaline oscillations. The original P-E field was obtained by relaxing the model to specified *steady* conditions during a spin-up run. This study uncovered a possible problem with this common technique; we found that relaxing to a steady condition resulted in

a flux field which drove the model to a steady condition. Climate models which use a unvarying "flux correction" term may be subject to the same problem. It seems critical to determine if these techniques lead models to consistently underestimate the degree of variability possible, as most, if not all, greenhouse simulations have used such schemes. In doing so they may have unknowingly affected the results.

5. *The detailed time evolution of the oscillations can be understood in terms of which Antarctic regions are convecting at any given moment*. Not every region of the Southern Ocean convects at the same time. In the LSG model, the Weddell and Enderby regions tend to convect together, some time before the Ross and Amundsen regions begin convecting. This sequence is reflected in the detailed time evolution of the maximum overturning streamfunction in the Southern Ocean.

It is important to note that the subsurface heating which drives oscillations is *not* a peculiarity of the particular water masses, bottom topographies, or model involved; rather, it is the mechanism for replacing heat lost to the atmosphere in the polar regions, and so will be present in the real world as well as in the LSG model. The only question is whether it will be sufficient to catastrophically destabilize the water column, initiating convection from below, or can be vented to the atmosphere by surface-induced convection or vertical mixing before that happens. This question cannot be answered without tighter observational constraints on the net heat flux across 60°S.

The surface fresh water flux field which has the effect of stabilizing the LSG model was obtained by relaxing the model to specified *steady* conditions during a spin-up run. This study uncovered a problem with this common technique; we found that relaxing to a steady condition resulted in a flux field which drove the model to a steady condition, rather than a flux field which was realistic. Climate models which use an unvarying "flux correction" term may be subject to the same problem. It seems critical to determine if these techniques lead models to consistently underestimate the degree of variability possible, as most, if not all, greenhouse simulations have used such schemes. In doing so they may have unknowingly affected the results.

Acknowledgments

The authors would like to thank Tim Osborn of the University of East Anglia and Paola Cessi of the Scripps Institution of Oceanography for valuable discussions, and Tony Tubbs of Scripps for his technical assistance. This research was supported by DOE CHAMMP contract DE-FG03-91-ER61215 and NSF contract ATM 9314495, and sponsored in part by the University of California and Digital Equipment Corporation under Research Grant No. 1243 (the Sequoia Project) and NOAA grants NA376P0518 and NA476P0188.

*List of References

Barnett, T. P., Chu, M., Wilde, R., and Mikolajewicz, U. (1994). Low frequency ocean variability induced by stochastic forcing of various colors. In *Proceedings of the Nat. Res. Council Workshop on Decade to CenturyTimescales of Climate Variability, Irvine, CA, Sep. 21-25, 1992*. In Press.

Broecker, W. S. (1979). A Revised Estimate for the Radiocarbon Age of North Atlantic Deep Water. *J. Geophys. Res.*, 84:3218–3226.

Broecker, W. S., Peteet, D. M., and Rind, D. (1985). Does the ocean-atmosphere system have more than one stable mode of operation? *Nature*, 315:21–26.

Bryan, F. (1986). High-latitude salinity effects and interhemispheric thermohaline circulations. *Nature*, 323:301–304.

Cai, W. (1994). Circulation driven by observed surface thermohaline fields in a coarse resolution ocean general circulation model. *J. Geophys. Res.*, 99:10163–10181.

Carissimo, B. C., Oort, A. H., Vonder, and Haar, T. H. (1985). Estimating the meridional energy transports in the atmosphere and oceans. *J. Phys. Oceanogr.*, 15:82–91.

Cox, M. D. (1975). A baroclinic numerical model of the World Ocean: preliminary results. In *Numerical Models of Ocean Circulation*, pages 107–120. National Academy of Sciences.

Cox, M. D. (1989). An idealized model of the world ocean. Part I: the global-scale water masses. *J. Phys. Oceanogr.*, 19:1730–1751.

Gordon, A. L. (1981). Seasonality of Southern Ocean sea ice . *J. Geophys. Res.*, 86:4193–4197.

Hastenrath, S. (1982). On meridional heat transports in the World Ocean . *J. Phys. Oceanogr.*, 12:922–927.

Hellerman, S. and Rosenstein, M. (1983). Normal monthly wind stress over the World Ocean with error estimates . *J. Phys. Oceanogr.*, 13:1093–1104.

Holland, W. R. (1973). Baroclinic and topographic influences on the transport in Western Boundar Currents. *Geophysical Fluid Dynamics*, 4:187–210.

Levitus, S. (1982). *Climatological atlas of the world ocean* . U.S. Dept. of Commerce, N.O.A.A.

Maier-Reimer, E., Mikolajewicz, U., and Hasselmann, K. (1993). Mean circulation of the Hamburg LSG OGCM and its sensitivity to the thermohaline surface forcing. *J. Phys. Oceanogr.*, 23:731–757.

Manabe, S., Stouffer, R. J., Spelman, M. J., and Bryan, K. (1991). Transient responses of a coupled ocean-atmosphere model to gradual changes of atmospheric CO_2. *Journal of Climate*, 4:785–818.

Mikolajewicz, U. and Maier-Reimer, E. (1990). Internal secular variability in an ocean general circulation model . *Climate Dynamics*, 4:145–156.

Mikolajewicz, U. and Maier-Reimer, E. (1991). One example of a natural mode of the ocean circulation in a stochastically forced ocean general circulation model. In Latif, M., editor, *Strategies for Future Climate Research*, pages 287–318. Max-Planck-Institut fur Meteorologie, Hamburg, Germany.

Nowlin, W. D. and Klinck, J. M. (1986). The physics of the Antarctic Circumpolar Current . *Reviews of geophysics*, 24:469–491.

Roads, J. O., Chen, S. C., Kao, J., Langley, D., and Glatzmaier, G. (1992). Global Aspects of the Los Alamos General Circulation Model Hydrologic Cycle. *J. Geophys. Res.*, 97:10051–10068.

Saunders, P. M. and Thompson, S. R. (1993). Transport, Heat, and Freshwater Fluxes within a Diagnostic Numerical Model (FRAM). *J. Phys. Oceanogr.*, 23:452–464.

Semtner, A. J. and Chervin, R. M. (1992). Ocean general circulation from a global eddy-resolving model . *J. Geophys. Res.*, 97:5493–5550.

Stocker, T. F. and Mysak, L. A. (1992). Climatic fluctuations on the century time scale: A review of high-resolution proxy data and possible mechanisms. *Climatic Change*, 20:227–250.

Stommel, H. (1961). Thermohaline convection with two stable regimes of flow . *Tellus*, 13:224–230.

Trenberth, K. E. (1979). Mean annual poleward energy transports by the ocean in the Southern Hemisphere. *Dynamics of Atmospheres and Oceans*, 4:57–64.

UNESCO (1981). *Unesco technical paper in marine science number 36* . UNESCO.

Weaver, A. J., Marotzke, J., Sarachik, E., and Cummins, P. (1993). Stability and variability of the thermohaline circulation . *J. Phys. Oceanogr.*, 23:39–60.

Welander, P. (1982). A simple heat-salt oscillator . *Dynamics of Atmospheres and Oceans*, 6:233–242.

Welander, P. (1986). Thermohaline effects in the ocean circulation and related simple models. In Willebrand, J. and Anderson, D. L. T., editors, *Large-Scale transport processes in oceans and atmosphere*. NATO ASI series, Reidel Pub. Co.

Winton, M. (1993). *Numerical Investigations of Steady and Oscillating Thermohaline Circulations* . University of Washington Ph. D. Thesis.

Winton, M. and Sarachik, E. S. (1993). Thermohaline oscillations induced by strong steady salinity forcing of ocean general circulation models. *J. Phys. Oceanogr.*, 23:1389–1410.

Woodruff, S. D., Slutz, R. J., Jenne, R. L., and Steurer, P. M. (1987). A comprehensive ocean-atmosphere dataset . *Bull. Amer. Meteor. Soc.*, 68:1239–1250.

Table Captions

Table 1. Estimated values of heat flux (PW) southward across 60°S. For methods, “obs.” indicates an estimate based on some sort of observations while “model” indicates an estimate from a numerical model. See the specific references for details.

Figure Captions

Fig. 1. Variability in the meridional overturning streamfunction is about ten times greater in the Southern Ocean (upper panel) than in the North Atlantic (lower panel). The oscillations have been excited by reducing the P-E field by 20% (case $\mathcal{F}(.8)$). Interval 'A' will be analyzed in detail.

Fig. 2. Fluctuations in the Drake Passage transport which develop in case $\mathcal{F}(.8)$. The correlation with the Southern Ocean overturning streamfunction is .949.

Fig. 3. Meridional circulation (Sv) for case $\mathcal{F}(.8)$ at time THC-ON, the peak of Southern Ocean thermohaline activity. Top panel: all oceans. Middle panel: Atlantic. Bottom panel: Pacific. Contour interval is 5 Sv.

Fig. 4. Meridional circulation (Sv) for case $\mathcal{F}(.8)$ at year model year THC-OFF, the minimum of Southern Ocean thermohaline activity. Top panel: all oceans. Middle panel: Atlantic. Bottom panel: Pacific. Contour interval is 5 Sv.

Fig. 5. The difference in surface salinity between years THC-ON (top panel) and THC-OFF (middle panel) is the presence of a strong, fresh cap in the Southern Ocean (bottom panel; negative values are shaded) when the thermohaline circulation is shut off. $\mathcal{F}(.8)$ case is shown.

Fig. 6. EOF mode 1 for the surface salinity field, $\mathcal{F}(.8)$ case. Top: normalized eigenvector. Bottom: eigenamplitude (ppt) as a function of time. EOF 1 corresponds to the fresh water cap which forms in the Ross sea during times of reduced convection.

Fig. 7. EOF mode 2 for the surface salinity field, $\mathcal{F}(.8)$ case. Top: normalized eigenvector. Bottom:

eigenamplitude (ppt) as a function of time. EOF 2 corresponds to activity in the Weddell and Enderby seas.

Fig. 8. Southern Ocean sea ice thickness at time THC-ON (top) and THC-OFF (bottom), $\mathcal{F}(.8)$ case. Contour levels are .1, .5, 1.0, 1.5, and 2.5 meters, and every 2.5 meters thereafter. The ice is considerably thinner when convection is bringing up heat from below.

Fig. 9. Time versus depth contours of model fields averaged over the Weddell Sea, $\mathcal{F}(.8)$ case. Top: Temperature (C), showing the subsurface (200-500 m) heating which destabilizes the water column. Middle: Salinity (ppt), showing the fresh water cap which forms, shutting off convection. Bottom: Frequency of convection (fraction of gridboxes).

Fig. 10. Evolution of the stability of the Weddell Sea water column, $\mathcal{F}(.8)$ case. Top: total N^2. Middle: the (destabilizing) contribution due to temperature. Bottom: the (stabilizing) contribution due to salinity. Negative regions are shaded; only the top 500 m are shown. Units are 10^{-6} s^{-2}.

Fig. 11. In the shaded regions, the depth of first convection after time THC-OFF is greater than 700 meters. In all other regions, the initial depth of convection is less than 75 meters. $\mathcal{F}(.8)$ case.

Fig. 12. Antarctic zones used for the analysis. All zones extend to 60°S.

Fig. 13. Top: average temperature (C) by Antarctic sector, as a function of time, $\mathcal{F}(.8)$ case. Also shown is the volume transport through the Drake Passage (lower panel) during the same time period, for comparison.

Fig. 14 The time evolution of convection around Antarctica for the $\mathcal{F}(.8)$ case. Plotted is the Convective

Adjustment Index (CAI), the fraction of time a gridbox is convected into from above.

Fig. 15. Average sea ice thickness (m) in the Weddell and Ross sectors as a function of time, $\mathcal{F}(.8)$ case.

Fig. 16. Transport through the Drake Passage (Sv), showing the change in model variability when convection in the North Atlantic is suppressed starting year 4600.

Fig. 17 Drake Passage transport (Sv) for the case with 2 mm/day of white noise added to the fresh water flux field. The section marked 'A' will by analyzed in detail.

Fig. 18. A: the normalized CAI for the Weddell and Enderby sectors. B: the normalized CAI for the Amundsen, Ross, and Wilkes sectors. C: sum of (A) and (B). D: Southern Ocean meridional overturning strength for the same period.

Fig. 19 The time evolution of convection around Antarctica for the $\mathcal{F}(noise)$case. Contours of CAI are shown.

