Eddy Parameterization Challenge Suite. I: Eady Spindown

S. Bachman and B. Fox-Kemper

Cooperative Institute for Research in Environmental Sciences (CIRES) and Dept. of Atmospheric and Oceanic Sciences (ATOC), University of Colorado, Boulder, CO, United States. bfk@colorado.edu.

Abstract

The first set of results in a suite of eddy-resolving Boussinesq, hydrostatic simulations is presented. Each set member consists of an initially linear stratification and shear as in the Eady problem, but this profile occupies only a limited region of a channel and is allowed to spin-down via baroclinic instability. The diagnostic focus is on the spatial structure and scaling of the eddy transport tensor, which is the array of coefficients in a linear flux-gradient relationship. The advective (antisymmetric) and diffusive (symmetric) components of the tensor are diagnosed using passive tracers, and the resulting diagnosed tensor reproduces the horizontal transport of the active tracer (buoyancy) to within ±7% and the vertical transport to within ±12%. The derived scalings are shown to be close in form to the standard Gent-McWilliams (antisymmetric) and Redi diffusivity (symmetric) tensors with a magnitude that varies in space (concentrated in the horizontal and vertical near the center of the frontal shear) and time as the eddies energize. The Gent-McWilliams eddy coefficient is equal to the Redi isopycnal diffusivity to within ±6%, even as these coefficients vary with depth. The scaling for the magnitude of simulation parameters is determined empirically to within ±28%. To achieve this accuracy, the eddy velocities are diagnosed directly and used in the tensor scalings, rather than assuming a correlation between eddy velocity and the mean flow velocity where ±97% is the best accuracy achievable. Plans for the next set of models in the challenge suite are described.

Keywords: eddies; parameterization; mesoscale; baroclinic instability; Gent-McWilliams; Redi; transport tensor; advective; diffusive; spin-down; Eady; scaling

1. Introduction

Computational power will always limit the resolution of ocean models. The oceanic mesoscale eddy field is important to the structure and sensitivity of the large-scale ocean (e.g., Danabasoglu and McWilliams, 1995; Eden et al., 2009; Grooms et al., 2011), and resolving the mesoscale eddy field is not routinely possible in oceanic general circulation models (OGCMs). Extrapolating current trends in computation predicts that mesoscale eddy parameterizations will be needed for some decades into the future, especially for use in high complexity earth system models and long duration or large ensemble scenarios needed for certain problems such as obtaining reliable statistics of tropical variability (Wittenberg, 2009; Stevenson et al., 2010), biogeochemistry spinup (Key et al., 2004), and paleoclimate (Jochum et al., 2012). As such, optimizing and evaluating eddy parameterizations is an important task, but few scenarios are sufficiently and repeatedly simulated to serve as a measure against which parameterizations may be tested.

All mesoscale eddy fluxes must be parameterized in present OGCMs and climate models where the oceanic horizontal grid resolution is O(100km)–insufficient for eddy growth in almost all regions. Grids of O(10km) or better are needed to adequately resolve most mesoscale motions (McClean et al., 2006), but even at these resolutions low stratification and polar regions remain poorly resolved. When eddies are partially resolved, parameterizations of the missing eddy fluxes are still needed (Large Eddy Simulation closures, e.g., Roberts and Marshall, 1998; Fox-Kemper and Menemenlis, 2008), but here the focus is on evaluating parameterizations designed for use when no eddies are resolved and all eddy fluxes are parameterized (Reynolds Averaged Model closures, e.g., Gent and McWilliams, 1990).

It is standard practice to parameterize the effects of subgridscale eddies by including extra terms in the equations of motion. At present, the greatest care is taken in modifying the active tracer equations, which analysis indicates to be the most likely eddy effects to be important (Grooms et al., 2011). Ideally, eddy parameterizations would be compared directly with observations, but the sheer number of observations required makes this rare or unfeasible. An evaluation using numerical models compares high-resolution models against eddy parameterizations. However, discretion is required in choosing which scenarios to simulate at high-resolution. An ideal set of tests would be representative of most likely scenarios and would accentuate the differences between parameterizations. A suite of such model results would be very useful in understanding and evaluating eddy parameterizations, including those not yet developed.

This paper discusses the first set of simulations in the construction and implementation of such a test bed, or *eddy parameterization challenge suite*. The transport of tracers in a large-scale, rotating, stratified turbulent flow is the focus of this set of challenges. The model scenario simulated here is similar to the Eady (1949) problem, in that eddies form from instabilities in constant stratification and shear. One reason for choosing this problem is the past work on related problems (e.g., Stone, 1972; Fox-Kemper et al., 2008). Other reasons are computational efficiency and a limited number of parameters to completely describe the simulation. As in Fox-Kemper et al. (2008) (hereafter FFH), a constant stratification and shear will be imposed in the initial conditions at the center of the domain and the shear will be smoothly tapered to zero toward the boundaries. This initial configuration will be allowed to evolve as freely as possible with slip, conservative, insulating boundary conditions. Unlike FFH, other variations in stratification to demarcate a mixed layer or pycnocline are neglected. While this scenario is unlike any particular ocean region, any parameterization capable of reproducing eddy tracer fluxes in more complex scenarios should accurately handle this simple case over a wide range of simulation parameters.

For situations where no eddies are resolved, parameterizations amount to approximation of processes at fine resolution (*subgrid*) in terms of coarse resolution (*resolved*) quantities. A subgrid flux to resolved gradient (*flux-gradient*) relationship is often assumed, which is of the form

$$\overline{\mathbf{u}'b'} = -\mathbf{R}\nabla\bar{b}, \tag{1}$$

$$\overline{\mathbf{u}'\tau'} = -\mathbf{R}\nabla\bar{\tau}. \tag{2}$$

These equations relate the subgrid eddy flux of buoyancy b or tracer concentration τ to the resolved gradient. In general, the operator $\overline{(\ )}$ could refer to a spatial, temporal, spectral, or other kind of average; *in this paper it will indicate a zonal, along-channel average*. The eddy component, denoted by the "prime" symbol, will be the local deviation away from this average. The proportionality between the flux vector and gradient vector is governed by a 3×3 *eddy transport tensor*, $\mathbf{R}$.

Throughout, it is assumed that the coordinate system is Cartesian and aligned with the geoid, as is appropriate for z-coodinate models. For models in other coordinates, conversions can be made (de Szoeke and Bennett, 1993; Hallberg, 2000). Furthermore, here all tracers are treated on equal footing; the popular method of averaging in density-coordinates singles out density as a special tracer, thus it is avoided despite any advantages it may bring. In principle, every tracer might have its own unique transport tensor $\mathbf{R}$, but theory (Taylor, 1921; Plumb and Mahlman, 1987; Dukowicz and Smith, 1997; Fox-Kemper et al., 2012b) connects $\mathbf{R}$ to the correlations of displacements of fluid parcels which are independent of tracer, so long as background tracer concentrations vary smoothly and local sources and sinks of the tracer are weak. Formulations to reduce the contribution of rotational and gauge indeterminacy of **R** have been proposed (Eden et al., 2007), but the **R** diagnosed by this method differs for each tracer and thus there is not a ready comparison to correlations of displacement theories (Taylor, 1921)[1]. Below, it will be shown that using one tensor serves to reproduce fluxes of all tracers studied, active and passive, to within a few percent error in these simulations. A single tensor for many tracers has been hypothesized or posed in many contexts (Andrews and McIntyre, 1978; Plumb, 1979; Plumb and Mahlman, 1987; Bratseth, 1998; Griffies, 1998) and can be used to represent a combination of advective and diffusive transport by eddies. Determining this tensor in the Eady-like problem is the prime analysis here.

This paper is organized as follows. Section 2 reviews the background and theory underlying diagnosis of the transport tensor using passive tracers. Section 3 details a scaling of the tensor elements based on fluid parcel exchanges. Section 4 contains diagnoses of the simulations including empirical scaling laws. The final section concludes. Appendices describe the simulation setup in detail, including the shear and stratification configuration (Appendix A), tracer initialization (Appendix B), and model setup (Appendix C).

2. Background and Theory

The simplest form for the transport tensor, **R**, is purely diagonal for downgradient Fickian diffusion and is often implemented with different vertical and horizontal diffusion rates.[2] This simple form is inadequate, as it spuriously mixes distinct watermasses (Veronis, 1975). The next simplest form is a symmetric tensor, representing diffusion with different rates along different directions, which may be any orthogonal set, not just horizontal and vertical. The Redi (1982) isoneutral diffusion parameterization (hereafter Redi) is an example. In incompressible flow, Taylor (1921) shows that a symmetric tensor is sufficient to capture diffusion by continuous movements.

However, a symmetric form is inadequate as well, as not all eddy transport is diffusive. Using a purely diffusive parameterization leaves advection unaffected, so the resolved velocity is the only advective transport. Observational and numerical model studies, especially of the stratosphere (Plumb, 2002) but also in the ocean (Gent et al., 1995; Marshall et al., 2006; Zika et al., 2010), have shown that on average tracers are *not* advected by the large-scale velocity. Indeed, Andrews and McIntyre (1978) and Dukowicz and Smith (1997) show that in compressible flow, or in flow along a two-dimensional subspace surface embedded within incompressible three-dimensional flow, an additional advective transport due to eddies is likely to arise. The Gent and McWilliams (1990) (hereafter GM) mesoscale eddy parameterization specifies this additional advective eddy tracer transport, consistent with a release of mean potential energy through baroclinic instability. In the context of the diagnosed eddy transport tensor **R**, such advective effects will lend an antisymmetric contribution.

2.1. Advective and Diffusive Fluxes

Griffies (1998) demonstrates that the transport tensor in the flux-gradient relationship can be interpreted as contributing both an advective and a diffusive component. This decomposition is uniquely equivalent to subdividing R_{ji} into antisymmetric and symmetric parts, respectively,

$$R_{ji} = S_{ji} + A_{ji}, \quad S_{ji} = \frac{R_{ji} + R_{ij}}{2}, \quad A_{ji} = \frac{R_{ji} - R_{ij}}{2} \tag{3}$$

[1] Although a personal communication with Eden and Griesel indicates that such a synthesis may be forthcoming.

[2] Recall that true vertical and horizontal are intended, not dianeutral and isoneutral mixing as in an isopycnal model.

Throughout, the symmetric tensor **S** will be referred to as the *diffusivity tensor*, the antisymmetric tensor **A** as the *advective tensor*, and the combination of both advection and diffusion **R** will be the *transport tensor*.

The diffusivity tensor, **S**, is symmetric and therefore will have all real eigenvalues. Each eigenvalue corresponds to a diffusivity in a particular (eigenvector) direction. Diffusion in this form is discussed elsewhere (Fox-Kemper et al., 2012a; Fox-Kemper et al., 2012b).

The tracer flux divergence due to the antisymmetric tensor **A** is identical to an advection by an incompressible velocity, $\mathbf{u}^\dagger$ (Griffies, 1998). The association with a velocity follows directly from the antisymmetry of A_{ji}, the symmetry of $\nabla_j \nabla_i$, and the fact that the inner product of a symmetric and an antisymmetric tensor is identically zero.[3]

$$\begin{aligned} u_j^\dagger &\equiv \nabla_i A_{ji}, \\ \nabla_j u_j^\dagger &\equiv \nabla_j \nabla_i A_{ji} = 0, \\ 0 = \nabla_j \nabla_i \left(A_{ji}\bar{\tau}\right) &= \nabla_j \left(A_{ji}\nabla_i \bar{\tau}\right) + \nabla_j \left(\bar{\tau}\nabla_i A_{ji}\right). \end{aligned} \tag{4}$$

So, using subscripts on scalars, such as tracer concentration or buoyancy, we find

$$\begin{aligned} \nabla \cdot \overline{\mathbf{u}'\tau'} &= -\nabla_j R_{ji} \nabla_i \bar{\tau}, \\ &= -\nabla_j (A_{ji} + S_{ji}) \bar{\tau}_i, \\ &= -\nabla_j \left(A_{ji} \nabla_i \bar{\tau}\right) - \nabla_j \left(S_{ji} \nabla_i \bar{\tau}\right), \\ &= \nabla_j \left(\bar{\tau} \nabla_i A_{ji}\right) - \nabla_j \left(S_{ji} \nabla_i \bar{\tau}\right), \\ &= \nabla_j \left(u_j^\dagger \bar{\tau}\right) - \nabla_j \left(S_{ji} \nabla_i \bar{\tau}\right), \\ &= \mathbf{u}^\dagger \cdot \nabla \bar{\tau} - \nabla \cdot (\mathbf{S} \cdot \nabla \bar{\tau}). \end{aligned} \tag{5}$$

It is convenient to associate this incompressible eddy-induced velocity $\mathbf{u}^\dagger$ with a streamfunction ($\nabla \times \psi^\dagger = \mathbf{u}^\dagger$). The components of $\psi^\dagger$ are just the reordered nonzero elements of the antisymmetric tensor **A** ($\epsilon_{ijk}\psi_k^\dagger = A_{ij}$, $\psi_k^\dagger = \frac{1}{2}\epsilon_{kij}A_{ij}$, where ϵ is the totally antisymmetric Levi-Civita symbol). In the case of interest here, the eddy-induced flow will be in the $y-z$ plane only, so only the x component of the streamfunction is nonzero. It is equal to $\mathbf{A}_{yz}$ and opposite $\mathbf{A}_{zy}$.

With the eddy-induced velocity and symmetric diffusivity tensor, the tracer equation is

$$\frac{\partial \bar{\tau}}{\partial t} + (\bar{\mathbf{u}} + \mathbf{u}^\dagger) \cdot \nabla \bar{\tau} = \nabla \cdot \mathbf{S} \cdot \nabla \bar{\tau}. \tag{6}$$

The usefulness of the symmetric / antisymmetric decomposition is dependent on how well the flux-gradient relationship is satisfied. If the flux-gradient inversion for **S** is underdetermined or noisy, then the symmetric / antisymmetric decomposition is not meaningful. If the flux-gradient relationship itself is not useful, e.g., if non-local effects are important, then there is no reason to think the symmetric / antisymmetric decomposition will improve matters. It is entirely appropriate to form separate parameterizations for the symmetric and antisymmetric tensors, as is done by Redi and GM, respectively. Here, scalings for the eddy fluxes based

[3] Einstein summation is implied, so repeated indices indicate a sum over all coordinate directions.

on problem parameters, and therefore scalings for each tensor element, are determined. These scalings are comparable to a diagnosis technique of model results detailed next.

