Abstract

Annual mean net heat fluxes from ocean general circulation models (OGCMs) are systematically too low in the tropical Indian Ocean, compared to observations. In the models, only some of the geostrophic inflow replacing southward Ekman outflow is colder than the minimum sea surface temperature (MINSST). Observed heat fluxes imply that much more inflow is colder than MINSST. Since inflow below MINSST can only join the surface Ekman transport after diathermal warming, the OGCMs must underestimate diathermal effects.

A crude analog of the annual mean Indian Ocean heat budget was generated, using a rectangular box model with a deep “Indo–Pacific” gap at 7°–10°S in its eastern side. Wind stress was zonal and proportional to the Coriolis parameter, so Ekman transport was spatially constant and equaled Sverdrup transport. For three experiments, zonally integrated Ekman transport was steady and southward at 10 Sv (Sv = 10⁶ m³ s⁻¹). In steady state, a 10 Sv “Indonesian Throughflow” fed a northward western boundary current of 10 Sv, which turned eastward along the northern boundary at 10°N to feed the southward Ekman transport. Most diathermal mixing occurred within an intense eddy in the northwest corner. Some of the geostrophic inflow was at temperatures colder than MINSST (found at the northeast corner of the eddy); it must warm to MINSST via diathermal mixing. Northern boundary upwelling exceeded the 10-Sv Ekman transport. The excess warms as it recirculates around the eddy, apparently supplying the heat to warm inflow below MINSST. In an experiment using the “flux-corrected transport” (FCT) scheme, diathermal mixing occurred in the strongly sheared currents around the eddy. However, the Richardson number never became low enough to drive strong diathermal mixing, perhaps because (like that of other published models) the present model’s vertical resolution was too coarse. In three experiments, the dominant mixing was caused by horizontal diffusion, spurious convective overturn, and numerical mixing invoked by the FCT scheme, respectively. All three mixing mechanisms are physically suspect; such model problems (if widespread) must be resolved before the mismatch between observed and modeled heat fluxes can be addressed. However, the fact that the density profile at the western boundary must be hydrostatically stable places a lower limit on the area-integrated heat fluxes. Results from the three main experiments—and from many published OGCMs—are quite close to this lower limit.

1. Introduction

This paper and the following one (Hu and Godfrey 2007, hereafter Part II) suggest that ocean general circulation models (OGCMs) are systematically underestimating the annual mean net heat flux (AMNHF) into the northern Indian Ocean, and explore possible reasons for this.

The magnitude of interannual variation of any quantity is often roughly proportional to its mean. If the AMNHF into this region is underestimated, its interannual variability—and hence the amount of heat that
can be stored in the ocean over one summer monsoon and restored to the atmosphere in following summers—may also be underestimated. If so, some predictability to the monsoon—and indeed, possibly also of the El Nino–Southern Oscillation phenomenon—may be being missed. For example, Meehl et al. (2003) have suggested that the tropospheric quasi-biennial oscillation may relate to storage of cold water created in a strong monsoon year in the tropical Indian Ocean, to cool the SST the following summer and cause a weak monsoon. This is hard to understand, if the heat content difference is confined only to the surface mixed layer—SST anomalies within this layer should not last a year. However, if diathermal mixing in the Indian Ocean extends well below the surface mixed layer, it provides a means by which Meehl et al.’s mechanism may work.

In this paper a simple model of the Indian Ocean heat budget system is developed, which may provide a useful, if crude, analog to the real world.

a. Comparison of observed and modeled heat fluxes

In a review of the seasonal circulation of the Indian Ocean, Schott and McCreary (2001) show available estimates of the area-integrated AMNHF into the tropical Indian Ocean north of a given latitude. Figure 1 shows a new version of this figure with several extra data sources, consisting of directly observed climatologies of surface heat fluxes (Table 1a; red lines in Fig. 1) and the outputs of several OGCM experiments (Table 1b; blue lines and symbols in Fig. 1, plus purple and green lines). The tables include comments on each source and values of area-integrated AMNHF north of 5°S.
All of the climatological flux observations (red; positive fluxes upward) are larger in magnitude than all the published OGCM results, with only one exception—the thin green line lies above the lowest climatology.

The blue lines and symbols show all published OGCM results where the area-integrated AMNHF north of some latitude in the tropical Indian Ocean is given, with two exceptions. Lee and Marotzke (1998) and Zhang and Marotzke (1999) used an adjoint model to adjust their ocean circulation to match the observed Oberhuber (1988) climatology and also match observed temperature and salinity profiles. They were successful in obtaining physically realistic solutions, thereby showing that the real Indian Ocean circulation may, indeed, be compatible with the large observed heat fluxes of Fig. 1—a very valuable result. However, such optimization by “observed” data is not possible in the usual, forward time stepping mode used in coupled climate prediction models, and the purpose of this work is to identify any problems in the ocean components of such models. Therefore the area-integrated AMNHFs from these two papers are omitted from Fig. 1.

The average of the heat fluxes north of 5°S due to these OGCMs, shown in Table 1b, is 0.34 PW (1 PW = 10^{15} W)—substantially less than half that from the climatologies (Table 1a), namely 0.85 PW. The difference, of order 0.51 PW, corresponds to an area average heat flux difference of order 39 W m^{-2}. Table 1b shows that among the models used there is a wide range of horizontal and vertical resolutions and parameterizations of horizontal friction and diapycnal mixing. Further, experiments differ by the presence or absence of the Red Sea Outflow, and, in two experiments, the OGCMs were driven solely with annual mean wind stresses, thereby completely removing the well-known, reversing monsoon flow patterns. Despite these large inter-model differences the climatologies suggest that the AMNHFs from the OGCMs are rather similar, but nearly all of them seriously underestimate the observed net heat flux.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Horizontal resolution (deg)</th>
<th>Vertical resolution</th>
<th>Lateral friction type</th>
<th>Diapycnal diffusivity type</th>
<th>Red Sea cell?</th>
<th>Wind seasonality?</th>
<th>Heat transport at 5°S (PW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wacongne and Pacanowski (1996)</td>
<td>1/3 × 1/3 near Africa</td>
<td>10 m, top 100 m</td>
<td>Laplacian</td>
<td>PP, b Horizontal Laplacian</td>
<td>No</td>
<td>Yes</td>
<td>-0.24</td>
</tr>
<tr>
<td>Garternicht and Schott (1997)</td>
<td>1/4 × 1/4</td>
<td>25 m, top 125 m</td>
<td>Biharmonic</td>
<td>PP</td>
<td>No</td>
<td>Yes</td>
<td>-0.32</td>
</tr>
<tr>
<td>Loschnigg and Webster (2000)</td>
<td>1/2 × 1/2</td>
<td>2 layers Laplacian</td>
<td>Constant, Horizontal Laplacian</td>
<td>No</td>
<td>Yes</td>
<td>-0.11</td>
<td></td>
</tr>
<tr>
<td>Part II, Expt 2</td>
<td>1/5 × 1/5</td>
<td>30 levels Laplacian</td>
<td>PP only (FCT advection)</td>
<td>No</td>
<td>No</td>
<td>-0.28</td>
<td></td>
</tr>
<tr>
<td>Part II, Expt 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UW, Control</td>
<td>2 × 1/2</td>
<td>15 m, top 100 m</td>
<td>Laplacian</td>
<td>PP, Horizontal Laplacian</td>
<td>Yes</td>
<td>Yes</td>
<td>-0.38</td>
</tr>
<tr>
<td>UW, 12MRM</td>
<td>2 × 1/2</td>
<td>15 m, top 100 m</td>
<td>Laplacian</td>
<td>PP, Horizontal Laplacian</td>
<td>Yes</td>
<td>No</td>
<td>-0.32</td>
</tr>
<tr>
<td>OFES1</td>
<td>0.1 × 0.1</td>
<td>54 levels Biharmonic</td>
<td>KPP and biharmonic horizontal</td>
<td>Yes</td>
<td>Yes</td>
<td>-0.42</td>
<td></td>
</tr>
<tr>
<td>OFES2</td>
<td>As above As above</td>
<td>As above As above</td>
<td>As above</td>
<td>Yes</td>
<td>Yes, plus daily forcing</td>
<td>-0.60</td>
<td></td>
</tr>
</tbody>
</table>

* OGCMs that have used adjoint methods to bring their heat fluxes close to observed climatologies are excluded; see text.