Fig. 20. Time versus depth contours of model fields averaged over the Weddell sector, $\mathcal{F}(.8)$ case. Top: Temperature (C), showing the subsurface (200-500 m) heating which destabilizes the water column. Middle: Salinity (ppt), showing the fresh water cap which forms, shutting off convection. Bottom: Frequency of convection (fraction of gridboxes).

Fig. 21. Time versus depth contours of model fields averaged over the Ross sector, $\mathcal{F}(.8)$ case. Top: Temperature (C). Middle: Salinity (ppt). Bottom: Frequency of convection (fraction of gridboxes).

Fig. 22 Drake Passage transport (Sv) for case $\mathcal{F}(1.1)$, where the P-E field has been multiplied by 1.1.

Fig. 23. The transport (Sv) through the Drake Passage for the case with averaged P-E in the eastern Enderby sector.

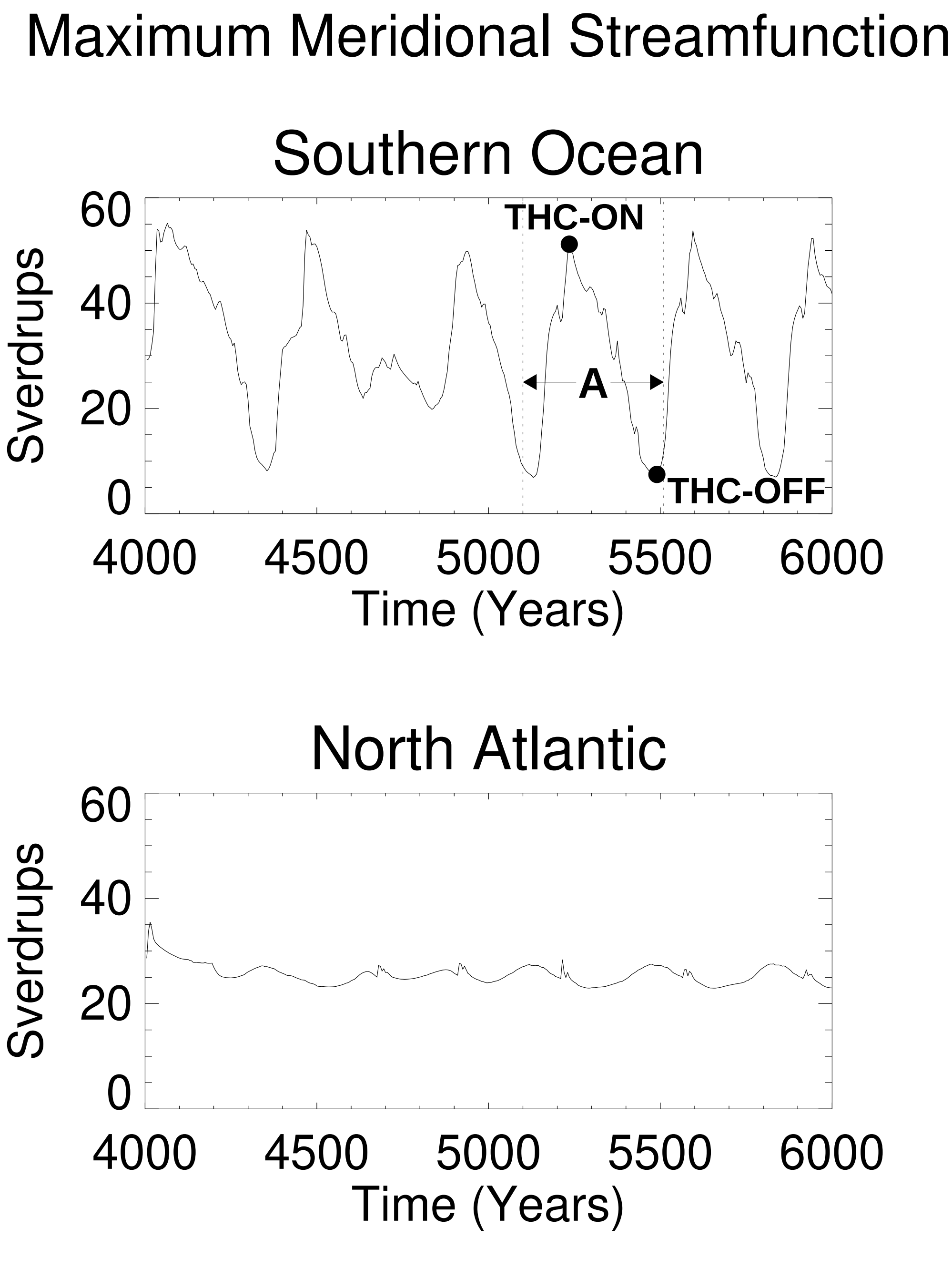
Fig. 24. Schematic of the “flip-flop” model. To maintain long-term balance, net heat lost through the surface must be returned at depth; all fresh water gained at the surface must be removed at depth.

Fig. 25. Drake Passage transport for 9 cases with a linearized equation of state: in none of the cases were the large-scale fluctuations seen.

Fig. 26. Drake Passage transport (Sv) for the test case using a quadratic equation of state. Unlike the cases with a linearized equation of state, oscillations develop readily.

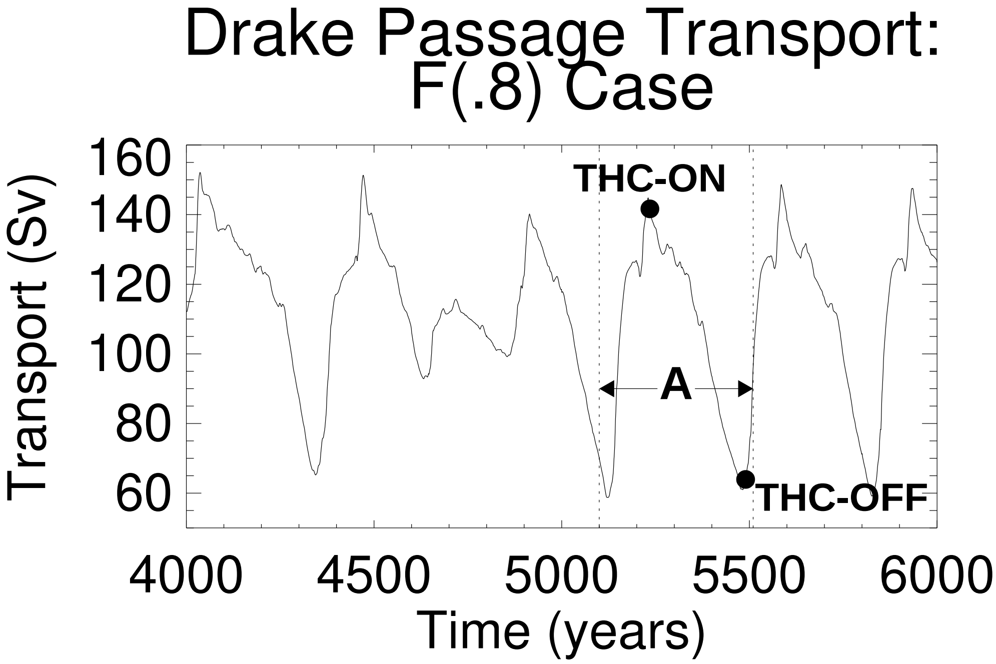
Estimates of Poleward Heat Flux Across 60° S

Source	*Method*	*Value (PW)*
Trenberth 1979	obs. (heat budget)	1.00
Gordon 1981	obs. (sea ice)	.54
Hastenrath 1982	obs. (heat budget)	.65
Carissimo *et al.* 1985	obs. (heat budget)	2.00
Manabe and Stouffer 1991	model (Atmos/Ocean)	.75
Semtner and Chervin 1992	model (GFDL)	.15
Saunders and Thompson 1993	model (FRAM)	.20



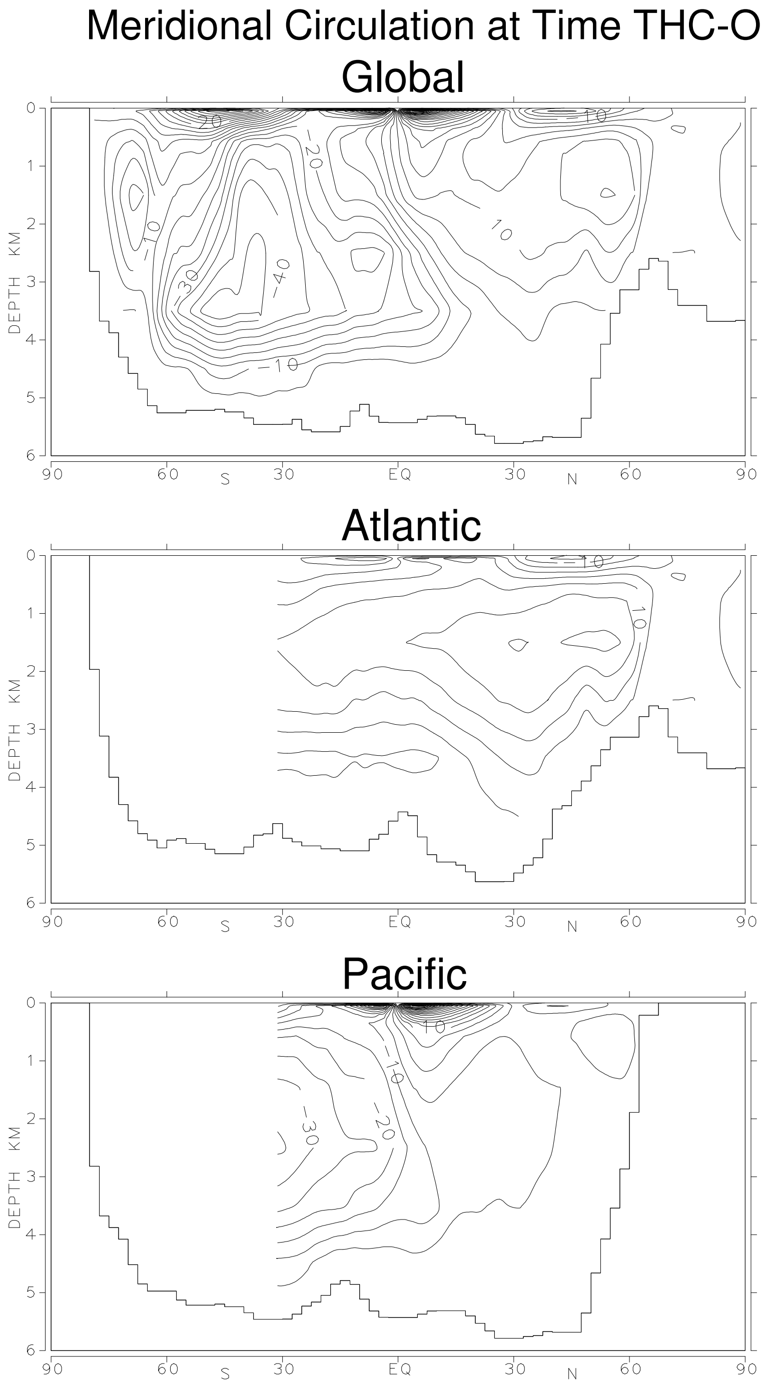
Maximum Meridional Streamfunction
Southern Ocean
60
40
20
0
Sverdrups
THC-ON
A
THC-OFF
4000
4500
5000
5500
6000
Time (Years)
North Atlantic
60
40
20
0
Sverdrups
4000
4500
5000
5500
6000
Time (Years)

Figure 1:



Drake Passage Transport:
F(.8) Case
Transport (Sv)
160
140
120
100
80
60
THC-ON
A
THC-OFF
4000
4500
5000
5500
6000
Time (years)

Figure 2:



Meridional Circulation at Time THC-ON
Global
Atlantic
Pacific
DEPTH KM
0
1
2
3
4
5
6
90
60
S
30
EQ
30
N
60
90
-20
-10
-30
-40
10

Figure 3:

Meridional Circulation at Time THC-OFF

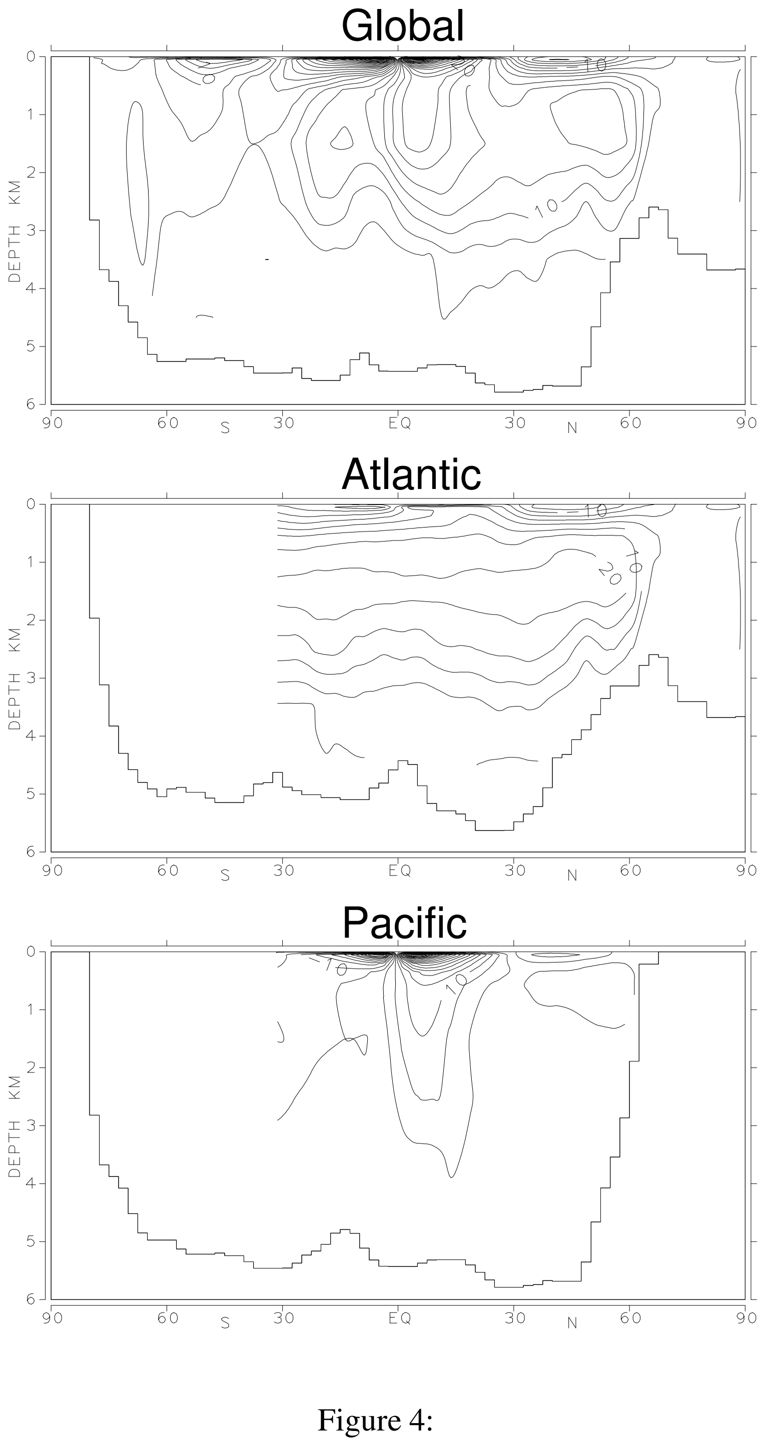


Figure 4:

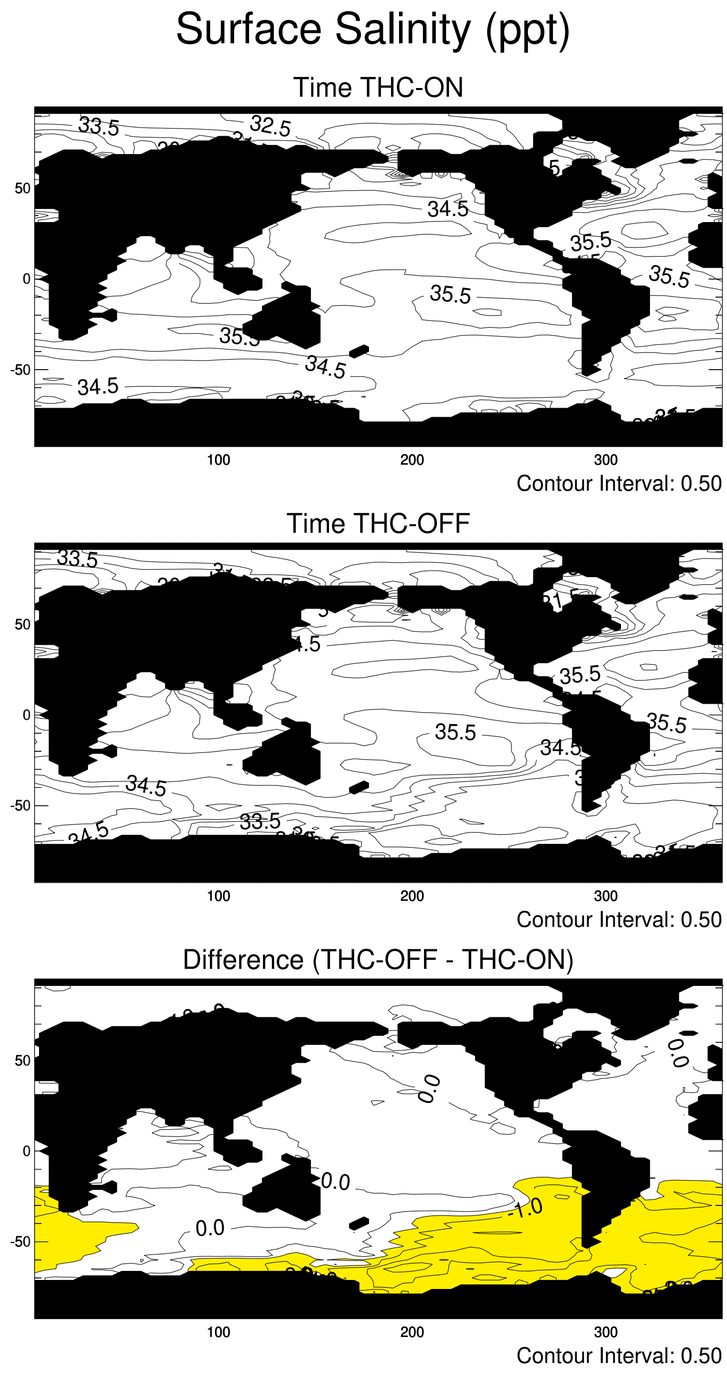
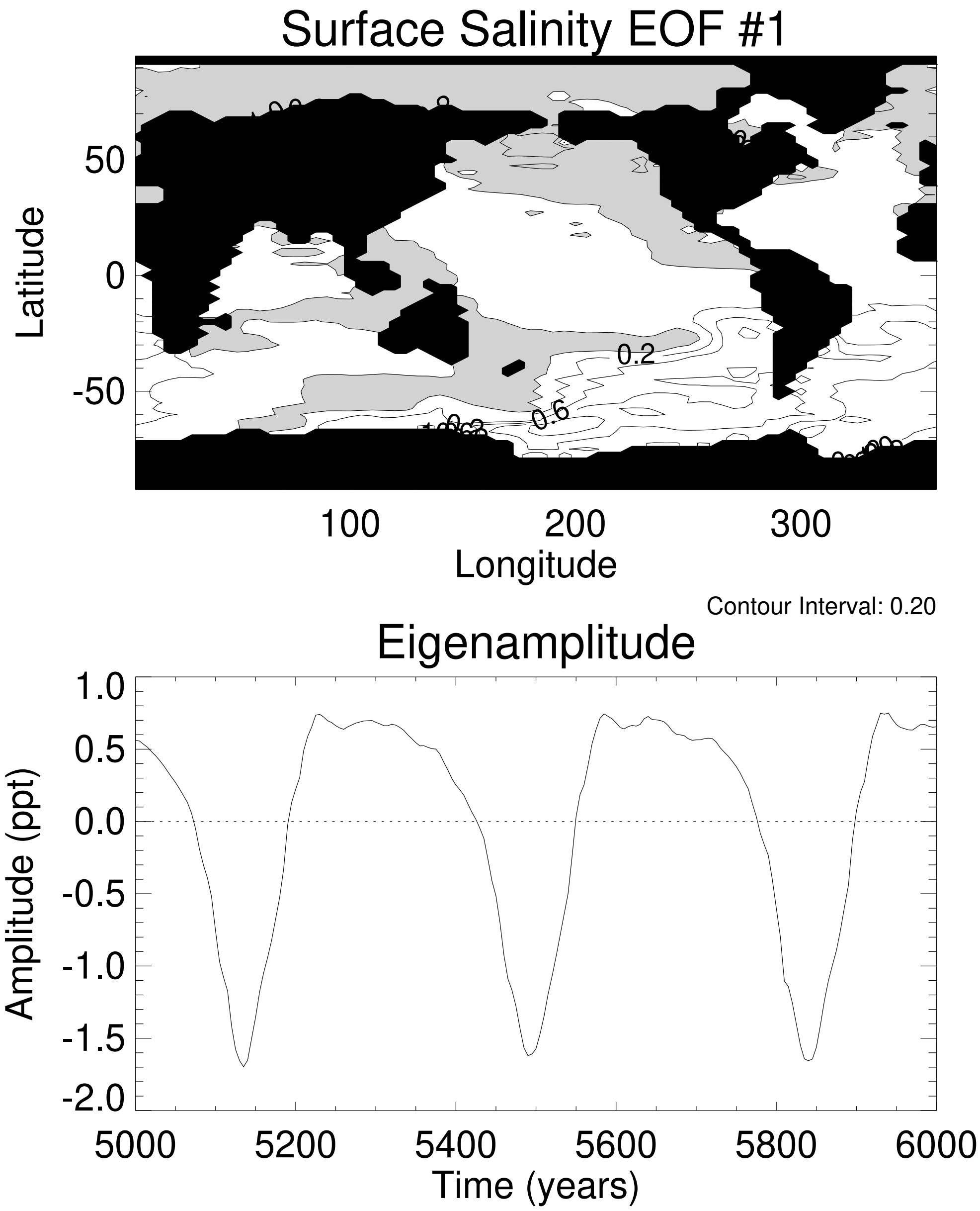

Surface Salinity (ppt)
Time THC-ON
33.5
32.5
34.5
35.5
35.5
35.5
35.5
34.5
34.5
50
0
-50
100
200
300
Contour Interval: 0.50
Time THC-OFF
33.5
35.5
35.5
34.5
35.5
34.5
33.5
50
0
-50
100
200
300
Contour Interval: 0.50
Difference (THC-OFF - THC-ON)
0.0
0.0
0.0
0.0
-1.0
50
0
-50
100
200
300
Contour Interval: 0.50

Figure 5:



Surface Salinity EOF #1
Latitude
50
0
-50
100
200
300
Longitude
0.2
0.6
Contour Interval: 0.20
Eigenamplitude
Amplitude (ppt)
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
5000
5200
5400
5600
5800
6000
Time (years)

Figure 6:

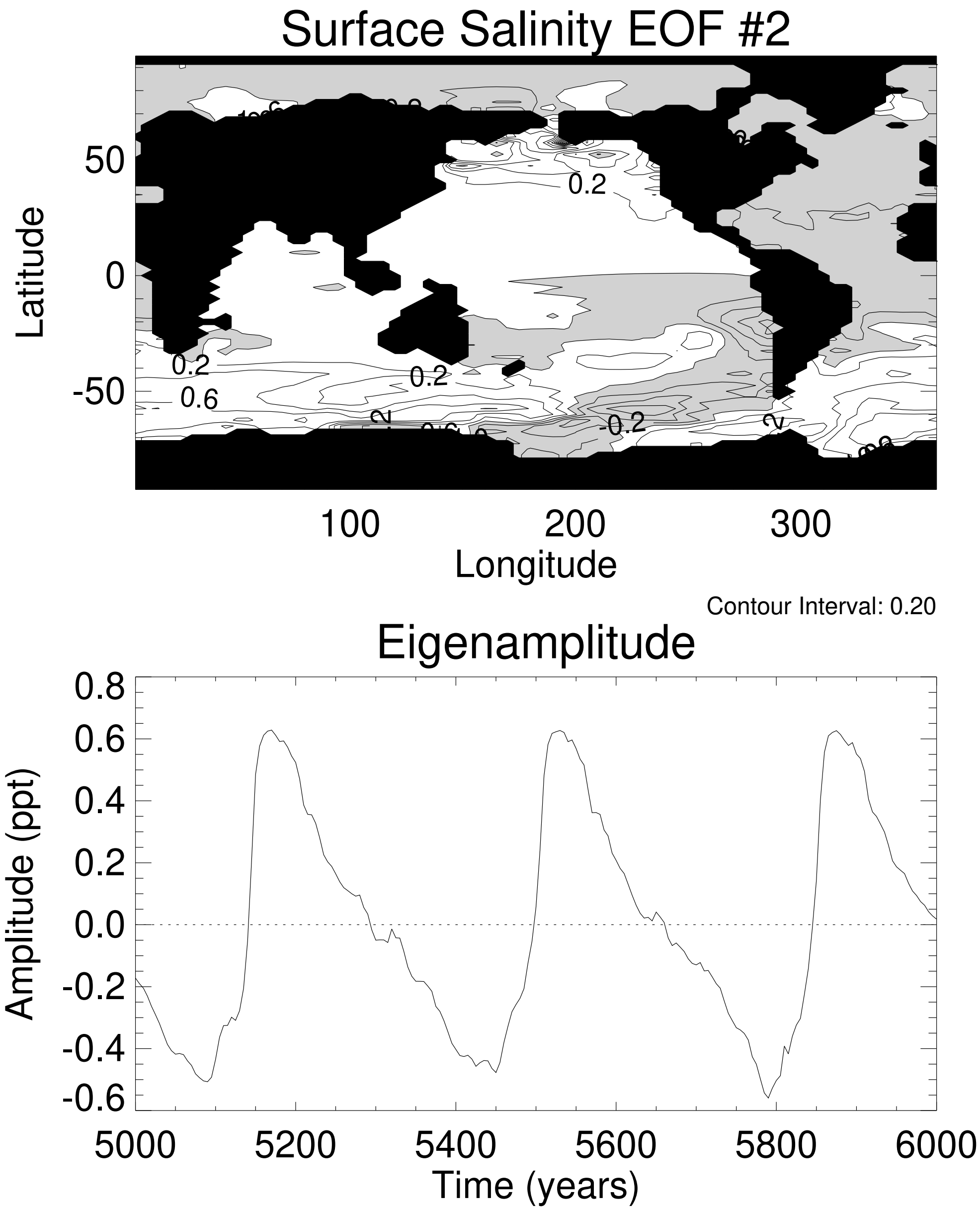

Surface Salinity EOF #2
Latitude
50
0
-50
100
200
300
Longitude
0.2
0.6
-0.2
Contour Interval: 0.20
Eigenamplitude
Amplitude (ppt)
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
5000
5200
5400
5600
5800
6000
Time (years)

Figure 7:

Southern Ocean Ice Thickness

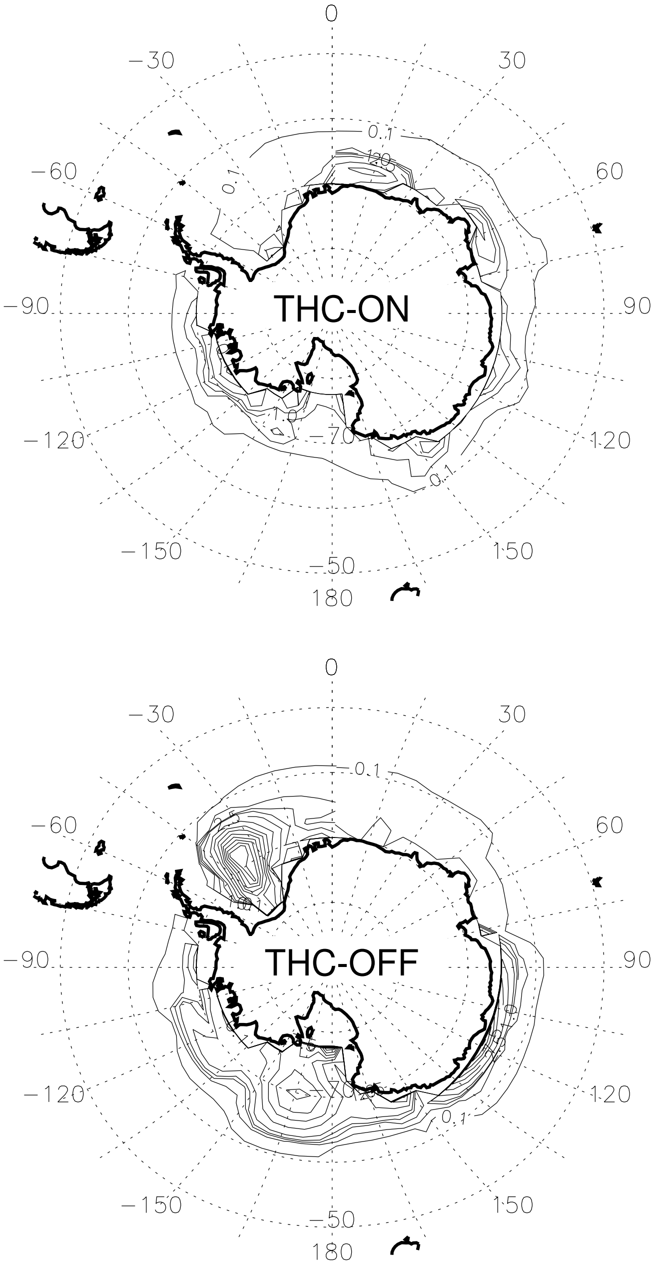
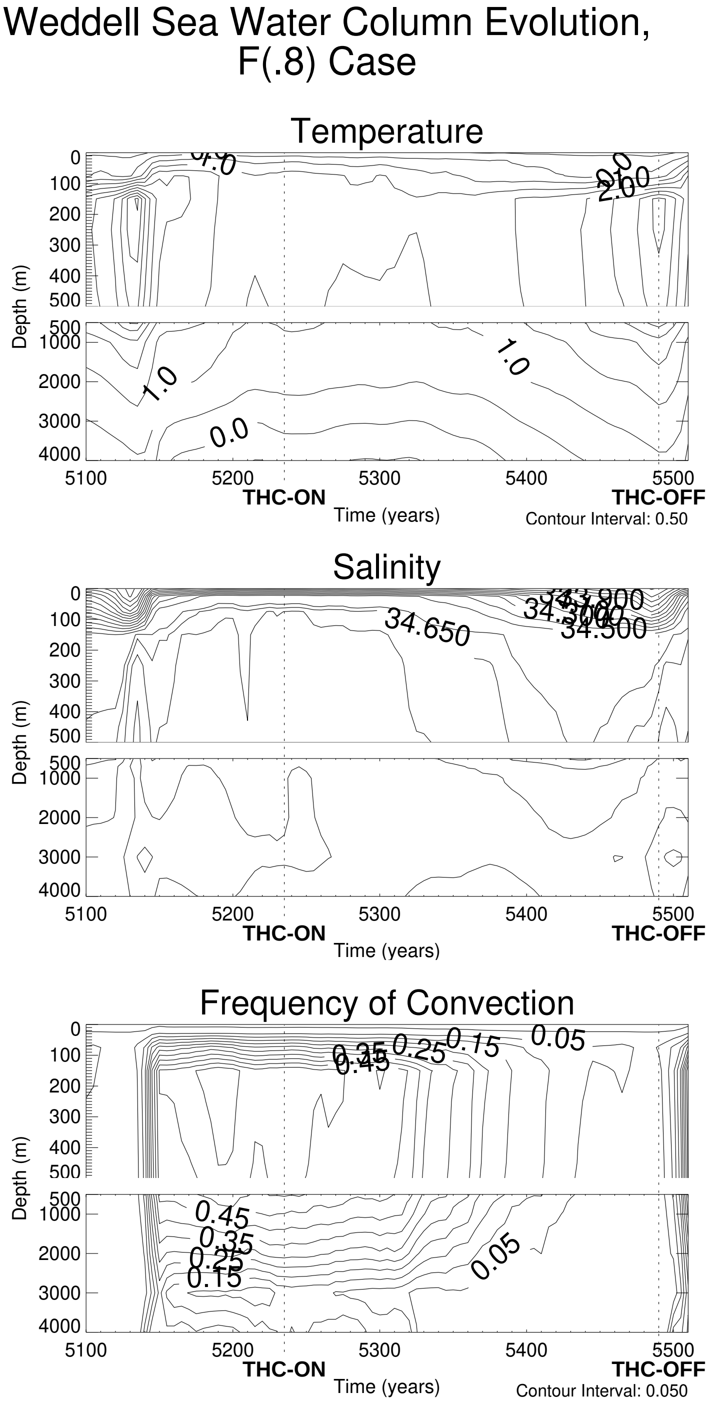
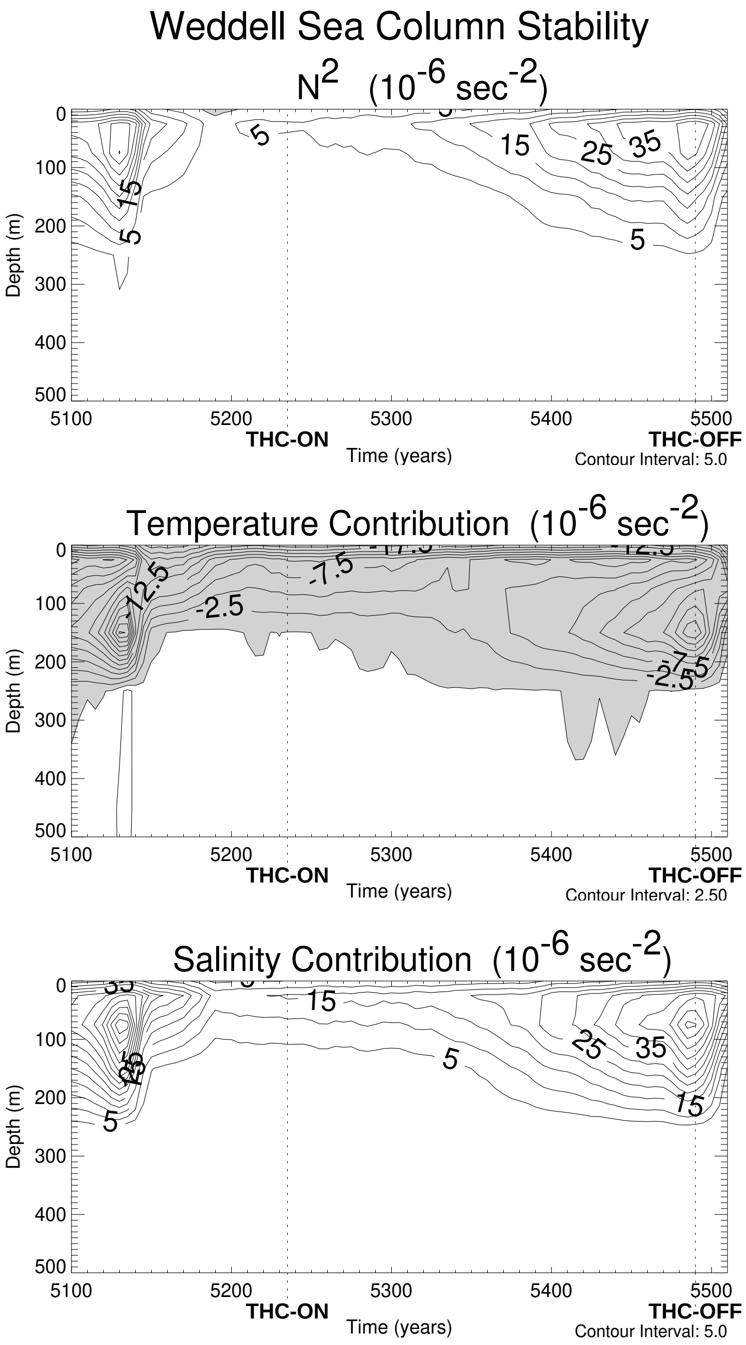


Figure 8:



Weddell Sea Water Column Evolution,
F(.8) Case
Temperature
Depth (m)
2.00
1.0
0.0
Time (years)
THC-ON
THC-OFF
Contour Interval: 0.50
Salinity
34.650
34.500
Depth (m)
Time (years)
THC-ON
THC-OFF
Frequency of Convection
0.45
0.35
0.25
0.15
0.05
Depth (m)
Time (years)
THC-ON
THC-OFF
Contour Interval: 0.050

Figure 9:



Weddell Sea Column Stability
N2 (10-6 sec-2)
Depth (m)
Time (years)
THC-ON
THC-OFF
Contour Interval: 5.0
Temperature Contribution (10-6 sec-2)
Depth (m)
Time (years)
THC-ON
THC-OFF
Contour Interval: 2.50
Salinity Contribution (10-6 sec-2)
Depth (m)
Time (years)
THC-ON
THC-OFF
Contour Interval: 5.0

Figure 10:

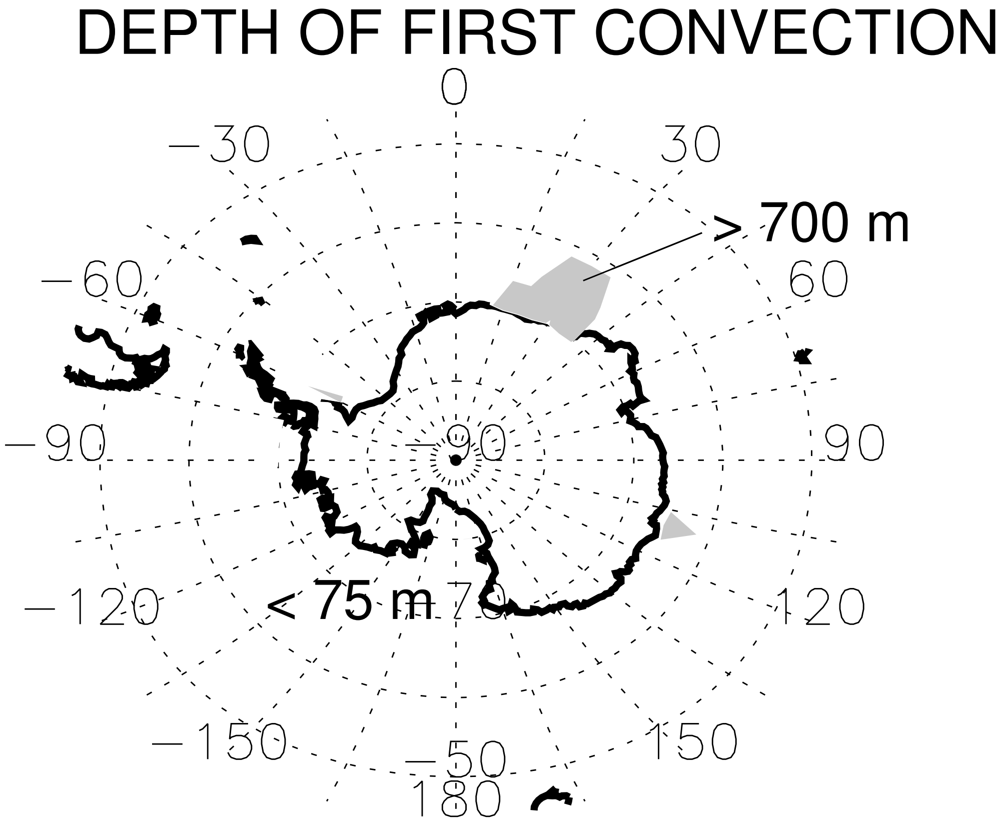

DEPTH OF FIRST CONVECTION
0
−30
30
> 700 m
−60
60
−90
−90
90
−120
< 75 m
−70
120
−150
−50
150
180

Figure 11:

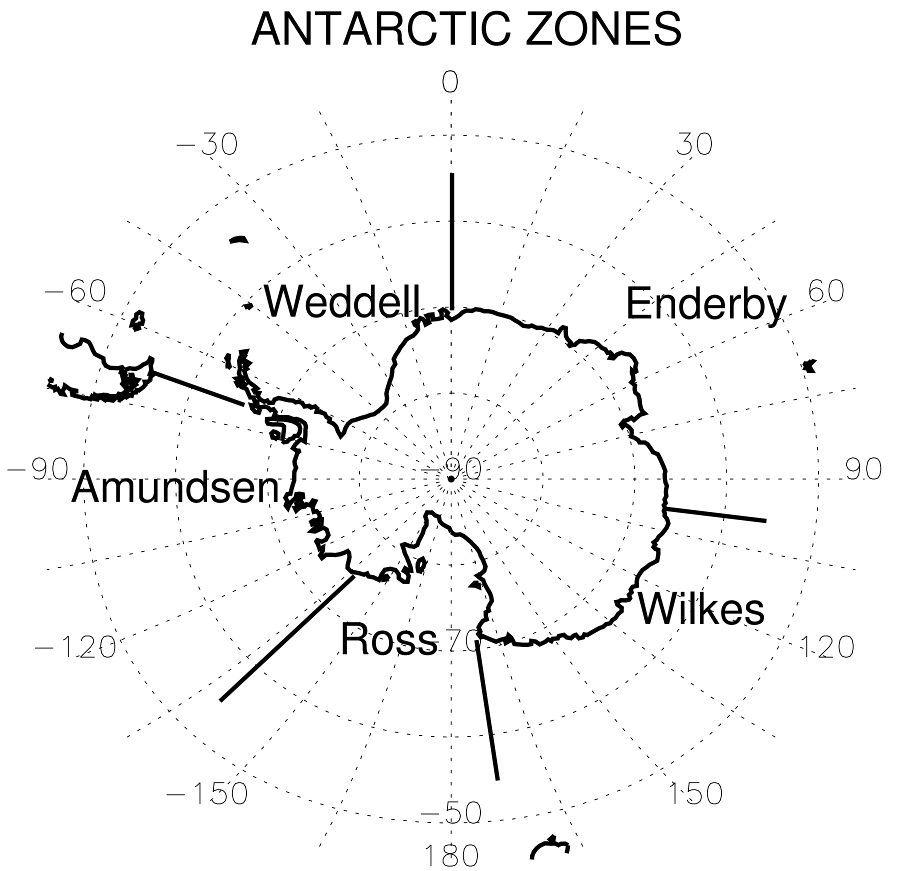
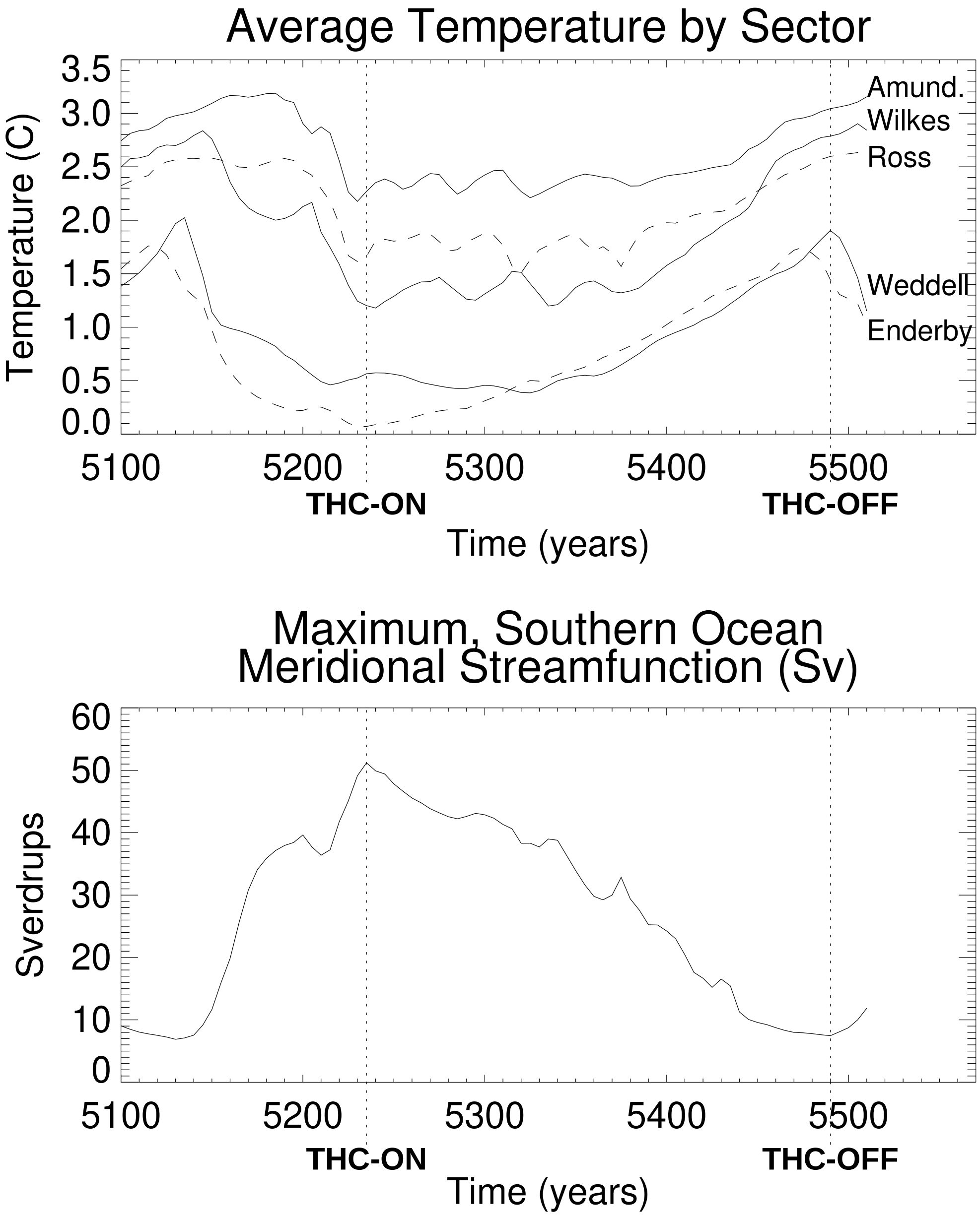

ANTARCTIC ZONES
0
−30
30
−60
60
−90
90
−120
120
−150
150
180
−50
−70
−90
Weddell
Enderby
Amundsen
Wilkes
Ross

Figure 12:



Average Temperature by Sector
Temperature (C)
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Amund.
Wilkes
Ross
Weddell
Enderby
5100
5200
5300
5400
5500
THC-ON
THC-OFF
Time (years)
Maximum, Southern Ocean
Meridional Streamfunction (Sv)
Sverdrups
60
50
40
30
20
10
0
5100
5200
5300
5400
5500
THC-ON
THC-OFF
Time (years)

Figure 13:

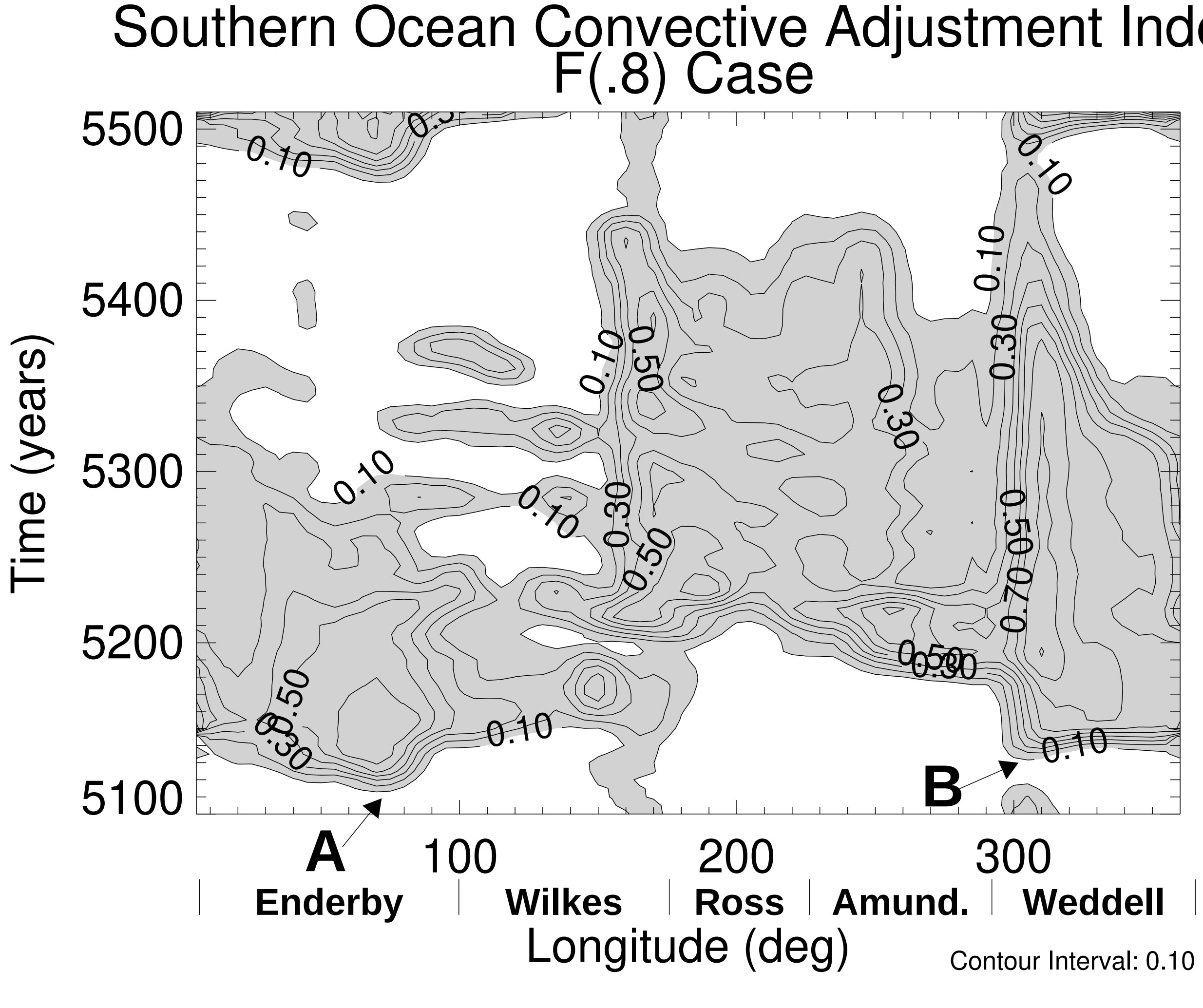

Southern Ocean Convective Adjustment Index
F(.8) Case
Time (years)
5500
5400
5300
5200
5100
100
200
300
A
B
Enderby
Wilkes
Ross
Amund.
Weddell
Longitude (deg)
Contour Interval: 0.10

Figure 14:

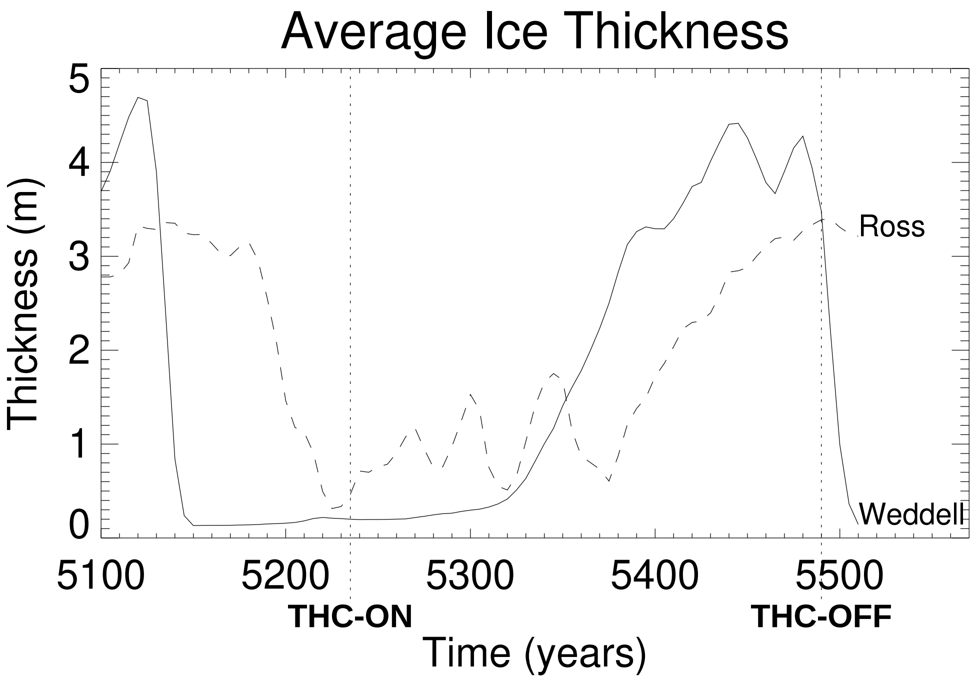

Average Ice Thickness
Thickness (m)
5
4
3
2
1
0
5100
5200
5300
5400
5500
THC-ON
THC-OFF
Time (years)
Ross
Weddell

Figure 15:

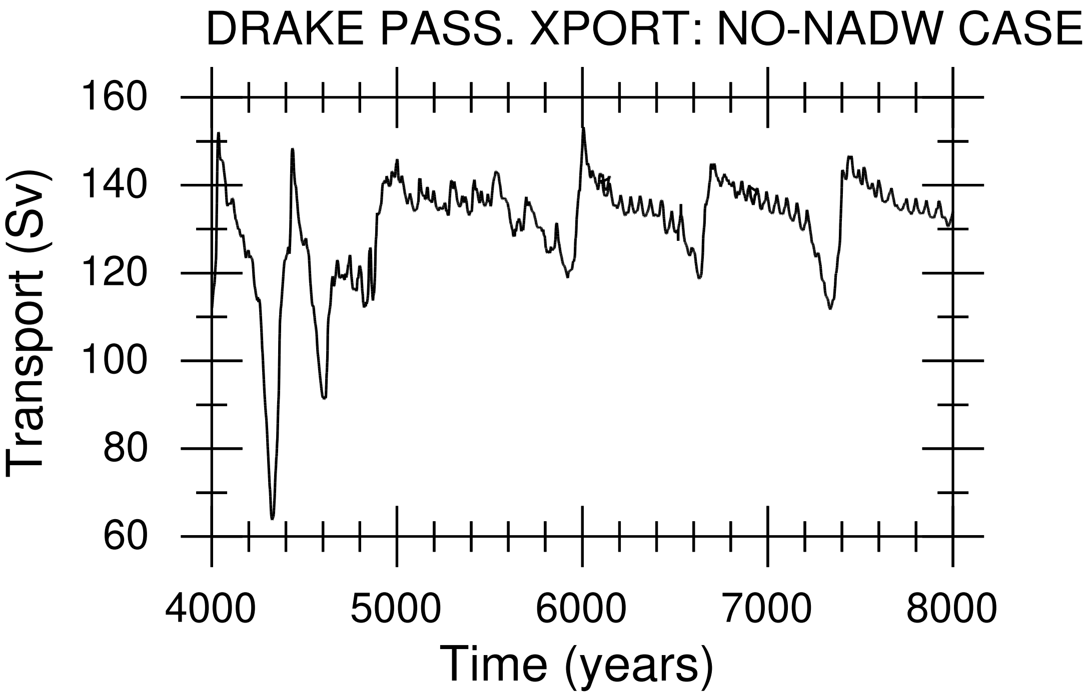
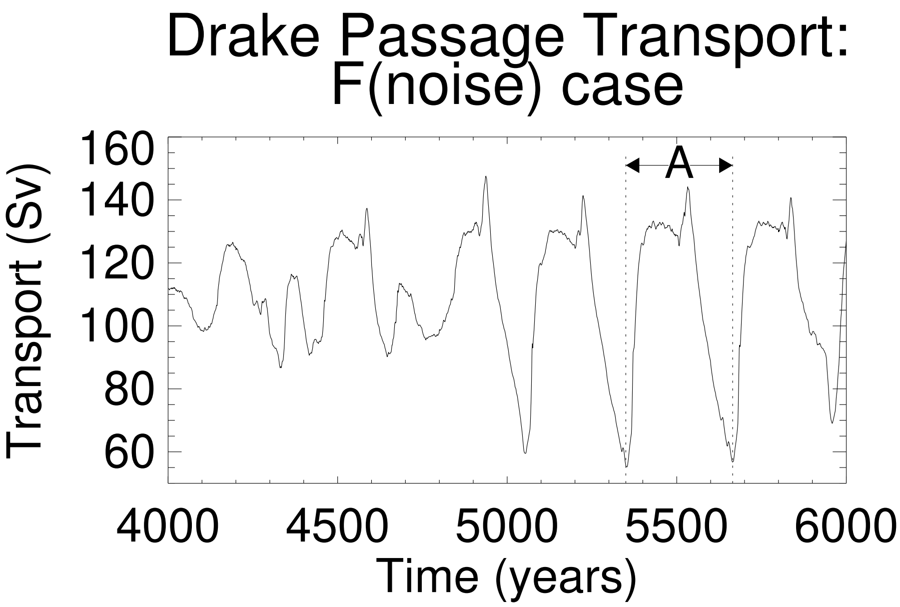

DRAKE PASS. XPORT: NO-NADW CASE
160
140
120
100
80
60
Transport (Sv)
4000
5000
6000
7000
8000
Time (years)

Figure 16:



Drake Passage Transport:
F(noise) case
A
Transport (Sv)
160
140
120
100
80
60
4000
4500
5000
5500
6000
Time (years)

Figure 17:

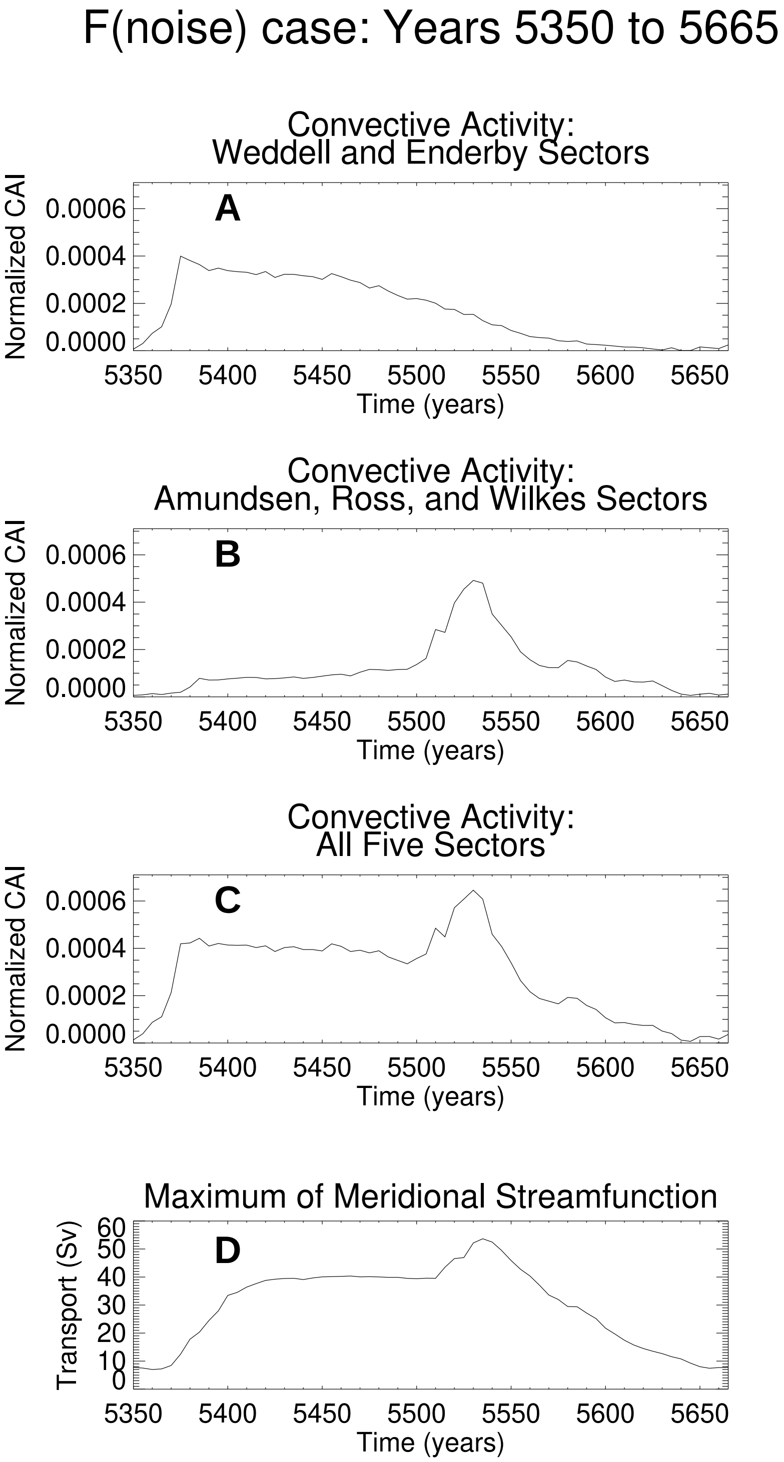
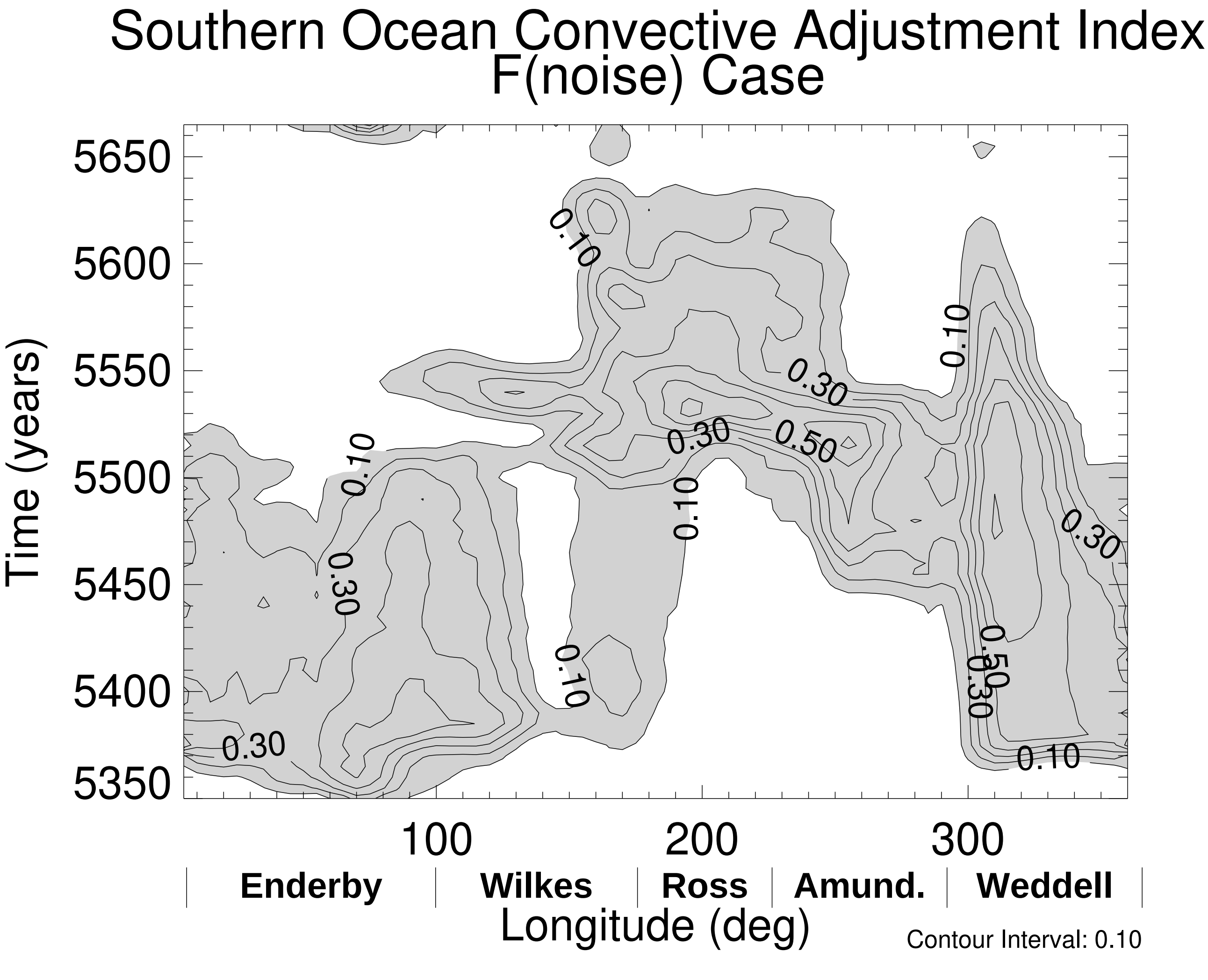
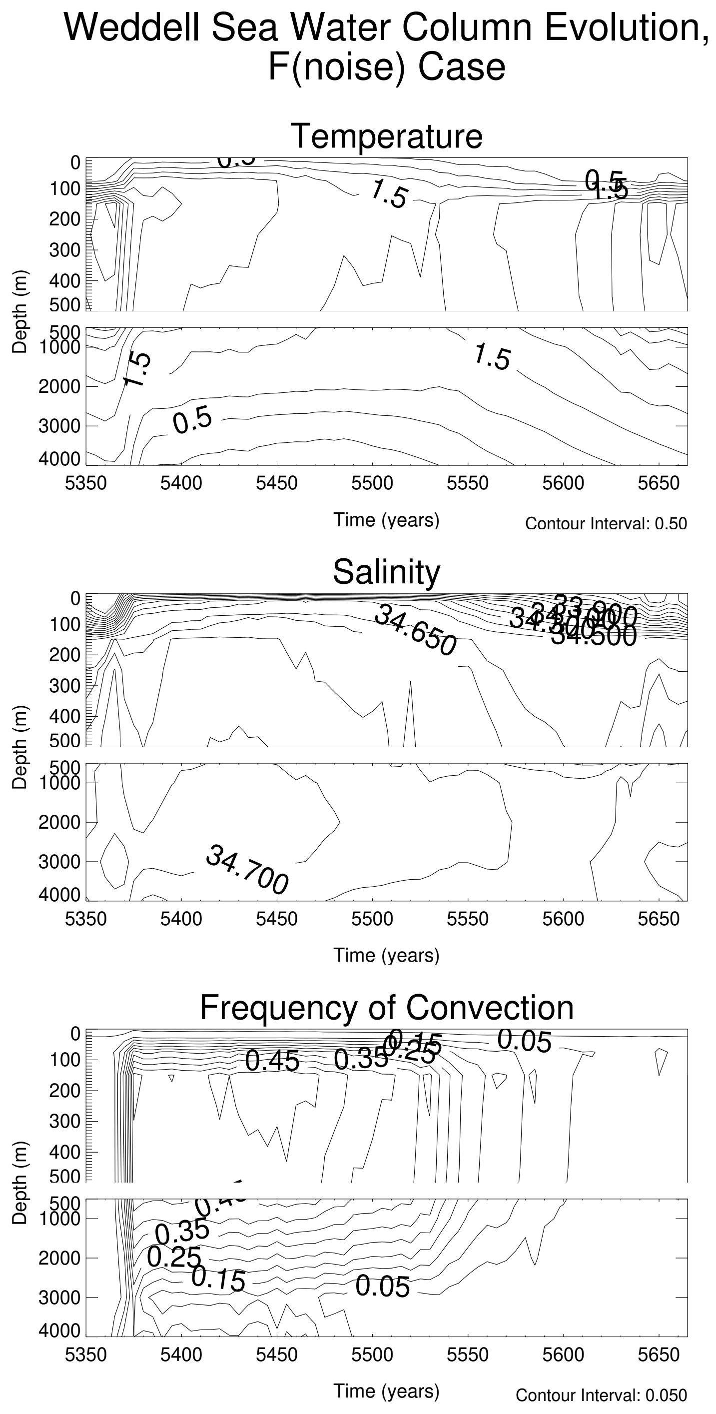

F(noise) case: Years 5350 to 5665
Convective Activity:
Weddell and Enderby Sectors
A
Normalized CAI
0.0006
0.0004
0.0002
0.0000
5350
5400
5450
5500
5550
5600
5650
Time (years)
Convective Activity:
Amundsen, Ross, and Wilkes Sectors
B
Normalized CAI
0.0006
0.0004
0.0002
0.0000
5350
5400
5450
5500
5550
5600
5650
Time (years)
Convective Activity:
All Five Sectors
C
Normalized CAI
0.0006
0.0004
0.0002
0.0000
5350
5400
5450
5500
5550
5600
5650
Time (years)
Maximum of Meridional Streamfunction
D
Transport (Sv)
60
50
40
30
20
10
0
5350
5400
5450
5500
5550
5600
5650
Time (years)

Figure 18:



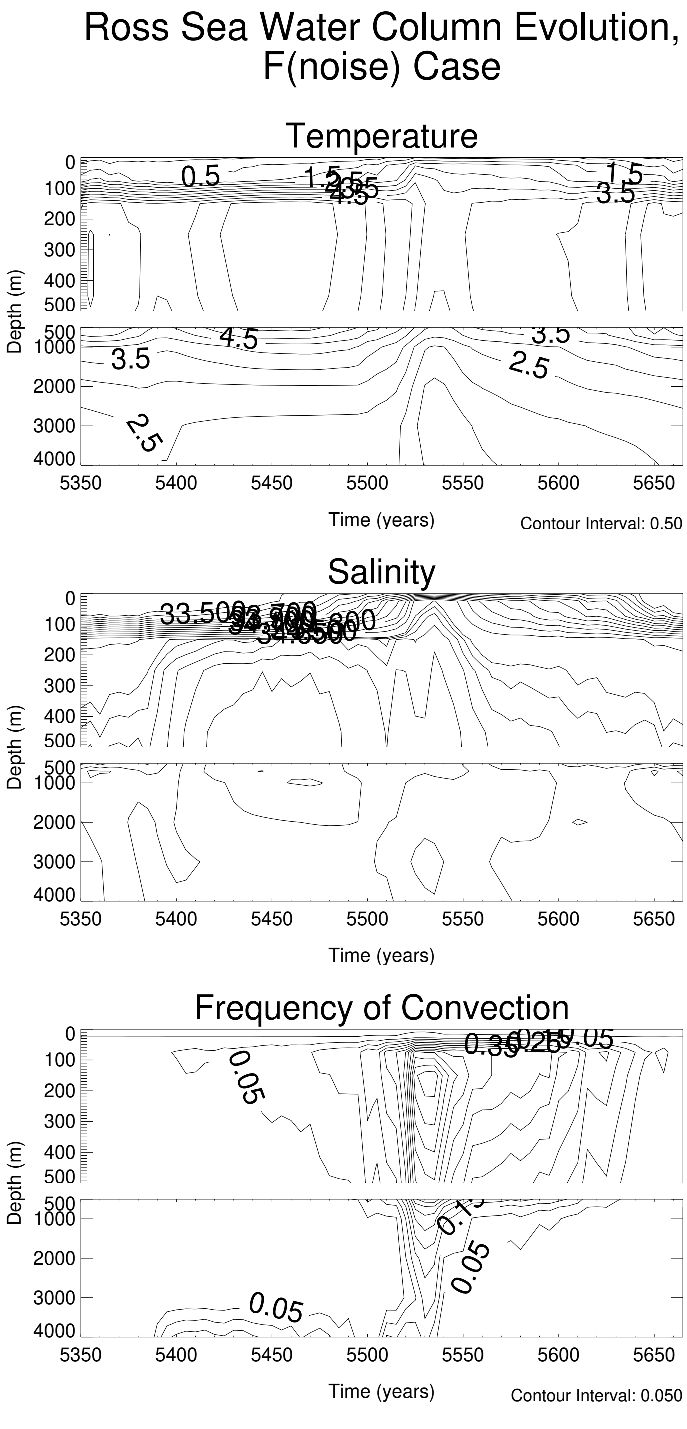
Southern Ocean Convective Adjustment Index
F(noise) Case
Time (years)
5650
5600
5550
5500
5450
5400
5350
100
200
300
Enderby
Wilkes
Ross
Amund.
Weddell
Longitude (deg)
Contour Interval: 0.10
0.10
0.30
0.50

Figure 19:



Weddell Sea Water Column Evolution,
F(noise) Case
Temperature
Depth (m)
Time (years)
Contour Interval: 0.50
1.5
0.5
Salinity
34.650
34.700
Frequency of Convection
0.45
0.35
0.25
0.15
0.05
Contour Interval: 0.050

Figure 20:



Ross Sea Water Column Evolution,
F(noise) Case
Temperature
Depth (m)
Time (years)
Contour Interval: 0.50
Salinity
Depth (m)
Time (years)
Frequency of Convection
Depth (m)
Time (years)
Contour Interval: 0.050

Figure 21:

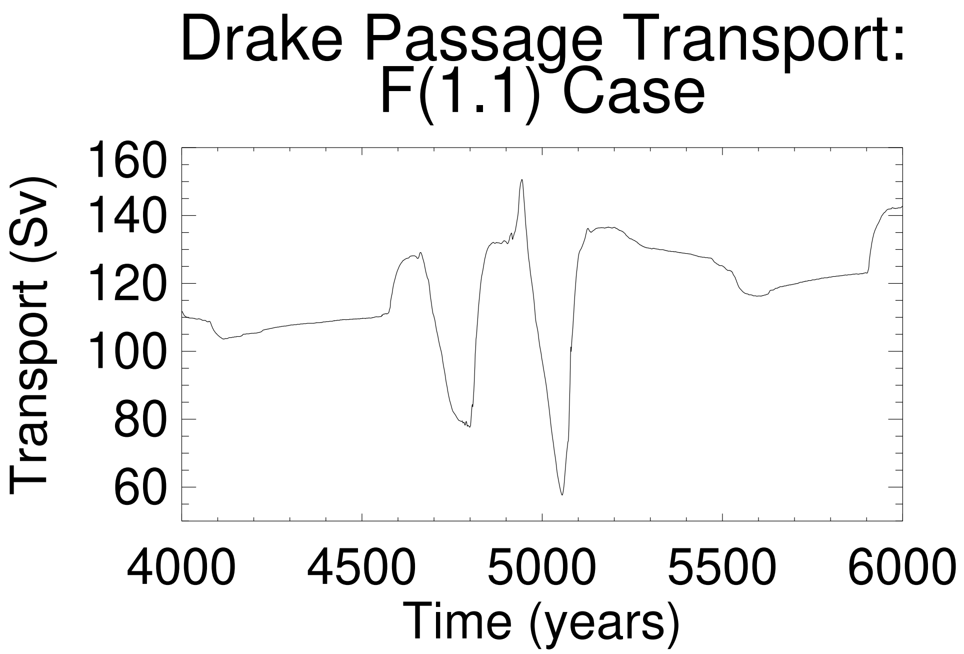

Drake Passage Transport:
F(1.1) Case
Transport (Sv)
160
140
120
100
80
60
4000
4500
5000
5500
6000
Time (years)

Figure 22:

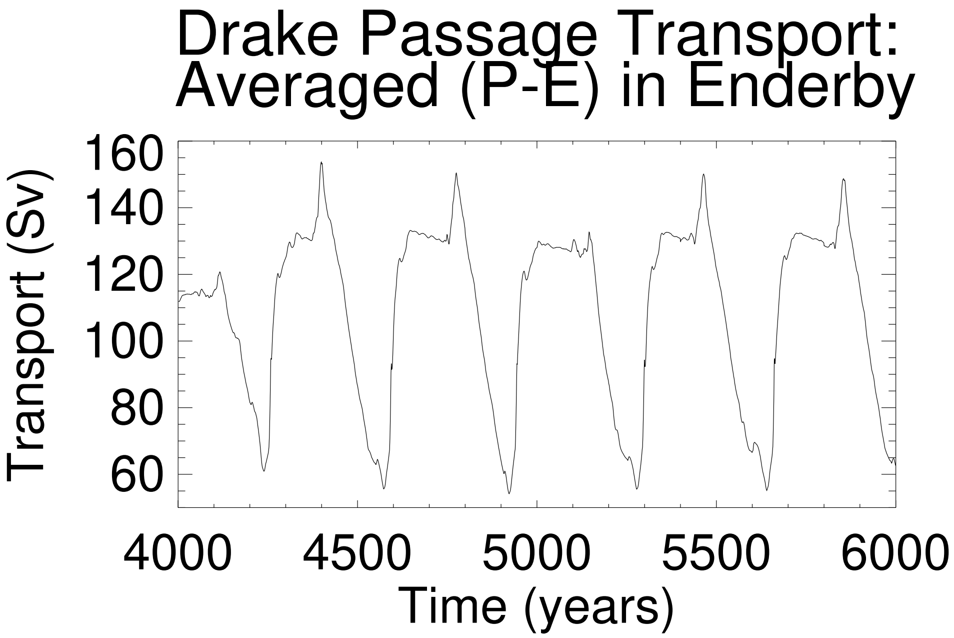

Drake Passage Transport:
Averaged (P-E) in Enderby
Transport (Sv)
160
140
120
100
80
60
4000
4500
5000
5500
6000
Time (years)

Figure 23:

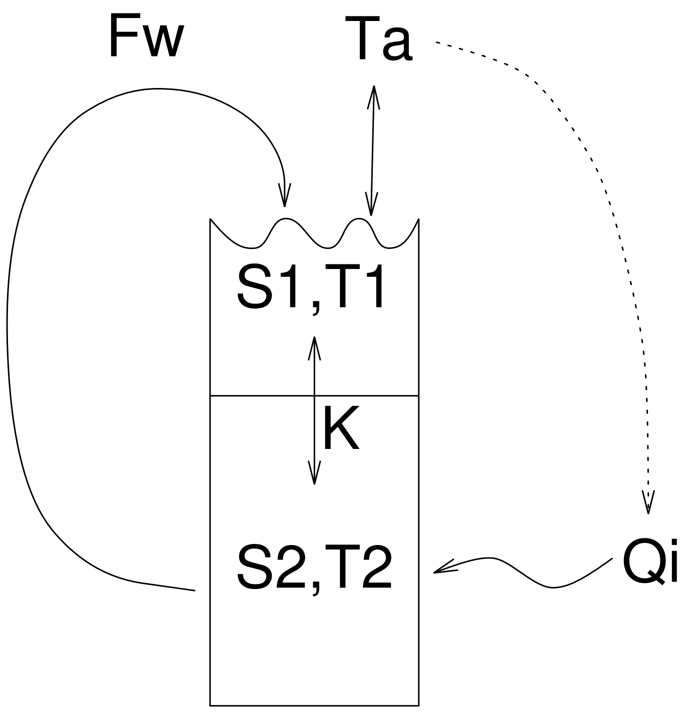
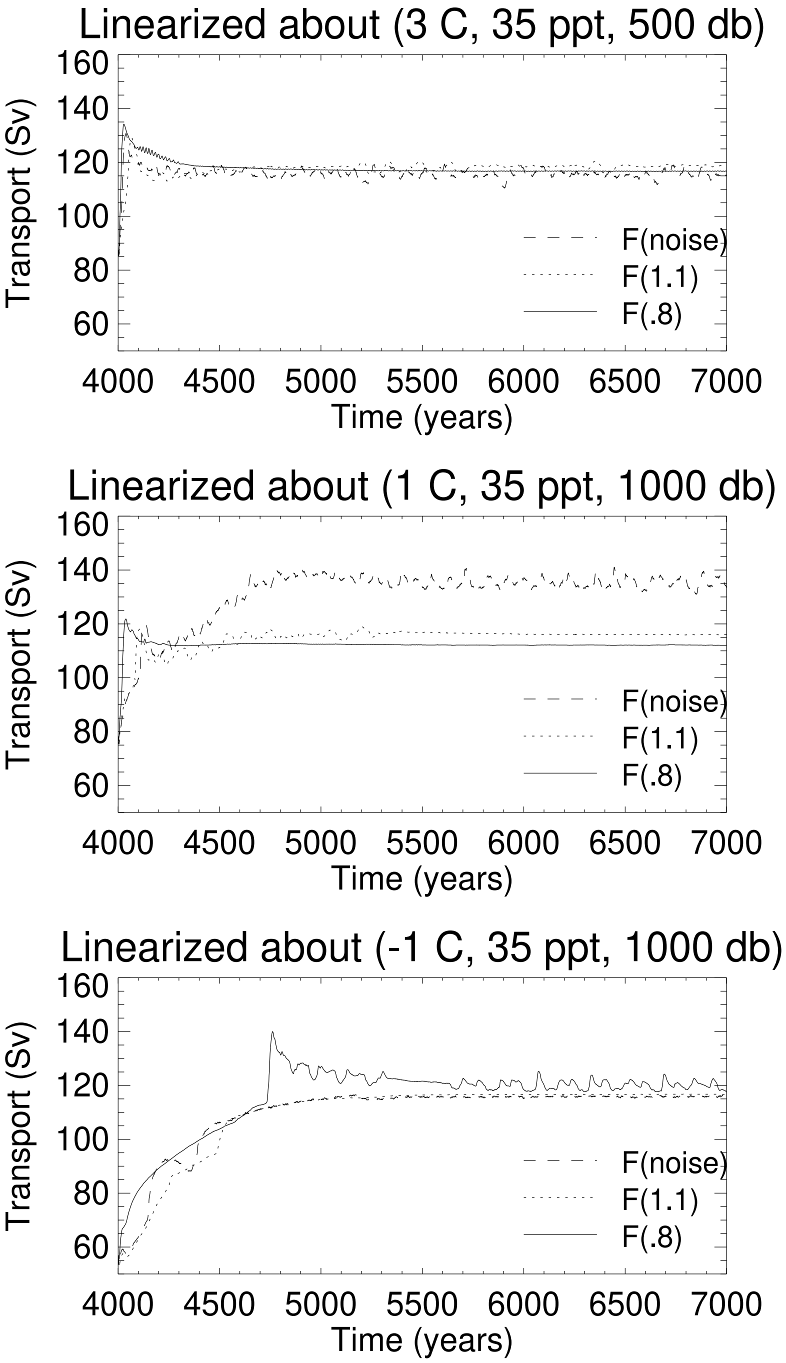


Figure 24:



Linearized about (3 C, 35 ppt, 500 db)
Transport (Sv)
Time (years)
F(noise)
F(1.1)
F(.8)
Linearized about (1 C, 35 ppt, 1000 db)
Transport (Sv)
Time (years)
F(noise)
F(1.1)
F(.8)
Linearized about (-1 C, 35 ppt, 1000 db)
Transport (Sv)
Time (years)
F(noise)
F(1.1)
F(.8)

Figure 25:

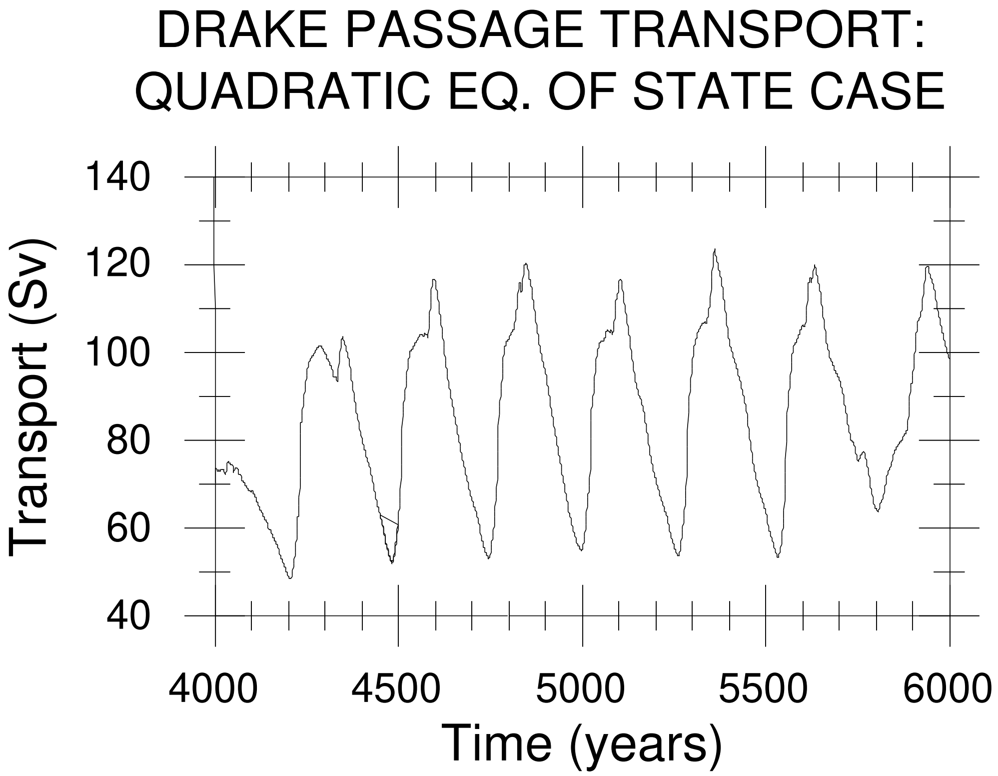

DRAKE PASSAGE TRANSPORT:
QUADRATIC EQ. OF STATE CASE
Transport (Sv)
140
120
100
80
60
40
4000
4500
5000
5500
6000
Time (years)

Figure 26:

Throw this page away

(Trenberth, 1979)

(Gordon, 1981)

(Hastenrath, 1982)

(Carissimo et al., 1985)

(Manabe et al., 1991)

(Semtner and Chervin, 1992)

(Saunders and Thompson, 1993)