Some authors (e.g., Eden et al., 2007) prefer not to consider the total, measurable physical flux $\overline{u'_j\tau'_\pi}$, but instead prefer to contaminate the measurable fluxes with rotational corrections that either are not unique (Fox-Kemper et al., 2003) or depend on which tracer in particular is used (e.g., Eden et al., 2007). Determining the rotational flux here halves the statistical value of each tracer, by introducing another unknown field along with each tracer. As the zonal-mean flux, and therefore the flux divergence, is readily predictable from the flux-gradient relationship here, there is little reason to introduce this extra noise into the determination.

2.2. Tensor Under- and Overdeterminacy

Often, models are diagnosed or theory is formulated to reproduce the evolution of a single tracer, usually potential temperature, buoyancy, or potential vorticity. The idealization is that small-scale turbulent motion acts similarly on all conserved tracers, and so the one can act as a "representative" tracer for all mixing. This assumption should be considered carefully.

A fundamental limitation is that for eddy transport in n dimensions, using one tracer in the flux-gradient relationship provides only n constraints on n^2 elements of the transport tensor. Thus, for $n > 1$, it is an underdetermined system. In a zonal average as here, $n = 2$. Each tracer gives two constraints (one for each flux component), but the transport tensor will consist of four elements. Thus, at least two tracers (giving four tracer flux components and four tracer gradient components) are needed to sufficiently determine the system. However, since zonal averaging effectively hides a large number of degrees of freedom, it is appropriate to overdetermine the system by using extra tracers and to solve for the transport tensor using a least-squares approach so that a "representative average" over all of the hidden degrees of freedom is found (Bratseth, 1998; Fox-Kemper et al., 2012c).

Why is the least-squares approach useful, rather than considering each tracer flux in turn? Consider for a moment a large ensemble of many different tracers that are initialized with the same zonal mean, but differ in their zonal variations. Consider also a large ensemble of velocity fields, that again agree in zonal mean but differ in zonal variations. Now consider all of the possible fluxes that result from advecting each tracer in the ensemble with each eddy field in the ensemble. The mean of all of the fluxes will constitute a weighted mean over different eddy features with different weights, because of coincidental correlations of the zonal eddy and tracer variations. Fundamentally, such chance correlations limit the reproducibility of tracer fluxes with any specific eddy field and tracer field to realize the ensemble flux to zonal-mean gradient relationship. However, if the zonal variations of tracer are chosen in a systematic and unbiased manner to sample across the eddy features evenly, then simultaneously considering them all (in the least-squares pseudoinverse sense) will converge rapidly to the ensemble mean, which is the transport tensor we seek. Appendix B describes the systematic method for sampling used here.

2.2.1. Tracer Inversion

In this research numerical simulations resolving mesoscale eddies are used to determine the eddy transport tensor. To do this, a set of passive tracers is initialized in each of the runs (Appendix B) and advected by the flow. There are no sources or sinks of these tracers, their explicit diffusivity is zero, and their numerical diffusivity is small, so it is assumed that a single, representative transport tensor R can be used for all of these tracers (Taylor, 1921; Dukowicz and Smith, 1997). The flux-gradient equation for such a system is

$$\overline{u'_j\tau'_\pi} = -R_{ji}\nabla_i\bar{\tau}_\pi, \tag{7}$$

$$R_{ji} = -\overline{u'_j\tau'_\pi}[\nabla_i\bar{\tau}_\pi]^{-1} \tag{8}$$

where in three spatial dimensions i and j run from 1 to 3 and π runs from 1 to the number of tracers used (Greek symbols are used for tracer number, since a Roman subscript would denote partial differentiation). The term $\nabla_i \bar{\tau}_\pi$ forms a matrix, which we will hereafter call the tracer gradient matrix. A two dimensional system would require two tracers, with each of their gradients misaligned with respect to one another, to uniquely solve for $\mathbf{R}$. In other words, the matrix formed by the gradients of the tracers $\bar{\tau}$ must be nonsingular for an ordinary inverse of the bracketed term on the right of (8).

Often, the system is underdetermined, as either fewer than three tracers are used or the tracer gradients are aligned in places, and in such a case the solution for $\mathbf{R}$ is not unique. A relevant example arises when buoyancy is the only tracer diagnosed. The sole tracer equation is then

$$\bar{b}_t + \nabla_j(\bar{u}_j \bar{b}) = \nabla_j \left[-\overline{u_j' b'} + \bar{B}_j \right], \tag{9}$$

where $\bar{B}_j$ represents small-scale diffusive fluxes, and boundary sources and sinks such as latent and sensible atmospheric heating. If $|\nabla \bar{b}| \neq 0$ everywhere, then multiplying by $1 = |\nabla \bar{b}|^2 / |\nabla \bar{b}|^2 \equiv (\bar{b}_i \bar{b}_i)/(\bar{b}_k \bar{b}_k)$ yields,

$$\bar{b}_t + \nabla_j \bar{u}_j \bar{b} = \nabla_j \left[-\frac{\overline{u_j' b'} \bar{b}_i}{\bar{b}_k \bar{b}_k} \bar{b}_i + \bar{B}_j \right]. \tag{10}$$

Here the transport tensor is identified with (identical to the isopycnal definition of Ferrari and Plumb, 2003)

$$R_{ji} = -\frac{\overline{u_j' b'} \bar{b}_i}{\bar{b}_k \bar{b}_k}, \tag{11}$$

and we can recover the form above by multiplying by $\bar{b}_i$:

$$-\overline{u_j' b'} = -\frac{\overline{u_j' b'} \bar{b}_i}{\bar{b}_k \bar{b}_k} \bar{b}_i = R_{ji} \bar{b}_i \tag{12}$$

Note that a choice was taken in multiplying by $(\bar{b}_i \bar{b}_i)/(\bar{b}_k \bar{b}_k)$, which is allowed by the underdetermination of this system. The solution for $\mathbf{R}$ is not unique; indeed we could add any components we like to R_{ji} for additional transport in the directions other than that spanned by the gradient of $\bar{b}$. Nor are the symmetric or antisymmetric parts unique, so diffusion, streamfunction, and eddy-induced velocity are untrustworthy when determined with only buoyancy[4].

Likewise, if one sought a unique solution for $\mathbf{R}$ using three tracers, the tracer gradients need to be everywhere misaligned to avoid singular matrices. Alignment is likely to occur occasionally no matter how the tracers are initialized, and indeed, straining by eddies tends to align tracer gradients. Therefore the methodology of Bratseth (1998) is adopted, which calls for overdetermination of the system by using more than three tracers. In this case, the tracer gradient matrix is inverted in the least squares sense using the Moore-Penrose pseudoinverse. If one considers the singular value decomposition of the tracer gradient matrix, then

[4]The exception is when fluxes are strictly adiabatic and steady, in which case the degrees of freedom are reduced (e.g. Colas et al. (2012)).

$$\nabla_i \bar{\tau}_\pi = U_{ik} \Sigma_{kj} V^*_{j\pi} \tag{13}$$

$$[\nabla_i \bar{\tau}_\pi]^{-1} = V_{\pi j} [\Sigma_{kj}]^{-1} U^*_{ki} \tag{14}$$

$$R_{ji} = -\overline{u'_j \tau'_\pi} V_{\pi j} [\Sigma_{kj}]^{-1} U^*_{ki} \tag{15}$$

In three spatial dimensions and with n tracers, U is a 3×3 unitary matrix, Σ a $3 \times n$ rectangular diagonal matrix, and V a $n \times n$ unitary matrix. In the pseudoinversion shown in (14), the new diagonal matrix $[\Sigma_{kj}]^{-1}$ is formed by taking the reciprocal of each of the non-zero diagonal values of Σ_{kj} and leaving the zero values[5] in place. This inverse equals a least-squares fit when the system is overdetermined, or a least-variance solution when the system is underdetermined.[6]

The present analysis assumes that all passive tracers are diffused similarly, which is not always the conclusion drawn by other authors using one-tracer-at-a-time diagnoses (e.g., Lee et al., 1997). Tracer fluxes may differ among tracers that have different diffusivities or different sources and sinks. Active tracers are particularly prone to variation along these lines, and different flow setups may differ widely in the sources and sinks of active tracers. It is beyond the scope of the present paper, which is focused on the Eady-like problem alone, to address or overly speculate about why fluxes may differ in different configurations such as that of Lee et al. (1997).

3. Hypothesized Parameter Scaling of Tensor Components

The simulations comprising the eddy parameterization challenge suite are run using Massachusetts Institute of Technology general circulation model (hereafter MITgcm) (Marshall et al., 1997). The hydrostatic, Boussinesq equations are solved to simulate a zonally reentrant channel on the f-plane, with a temperature front oriented in the cross-channel direction. In this set of simulations, the density gradient is constant in z and y inside the front, akin to the Eady (1949) model. The use of the Eady model on the f-plane is simpler for the purpose of this research than β-plane models used in previous studies (Eden, 2010, 2011), in that it has one fewer parameters and does not form jets.

The velocity fields are initialized in geostrophic balance, to minimize ageostrophic waves. Stratification (N^2), rotation (f), and front dimension (L_f) and velocity (U) are set according to the desired nondimensional parameters: Rossby ($Ro = U/fL_f$) and Richardson ($Ri = N^2/(\partial U/\partial z)^2 \approx N^2 f^2/|\nabla \bar{b}|^2$) numbers. Sixty-nine simulations are performed spanning a range of these parameters. Each simulation depicts the baroclinic spindown of the temperature front (Fig. 1). A few inertial periods after the beginning of the model run the alongfront geostrophic shear goes baroclinically unstable[7]. Restratification will begin as the instabilities reach finite amplitude and begin to slump the isopycnals, akin to FFH. The eddies grow out of the frontal region and will spread meridionally throughout the domain. The simulation is stopped just before the buoyancy perturbation of the front reaches the lateral walls in order to prevent sidewall boundary effects. More details about the model setup and diagnostic methods can be found in Appendices A and C.

A new approach taken in this research is the use of transient snapshots in the collection of the eddy statistics. Despite the ever-changing nature of the ocean, an equilibrated eddy field is more commonly used

[5]Zero is treated numerically as any value less than $\max(k, j) \times \|\Sigma\|_2 \times \epsilon$, where $\|\cdot\|_2$ is the L_2 matrix norm and ϵ is the machine precision. In practice this value tended to be between 10^{-18} and 10^{-17}.

[6]Indeed, the buoyancy-only inversion in (12) was a Moore-Penrose pseudoinverse in disguise!

[7]An initial Richardson number greater than one precludes other instabilities, e.g. symmetric.

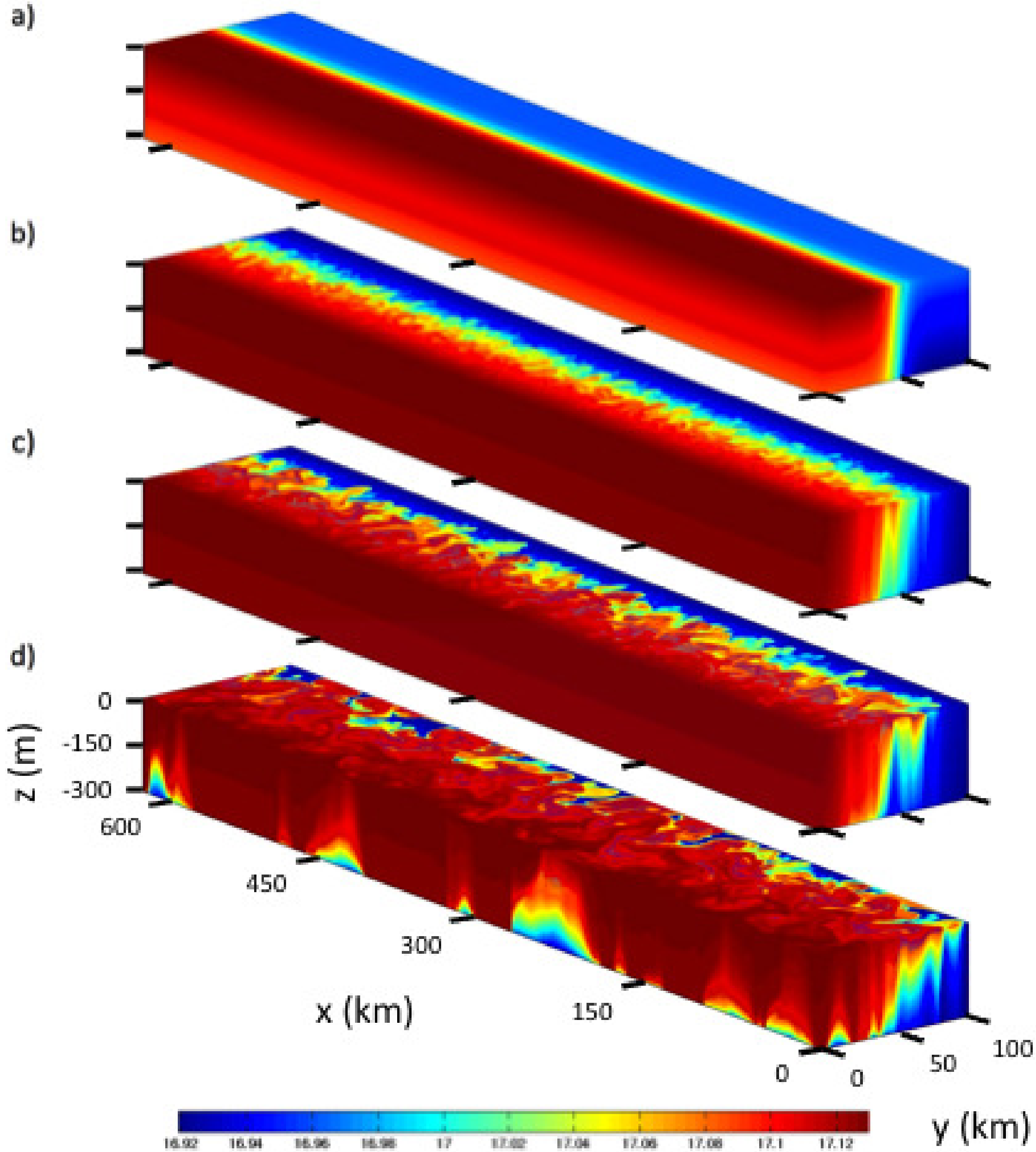

Figure 1: Potential temperature during a typical frontal spindown simulation. Baroclinic instability causes the front to slump towards the horizontal, releasing potential energy in the process. The eddies grow from this potential energy release as the front slumps from its initial configutation (a), through a fully nonlinear turbulent state (b-c), until the simulation is complete (d).