KPP: Large et al. (1994) mixing.
In an unpublished work, referred to below as UW, we and others explored long-term mean northward heat transports in a coarse grid global OGCM. In one experiment (UW — Control) the observed seasonality of all atmospheric forcing variables was retained. In the other (UW — 12MRM), the wind stress (only) was replaced at every point by its 12-month running mean (12MRM). The area-integrated AMNHF from both experiments lay within the general group of OGCMs (thin purple lines in Fig. 1). Lee and Marotzke (1998) pointed out that, when seasonal winds are added to the annual mean winds, the time-mean current speeds and shears will increase at each point; more shear-driven diathermal mixing must result by a “rectification” effect. We show in appendix A that more diathermal mixing implies stronger AMNHF. The AMNHF from our unpublished work (thin purple lines) has the sign expected from this; that is, the more negative (stronger) result comes from the UW — Control run. We and Lee and Marotzke (1998) both found that AMNHF did not increase strongly when wind seasonality was included, but this may be because both our model and theirs had coarse grids (of order 2° × 2°).

Part II repeated the experiments of UW, with a finer-resolution (0.33° × 0.33°) model (thick purple lines in Fig. 1). Part II’s experiment with seasonal winds again had the larger (more negative) heat transport, but the area-integrated AMNHF difference north of 5°S between the experiments with and without seasonal winds with a fine grid (0.14 PW: Part II experiments 2 and 3) is more than twice that in the coarse-grid case (0.06 PW: experiments UW — Control and UW — 12MRM). Note that both UW and Part II used the Florida State University (FSU) wind stress climatology (Legler et al. 1989), so the difference in grid resolution seems the most likely cause of this difference.

Time-dependent eddies also generate current shears, which may increase mixing via a similar rectification effect. Figure 1 cannot provide definite evidence on this point. The short-dashed blue line comes from a model that permitted eddies (Garternicht and Schott 1997) due to their use of biharmonic friction. Their area-integrated AMNHF of 0.32 PW is less than that of Part II experiment 3 (more negative thick purple line: 0.42 PW), even though eddies did not develop in the Part II model. However, Garternicht and Schott’s (1997) annual mean Ekman fluxes appear to have been substantially smaller than those used by Part II. On the other hand, Jochum and Murtugudde (2005) found that eddies drive a net heat flux divergence out of a box in the western equatorial Indian Ocean, equivalent to a surface heat flux of order 80 W m⁻².

The green lines in Fig. 1 are from two OGCM for Earth Simulator experiments (OFES1 and OFES2) of the global 0.1° by 0.1° resolution OFES model (Masumoto et al. 2004). The OFES1 result (thick green line) almost coincides with one of our own results (experiment 2 in Part II; Table 1b; nearby thick purple line), with an area-integrated AMNHF north of 5°S of 0.42 PW. Both of these experiments were driven by seasonal wind stress climatology.

However, the OFES2 result (thin green line, with 0.60 PW heat transport through 5°S) is higher than OFES1 by 0.18 PW. It is the only one of all these OGCM experiments whose heat flux exceeds one of the climatologies (Hastenrath and Greischar 1993: 0.57 PW). It is also the only experiment in Table 1 in which observed daily wind forcing was used, suggesting that the annual mean net heat flux in the OFES model may be very sensitive to the storms that are smoothed out in climatological forcing. If mixing and AMNHF are increased by a rectification effect when seasonality of wind stress is added to the forcing, they may also be increased when the transient currents generated by daily winds—including storms—are added to the mean seasonal forcing. It is also worth noting that the OFES2 experiment was forced with wind stresses from the National Centers for Environmental Prediction—National Center for Atmospheric Research (NCEP–NCAR), which smoothes out storm activity. The increase in AMNHF may be still stronger, in a model forced with satellite-based wind observations.

In general, however, a trend can be seen (among finer-resolution models) to greater heat fluxes as more high-frequency wind stress activity is included in the forcing. This may well offer a route for solving the mismatch between OGCMs and observations seen in Fig. 1, but the oceanography of how the increase in mixing happens—at least in the models—remains to be explored.

b. Deductions from an “overturning streamfunction in temperature coordinates”

The consequences of the above for ocean circulation become clearer when an overturning streamfunction in temperature coordinates is defined. This has the great advantage that the annual mean heat transport is simply expressible in terms of the annual mean value of the overturning streamfunction, with no need for “eddy fluxes” (appendix A). Therefore, following Levitus (1987), the annual mean temperature transport is expressible as the annual mean Ekman transport times the “transport-weighted temperature difference” between this Ekman transport and that of the equal and opposite geostrophic flow that replaces it. If the average heat transport through 5°S from the climatologies
(0.85 PW) is divided by \((\rho C_p)\) times the annual mean Ekman transport through 5°S [13 Sv (Sv = 10^6 m³ s⁻¹)] according to FSU winds; 11 Sv from the Quick Scatterometer (QuikSCAT) climatology, a transport-weighted temperature difference of about 17°C is obtained. Since the Ekman transports exit across 5°S with an annual mean temperature of order 28°C, the equal and opposite geostrophic transport should have a transport-weighted temperature of about 28° – 17° = 11°C. By contrast, if the heat flux north of 5°S were only 0.34 PW as suggested by the average of the OGCMs, the transport-weighted temperature of the geostrophic inflow would be 20.5°C. This is only slightly lower than the observed minimum sea surface temperature (MIN SST) of 22°C found in the northern Indian Ocean (e.g., Hastenrath and Lamb 1979)—in summer off Somalia or in winter off northwest India. Interestingly, Schott et al. (1990) report a transport-weighted temperature for cross-equatorial western boundary flow (observed over 1984–86) of 15.4°C, higher than the value needed for compatibility with the mean of available heat fluxes, but well below observed MIN SST.