(Lee et al., 1997; Eden, 2010, 2011) to test eddy parameterizations than snapshots. However, in reaching equilibrium, the eddy fluxes are often constrained to reach a balance by satisfying viscous integral budgets or by balancing production and dissipation. Such balances depend explicitly and sensitively on unknown subgrid parameters and drag coefficients (Fox-Kemper and Pedlosky, 2004; Thompson and Young, 2007). The transient simulations used in this research may not be representative of all situations that occur in the ocean, but they are not strongly dependent on poorly-known subgrid parameters.

Following the methods of FFH, statistics are gathered at each snapshot after the vertical eddy kinetic energy saturates. The criteria used for this saturation was that the vertical EKE did not change by more than 3% of its value at the previous snapshot. This criteria is essentially a proxy to ensure that the eddies have reached finite amplitude and that the eddy interactions have become saturated by nonlinearity. At each time snapshot the velocity, temperature, and tracer fields are zonally averaged and written to file. One iteration of an unweighted, sliding-average smoothing algorithm is applied to all fields to reduce biases in the averaging from transient, powerful eddies. Finite differencing is used to create the tracer gradient matrices, and the inversion for the transport tensor is carried out using the method described in Section 2. Thus, in each run a time series of values is generated for each element of the transport tensor at each point in the zonally averaged field. Variability in these time series arises from the presence of internal gravity waves as well as temporal fluctuations in the eddy statistics, neither of which significantly affect the results of this paper.

A robust method for initializing the tracers was found by experimentation to require six tracers, initialized in orthogonal directions and orthogonal functions (see Appendix B). After zonal averaging, the fluxes and gradients of the tracer fields are used to solve for **R** using the tensor inversion method. The accuracy of this method is verified by comparing the diagnosed buoyancy fluxes with a set of "reproduced" fluxes derived by multiplying the inverted transport tensor **R** by the buoyancy gradient. That is, the relative error is given by

$$E(\overline{u_j'b'}) = \frac{\left|\overline{u_j'b'} + R_{ji}\nabla b_i\right|}{\left|\overline{u_j'b'}\right|}, \tag{16}$$

and is evaluated for both v' and w' (Fig. 2a-b). These errors are calculated by averaging over the region defined in Appendix C. Even though buoyancy fluxes and gradients are not used in calculating R_{ji} (only those of passive tracers are), the 95% confidence interval in reproducing buoyancy fluxes has less than 7% relative error for the horizontal fluxes and 12% for the vertical fluxes. Thus, *even though buoyancy and buoyancy fluxes are not used to constrain the calculation of* **R**, *the buoyancy fluxes are reproduced from the mean buoyancy gradient to high accuracy*. While passive and active tracers differ in whether they affect the evolution of the flow, the comparison here is whether they differ *given the same flow*. Thus, *all* components of **R** are constrained to agree with the eddy-induced evolution of the 6 tracers, and buoyancy advection is accurate for free. Ideally, fluxes of potential vorticity, salinity, or any other nearly-materially-conserved scalar tracers active or passive would be similarly accurate, so long as the diagnosis of **R** has converged. However, numerically buoyancy and passive tracers are *exactly conserved* by the finite-volume code used, while potential vorticity is not. Indeed, reconstructed potential vorticity fluxes are noisier than those of buoyancy, so are not used as a check on the accuracy of **R**.

Previous numerical experiments (Rix and Willebrand, 1996; Roberts and Marshall, 2000; Eden et al., 2007) have found low correlations between the diagnosed eddy diffusivity and the actual eddy fluxes, with the distributions of diffusivity often being noisy and permitting unphysically large or even negative values. This can be due in part to the presence of a large rotational component that does not affect the dynamics, but has the potential to contaminate the diagnosis of the eddy diffusivity (Gille and Davis, 1999; Bryan et al., 1999; Eden, 2006; Griesel et al., 2009). A good diagnostic method for $\mathbf{R}$ should therefore yield both good quantitative accuracy via (16) as well as an excellent representation of the dynamics in the flux divergence equation

$$\nabla \cdot \overline{\mathbf{u}'b'} = -\nabla \cdot \mathbf{R}\nabla\bar{b}. \tag{17}$$

The latter can be compared in a fashion similar to (16), except now the error of the flux divergence is measured as

$$E(\nabla_j \overline{u_j' b'}) = \frac{\left|\nabla_j \overline{u_j' b'} + \nabla_j R_{ji} \nabla b_i\right|}{\left|\nabla_j \overline{u_j' b'}\right|}, \tag{18}$$

and is shown in panel c) of Fig. 2. The error in divergence is larger than the error in flux components, consistent with the added derivatives in (18) over (16). However, it is clear from the results based on (18) that the reproduction of fluxes by $\mathbf{R}$ also has skill in reproducing the flux divergences, not just rotational components of the fluxes.

Since Marshall and Shutts (1981), it has been appreciated that large rotational fluxes can arise when tracer gradients and variations in eddy variability take on specific configurations. The use of multiple tracers to determine a single transport operator alters this connection, and a detailed discussion of such alterations while maintaining a connection to the displacement theory of Taylor (1921) requires substantial mathematical detail intended for a future publication. For this reason, because the fluxes are ultimately not unique (Fox-Kemper et al., 2003), and because (17) and Fig. 2 make it clear that the diagnosis here correctly captures the flux divergence, separate rotational and divergent fluxes are not presented.

3.1. Tensor Scaling

Because of the high accuracy in the buoyancy flux reconstruction (Fig. 2), a scaling for the inverted tensor components can follow the dimensional scaling of the transport tensor for buoyancy. The dependence of each tensor element on both dimensional and nondimensional parameters is sought. This can be considered an extension of FFH, who used dimensional arguments to scale the buoyancy fluxes in terms of coarse-resolution gradients. First, a rough estimation of the expected scalings is presented, then the experimental results are used to look for numerical constants and corrections to the scaling. In the FFH approach, the gradients and fluxes were time-averaged after the eddy kinetic energy $\overline{\mathbf{u}'^2}$ became saturated in the domain. Here spatially-averaged results are taken at each snapshot instead of temporal averaging, allowing transient eddy effects and subtle scaling dependencies to be found.

For the tensor components, proportionalities are sought of the form

$$R_{ji} \propto F(N^a, M^b, H^c, f^d, \overline{v'^2}^e, \overline{w'^2}^f, \ldots), \tag{19}$$

so that all dependencies are on coarse-grid quantities. Here F represents some multiplicative function of the vertical buoyancy frequency N, horizontal buoyancy frequency M, fluid depth H, Coriolis parameter f, Reynolds stresses, and possibly other flow variables. Lowercase letters represent unknown exponents. FFH use kinematic arguments to derive scalings for the buoyancy fluxes in terms of these coarse-grid quantities, and these scalings can be used along with the flux-gradient relationship to scale the elements of the transport tensor. That is,

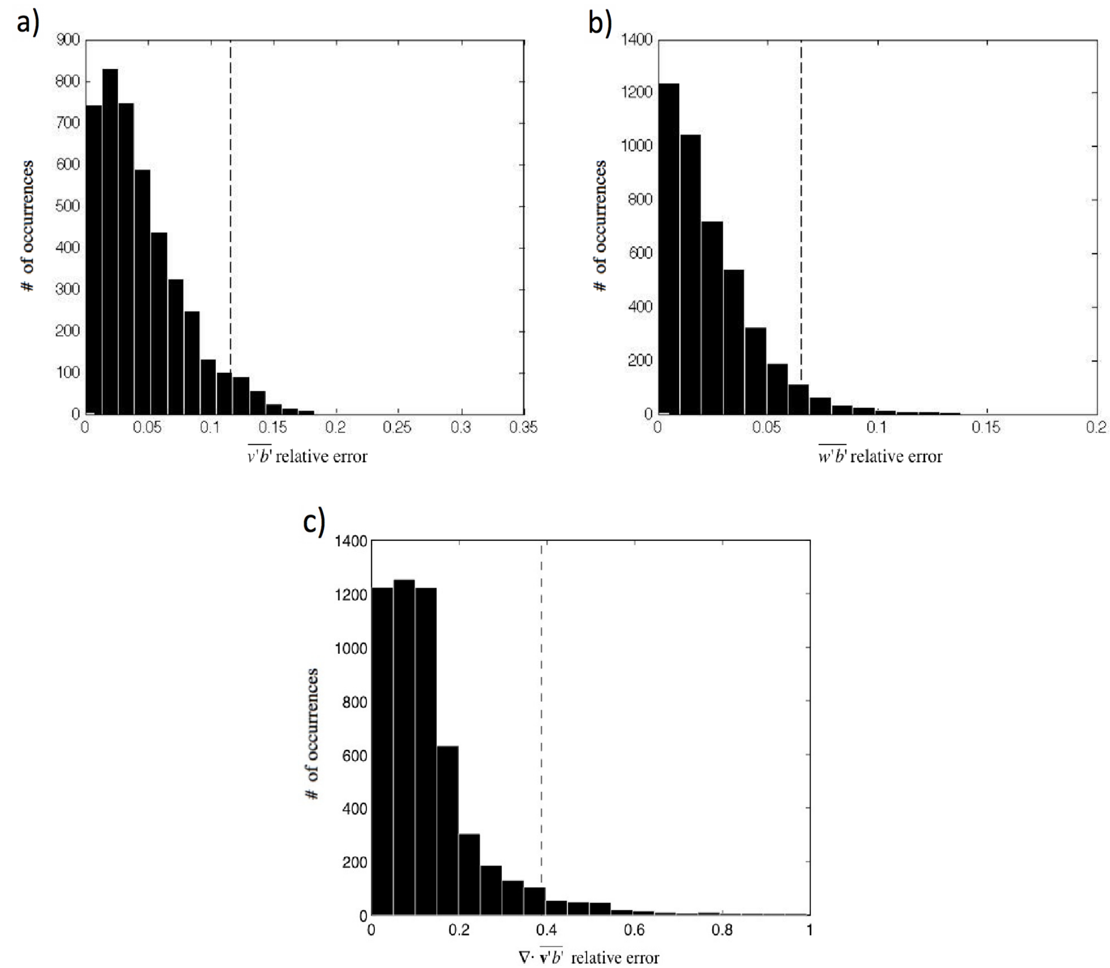

Figure 2: (a-b) The relative error in reconstructing the horizontal and vertical buoyancy fluxes, spanning all snapshots taken from all 69 simulations. Higher relative errors tend to occur at lower values of *Ri*. c) The relative error in reconstructing the eddy flux divergence. The dashed vertical lines in all panels indicate the 95% confidence interval.

$$\begin{bmatrix} \overline{v'b'} \\ \overline{w'b'} \end{bmatrix} = \begin{bmatrix} R_{yy} & R_{yz} \\ R_{zy} & R_{zz} \end{bmatrix} \begin{bmatrix} \bar{b}_y \\ \bar{b}_z \end{bmatrix} \propto \begin{bmatrix} \frac{N^2H^2M^2}{|f|} \\ \frac{M^4H^2}{|f|} \end{bmatrix}, \tag{20}$$

$$R_{yy}\bar{b}_y + R_{yz}\bar{b}_z \propto \frac{N^2H^2M^2}{|f|}, \tag{21}$$

$$R_{zy}\bar{b}_y + R_{zz}\bar{b}_z \propto \frac{M^4H^2}{|f|}, \tag{22}$$

$$\begin{bmatrix} R_{yy} & R_{yz} \\ R_{zy} & R_{zz} \end{bmatrix} \propto \begin{bmatrix} \frac{N^2H^2}{|f|} & \frac{M^2H^2}{|f|} \\ \frac{M^2H^2}{|f|} & \frac{M^4H^2}{N^2|f|} \end{bmatrix}, \tag{23}$$

where the constant of proportionality can be different for each element. Note that the scaling for the off-diagonal elements are found from (21-22) by assuming that both terms in those equations contribute equally and that $\bar{b}_y \propto M^2$ and $\bar{b}_z \propto N^2$.

Breaking the tensor into its symmetric and antisymmetric parts preserves the scale of the off-diagonal elements, so

$$\psi \propto \frac{M^2H^2}{|f|}, \qquad S \propto \begin{bmatrix} \frac{N^2H^2}{|f|} & \frac{M^2H^2}{|f|} \\ \frac{M^2H^2}{|f|} & \frac{M^4H^2}{N^2|f|} \end{bmatrix}, \tag{24}$$

where ψ is the eddy advection streamfunction from section 2.1. The scaling for ψ is identical to that of FFH. The eddy diffusivities may be associated with $\mathbf{S}$, the symmetric part of $\mathbf{R}$, which for the zonally-averaged case will have two real eigenvalues with orthogonal eigenvectors (for 3 degrees of freedom). The eigenvalues represent diffusivities in different directions, so they will be denoted κ_1 and κ_2.

The transport, streamfunction, and diffusivities tend not to be constant in the vertical (Fig. 3). One expects Ekman effects near the surface and bottom boundaries, but perhaps it is less obvious that variations will occur in the interior of the fluid even though stratification and shear are constant. At the surface and bottom, $w' = 0$, which impacts the vertical fluxes, and interior values are smoothly connected to these boundary values. Griffies (2004) summarizes a variety of ways that the eddy transport streamfunction is thought to vary in the vertical, noting that many such methods require an ad hoc upper limit on the isopycnal slope in order to preserve numerical stability. More recently, methods have been developed to match the streamfunction between the boundary layers and the interior (Ferrari et al., 2008, 2010) and have been shown to greatly improve simulation results compared to other methods (Danabasoglu et al., 2008).

Physical rationales for the vertical structure of ψ in the Eady problem have appeared in a variety of locations (Stone, 1972; Branscome, 1983a,b; Fox-Kemper et al., 2008; Ferrari et al., 2008, 2010), but vertical variations of diffusive fluxes occur as well. Indeed, here the effect is to give both fluxes a matching vertical profile. In this research the vertical structures are calculated by taking the cross-channel, basinwide averages of each quantity and plotting these averages as functions of z. Fig. 3d shows the vertical structures of ψ, κ_1, and κ_2 averaged over all of the snapshots from all runs, each normalized by their maximum value in the vertical. The near-parabolic vertical structure for ψ is in agreement with those appearing in the aforementioned papers concerning the Eady problem. The diapycnal diffusivity κ_2 tends to have a similar vertical structure while the larger eigenvalue κ_1 tends to be uniform in the vertical.

From here on multiplicative separability in the vertical is assumed for the tensor component scalings. That is, it is assumed that one can recover the value of any tensor component at any depth level by multiplying

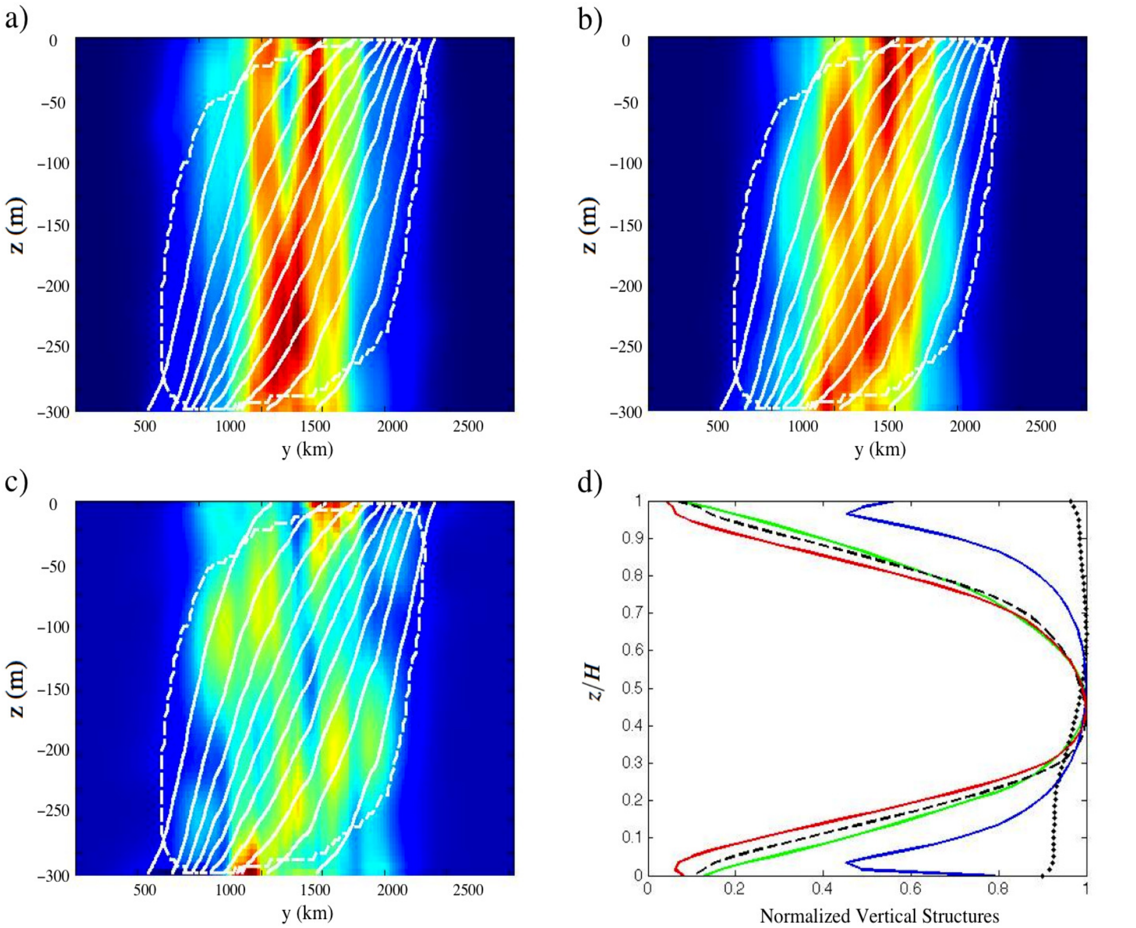

Figure 3: (a-c) Example snapshots during a typical simulation, taken after x-averaging. Shown here are mean isopycnals (solid white lines) and eddying region (enclosed within the dashed white line). The colored backgrounds represent fields for a) R_{yy}, b) R_{zy}, and c) R_{zz}. d) Vertical structures of ψ (green), κ_1 (isopycnal eddy diffusivity, black dotted line), κ_2 (diapycnal eddy diffusivity, black dashed line), isopycnal slope (blue), and κ_1 multiplied by isopycnal slope (red), averaged over all time snapshots and all runs. Each is normalized to have a maximum value of 1.

that component's scaling by its vertical structure function. Also, the y-dependence of eddy variables is taken to be similar across eddy variables, as it decays to zero outside of the region of eddy activity and occur over roughly same portion of the front (Fig. 3b-d). Therefore, averages in y are taken over the region where the value of $\overline{w'b'}$ at that y is greater than one-tenth of the maximum value of $\overline{w'b'}$ found anywhere in the domain. This region is taken to represent the eddying zone. Values of M^2 are also obtained by averaging over this region. With these assumptions, scaling relations for the tensor components independent of y and z may be sought, and vertical dependences of a parameterization are recovered by multiplying the scaling relations by their corresponding structure functions.

3.2. Diffusivity Tensor: Mixing Length Scales

A purely diffusive tensor can be cast in terms of autocorrelation and cross-correlation functions of Lagrangian parcel displacements (Taylor, 1921; Plumb, 1979; Plumb and Mahlman, 1987). Defining the horizontal and vertical displacements (η, ξ), this tensor is written

$$\mathbf{S} = \begin{bmatrix} \frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\eta^2}\right) & \frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\eta\xi}\right) \\ \frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\xi\eta}\right) & \frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\xi^2}\right) \end{bmatrix}. \tag{25}$$

It is appropriate to approximate $\frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\eta^2}\right) = \overline{v'\eta}$ by either of the expressions $\overline{v'^2}\delta t$ or $\sqrt{\overline{v'^2}}\delta y$, where $\overline{v'^2}$ is the Eulerian RMS eddy velocity and $(\delta t, \delta y)$ are unknown time and length scales, respectively. An unambiguous measurement of δt or δy is unlikely, since eddies here span an entire spectrum of wavelengths and time scales. One possibility is the statistical moments of the spectra to "measure" the energy-containing scales (Stammer, 1997; Scott and Wang, 2005; Tulloch et al., 2011). However, even if such measurements are made, none of this information about eddy scales would be available in a coarse OGCM run. Lacking turbulence statistics, geometric scaling considerations are needed to reproduce the missing quantities. The focus of the next section will be on an appropriate choice for the eddy length scale δy.

These mesoscale eddies are dominated by baroclinic instability, so one might assume that a reasonable length scale that approximates δy is the first baroclinic deformation radius, which is proportional to NH/f (Stone, 1972). The deformation radius appears as a crucial dimension in linear baroclinic instability, with unstable modes appearing near this scale. In nonlinear calculations, however, the zone of eddy activity quickly expands past the deformation radius. Green (1970) argues that a more suitable choice of length scale is the baroclinic zone width, which here is N^2H/M^2, since the horizontal scale of the eddies is limited by the availability of mean PE. In fact, in the real ocean the inverse cascade of horizontal KE is halted somewhere wider than the instability scale yet narrower than climatological gradients (Scott and Wang, 2005; Thompson and Young, 2007; Tulloch et al., 2011).

Therefore, given that the majority of the eddy kinetic energy (EKE) is trapped in a wavenumber range between the deformation radius and the time-evolving front width, it is convenient to choose one of these length scales and rely on nondimensional parameters to reconcile the difference. For the remainder of this paper the assumed scaling will be $\delta y \propto \frac{N^2H}{M^2}$, with a focus on using a power of Ri to improve the approximation. The front width, N^2H/M^2 is larger than the deformation radius by a factor of $\sqrt{Ri}$, so it is reasonable to expect that a good scaling for δy might involve a power of Ri between -0.5 and 0. Finally, since the eddies here (in constant stratification) extend the full depth of the water column, the vertical length scale δz is chosen be proportional to the full fluid depth H.

It is now straightforward to replace $(\eta, \xi, \frac{\partial}{\partial t})$ with the scalings above. In summary, the following choices have been made for these scales:

$$\eta \sim \delta y \propto \frac{N^2 H}{M^2}, \qquad \frac{\partial \eta}{\partial t} \sim \frac{\delta y}{\delta t} \propto \sqrt{\overline{v'^2}}, \tag{26}$$

$$\xi \sim \delta z \propto H, \qquad \frac{\partial \xi}{\partial t} \sim \frac{\delta z}{\delta t} \propto \sqrt{\overline{w'^2}}. \tag{27}$$

The resulting scaling for each component of $\mathbf{S}$ is

$$S_{yy} = \frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\eta^2}\right) \propto \frac{N^2 H}{M^2}\left(\sqrt{\overline{v'^2}}\right), \tag{28}$$

$$S_{yz} = S_{zy} = \frac{1}{2}\left(\overline{\eta\frac{\partial \xi}{\partial t}} + \overline{\xi\frac{\partial \eta}{\partial t}}\right) \propto \frac{H}{M^2}\left(\sqrt{\overline{v'^2}}M^2 + \sqrt{\overline{w'^2}}N^2\right), \tag{29}$$

$$S_{zz} = \frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\xi^2}\right) \propto H\left(\sqrt{\overline{w'^2}}\right), \tag{30}$$

so that

$$\mathbf{S} \propto \begin{bmatrix} \frac{N^2 H}{M^2}\left(\sqrt{\overline{v'^2}}\right) & \frac{H}{M^2}\left(\sqrt{\overline{v'^2}}M^2 + \sqrt{\overline{w'^2}}N^2\right) \\ \frac{H}{M^2}\left(\sqrt{\overline{v'^2}}M^2 + \sqrt{\overline{w'^2}}N^2\right) & H\left(\sqrt{\overline{w'^2}}\right) \end{bmatrix}. \tag{31}$$

The scaling (31) differs from the FFH-based scaling (24) in that eddy velocity scales appear explicitly.

3.3. Advective and Diffusive Tensor Scalings

FFH focuses on formulating a ψ such that the vertical buoyancy flux and extraction rate of mean potential energy (PE) are captured. They form a scaling law for $\overline{w'b'}$ by considering an exchange of fluid parcels over a decorrelation distance $(\Delta y, \Delta z)$ in time Δt, so that

$$\overline{w'b'} \propto \frac{\Delta z(\Delta y M^2 + \Delta z N^2)}{\Delta t}, \qquad \overline{v'b'} \propto \frac{\Delta y(\Delta y M^2 + \Delta z N^2)}{\Delta t}. \tag{32}$$

This approach is appropriate for scaling ψ only, since isoneutral diffusion does not affect the APE of the system (Griffies, 1998). Therefore, it suffices to consider the symmetric Taylor tensor to be contributing the diffusive part of the eddy flux. For the reasons explained above, FFH did not uniquely diagnose the symmetric and antisymmetric parts of the tensor (they used only buoyancy as a tracer), but Fox-Kemper and Ferrari (2008) did find that the residual horizontal flux, which could not be explained by ψ alone, was of similar magnitude to the flux capture by ψ. This implies that the symmetric and antisymmetric tensors might have similar off-diagonal components. Griffies (1998) notes that this occurs when the GM coefficient and the Redi isopycnal diffusivity are equal, a result consistent with the stochastic theory of Dukowicz and Smith (1997). The two-dimensional eddy transport tensor is thus

$$\mathbf{R} = \mathbf{S} + \mathbf{A} = \begin{bmatrix} \frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\eta^2}\right) & \frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\eta\xi}\right) \\ \frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\xi\eta}\right) & \frac{\partial}{\partial t}\left(\frac{1}{2}\overline{\xi^2}\right) \end{bmatrix} + \begin{bmatrix} 0 & \psi \\ -\psi & 0 \end{bmatrix}. \tag{33}$$

The scalings above assuming equal GM and Redi coefficients lead to

$$\mathbf{R} \propto \begin{bmatrix} \frac{N^2 H}{M^2}\left(\sqrt{v'^2}\right) & 0 \\ \frac{H}{M^2}\left(\sqrt{v'^2}M^2 + \sqrt{w'^2}N^2\right) & H\left(\sqrt{w'^2}\right) \end{bmatrix}, \tag{34}$$

$$\overline{v'b'} \propto N^2 H\left(\sqrt{v'^2}\right). \tag{35}$$

$$\overline{w'b'} \propto H\left(\sqrt{v'^2}M^2 + 2\sqrt{w'^2}N^2\right). \tag{36}$$

It may not be the case that GM and Redi coefficients exactly equal one another, but it will be assumed that the scaling for all the off-diagonal component of $\mathbf{S}$ and ψ is nonetheless the same. To confirm, it suffices to show that either R_{zy} dominates R_{yz}, or that $R_{zy} \approx -2\psi$. Fig. 4 shows that indeed the R_{yz} component is quite small and that the relationship between ψ and $\kappa_1\mathbf{S}$ is as predicted by GM, Redi, and Dukowicz and Smith (1997). Furthermore, Fig. 3 shows that the vertical structure function for ψ and $\kappa_1\mathbf{S}$ are very similar, *so the GM coefficient equals the Redi isopycnal diffusivity at every depth as well as in scale*. Below, the scaling for tensor element R_{yz} is taken to be the same as for R_{zy}, only the leading coefficient is found to be near zero. The results from the models support this point except at small Richardson numbers (< 2500), when we anticipate greater vertical excursions due to coherent vortices crossing the density surfaces (McDougall, 1987b). The above scalings form the basis for the scaling laws sought in the model runs, anticipating that a nondimensional numerical constant and potentially small powers of *Ri* or *Ro* will be necessary to optimize them.

The same logic provides the scalings here and in FFH, except here it is not assumed that the horizontal eddy velocity scales as the mean thermal wind velocity. Fig. 5 shows the time-evolving ratio of eddy to mean velocity varies during simulations and according to the initial Richardson number. Likewise, Fig. 5b shows that the eddy velocity slope differs from the isopycnal slope and with Richardson number. Therefore, EKE and *Ri* contain *distinct* information about the instantaneous state of the turbulence in a way that a naive scaling using only M^2 and N^2 does not. Using eddy velocity statistics in conjunction with powers of *Ri* lends accuracy to the scalings, but to realize this extra accuracy in a model requires successful parameterization of the EKE. Some authors have proposed prognostic methods for EKE (Eden and Greatbatch, 2008; Eden et al., 2009), and such approaches are common in engineering applications (Pope, 2000). The next section concludes with suggestions about how to incorporate these scaling ideas into a model depending on which diagnostics are available, and the implications that each has on the robustness of a parameterization.