If a parcel of geostrophic inflow enters at temperatures sometimes found in the surface mixed layer, then no entrainment is needed. The inflow can flow adiabatically to these sites and upwell there; surface warming is all that is then needed to warm water to 28°C. However, it is shown in appendix A that on the long-term mean, the net southward volume inflow through latitude \(y\) and below a certain temperature \(T\) (which is colder than MIN SST) can only upwell by “diathermal entrainment”; that is, the water must warm (cool) via diathermal heating/cooling as it upwells (downwells). This can be due either to mixing along isopycnals, or across isopycnals but, in either case, diabatic mixing processes are involved in an essential way. High eddy diffusivities (diapycnal and/or isopycnal) are therefore needed somewhere along the pathway of any incoming water, which was initially colder than MIN SST. Therefore Fig. 1 suggests that, for some reason, present OGCMs are underestimating the strength of the mixing somewhere in the tropical Indian Ocean.

c. Outline of this paper

The present paper and its sequel illustrate the role of mixing in the tropical Indian Ocean, in OGCMs with typical present-day (coarse) vertical resolution below the upper thermocline. This paper describes experiments with an idealized rectangular-box model of the tropical Indian Ocean. The model and the three main experiments performed with it are described in section 2. In section 3 broad aspects of flow from the various experiments are described. Appendix B derives a principle involving geostrophy and the hydrostatic stability of the density profile at the western boundary, which places a lower limit (“minimum depth”) on how shallow the flow can become, for steady forcing. The expression for this limit is independent of the grid resolution, and of the parameterization of diapycnal mixing, that is used. This also sets a finite (and quite substantial) lower limit to the magnitude of the heat transport and area-integrated AMNHF, set simply by the minimum depth principle, even when (as in this paper) only steady winds are used to drive the system. The area-integrated AMNHFs from our three main experiments—and many of the OGCM results shown in Fig. 1—are quite close to the lower limit from the minimum-depth principle.

2. The idealized model

a. Basin shape, winds, architecture, and initial conditions

The idealized ocean basin (Fig. 2) is rectangular with meridional walls 50° longitude apart. All boundaries are vertical and the bottom is flat, allowing the use of linear mode theory (e.g., Gill 1982) to analyze flow behavior (until flows become fast enough for nonlinearities to occur). There is a northern boundary at a latitude \(y_N = 10°\ N\), roughly the latitude where upwelling occurs off Somalia and Arabia. There is a gap through the full depth of the eastern wall at 7°–10°S, representing the entry point of the Indonesian Throughflow. A southern wall is added at 14°S, but it has a gap in the top 50 m. Ekman transport is allowed to flow through this gap.

The zonal vectors in midbasin in Fig. 2 represent the assumed (zonally constant) pattern of wind stresses. The meridional component \(\tau^{(y)}\) of wind stress is zero, while the zonal component \(\tau^{(x)}\) is zero at the equator and is strictly proportional to the Coriolis parameter, \(f = \beta y\). Thus \(\tau^{(x)}\) is given everywhere by \(B(t)/\beta\), where \(B(t) = \partial^4/\partial y^4\) may vary with time but is spatially constant. (For ease of discussion, flows are described as if they were on a \(\beta\) plane, ignoring the other small effects of curvature in the actual model domain.) With this choice, the Ekman transport \(-\tau^{(x)}(\rho f)\) is also spatially constant at \(-B(t)/\beta\), is well-defined at the equator (e.g., Miyama et al. 2003; Godfrey et al. 2001), and is everywhere identical to the Sverdrup transport \([-\partial^4/\partial y^4]/\beta\]). No Ekman pumping occurs anywhere in the ocean interior at any time. There is coastal upwelling (or downwelling) at the northern boundary, and by construction no divergence of Ekman flow occurs at the southern boundary. (In other words, linear theory predicts that the Ekman transport flows like a uniform
carpet, into or out of the northern boundary; at the southern end, the 50-m gap allows it to flow out of or into the domain.\)

This wind distribution captures much of the variance of the observed winds, both seasonally and on annual mean, near the equator. Figure 3 shows the cross-equatorial Ekman flux \(\frac{-\sigma \alpha}{\beta} \) in the Indian Ocean, obtained by discretizing this gradient in the FSU wind stress product. Figure 3 shows that this estimate is nearly independent of the discretization distance \(\Delta y\) used in this estimate, up to \(\Delta y = 3\). Since area-integrated AMNHF was found in UW to be rather insensitive to whether or not seasonal winds were added to annual mean winds, the total Ekman flow in experiments 1–3 was assumed to be steady at the average FSU equatorial value, namely a southward flow of 10 Sv (dotted line in Fig. 3). Experiments have also been performed with the fully seasonal winds from Fig. 3. As noted earlier, they cause a fairly large increase in net heat flux compared to our experiments with steady, annual mean wind stresses. Nevertheless the focus in this paper is only on experiments with steady winds, because a minimum-depth argument, which applies in this case, seems to explain the insensitivity that we find of area-integrated AMNHF to mixing parameterization.

The model architecture was the Modular Ocean Model version 2 (MOM2; e.g., Pacanowski 1995). There were 30 grid points in the vertical, with 10-m spacing in the top 100 m; the basin was 3000 m deep. The horizontal grid resolution was 1/3° by 1/3°. A Laplacian horizontal eddy viscosity \(A_M\) was held fixed at 2000 m\(^2\) s\(^{-1}\), giving a constant Munk western boundary current width \(A_M/\beta\)\(^{1/3}\) of 46 km in all experiments—just resolvable with the model grid. There was no bottom friction. In a “rigid lid” model, the barotropic flow through open boundaries must be specified. Barotropic flow through the shallow gap in the southern boundary was set at each longitude to the instantaneous value of Ekman transport, that is, \(B(t)/\beta\). The streamfunction along the eastern boundary has the value \(B(t)\alpha/\beta\), between 10° and 14°S. It falls linearly to zero from 10° to 7°S across the (full depth) opening for the Indonesian Throughflow.

The initial temperature profile was horizontally uniform and was set at each depth to the observed area-averaged, annual mean temperature profile for the Indian Ocean north of 7°S (Levitus 1982). Everywhere south of 10°S—the southern edge of the gap in the eastern boundary—the temperature was relaxed strongly toward the initial profile, throughout each experiment. Salinity was held constant at 35 psu, so buoy-
Fig. 3. The solid line shows the mean seasonal cycle of zonally integrated Ekman–Sverdrup transport across the equator in the Indian Ocean (positive northward, in Sv). FSU winds (Legler et al. 1989) have been used. The three lines represent estimates based on taking the average of Ekman transports at 3°N and 3°S (solid line); 2°N and 2°S (dashed line); and 1°N and 1°S (dashed–dotted line). The long-term mean (straight line) of the solid line is close to −10 Sv. The steady wind stress amplitude is chosen to give this Ekman transport at all latitudes. The seasonal cycle adds an annual frequency sinusoid of about 25-Sv amplitude to this. Semiannual components are present in the individual Ekman transports, but they cancel out in the three averages above.

b. Surface heat flux boundary condition

North of 10°S, downward net surface heat fluxes $Q$ were calculated from $Q = -\lambda(SST - SST_0)$, where SST$_0$ is the undisturbed surface temperature of 28.14°C from Levitus (1982); $\lambda$ was taken as 15 W m$^{-2}$ °C$^{-1}$ in all but one experiment. This is less than the typical values of 30 W m$^{-2}$ °C$^{-1}$ recommended by Haney (1971) or Oberhuber (1988), but these were obtained making the unrealistic assumption that air temperature and humidity do not change when SST (only) changes. With more realistic assumptions that (air temperature – SST) and relative humidity are constant (e.g., Seager et al. 1988), $\lambda$ is found to be about 15 W m$^{-2}$ °C$^{-1}$.

c. Mixing parameterizations

This paper is largely about diathermal mixing (which with constant salinity is the same as diapycnal mixing), in a model whose vertical resolution has the coarseness typical of most present-day OGCMs. We found in UW that the physically based Pacanowski–Philander (PP mixing; Pacanowski and Philander 1981), although implemented, was a minor contributor to the mixing that actually occurred; most mixing was due either to horizontal eddy diffusion or to spurious convective overturn.