4. Empirical Parameter Dependence of Tensor Components

As noted previously, dimensional scalings for each of the diffusivities do not rule out the possibility of dependence on nondimensional parameters. In particular, nondimensionalization of the Boussinesq, hydrostatic primitive equations (see Appendix A) reveals that the Rossby, and Richardson numbers are relevant in this problem. At the oceanic mesoscale the Reynolds and Peclet numbers are dynamically unimportant due to small viscosity and molecular diffusivity, and are not included in this analysis.

FFH suggest that the Rossby number inside the front becomes irrelevant as soon as the eddies expand beyond the front width in the horizontal, which occurs not long after finite amplitude is achieved. The results from these models results agree with this claim (not shown). The Richardson number, however, is central to the problem, and can be used to properly scale the buoyancy fluxes and elements of the transport tensor. Fig. 6 shows the dimensional scalings of Section 3 compared to the diagnosed values of each tensor element. The leading constant and power of Richardson number on each term are obtained by performing a logarithmic least-squares fit of the dimensional scalings to the true values of each element.

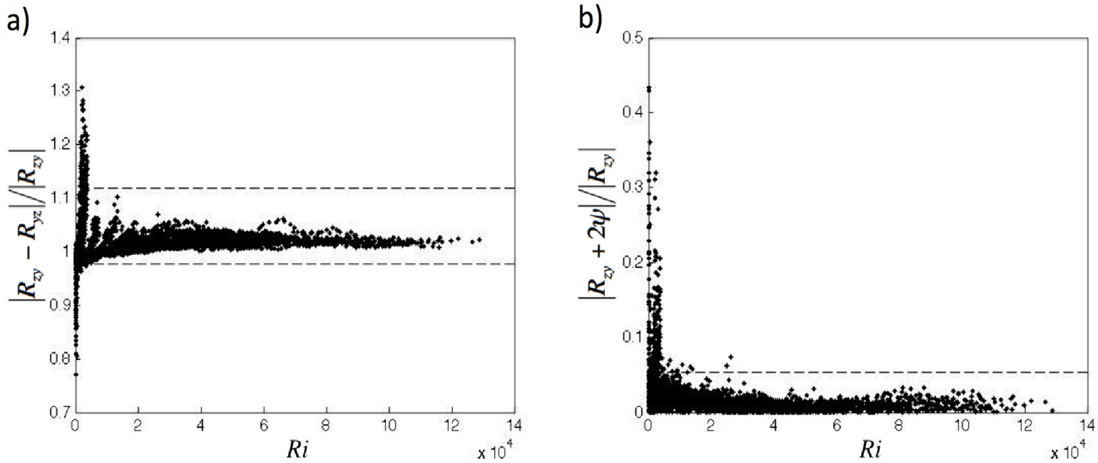

Figure 4: a) Comparison of R_{yz} and R_{zy}, confirming that R_{yz} is essentially equal to zero relative to R_{zy}. b) The equality suggested in Dukowicz and Smith (1997), $\psi = \kappa S$, is true to within 6%. in all snapshots except at small *Ri*. Dashed lines indicate the 95% confidence intervals.

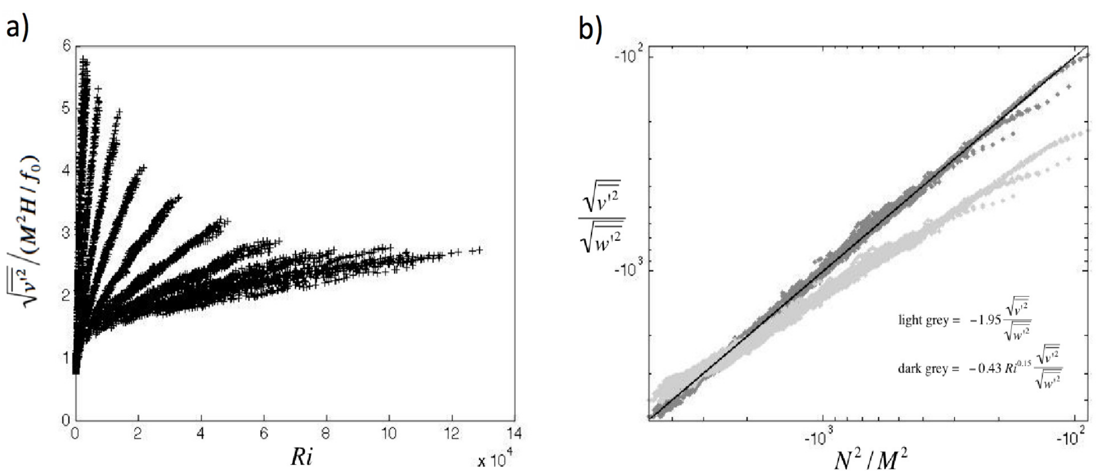

Figure 5: a) The horizontal RMS eddy velocity divided by the mean thermal wind. The degree to which a parameterization would suffer from approximating the eddy velocity with the mean velocity depends on the initial conditions. b) $\frac{\sqrt{v'^2}}{\sqrt{w'^2}} / \frac{N^2}{M^2} \propto Ri^{-0.15}$.

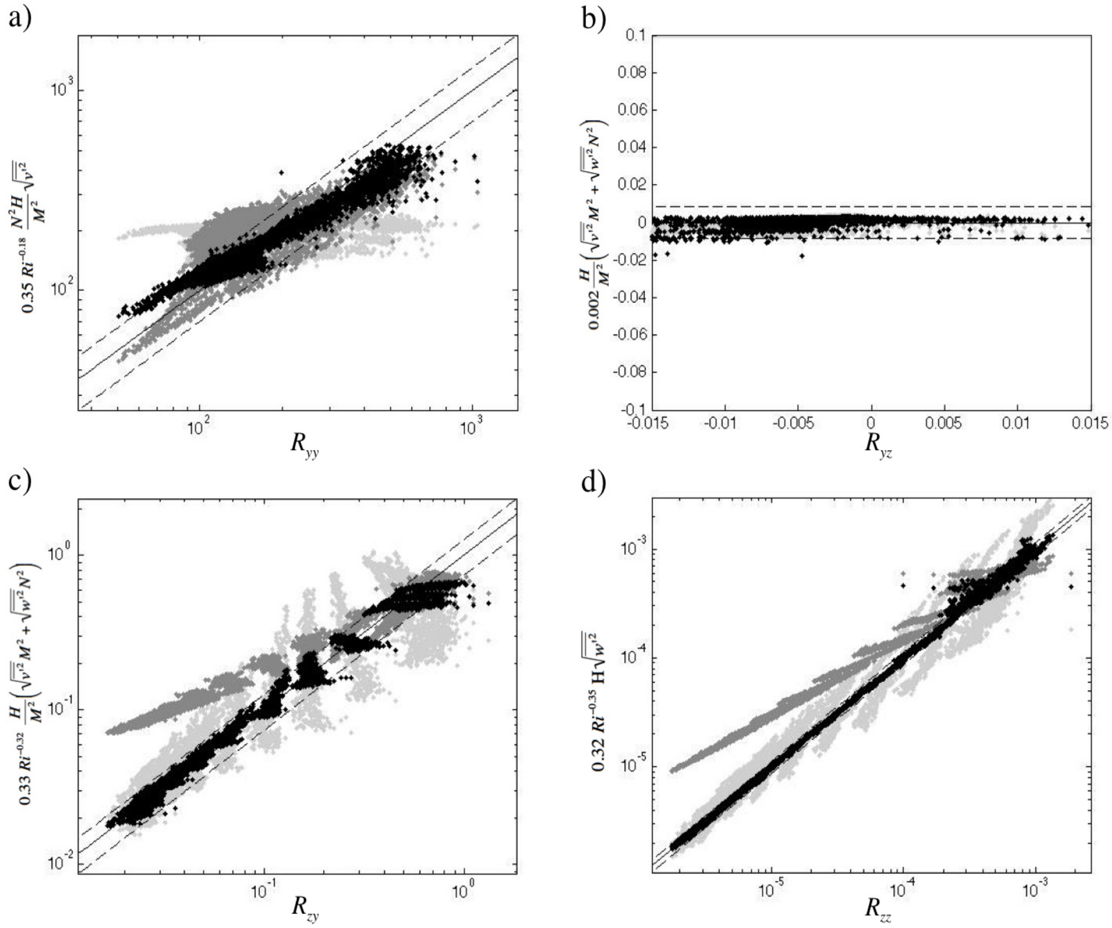

Figure 6: Diagnosed tensor components versus those parametrically scaled. a) $R_{yy,s}$, b) $R_{yz,s}$, c) $R_{zy,s}$, d) $R_{zz,s}$. In all panels black shows (34) with an empirical Ri correction, dark grey shows (34) without an empirical Ri correction, and light grey shows FFH scalings (24). Scalings are given in Tables 1-2. Dashed lines indicate 95% confidence intervals.

Quantity	Optimized (34) with Ri	Optimized (34) without Ri
$R_{yy,s}$	$(0.35 \pm 0.10)Ri^{-0.18\pm0.06}\frac{N^2H}{M^2}\left(\sqrt{v'^2}\right)$	$(0.07 \pm 0.05)\frac{N^2H}{M^2}\left(\sqrt{v'^2}\right)$
$R_{yz,s}$	$(0.002 \pm 0.01)\frac{H}{M^2}\left(\sqrt{v'^2}M^2 + \sqrt{w'^2}N^2\right)$	$(0.002 \pm 0.01)\frac{H}{M^2}\left(\sqrt{v'^2}M^2 + \sqrt{w'^2}N^2\right)$
$R_{zy,s}$	$(0.33 \pm 0.08)Ri^{-0.32\pm0.10}\frac{H}{M^2}\left(\sqrt{v'^2}M^2 + \sqrt{w'^2}N^2\right)$	$(0.03 \pm 0.07)\frac{H}{M^2}\left(\sqrt{v'^2}M^2 + \sqrt{w'^2}N^2\right)$
$R_{zz,s}$	$(0.32 \pm 0.03)Ri^{-0.35\pm0.03}H\left(\sqrt{w'^2}\right)$	$(0.03 \pm 0.08)H\left(\sqrt{w'^2}\right)$

Table 1: Empirical optimized tensor element scalings with eddy statistics corresponding to Fig. 6, Fig. 7, and (34), with 95% confidence intervals. Note that the reported error of the component diagnosis is 28% from the $R_{yy,s}$ leading coefficient uncertainty. Note also that based on the observed scatter there is no dependence on Ri for $R_{yz,s}$, and so we fix the exponent on Ri to be zero for that scaling.

Quantity	Optimized (24) with Ri	Optimized (24) without Ri
$R_{yy,s}$	$(0.77 \pm 0.75)Ri^{-0.18\pm0.18}\frac{N^2H^2}{\lvert f\rvert}$	$(0.17 \pm 0.48)\frac{N^2H^2}{\lvert f\rvert}$
$R_{yz,s}$	$(0.01 \pm 0.09)\frac{M^2H^2}{\lvert f\rvert}$	$(0.01 \pm 0.09)\frac{M^2H^2}{\lvert f\rvert}$
$R_{zy,s}$	$(0.85 \pm 0.77)Ri^{-0.22\pm0.23}\frac{M^2H^2}{\lvert f\rvert}$	$(0.17 \pm 0.68)\frac{M^2H^2}{\lvert f\rvert}$
$R_{zz,s}$	$(0.30 \pm 0.21)Ri^{-0.20\pm0.14}\frac{M^4H^2}{N^2\lvert f\rvert}$	$(0.06 \pm 0.07)\frac{M^4H^2}{N^2\lvert f\rvert}$

Table 2: Empirical optimized tensor element scalings without eddy statistics, i.e., following FFH, corresponding to Fig. 6, Fig. 7, and (24), with 95% confidence intervals. Note that the $R_{zy,s}$ agrees with the FFH estimate of $0.06 - 0.06M^2H^2/|f|$ if GM and Redi coefficients are the same so that $R_{zy,s} = 2\psi$. Note also that the coefficient for the FFH scaling is uncertain beyond ±97%.

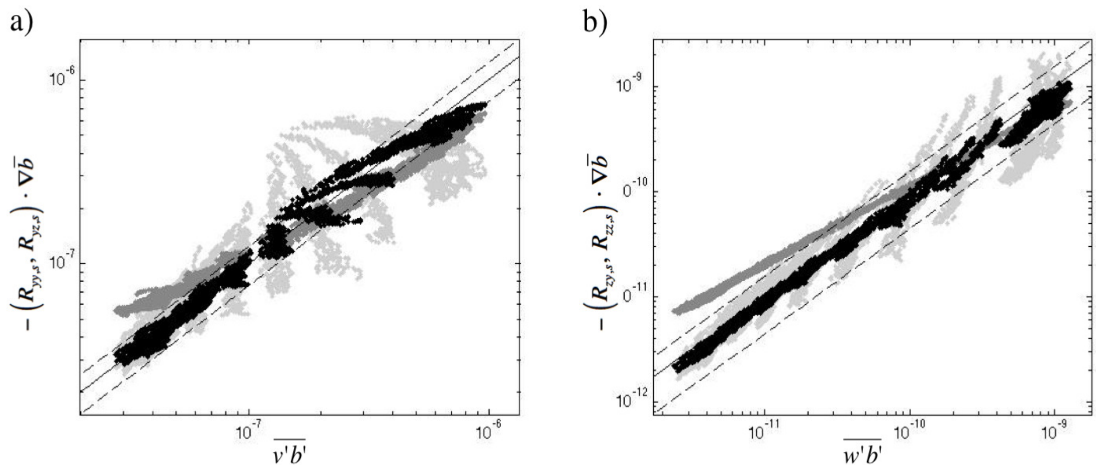

Figure 7: Reconstructions of a) $\overline{v'b'}$ and b) $\overline{w'b'}$ using the tensor scalings from Fig. 6 and Table 1-2. The solid lines indicate the true values of each flux. Black indicates reconstructions from the scalings from (34), with a power of *Ri*. 95% of the black data points are contained in the region bounded by the dashed lines. Dark grey indicates reconstructions from the scalings from (34), without a power of *Ri*; light grey indicates reconstructions from the FFH scalings (24) with a power of *Ri*.

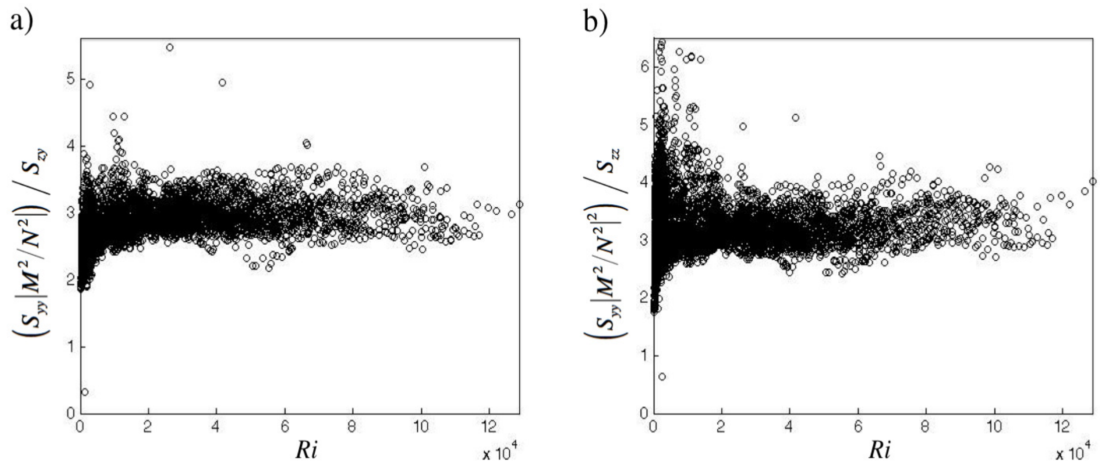

Figure 8: The ratio between S_{yy} multiplied by isopycnal slope and the offdiagonal elements of **S**. The Redi form of **S** would have that this ratio would be equal to one in both plots, which is not the case here. The observed ratio is greater than expected in comparison to the Redi along-isopyncal scaling by a factor of 3 on average. Thus, the eigenvectors of S are oriented at a slightly shallower angle than the isopycnal slope.

The symmetric part of the diagnosed tensor does not exactly agree with the Redi isoneutral diffusion tensor, which is, in 2D and making the small angle approximation,

$$\mathbf{S}_{Redi} \propto \begin{bmatrix} \kappa_1 & \kappa_1 \mathcal{S} \\ \kappa_1 \mathcal{S} & \kappa_1 \mathcal{S}^2 \end{bmatrix}, \tag{37}$$

where κ_1 is the along-isopycnal diffusivity and $\mathcal{S} = \left|M^2/N^2\right|$ is the absolute value of the local isopycnal slope. The S_{yy} component obtained from the model results tends to be a factor of 3 larger (Fig. 8), so that

$$\mathbf{S} \approx \begin{bmatrix} 3\kappa_1 & \kappa_1 \mathcal{S} \\ \kappa_1 \mathcal{S} & \kappa_1 \mathcal{S}^2 \end{bmatrix} \tag{38}$$

The time dependence of the spin-down problem can be used to understand both the excess of S_{yy} and the small diapycnal diffusivity κ_2. The buoyancy variance equation is

$$\frac{D\left(\frac{1}{2}\overline{b'^2}\right)}{Dt} = -\overline{\mathbf{u}'b'} \cdot \nabla\bar{b} - \overline{\mathbf{u}' \cdot \nabla\left(\frac{1}{2}\overline{b'^2}\right)} \tag{39}$$

Assuming the term on the far right is small and using the preceding nondimensionalization and scaling results, as well as a proportionality for the flux direction compared to the isopycnal slope direction (not shown),

$$\overline{v'b'} \cdot \nabla\bar{b} \propto N^2 M^2 H \sqrt{\overline{v'^2}}, \quad \overline{w'b'} \cdot \nabla\bar{b} \propto N^2 M^2 H \sqrt{\overline{v'^2}} Ri^{-0.09} \tag{40}$$

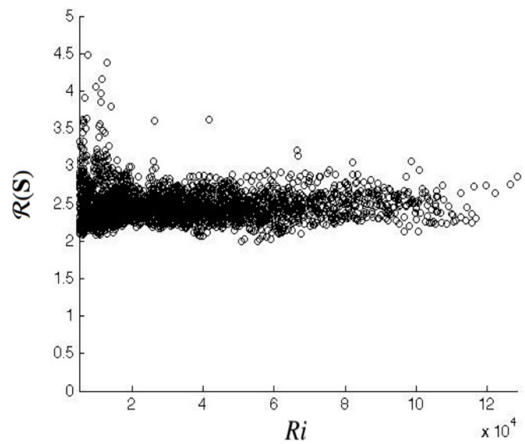

Figure 9: The ratio $\mathcal{R}(\mathbf{S})$ between the isopycnal slope and the slope of the diffusive flux. $\mathcal{R}(\mathbf{S})$ is above the value of 2 predicted by linear theory across the full range of *Ri* in the simulations.

Thus, the resulting scaled buoyancy variance budget is unable to be exactly along isopycnals for all *Ri*. The $\overline{w'b'}$ term's contribution to the variance decreases relative to the $\overline{v'b'}$ term as *Ri* increases. For the range of *Ri* in our simulations (from 136 to 128,760), the relative contribution of the $\overline{w'b'}$ term is from 1.6 to 2.9 times smaller, with a mean of 2.5. This is close to the excess of S_{yy} over Redi of 2.8.

It is often argued that κ_2 should be zero in steady, adiabatic situations (McDougall and McIntosh, 1996). In such situations the Redi flux should also be exactly along the isopycnals, matching the form of (37). The solutions here are nearly adiabatic, but they are not steady as both eddy variance and isopycnal slope evolve during the course of each simulation.The buoyancy flux clearly has some diapycnal flux, presumably associated with the neglected triple correlation and time dependence of eddy variance, since the terms on the right hand side do not balance exactly.

Linear theory suggests that the diffusive flux should be oriented at half the isopycnal slope to maximize potential energy extraction (Haine and Marshall, 1998). However, this result cannot be expected to hold precisely in a nonlinear, time-evolving simulation set. The model results suggest that the ratio of the isopycnal slope to the diffusive flux slope, calculated as

$$\mathcal{R}(\mathbf{S}) = \frac{\langle S \rangle}{\left\langle \frac{S_{zy}M^2 + S_{zz}N^2}{S_{yy}M^2 + S_{yz}N^2} \right\rangle}, \tag{41}$$

remains consistently close to 2.5 across the full spectrum of Richardson numbers in the simulations (Fig. 9). Here the angle brackets indicate an average over the eddying region defined in Appendix C. This value is consistent with the above finding that the growth of eddy variance is responsible for the small diapycnal component of the diffusive flux. It also suggests that the excessively large value of S_{yy} can largely be attributed to the flux being directed below the true isopycnal slope.

Many current OGCM's use a combination of the GM and Redi tensors in calculating diffusivities. The diffusivities are often set to be equal to each other (Griffies, 1998) for numerical convenience, leaving a full mixing tensor of the form

Option	$E(\overline{\mathbf{u}'b'})$	$\kappa_{GM-Redi}$
1	0.188	Table 1, second column, or (44)
2	0.270	GM-Redi with $\kappa_1 = 0.32 Ri^{-0.31} \frac{N^2 H}{M^4} \left(\sqrt{v'^2} M^2 + \sqrt{w'^2} N^2 \right)$, or (46)
3	0.551	Table 2, second column, or (45)
4	0.703	GM-Redi with $\kappa_1 = 0.58 Ri^{-0.22} \frac{N^2 H^2}{\lvert f \rvert}$, or (47)

Table 3: The relative errors based on different approximations to the scalings for **R**. These errors are calculated by averaging over the region defined in Appendix C.

$$\mathbf{R} = \mathbf{S} + \mathbf{A} \approx \begin{bmatrix} \kappa_1 & 0 \\ 2\kappa_1 \mathcal{S} & \kappa_1 \mathcal{S}^2 \end{bmatrix}. \tag{42}$$

Note that in the GM-Redi formulation κ_1 is described fully by the R_{yy} tensor element. All other elements are obtained by multiplying this value by some power of the isopycnal slope.

Overall, the principle result of this paper is a set of scalings for the eddy diffusivity tensor that vary according to the coarse-grained stratification. Each scaling represents the sum of classical Taylor diffusion with a GM-style skew diffusion, with leading constants and powers of Richardson numbers optimizing the results across a wide range of scales. A modeler wishing to incorporate the results of this research into a flow model is left with multiple choices: 1) use the "best" diagnosed **R** suggested by these model results, which would require some prescription of $\sqrt{v'^2}$ and $\sqrt{w'^2}$; 2) use a GM-Redi optimized version of these results, wherein a modeler could simply plug in a choice for κ calibrated from these results; 3) use an FFH-style κ, which would not require calculating $\sqrt{v'^2}$ or $\sqrt{w'^2}$; or 4) use a GM-Redi optimized version with an FFH-style κ. These options, with their corresponding relative errors as defined in (43), are shown in Table 3, with

$$E(\overline{\mathbf{u}'b'}) = \frac{\left| \overline{\mathbf{u}'b'} + \mathbf{R}\nabla\bar{b} \right|}{\left| \overline{\mathbf{u}'b'} \right|} \tag{43}$$

Option 1 is clearly optimal insofar as accuracy is concerned, but a GM-Redi diffusivity using the κ from option 2 is a good alternative. Any FFH-style implementations carry with them a nontrivial loss of accuracy, but the diagnostics required for these would be readily available in an OGCM.

In summary, the optimal full tensor, which uses the RMS eddy velocities $\sqrt{\overline{v'^2}}$ and $\sqrt{\overline{w'^2}}$, is

$$\mathbf{R} \propto \begin{bmatrix} 0.35 Ri^{-0.18} \frac{N^2 H}{M^2} \left(\sqrt{\overline{v'^2}} \right) & 0 \\ 0.33 Ri^{-0.32} \frac{H}{M^2} \left(\sqrt{\overline{v'^2}} M^2 + \sqrt{\overline{w'^2}} N^2 \right) & 0.32 Ri^{-0.35} H \left(\sqrt{\overline{w'^2}} \right) \end{bmatrix} \tag{44}$$

The best FFH-style full tensor, which uses only the buoyancy gradients, is

$$\mathbf{R} \propto \begin{bmatrix} 0.77Ri^{-0.18}\frac{N^2H^2}{|f|} & 0 \\ 0.85Ri^{-0.22}\frac{M^2H^2}{|f|} & 0.30Ri^{-0.20}\frac{M^4H^2}{N^2|f|} \end{bmatrix} \tag{45}$$

For codes using the GM/Redi tensor, the optimal choice for the diffusivity is given by

$$\kappa_1 = 0.32Ri^{-0.31}\frac{N^2H}{M^4}\left(\sqrt{\overline{v'^2}}M^2 + \sqrt{\overline{w'^2}}N^2\right). \tag{46}$$

If the eddy velocities required for this κ are not available, a good alternative is

$$\kappa_1 = 0.58Ri^{-0.22}\frac{N^2H^2}{|f|}. \tag{47}$$

5. Conclusion and Future Work

Spin-down of a frontal feature by geostrophic turbulence in the Eady problem has been simulated as part of a suite of challenges for parameterizations. Using a diagnostic method (Plumb and Mahlman, 1987) allows each component of the eddy transport tensor to be directly measured by utilizing a set of passive tracers in the flow field. Scalings have been derived for the advective and diffusive components of the transport tensor, and the vertical structure of each of these pieces has been found. We contend that all parameterizations used in GCMs should give similar fluxes when configured for a similar scenario.

The optimal scaling arguments used in this paper are motivated by classical Taylor diffusivity, with large diffusivity axis oriented along isopycnal slope as proposed by Redi, and a matching Gent-McWilliams skew flux with the same coefficient as the Redi isoneutral diffusivity as suggested by Dukowicz and Smith (1997). Using the eddy velocities directly in the scaling, rather than the mean velocities, reduces scatter about the scalings to within 28% rather than 97% found for the FFH scaling. However, a separate prognostic parameterization for the eddy velocities is required to provide these eddy statistics in coarse GCMs. These scalings may be useful in the Mesoscale Ocean Large Eddy Simulation (MOLES) regime, where eddy velocities can be directly diagnosed from the resolved eddies (Fox-Kemper and Menemenlis, 2008). Alternative scalings presented here, which are based on FFH and *do* approximate the eddy velocity by the mean velocity, offer the modeler a set of scalings which are likely easier to implement. That is, they use quantities readily available in GCMs, namely the horizontal and vertical buoyancy frequencies and total fluid depth. The RMS eddy velocities used in the optimized scalings offer improved accuracy if they are available, but forecasting these quantities in a model is an area of ongoing research (Eden and Greatbatch, 2008; Eden et al., 2009).

Scatter is further reduced by reducing temporal averaging of the eddy quantities during the simulations. Once the eddies are finite amplitude, their behavior is fully nonlinear. Utilizing these scalings in an OGCM approximates a subgridscale eddy field that is never in the linear growth phase, or whose kinetic energy does not dominate the mean. Conceptually, this regime is more familiar in comparison to observations where eddies are ubiquitous and powerful. The advantage of nonlinear simulation scalings is realism, but unlike linear theory the nonlinear results here favor empiricism over derivation.

This work leaves open many avenues for future research. More complex physical interactions involving submesoscale eddies, Rossby wave formation, frontogenesis, and deep convection might be explored. In the near future, the challenge suite will be expanded to include flows that better approximate the stratification and shear of the real ocean, whereupon the scalings shown here will be compared and re-evaluated.

Of principle interest in future research is the vertical structure of the advective and diffusive parts of the tensor. Here the scalings are derived by factoring out the vertical structures before the averaging, but they must be incorporated if a full picture of depth-dependent diffusivity is to emerge. Recent progress proposing vertical structures of eddy parameterizations (Ferrari et al., 2008, 2010) has brought physical thinking to dominate over the numerically-motivated tapering schemes of the past, but evaluation of these vertical structures in eddy-resolving models is insufficient at present.

The generality of the scalings here is limited to only the Eady-like flow configuration; only after a broader challenge suite is completed with many different configurations can such scalings be deemed robust. If these prove to be consistent across all the simulations in the suite, they may form the basis of a mesoscale eddy parameterization that goes beyond the traditional GM and Redi forms by specifying a flow-dependent κ_1. Ultimately, this series of fine-resolution models will provide a set of reliable scaling rules for many flow regimes that any new parameterization must obey.