Pacanowski–Philander mixing was implemented in all experiments of this paper, plus another mixing mechanism. The latter dominated in all cases. In experiment 1, horizontal diffusion was added, which generates cross-isopycnal mixing in the presence of sloping isopycnals. It is known to create unphysical problems (e.g., McDougall and Church 1986) but was used by Wacongne and Pacanowski (1996) and in our unpublished work. To try to remove these problems, this diffusivity was set to zero in experiment 2. This generated an extreme case of the situation found in our unpublished work, namely that most of the mixing was generated by spurious convective overturn below about 100 m. The resulting circulation was physically more realistic in the sense that flow was more closely adiabatic; but solutions were numerically very noisy, and there was no means in the model code to make the essential estimates of the diapycnal heat flux. Finally in experiment 3 horizontal eddy diffusivity was again chosen to be zero, but a different numerical scheme was chosen for the advection of tracers—namely the “flux-corrected transport” (FCT) scheme (Griffies et al. 2000, and references therein). This ingenious scheme is specifically designed to deal more effectively with situations in which “two-gridpoint noise” becomes so severe that convective overturn can result. The “centered difference” advection scheme usually used in MOM models—while it avoids numerical diffusion completely—can also generate two-gridpoint noise, whenever the Peclet number $\Psi = w\Delta z/\kappa$ exceeds 2. The FCT scheme permits some two-gridpoint noise but, if convective overturn is imminent, the centered difference advective scheme is replaced with the (numerically diffusive) “upwind” scheme. This avoids gridpoint noise from becoming too large. A further major advantage of this scheme is that it returns a value of eddy diffusivity when in the diffusive upwind mode, so the diapycnal fluxes of buoyancy and heat it generates can be quantified. However, it is important to note that the FCT scheme does generate numerical diffusion; so, if FCT mixing dominates in this experiment, the mixing is again physically suspect.

The three main experiments just mentioned covered four years of model time. Equilibrium was nearly reached by the end of these experiments. Two others were run for two years only. Details of these experiments appear in Table 2a.

3. Results

a. Depth-integrated and overturning streamfunctions

According to a linear model of the configuration of Fig. 2, the first response to onset of steady winds should
be barotropic; that is, the currents should initially be basically depth independent beneath the surface Ekman–Sverdrup flow. Examination of flow patterns 7 days into each experiment qualitatively confirmed this picture (not shown). Flow patterns should then shoal, as more and more of the baroclinic modes reach Sverdrup equilibrium—until the fixed western boundary transport is shallow and intense enough for nonlinear effects to become important.

Figures 4a–c show the horizontal streamfunction after four years, at the end of experiments 1–3, respectively. In all three figures a zonal jet at 7°–10°S (with a southward Ekman transport superimposed), a northward western boundary current, and an oppositely directed zonal jet at 8°–10°N are apparent, as expected near steady state from Sverdrup balance. The intense eddy near the northwest corner is stronger in experiments 2 and 3 than in experiment 1, due presumably to the removal of horizontal eddy diffusion. Linear theory suggests that the interior flow should consist solely of Ekman flow, moving uniformly southward from the northern boundary like a “carpet.” This is not correct; zonal flow is quite evident between about 5°S and 5°N in the midocean in Figs. 4a–c. To investigate the reasons for this, the depth-integrated nonlinear terms in the momentum equations near the end of each experiment were treated [a posteriori, as described by Godfrey (1973) and Kessler et al. (2003)] as if they were known “nonlinear stresses”; their curls were added to the (spatially constant) wind stress curl; and the resulting modified Sverdrup relation integrated westward from the eastern boundary to calculate an approximate streamfunction. The patterns of Figs. 4a–c look like smoothed versions of this Sverdrup calculation (not shown). The smoothing is expected since the Sverdrup calculation omits lateral friction. These nonlinear stresses are important to generation of the recirculating eddy.

The vertical distribution of the overturning streamfunction (in depth coordinates), at the end of four

![Figure 4](image-url)
years, is seen in Figs. 5a–c for experiments 1–3. The inflow feeding the Ekman transport has a similar depth distribution across the equator in the three experiments, despite removal of horizontal eddy diffusivity in experiments 2–3. In the latter two experiments, the inflow shoals and becomes more intense as it flows northward. Even in these experiments, however, the flow still extends to a depth of order 200 m at 8°N—much deeper than the surface mixed layer. Figure 6 shows the overturning streamfunction through 8°N as a function of depth and time for experiments 1–3. It can be seen that the shallow but finite depth of the inflow is already in place after 4 months, after which only minimal changes occur in the overturning streamfunction at this latitude.

b. Overturning streamfunction in temperature coordinates

Since area-mean temperatures decrease with depth in much the same way in experiments 1–3, one might suppose that the deeper velocity profile of experiment 1 would result in more heat transport than experiment 3 (Figs. 5a,c). However, the quantity that determines the surface heat flux into the northern Indian Ocean is not the distribution of the western boundary inflow by depth, but by temperature—and, at least above 200 m, temperatures at 8°N are markedly warmer near the western boundary (where the northward flow occurs) in experiment 1 than experiment 3. The overturning streamfunction in temperature coordinates is the appropriate variable for examining the distribution of meridional flow by temperature; it is shown at the end of four years for the three experiments (Fig. 7).

In all three experiments, and especially in experiments 2 and 3, there is very little flow below the 10°C isotherm. The inflow is fairly adiabatic (i.e., it nearly flows along lines of constant temperature in Fig. 7) until it reaches the “surface mixed layer”—well defined by the intense Ekman flow of surface water seen to arc upward (in temperature space) and southward at the top of each panel in Fig. 7. The big differences between streamfunctions in depth coordinates seen in experiments 1, 2, and 3 in Fig. 5 have surprisingly little effect on the overturning streamfunction in temperature coordinates. This result has strong consequences for the area-integrated AMNHF into the ocean, in the three experiments.
c. Heat budget

As shown in appendix A, the northward heat transport $VT$ through a given latitude $y$ is given by

$$VT = \rho C_p \int_{T_{\min}}^{T_{\max}} (T \phi T) \, dT,$$

where $\phi$ is the overturning streamfunction in temperature coordinates. In a near-steady-state situation, the heat transport through $y$ should nearly equal the area-integrated AMNHF north of $y$. Table 2b shows that this is well-satisfied at the end of the 4-yr experiments, at the equator. Equation (1), and the similarity (south of the equator) of the overturning streamfunctions in temperature coordinates of Fig. 7, suggests that area-integrated AMNHFs from $0^\circ$ to $5^\circ$S should be much the same in each experiment. Table 2b confirms this result. However, it should be noted that there are quite large differences in the spatial pattern of surface heat fluxes between the three experiments (Fig. 8). This similarity of area-integrated AMNHFs, despite the fact that the active eddy diffusion mechanism in each experiment is very different (and is physically unrealistic in each case), may be analogous to the similarity of OGCM fluxes in Fig. 1. Indeed, the actual value of the area-integrated AMNHF in Table 1 of about 0.25 PW is surprisingly close to the value from the two OGCMs in Fig. 1, which were driven by annual mean winds, namely 0.32 PW (UW – 12MRM) and 0.28 PW (Part II, experiment 1).

d. The nonlinear minimum-depth hypothesis

Why are area-integrated AMNHFs from the present three experiments so similar despite the large differences between the diffusion mechanisms that generated them? A possible solution is found in the concept that, in all three experiments, the flow is nearly confined above a minimum depth defined by a nonlinear constraint based on geostrophy and hydrostatic stability.