6. Acknowledgement

BF-K and SB were supported by NSF OCE 0825614. Useful discussions with Frank Bryan, Peter Gent, and Keith Julien are happily acknowledged as improving this work.

Appendix A. Eady Spindown Configuration

The MITgcm (Marshall et al., 1997) is used to solve the Boussinesq, hydrostatic primitive equations on a $900 \times 150 \times 60$ grid. The dimensional form of these equations are

$$\frac{D\vec{\mathbf{v}}_h}{Dt} + f\vec{\mathbf{k}} \times \vec{\mathbf{v}}_h + \nabla_z \phi = 0 \tag{A.1}$$

$$\frac{Db}{Dt} = 0 \tag{A.2}$$

$$\nabla_z \cdot \vec{\mathbf{v}}_h + \frac{\partial w}{\partial z} = 0 \tag{A.3}$$

$$\frac{\partial \phi}{\partial z} = b \tag{A.4}$$

$$\rho = \rho_0 + \alpha(\theta - \theta_0) \tag{A.5}$$

where the material derivative $D/Dt = \partial/\partial t + u\partial/\partial x + v\partial/\partial y + w\partial/\partial z$ and the operator $\nabla_z = (\partial/\partial_x, \partial/\partial_y)$. In this equation set $\mathbf{v}_h$ is the horizontal velocity, f is the Coriolis parameter, $\phi = p/\rho_0$ is the density-normalized pressure, w is the vertical velocity, θ is the potential temperature, and α is a nondimensional constant. Consider the nondimensional scalings

$$x, y \sim L \qquad z \sim H \tag{A.6}$$

$$\phi \sim \phi_0 = \max(UfL, U^2) \qquad b \sim \frac{\phi_0}{H} \tag{A.7}$$

$$u, v \sim U \qquad t \sim \frac{L}{U} \tag{A.8}$$

$$w \sim \frac{\phi_0 U}{N^2 HL} \tag{A.9}$$

and the nondimensional parameters

$$Ro = \frac{U}{fL} \qquad\qquad Ri = \frac{N^2 f^2}{M^4}, \tag{A.10}$$

where $N^2 = b_z$ and $M^2 = b_y$. With the above scalings the nondimensionalized form of A.1 - A.3 becomes

$$\frac{\partial \tilde{\mathbf{v}}_h}{\partial \tilde{t}} + \tilde{\mathbf{v}}_h \cdot \tilde{\nabla}\tilde{\mathbf{v}}_h + \frac{1}{Ro\,Ri}\tilde{w}\frac{\partial \tilde{\mathbf{v}}_h}{\partial \tilde{z}} + \frac{1}{Ro}\tilde{f}\vec{\mathbf{k}} \times \tilde{\mathbf{v}}_h = -\frac{1}{Ro}\tilde{\nabla}\tilde{\phi} \tag{A.11}$$

$$\frac{\partial \tilde{b}}{\partial \tilde{t}} + \tilde{\mathbf{v}}_{\mathbf{h}} \cdot \tilde{\nabla}_z \tilde{b} + \frac{1}{Ro\,Ri}\tilde{w}\frac{\partial \tilde{b}}{\partial \tilde{z}} = 0 \tag{A.12}$$

$$\tilde{\nabla}_z \cdot \tilde{\mathbf{v}}_h + \frac{1}{Ro\,Ri}\frac{\partial \tilde{w}}{\partial \tilde{z}} = 0, \tag{A.13}$$

thus opening the possibility of scaling dependencies on Ro and Ri.

The Rossby adjustment simulations begin with a temperature front above a stratified interior. The initial stratification is

$$\bar{b} = N^2(z + H) + \frac{L_f M^2}{2}\tanh\left[\frac{2(y - y_0)}{Lf}\right] + b_0 \tag{A.14}$$

The channel is 300 m deep. The initial vertical stratification has parameters H,M, N, and L_f. The front width varies inversely with the Rossby number in the front, which itself varies from $O(10^{-4})$ to $O(10^{-1})$. Values for the time-varying Richardson number are found by averaging the local values of $N^2 f^2/M^4$ over the eddy activity region, which is defined as the part of the flow where $\overline{w'b'}$ is greater than one-tenth of the largest $\overline{w'b'}$ found at any y, z location in the domain. The actual values of Richardson number used in the analysis ranged from $O(10^3)$ to $O(10^5)$.

$$Ro = \frac{U}{fL_f} \tag{A.15}$$

$$L_f = \frac{U}{fRo} = \frac{M^2 H}{fRo} = \frac{H}{Ro} \tag{A.16}$$

QG linear instability solutions are used to tune the parameters so that the most unstable modes fit in the domain. The horizontal resolution tends to be very close to the first baroclinic deformation radius, consistent with the fact that baroclinic instability occurs at scales larger this (Nakamura, 1993).

Appendix B. Tracer Initialization

A zonal average is taken over all of the eddy variability in the along-channel direction and thus allows one to consider the results in the yz-axis only. The tracers are initialized pairwise so that the tracer gradients for each pair are orthogonal at time $t = 0$. Some experimentation was required to determine how to initialize the tracer fields so that their gradients would remain misaligned through the duration of each run. Bratseth (1998) initialized the tracers to be products of Chebyshev polynomials because of their wavelike characteristics on the interval $[-1, 1]$, and because higher-order polynomials would effectively sample smaller eddies.

Loosely following his approach, the tracer fields in this research are initialized sinusoidally. That is, the tracer concentrations at time $t = 0$ satisfy

$$\tau_{2i-1}(y,z) = \sin\left(i\frac{\pi y}{L_y}\right) \qquad \tau_{2i}(y,z) = \sin\left(i\frac{\pi z}{H}\right) \tag{B.1}$$

These tracer fields are qualitatively similar to Chebyshev polynomials in their obvious wavelike characteristics, but also because both sets of functions are orthogonal on a closed subset of $\mathbb{R}$. A set of tracers initialized with the gravest wavenumbers $i = \{1, 2, 3\}$ (Fig. B.1) was found to produce the best results. Initializing the tracers in this way maintained misalignment of the tracer gradients until the simulation stopping criterion was reached. This misalignment was measured by the accuracy to which the reconstructed buoyancy fluxes approximated the actual buoyancy fluxes. Initializing the tracers sinusoidally led to no significant degradation of this approximation by the time the front had slumped to the lateral wall.

The authors also experimented with initializing the tracers linearly, so that

$$\tau_{2i-1}(y,z) = \frac{i\,y}{L_y} \qquad \tau_{2i}(y,z) = \frac{i\,z}{H}. \tag{B.2}$$

However, this did not do as well to keep the tracer gradients misaligned through the duration of the runs. As a result, the reconstructed buoyancy fluxes were not as good of an approximation to the true buoyancy fluxes (not shown).

There is some subtlety in choosing how many tracers to implement. Using a large number of tracers, which by (B.1) initializes sine functions of large wavenumber, tends to accentuate progressively smaller eddies. Because the buoyancy is transported primarily by the largest eddies, such sampling suffers a loss of accuracy with respect to reconstructing the buoyancy fluxes (Fig. B.2). Ensemble averaging over all possible sets of n tracers did show an improvement in accuracy as more tracers were used; however, the postprocessing time increased substantially and no ensemble was as accurate as simply using a modest number of tracers at low wavenumber. Six tracers were determined to be sufficient for a good approximation of the buoyancy fluxes.

Appendix C. Model Setup and Diagnostics

To initialize the grid, a numerical solver is used to determine the linear growth rates of the baroclinic waves for the chosen stratification. The grid spacing is set roughly equivalent to the largest possible wavelength that is anticipated to play a role on in the slumping of the front; this is taken to be the smallest wave whose growth rate is at least 10% of the maximum growth rate in the domain. The front width is automatically scaled so that at least ten deformation radii lie within the front. This essentially guarantees that the effects of barotropic instability will be avoided, and that several gridpoints will be available to resolve each mesoscale eddy.

To be fair in the informing of an eddy parameterization, pointwise values of the transport tensor are not as useful as domain-averaged values, since the entire domain in these fine-resolution runs is *ipso facto* below the grid scale of a non-eddy-resolving model. Care must be taken in how one chooses to average these quantities, however. Averaging over the entire basin is the most realistic method for the sake of a parameterization, in that the horizontal grid is fixed in time. However, this leads to inconsistencies between quantities that are nonzero over the whole domain (such as M^2 and N^2) and those that are nonzero only in the ”eddying region” (such as the tensor components and tracer fluxes). At any given timestep the eddies will

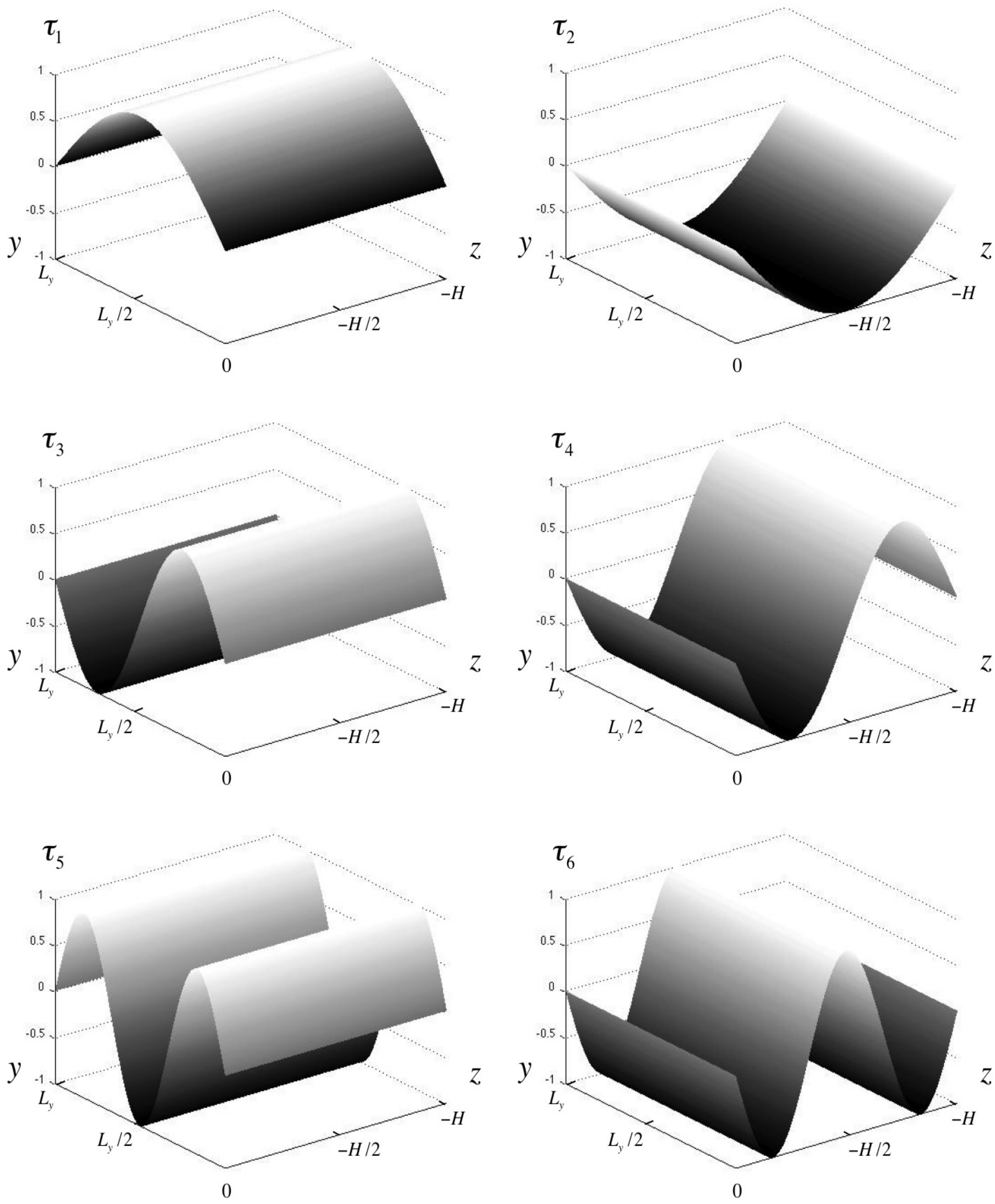

Figure B.1: Initialization of tracer fields in y and z. All fields were constant in the zonal direction.

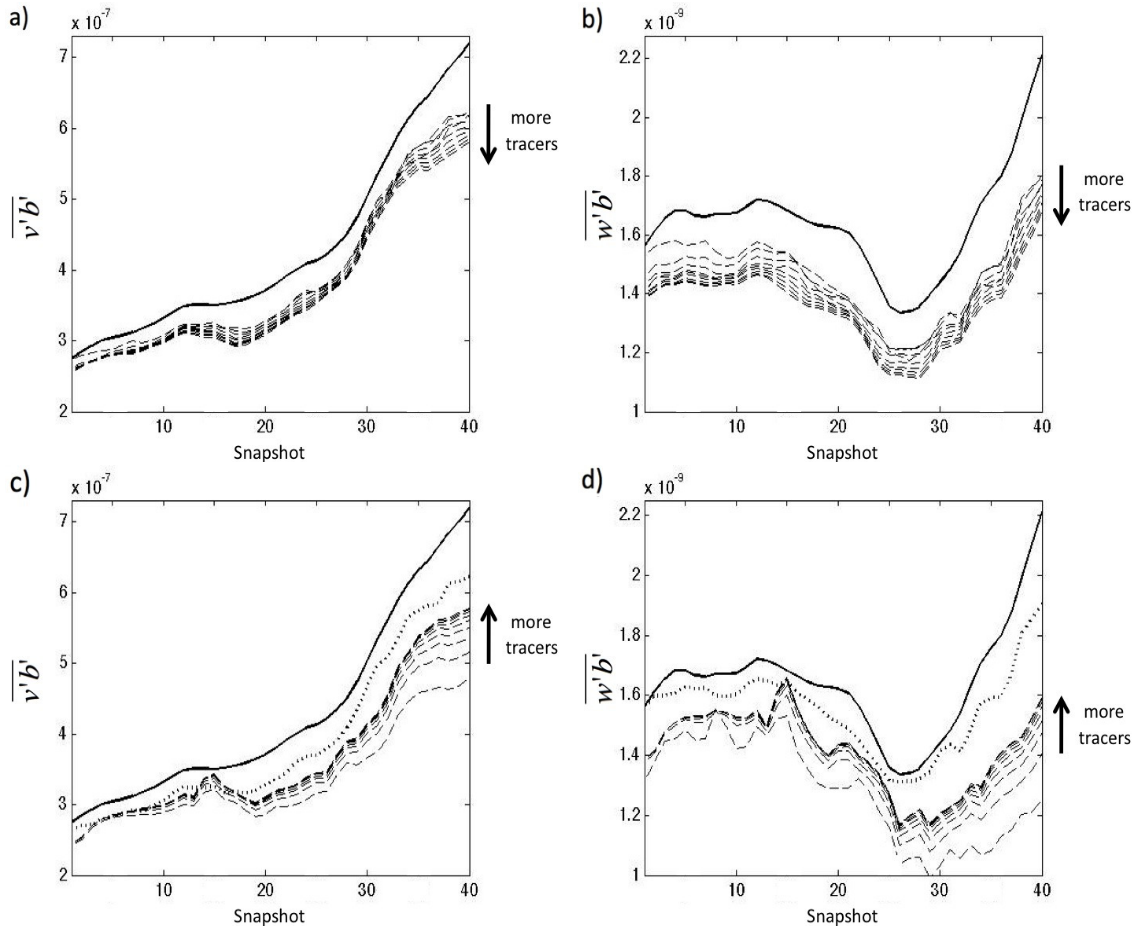

Figure B.2: (a-b) Example of a time series reconstruction of a) $\overline{v'b'}$ and b) $\overline{w'b'}$ during one simulation, using $\{6, 8, 10, \ldots, 20\}$ tracers. The solid black line is the true value of the flux, while the dashed lines are the reconstructions. The reconstruction for both fluxes becomes less accurate as more tracers are used. (c-d) The same time series reconstruction using an ensemble average over all possible sets of n tracers out of 20 total. In this case the reconstructions become more accurate as more tracers are used, but none are as accurate as using six tracers initialized at lowest wavenumber (dotted line).