The basic idea behind this constraint is that the northward geostrophic transport, being equal and opposite to the Ekman transport $V_{Ek}$, is of known magnitude. The geostrophic transport is the vertical integral of the zonally integrated flow at each depth $z$; the Coriolis force on this transport is balanced by the cross-basin pressure difference at $z$. But this pressure difference can be expressed, through the hydrostatic relation,
as the vertical integral of the density difference across the basin, from an assumed “depth of no motion” $Z$ up to $Z$. Therefore the known transport is a double vertical integral of cross-basin density difference. If this geostrophic transport is to be confined above a shallow depth—that is, if $Z$ is to be small—the cross-basin density difference must be large; but for a given density profile at the eastern boundary (purple line in Fig. 9a), the largest this density difference can get is for the potential density profile at the western boundary to be like the black line in Fig. 9a. The black and purple lines are identical below depth $Z$; the black line is constant above $Z$. Any larger difference in potential density between eastern and western boundaries will result in hydrostatic instability at the western boundary. The mathematics of these ideas are explained in greater detail in appendix B, where it is shown that the rule $Z = (f^2 \Delta N^2)^{1/2}$ provides a good estimate for the minimum depth $Z$, where $f$ is the zonally integrated Ekman transport; $f$ is the Coriolis parameter; and $N^2$ is the squared Brunt–Väisälä frequency at the eastern boundary, averaged over the depth $Z$. For the purple density profile in Fig. 9a, a geostrophic transport $V$ of 10 Sv at latitude 8°N, where flow separates from the western boundary (Fig. 4), this rule yields $Z = 205$ m (a more exact figure is 212 m). This minimum depth is seen to agree well with the depth of the inflow, especially in Fig. 6c, where final net flow is less than 1 Sv everywhere below 200 m. The potential density at depth $Z$ is that of the neutrally buoyant layer (black line)—that is, 26.41, with a potential temperature of 12.9°C. Inspection of Figs. 7b,c shows that 12.9°C is close to the lowest temperature for which significant flow occurs at 8°N, confirming that the flow in these two experiments conforms well with the minimum-depth principle.

This principle almost permits a minimum possible area-integrated AMNHF north of 5°S to be estimated, if two further assumptions are made: namely that inflow is adiabatic, that is, north-flowing streamfunction contours follow isotherms in diagrams like Fig. 7; and (as is roughly true in Fig. 7b,c) $\partial \psi/\partial T$ is constant above the minimum temperature of $T_{min} = 12.9°C$, up to just below the Ekman layer at SST$_0$ = 28°C. The heat transport is then estimated as $\rho C_p V_{Ek} (SST_0 - T_{min})/2 = 0.31$ PW. However, this is an overestimate because the Ekman transport through a given latitude has a range of temperatures overlapping those of the warmest geostrophic inflow. The maximum overturning streamfunction in Fig. 7 is about 7.5 Sv in all three experiments due to this overlap. Replacing $V_{Ek}$ by 7.5 Sv instead of 10 Sv gives a minimum area-integrated AMNHF estimate of 0.23 PW, slightly smaller than those from the three experiments (Table 2b). Note that no assumption of a mixing mechanism is needed in making this minimum area-integrated AMNHF estimate.

More detailed tests of the minimum-depth principle are given in appendix B, including discussion of the profiles of potential density at the western boundary seen in Fig. 9a, and the profiles of zonally integrated geostrophic flow calculated from them, seen in Fig. 9b. The essential points are that, despite the fact that the western boundary density profiles from the three experiments are very different, the resulting zonally integrated flow is quite similar and that the results from experiment 3 are markedly closer to those predicted by the minimum-depth principle than those from the other two experiments.

e. Notes on diapycnal mixing beneath the mixed layer, in experiment 3

Figures 10b and 10c show the temperature along the western and northern boundaries, respectively, at the end of experiment 3. Looking northward in Fig. 10b, isotherms between 11.5° and 12.5°C rise steeply from near 200 m, then fall almost “vertically,” indicating that temperature (density) is almost independent of depth. In other words, deep subsurface mixed layers are to be found from 4° to 8°N along the western boundary. A similar deep internal mixed layer is seen at 53°E along the northern boundary (Fig. 11c), with less dramatic examples farther east. The stratification is much stronger above the 14°C, so a coarser contour interval was used for $T > 14°C$.

The black contours and gray shades in Fig. 10a show the depth of the 13°C isotherm at the end of experiment 3 in the northwest corner; the steep isotherm slopes are associated with the anticyclonic eddy seen in Fig. 4c. The 13°C isotherm is well below the MIN SST of 14.71°C (Table 2b), so diathermal mixing is essential for any flow below this level to join the surface Ekman layer. The red contours show the diathermal heat flux $\rho C_p \alpha_T \partial T/\partial z$ through the isopycnal; $\alpha_T$ is the eddy diffusivity due to FCT numerical diffusion. The (vigorous) two-gridpoint noise in this flux has been smoothed by averaging adjacent pairs of points in both x and y directions. Strong downward heat flux is seen just offshore from the western boundary near 7.5°N, 50.5°E. This flux patch reaches a maximum of nearly 400 W m$^{-2}$. A weaker patch of heat flux—this time upward—is seen near 9.0°N, 52°–54°E. A third patch—downward heat flux again—is seen near 8°N, 55°E. Note that these patches all occur near the maximum slope $S$ of the isopycnal, far enough from the western boundary for flow along the isopycnal to be in geostrophic balance. Under these circumstances, the Richardson number $Ri$ is given by $f^2 (N^2 S^2)$ (e.g., Godfrey et al. 1980, 1986):
that is, we might expect low Ri. However, as illustrated in Figs. 10b,c, the stratification has a very strong discontinuity near the 13°C isotherm. If a value of $N^2$ were chosen from just above 13°C, the Richardson number is below 0.25, and very strong PP mixing would be expected. If a value of $N^2$ were chosen from just below 13°C, Ri is well above 0.25. In our model, the numerics “choose” a small value of $N^2$, so Ri is large enough that
PP mixing (not shown) is small by comparison with FCT mixing. Thus one reason why the model chooses FCT rather than PP mixing may be that the vertical resolution is inadequate to resolve the near-discontinuity in $N^2(z)$. Another possible numerical problem is that $Ri$ is calculated in MOM using data from nine adjacent temperature points and four velocity points, that is, over a square region 66 km wide on our grid. This is larger than the Munk width, so horizontal resolution may also be a problem.

The black curve in Fig. 11 shows the area integral of estimates of entrainment, obtained via appendix A as $\int [\kappa_{FCT} \frac{\partial T}{\partial z}] \frac{\partial z}{\partial T}$, where $\rho C_p \kappa_{FCT} \frac{\partial T}{\partial z}$ is the heat flux plotted in the red contours in Fig. 10a. The integral is over the area north of the equator, through isotherms less than the MINSST. The red curve shows the same from PP mixing; both are plotted as a function of $T$ at the end of experiment 3. (Here $\kappa_{FCT}$ is the eddy diffusivity due to upwind advection when this is invoked by the FCT scheme.) The plot is restricted to temperatures less than the MINSST so that heating is only by entrainment processes. The sum of temperature fluxes due to the PP and FCT schemes is shown by the green line. The overturning streamfunction in temperature at the equator is also shown (blue line) from the coldest temperatures up to $T$. As shown in appendix A, this should be the total mass flux through the $T$ isotherm, so it should be the same as the sum of the first two curves in Fig. 11. Evidently, it is not. The numerical problems mentioned earlier may be responsible for this mismatch.

f. The Great Whirl as a mixing organizer

Figure 10 suggests that most of the FCT mixing, which warms water below the MINSST at the expense of the heat content of shallower water, occurs within

---

Fig. 10. (a) Gray shading and solid contours show the depth of the 13°C isotherm in meters, at the end of experiment 3. Estimated horizontal distribution of downward diapycnal heat flux due to FCT numerical diffusion at the end of experiment 3, through the 13°C isotherm, is shown in solid red contours; upward is shown in dashed red contours. Contour is 50 W m$^{-2}$. (b),(c) Temperature section along the (b) western and (c) northern boundaries at the end of experiment 3, above 500 m (depth scale in cm). Contours are every 0.5°C, up to 14°C (black), and are every 2°C, above 14°C (red). Note that vertical temperature gradients show strong discontinuities near 13°C.
the energetic “Great Whirl”–like eddy in the northwest corner. This was also true of the two other experiments. On the other hand, Fig. 4 shows that, in all experiments, the maximum value of the horizontal streamfunction in the eddy in the northwest corner—which may be regarded as the model version of the observed Great Whirl (e.g., Schott and McCreary 2001)—is greater than the imposed Ekman flow of 10 Sv. The difference between them, namely 4 Sv in experiment 1, 8 Sv in experiment 2, and 6 Sv in experiment 3, recirculates around the Great Whirl, joining with the northward western boundary flow into this region. In particular, Fig. 4 shows that in every experiment some flow on the outermost edges of the eddy flows southward across the equator in the recirculation before joining the main (northward) western boundary current. These waters are thus warmed substantially by the surface heat flux before returning to the inside edge of the Great Whirl.

The heat to warm inflow below MINSST may be generated as follows: Some of this cold water upwells to the surface, near the northeast corner of the eddy where the MINSST is found (e.g., Fig. 8c, recalling that heat flux is proportional to SST). For a short time it is heated intensely by the strong surface heat flux there, then sinks below the surface mixed layer to recirculate around the Great Whirl. On return to the northwest corner, mixing of this warmed water with some new, cold inflow allows the resulting mixture to exceed MINSST; this will then join the Ekman transport and eventually exit through 14°S.

g. Processes setting the temperature distribution of the Indonesian Throughflow

Figure 6 shows that flow toward the northwest corner shoals over some months until it approaches the minimum depth, at which time entrainment processes must become active in this region. This can explain the steadiness of the inflow toward the northwest corner after the first few months: but, a long way from the site of the entrainment, the Indonesian Throughflow also settles toward an equilibrium state in which the distribution of Indonesian inflow by temperature is much the same as through 8°N. We do not show this explicitly, but it is suggested by Fig. 7, which shows that the distribution of the western boundary current by temperature is much the same at all latitudes south of 8°N. This suggests that—just as nonlinear stress curls can have distant effects on circulation via an “extended Sverdrup relation” (e.g., Kessler et al. 2003)—the nonlinear entrainment must somehow generate signals that “radi-
ate” from the northwest corner to affect the Indonesian Throughflow.

Gill (1982) details the theory of waves driven by a sum of wind-driven upwelling and precipitation minus evaporation ($P - E$). Here $P - E$ is a kind of “entrainment”; as seen in Fig. 11, entrainment across density levels amounts to over 1 Sv in our model. The strongest diapycnal fluxes (and associated entrainment) in Fig. 10a occur on the north and west sides of the eddy, that is, squarely within the coastal Kelvin wave pathway. It is reasonable to think that internal Kelvin and Rossby waves will be generated by nonlinear entrainment, just as they are by wind-driven upwelling or $P - E$, and will propagate through to the north end of the Indonesian gap in the familiar way, thereby changing pressure distributions there. This will alter the distribution of the Indonesian Throughflow by temperature class. Thus—at least in the present model geometry—entrainment processes within the northern Indian Ocean may affect the vertical profile of the Indonesian Throughflow (rather than the other way around). However, a detailed test of this idea is outside the scope of this paper; and, in the real ocean, this Indian Ocean effect on the throughflow is only one among many such influences (e.g., Godfrey 1989, 1996; Spall 2003).

4. Summary

Figure 1, the theorems on heat fluxes and transports given in appendix A, and conclusions from an unpublished OGCM study of Indian Ocean heat fluxes all suggested that present-day OGCMs are somehow underestimating the diathermal mixing that must occur somewhere within the tropical Indian Ocean. This motivated us to devise an idealized model of the tropical Indian Ocean. This simple enough that much of what happens within it is readily understood analytically, and many problems that complicate heat budget studies in more “realistic” models are avoided. Analysis of diathermal mixing (a strongly nonlinear quantity) is very difficult in a model of the real Indian Ocean, with its huge seasonal variation of winds and currents. These are avoided by omitting the large observed seasonal cycle in the winds (Fig. 3). Problems associated with the “staircase” profile of the African coastline are also avoided in our simple model. Salinity was held constant at 35 psu so that “diathermal” and “diapycnal” mixing were identical. The resulting model showed that

- Area-integrated AMNHF had similar values of heat transport in all three models, which were quite close to the values from the OGCMs in Fig. 1 from experiments that used only annual mean wind stresses. In the present model, this area-integrated AMNHF/heat transport appears to be due to a nonlinear effect associated with the finite amplitude of the cross-equatorial Ekman flow ($|V_{Ek}| = 10$ Sv). As explained in the text, the flow must be confined above a depth $Z$ whose minimum value is close to $(6 f V / N^2)^{1/3}$, or about 200 m at 8°N. The resulting estimate of the area-integrated AMNHF as $\rho C_p (6 f (8°N) V_{Ek}^2 / N^2)^{1/3} \partial T_E / \partial z$, where $T_E (z)$ is the vertical temperature stratification at 8°N, is 0.23 PW—which is similar to the values obtained in our three main experiments (Table 2b), and in a number of models in Fig. 1 (Table 1b), especially the two (UW – 12MRM and Part II’s HG2) driven only with annual mean wind stresses.

- Some inflow did occur below MINSST. In all three experiments, but especially in experiment 3, the diathermal mixing that warmed this inflow up to MINSST was confined to the intense flow around the eddy in the northwest corner (Fig. 10a and text). The mixing in experiment 3 is due to FCT mixing, but another experiment with higher resolution may well generate similar results with PP mixing.

- Total upwelling east of the MINSST point, found just northeast of the eddy at 10°N, 54°E (Fig. 8c) exceeds the wind-driven Ekman transport. The excess upwelled water warms as it recirculates around the eddy. This may compensate for the heat loss of water above MINSST due to mixing with inflow below MINSST. In effect, the intense eddy allows inflow below MINSST to recirculate more than once, warming as it goes, to finally escape through 14°S in the Ekman transport.

- The time evolution of the model “Indonesian Throughflow” suggested that the entrainment generated by this mixing must influence the rest of the model domain through Kelvin and Rossby waves, just as wind-driven upwelling does.

Observational studies of mixing in the Gulf Stream have so far failed to find strong suggestions of mixing in the lower thermocline water, where it is needed to resolve the problem of Fig. 1 (J. Toole 2005, personal communication). However, this may not be true of all western boundary currents. The present model provides a counterexample, possibly relevant to the real Somali Current, in which substantial mixing and entrainment occurs even when it is driven only by annual mean winds. As discussed in the introduction, allowance for seasonal and wind stress variations, and inclusion of eddy variability via use of biharmonic friction, is likely to increase the mixing/heat transport/area-integrated AMNHF by a “rectification” mechanism (Lee
and Marotzke 1998); evidence from Fig. 1 suggests that these effects may be substantial. Of particular interest is the fact that there is a maximum in wind-driven energy input at near-inertial frequencies in summer, directly over the Somali Current (Alford 2003). This may break geostrophy, forcing some of the immense store of large-scale available kinetic energy in the Somali Current to cascade downscale and generate small-scale mixing. The existence of this maximum may explain the sensitivity of the OFES results to the inclusion of daily winds.

Part II of this study explores similar issues in a model with realistic coastlines and wind stresses. The “minimum depth” principle is generalized to accommodate this more realistic case, and we explore the effects of adding seasonal wind stress variations to the same annual mean winds.

Acknowledgments. This paper has been 10 years in preparation. A complete acknowledgement to all who have helped is impossible. Many comments by anonymous reviewers and others on previous versions of this paper have resulted in several massive, and beneficial, overhauls. This study was supported by a program from the National Natural Science Foundation of China (NSFC) under Grant 40576006.

APPENDIX A

Theorems on Overturning Streamfunctions in Temperature Coordinates

Following Nurser et al. (1999), Marsh et al. (2000), and Schiller (2003), the meridional flow through a zonal section at latitude \( y \) and time \( t \), above the isotherm for potential temperature \( T \), is defined to be the “overturning streamfunction in temperature coordinates”: \( \Psi(y, T, t) \). Provided that the potential temperature is a single valued function of depth \( z \) (not always true in, e.g., the Bay of Bengal), the flow between temperatures \( T \) and \( T + dT \) is \( (\partial \psi / \partial T) dT \), and this expression gives the magnitude of all meridional flow within that temperature range. If the section spans a closed basin, such as the Indian Ocean north of 7°S, then the net mass transport through the section is zero, and the instantaneous heat transport through the section is well defined at

\[
\int_{T_{\min}}^{T_{\max}} \rho C_p \left[ T \partial \Psi(y, T, t)/\partial T \right] dT.
\]

In the section \( T_{\max} \) and \( T_{\min} \) are the highest and lowest temperatures encountered. This is a useful formulation in the northern Indian Ocean for two main reasons.

First, through heat conservation, the long-term mean net heat flux \( \overline{Q} \) north of latitude \( y \) equals the long-term mean heat transport through \( y \), namely,

\[
\int_{T_{\min}}^{T_{\max}} \rho C_p \left( T \overline{\Psi}/\partial T \right) dT,
\]

where an overbar indicates a time average. Because the temperature \( T \) is the variable of integration, the splitting of the heat transport into a part driven by the time-mean circulation, and a second part due to eddy heat fluxes is avoided. In a closed basin \( \overline{\Psi} \) consists of a sum of two parts: an annual mean Ekman transport of \( E \), determined solely by the applied wind stresses, and an equal and opposite geostrophic transport. The time-mean Ekman and geostrophic flows above the isotherm \( T \) can be defined separately as \( \overline{\Psi}_{\text{Ek}}(y, T) \) and \( \overline{\Psi}_{\text{geostrophic}}(y, T) \):

\[
\overline{Q} = \rho C_p \overline{E}(T_{\text{Ek}} - T_{\text{geo}}),
\]

where

\[
T_{\text{Ek}} = \int_{T_{\min}}^{T_{\max}} \rho C_p \left( T \overline{\Psi}_{\text{Ek}}/\partial T \right) dT/E
\]

and a similar expression for \( T_{\text{geo}} \) are “transport-weighted temperatures.” For a given \( E \), \( \overline{Q} \) can only be increased in magnitude by increasing the time-averaged temperature difference \( (T_{\text{Ek}} - T_{\text{geo}}) \) between the Ekman outflow and the colder, geostrophic inflow.

The second reason that the overturning streamfunction in temperature coordinates is useful is that for isotherms below MINST, mass conservation takes the form:

\[
\psi(y, T, t) = \partial \Psi / \partial t - \int w^*(x, y, T) dx \ dy,
\]

where \( V \) is the volume of water above the \( T \) isotherm north of \( y \), \( w^* \) is the “entrainment velocity” through the isotherm, and the integral is over the area north of \( y \). In the application of this paper, where there are no salinity effects, \( w^* \) is just \( \partial (\kappa \partial T / \partial z) / \partial z \) with \( \kappa \) the diapycnal diffusivity. In more general applications, an extra term of \( A \nabla^2 T / \partial \partial z \) must be added, where the gradients of \( T \) are taken along isopycnal surfaces and \( A \) is the isopycnal diffusivity.

On long-term mean the time-dependent term vanishes, so a nonzero value of the time average streamfunction above a given isotherm can only be balanced by processes proportional to the eddy diffusivities \( \kappa \) and \( A \). This establishes the statement made in the introduction.
APPENDIX B

Nonlinear Self-Regulation of Inflow Depth

In the present vertical-walled model, the constraint of hydrostatic stability, plus the requirement that the total inflow replacing Ekman transport be geostrophically balanced, places a theoretical lower limit on the depth of the northward flowing geostrophic current. A limited depth of no motion assumption is made: that there is some depth \(-Z\) (with pressure \(p_0\)) such that for all \(|z| > |Z|\) (or \(p > p_0\)) the net basinwide meridional geostrophic flow \(\nu(p) = f v(x, p)\) \(dx\) is zero. Note that \(v(x, p)\) itself need not be zero for \(|p| > |p_0|\)—only its zonal integral; so, for example, stationary eddies can extend below depth \(Z\). Then the net northward geostrophic transport \(V = \int \nu(p) \, dp/\rho g\) through a given latitude \(\theta\) is given by \(g \delta P f(\theta)\), where

\[
\delta P = \left[ \int_{p_0}^p \int_P \delta P(p')/\rho \right] (dp'/\rho g) (dp/\rho g)
\]

\[
= \int_Z^0 z \rho_0 [\Delta W(z) - \Delta_E(z)] \, dz
\]

(B1)

is the depth-integrated steric height (DISH) difference from east to west at a given latitude and \(\rho g\) times the steric height is the pressure calculated from inverse density, or specific volume anomaly (SVA) data relative to a depth of no motion; thus \(\rho g\) times the steric height difference between two points at a given depth balance geostrophic flow between the points, at that depth. Therefore the DISH difference balances the depth-integrated geostrophic flow between them (e.g., Tomczak and Godfrey 2003). The second equation is obtained by changing orders of integration in \(p'\) and \(p\) and using the Boussinesq approximation. Here \(\Delta(z)\) is the SVA, and \(\delta \Delta = \Delta_W(z) - \Delta_E(z)\) is the difference in SVA between points next to the eastern and western walls at the chosen latitude, at depth \(z\). The result is an integral constraint on the in situ density difference between the two basins of the sides:

\[
\int_Z^0 z \rho_0 [1/\rho(x_E, y, z) - 1/\rho(x_W, y, z)] \, dz = f(\theta)V/g
\]

(B2)

(here \(\rho_0\) is a typical mean water density). But for the shallow depths and relatively high temperatures involved in our tropical application, the compressibility of water is nearly independent of temperature and salinity; so in the density difference appearing in (B2), the in situ density can be replaced by potential density. In the future, unless otherwise specified, “density” will mean potential density, and “temperature” means “potential temperature.”

By geostrophy and the limited depth of no motion assumption, the density profile \(\rho_E(z)\) at the western boundary must equal \(\rho_W(z)\), below depth \(Z\). Therefore, in (B2), the shallower \(Z\) becomes, the greater the inverse density difference \((1/\rho_E - 1/\rho_W)\) must become. But, for given \(Z\) there is a well-defined “dendest possible density profile” at the western boundary, for which water at every depth above \(Z\) is at neutral stability (Fig. 9a)—that is, its potential density is constant above depth \(Z\), at \(\rho_E(Z)\). The minimum possible depth \(Z\) is then found by choosing \(Z\) so that

\[
\int_Z^0 z \rho_0 [1/\rho_E(Z) - 1/\rho_W(z)] \, dz = f(\theta)V/g
\]

\[
= \int_{x_E}^{x_W} \tau^{(\infty)}(x, y) \, dx / \rho g.
\]

(B3)

The second equality comes from the fact that the northward geostrophic flow must balance the southward Ekman flow. Some physical insight into (B3) is obtained by assuming that the squared Brunt–Väisälä frequency \(N^2 = (g/\rho)(\partial p/\partial z)\) at the eastern boundary is constant above the depth \(Z\); then (B3) yields

\[
Z = (6fV/N^2)^{1/3} = \left[ 6 \int_{x_E}^{x_W} \tau^{(\infty)}(x, y) \, dx / \rho g \right]^{1/3};
\]

(B3')

\(Z\) increases rather slowly with increasing transport \(V\) and with latitude through \(f = \beta y\). It decreases with increasing stratification, \(N^2\). For \(V = -10\) Sv and for \(f\) at 8°N, and \(N^2\) from the average gradient over the top 200 m of the black line in Fig. 9a, (B3') yields \(Z = 205\) m. Note that this is much deeper than a typical mixed layer depth in the present model.

The purple line in Fig. 9a shows the density profile \(\rho_E(z)\) at 8°N at the eastern wall of the model basin, at the end of experiment 3, described below (these \(\rho_E\) profiles are virtually identical in all experiments). The vertical black line in Fig. 9a shows the densest possible density profile at the western boundary, with depth \(Z\) chosen for \((8^\circ \text{N})\) and \(V = -10\) Sv so that \(\Delta P\) is given by \(f(8^\circ \text{N})V/g\), or 20.2 m². The resulting depth \(Z\) is 212 m [only 7 m deeper than the estimate from (B3')]. The density of the deep mixed layer on the inshore edge of the western boundary current is 26.4, with a corresponding temperature of 13.1°C.

Tests of the minimum depth hypothesis

The green, blue, and red lines in Fig. 9a show the potential density profiles \(\rho_W(z)\) at the western boundary at 8°N, obtained at the end of four years in experi-
ments 1–3, respectively. Unlike their eastern boundary counterparts \( \rho_E(z) \), which were virtually identical to that from experiment 3 (purple line in Fig. 9a), there is a large difference between the three profiles \( \rho_W(z) \). The potential density profile from experiment 1 (green) is highly stratified at all depths and departs from \( \rho_E(z) \) well below \( Z \). That from experiment 2 (blue) is at near-neutral stratification down to 400 m and it also departs from \( \rho_E(z) \) to well below 212 m. However, the western boundary profile from experiment 3 (red line) matches the minimum depth density profile closely, at all depths beneath a pycnocline only 40 m deep.

As seen in Figs. 5b,c, contours of the overturning streamfunction shallow northward near the coast. The current is flowing down a pressure gradient; if there is a valid depth of no motion, then lower pressures in the upper ocean mean shallower isotherms (at least in baroclinic flows). As seen in Fig. 7, the flow is nearly adiabatic south of 8°N; this means that the flow must also shallow as it moves poleward. On the other hand, the minimum depth from (B3) [or (B3')] increases monotonically with distance poleward. It seems that in experiment 3, the imposed transport V shoals until—at the separation point at 8°N—the density profile comes close to matching that deduced from (B3) (cf. black and red lines in Fig. 9a) without actually violating static stability there.

The green, blue, and red lines in Fig. 9b show the vertical profile of zonally integrated northward geostrophic flow inferred by upward vertical integration of the difference \( g \rho_0(1/\rho_E(z) - 1/\rho_W(z))/f(8°N) \) from experiments 1–3. The black line shows the zonally integrated velocity profile that would be obtained if \( \rho_W(z) \) were the same as \( \rho_E(z) \) for \( z \) deeper than \( Z \), and was \( \rho_E(Z) \) for \( z \) shallower than \( Z \). This minimum depth current profile in Fig. 9b is seen to be a fairly useful guide to those actually obtained, even in experiments 1 and 2—but more strongly in experiment 3. Figure 9b shows that the depth profile of zonally integrated geostrophic flow \( \mathbf{u}(z) \) through 8°N is markedly more surface intensified in experiment 3 compared to experiment 1. The southward counterflow seen between 200 and 500 m in experiment 2 (Fig. 9b) is associated with an “overshoot” of convective overturn near the western boundary to well below the minimum depth of 212 m.

Two other experiments (experiments A and B) were performed; they are summarized in Table 2. They were both identical to experiment 1 except that in experiment A a strong “summertime” value of cross-basin Ekman transport was taken (50 Sv southward; see Fig. 3) and the wind stress was held steady at that large value. In experiment B the same was done with a strong “wintertime” Ekman transport of 30 Sv northward. In the first of these extra experiments, the inflow extended substantially deeper as well as being five times stronger; the modeled depth profile through 8°N was somewhat deeper than the minimum depth profile (which in this case was nonzero only above a depth \( Z \) of some 765 m). However the net heat flux was correspondingly only about five times that from experiment 1 because the surface mixed layer was also substantially colder than in experiments 1–3, bringing the temperature difference between inflow and outflow roughly back to what it had been in those experiments. In experiment B, essentially no surface heat flux occurred north of the equator at the end of two years; the northward transport of heat was balanced by a steady deepening of the surface mixed layer, so no steady state was reached after two years.

**REFERENCES**


Hu, R.-J., and J. S. Godfrey, 2007: Explorations of the annual