be most active in those parts of the domain experiencing potential energy release; these quantities quickly taper to zero as one travels horizontally away from the center of the front, or vertically towards the top or bottom boundaries. The chosen solution was to average over those points of the domain where the vertical buoyancy flux $\overline{w'b'}$ is greater than 10% of its global maximum value. All averaging operations in this paper are taken over this region. In this way, the averaging only includes regions undergoing significant PE release, the overall area of which grows as energy cascades to graver scales.

Overall, sixty-nine model runs were conducted to generate the results in this paper. The time interval for each snapshot was chosen empirically based on the initial stratification so that each model run would generate about one hundred snapshots before completion. The stopping criteria for a given run was satisfied when the front slumped far enough so that the leading edges of it nearly reached the lateral boundary. This criteria was tracked at each snapshot by calculating the buoyancy perturbation, $\hat{b} = \bar{b} - \bar{b}_y$. The value of $\hat{b}$ at the center of the front was defined to be identically zero at the start of the run, so the frontal passage at any location was imminent when $\hat{b}$ began to approach zero there. The runs were halted when $\hat{b}$ became zero at a point ten gridpoints from the edge of the domain and along the bottom boundary.

These runs used free slip boundary conditions along the lateral and bottom boundaries. The strain-tension form of the viscous terms in the primitive equations was used, with the Laplacian Smagorinsky viscosity was set to be $\nu = (\Delta x/\pi)^2 \sqrt{(\nabla_k u_i + \nabla_i u_k)(\nabla_k u_i + \nabla_i u_k)/4}$. Implicit numerical diffusion and implicit viscosity were both used.

References

Andrews, D. G., McIntyre, M. E., 1978. An exact theory of nonlinear waves on a Lagrangian-mean flow. Journal of Fluid Mechanics 89, 609–646.

Branscome, L. E., 1983a. The charney baroclinic stability problem: Approximate solutions and modal structures. Journal of the Atmospheric Sciences 40, 1393–1409.

Branscome, L. E., 1983b. A parameterization of transient eddy heat flux on a beta-plane. Journal of the Atmospheric Sciences 40, 2508–2521.

Bratseth, A. M., 1998. On the estimation of transport characteristics of atmospheric data sets. Tellus 50A, 451–467.

Bryan, K., Dukowicz, J. K., Smith, R. D., 1999. On the mixing coefficient in the parameterization of bolus velocity. Journal of Physical Oceanography 29, 2442 - 2456.

Colas, F., Capet, X., McWilliams, J.C., Li, Z., 2012. Mesoscale eddy buoyancy flux and eddy-induced circulation in eastern-boundary upwelling systems. Journal of Physical Oceanography, submitted.

Danabasoglu, G., Ferrari, R., McWilliams, J. C., Mar. 2008. Sensitivity of an ocean general circulation model to a parameterization of near-surface eddy fluxes. Journal of Climate 21 (6), 1192–1208.

Danabasoglu, G., McWilliams, J., Dec. 1995. Sensitivity of the global ocean circulation to parameterizations of mesoscale tracer transports. Journal of Climate 8 (12), 2967–2987.

de Szoeke, R. A., Bennett, A. F., Oct. 1993. Microstructure fluxes across density surfaces. Journal of Physical Oceanography 23 (10), 2254–2264.

Dukowicz, J., Smith, R., 1997. Stochastic theory of compressible turbulent fluid transport. Physics of Fluids 9, 3523–3529.

Eady, E. T., 1949. Long waves and cyclone waves. Tellus 1, 33–52.

Eden, C., 2006. Thickness diffusivity in the Southern Ocean. Geophysical Research Letters 33, doi:10.1029/2006GL026 157.

Eden, C., Greatbatch, R. J., 2008. Towards a mesoscale eddy closure. Ocean Modelling 20 (3), 223–239.

Eden, C., Greatbatch, R. J., Olbers, D., 2007. Interpreting eddy fluxes. Journal of Physical Oceanography 37, 1282–1296.

Eden, C., Jochum, M., Danabasoglu, G., 2009. Effects of different closures for thickness diffusivity. Ocean Modelling 26 (1-2), 47–59.

Eden, C., 2010. Anisotropic rotational and isotropic residual isopycnal mesoscale eddy fluxes. Journal of Physical Oceangraphy 40 (11), 2511 - 2524.

Eden, C., 2011. A closure for mesoscale eddy fluxes based on linear instability theory. Ocean Modeling 39 (3-4), 362 - 369.

Ferrari, R., McWilliams, J. C., Canuto, V. M., Dubovikov, M., JUN 15 2008. Parameterization of eddy fluxes near oceanic boundaries. Journal of Climate 21 (12), 2770–2789.

Ferrari, R., Griffies, S. M., Nurser, A. J. G., Vallis, G. K., 2010. A boundary-value problem for the parameterized mesoscale eddy transport. Ocean Modelling 32 (3-4), 143–156.

Ferrari, R., Plumb, R. A., 2003. Residual circulation in the ocean. Proceedings of the 13th 'Aha Huliko'a Hawaiian Winter Workshop 13, 219–228.

Fox-Kemper, B., Ferrari, R., and Pedlosky, J., 2003. On the indeterminancy of rotational and divergent eddy fluxes. Journal of Physical Oceanography 33, 478-483.

Fox-Kemper, B., and Pedlosky, J., 2004. Wind-driven barotropic gyre I: Circulation control by eddy vorticity fluxes to an enhanced removal region, Journal of Marine Research 62, 169-193.

Fox-Kemper, B., Ferrari, R., 2008. Parameterization of mixed layer eddies. Part II: Prognosis and impact. Journal of Physical Oceanography 38 (6), 1166–1179.
URL `http://ams.allenpress.com/perlserv/?request=get-abstract&doi=10.1175%2F2007JPO3788.1`

Fox-Kemper, B., Ferrari, R., Hallberg, R., 2008. Parameterization of mixed layer eddies. Part I: Theory and diagnosis. Journal of Physical Oceanography 38 (6), 1145–1165.
URL `http://ams.allenpress.com/perlserv/?request=get-abstract&doi=10.1175%2F2007JPO3792.1`

Fox-Kemper, B., Bryan, F. O., Dennis, J. M., 2012a. Global diagnosis of the mesoscale eddy tracer flux-gradient relation. Ocean Modelling. In preparation.

Fox-Kemper, B., Lumpkin, R., Bryan, F. O., 2012b. Lateral transport in the ocean interior. In: Siedler, G., Church, J., Gould, J., Griffies, S. (Eds.), Ocean Circulation and Climate – Observing and Modelling the Global Ocean. Elsevier, submitted.
URL `http://tinyurl.com/FLB12`

Fox-Kemper, B., McWilliams, J. C., Bachman, S., Bryan, F. O., 2012c. The geometry of eddy diffusion and advection, in preparation.

Fox-Kemper, B., Menemenlis, D., 2008. Can large eddy simulation techniques improve mesoscale-rich ocean models? In: Hecht, M., Hasumi, H. (Eds.), Ocean Modeling in an Eddying Regime. Vol. 177. AGU Geophysical Monograph Series, pp. 319–338.

Gent, P. R., McWilliams, J. C., 1990. Isopycnal mixing in ocean circulation models. Journal of Physical Oceanography 20, 150–155.

Gent, P. R., Willebrand, J., McDougall, T. J., McWilliams, J. C., 1995. Parameterizing eddy-induced tracer transports in ocean circulation models. Journal of Physical Oceanography 25, 463–474.

Gille, S. T., Davis, R. E., 1999. The influence of mesoscale eddies on zonally averaged density: An examination of subgrid-scale parameterization. Journal of Physical Oceanography 29, 1109 - 1123.

Green, J. S. A., 1970. Transfer properties of the large-scale eddies and the general circulation of the atmosphere. Q. J. Roy. Meteor. Soc. 96, 157–185.

Griesel, A. S., Gille, S. T., Sprintall, J., McClean, J. L., Maltrud, M. E., 2009. Assessing eddy heat flux and its parameterization: A wavenumber perspective from a $1/10°$ ocean simulation. Ocean Modelling 29, 248 - 260. doi:10.1016/j.ocemod.2009.05.004.

Griffies, S. M., May 1998. The Gent-McWilliams skew flux. Journal of Physical Oceanography 28 (5), 831–841.

Griffies, S. M., 2004. Fundamentals of ocean climate models. Princeton University Press, Princeton, N.J. URL `http://www.loc.gov/catdir/enhancements/fy0654/2004043409-d.html`

Grooms, I., Julien, K., Fox-Kemper, B., 2011. On the interactions between planetary geostrophy and mesoscale eddies. Dynamics of Atmospheres and Oceans 51, 109–136. URL `http://dx.doi.org/10.1016/j.dynatmoce.2011.02.002`

Haine, T. W. N., Marshall, J.C., 1998. Gravitational, symmetric and baroclinic instability of the ocean mixed layer. Journal of Physical Oceanography 28, 634 - 658.

Hallberg, R., 2000. Time integration of diapycnal diffusion and richardson number-dependent mixing in isopycnal coordinate ocean models. Monthly Weather Review 128 (5), 1402–1419.

Jochum, M., Jahn, A., Peacock, S., Bailey, D. A., Fasullo, J. T., Kay, J., Levis, S., Otto-Bleisner, B., 2012. True to Milankovitch: glacial inception in the new community climate system model, Journal of Climate 25 (7), 2226-2239.

Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., Feely, R.A., Millero, F. J., Mordy, C., Peng, T.-H., 2004. A global ocean carbon climatology: Results from Global Data Analysis Project (GLODAP), Global Biogeochem. Cycles, 18, GB4031, doi:10.1029/2004GB002247.

Lee, M.-M., Marshall, D. P., Williams, R. G., 1997. On the eddy transfer of tracers: advective or diffusive? Journal of Marine Research 55, 201 - 222.

Marshall, J., Shutts, G., 1981. A note on rotational and divergent eddy fluxes. Journal of Physical Oceanography 11, 1677-1679.

Marshall, J., Adcroft, A., Hill, C., Perelman, L., Heisey, C., 1997. A finite-volume, incompressible Navier-Stokes model for studies of the ocean on parallel computers. Journal of Geophysical Research 102 (C3), 5753–5766.

Marshall, J., Shuckburgh, E., Jones, H., Hill, C., 2006. Estimates and implications of surface eddy diffusivity in the southern ocean derived from tracer transport. Journal of Physical Oceanography 36 (9), 1806–1821.

McClean, J. L., Maltrud, M. E., Bryan, F. O., 2006. Measures of the fidelity of eddying ocean models. Oceanography 19, 104–117.

McDougall, T. J., 1987a. Neutral surfaces. Journal of Physical Oceanography 17, 1950–1964.

McDougall, T. J., 1987b. The vertical motion of submesoscale coherent vortices across neutral surfaces. Journal of Physical Oceanography 17, 2334–2342.

McDougall, T. J., Jackett, D. R., 1988. On the helical nature of neutral trajectories in the ocean. Progress In Oceanography 20 (3), 153–183.

McDougall, T. J., McIntosh, P. C., 1996. The temporal-residual-mean velocity. Part I: Derivation and the scalar conservation equations. Journal of Physical Oceanography 26, 26532665.

Nakamura, N., 1993. An illustrative model of instabilities in meridionally and vertically sheared flows. Journal of Atmospheric Sciences 45, 3253 - 3267.

Nycander, J., Jan. 2011. Energy conversion, mixing energy, and neutral surfaces with a nonlinear equation of state. Journal of Physical Oceanography 41 (1), 28–41.

Plumb, R., 2002. Stratospheric transport. Journal of the Meteorological Society of Japan 80, 793–809.

Plumb, R. A., 1979. Eddy fluxes of conserved quantities by small-amplitude waves. Journal of the Atmospheric Sciences 36, 1699–1704.

Plumb, R. A., Mahlman, J. D., 1987. The zonally averaged transport characteristics of the GFDL general circulation/transport model. Journal of the Atmospheric Sciences 44, 298–327.

Pope, S. B., 2000. Turbulent flows. Cambridge University Press, Cambridge.
URL http://www.loc.gov/catdir/description/cam0210/99044583.html

Redi, M. H., 1982. Oceanic isopycnal mixing by coordinate rotation. Journal of Physical Oceanography 12, 1154–1158.

Rix, N. H., Willebrand, J., 1996. Parameterization of mesoscale eddies as inferred from a high-resolution circulation model. Journal of Physical Oceanography 26, 2281-2285.

Roberts, M. J., Marshall, D. P., 1998. Do we require adiabatic dissipation schemes in eddy-resolving models? Journal of Physical Oceanography 28, 2050–2063.

Roberts, M. J., Marshall, D. P., 2000. On the validity of downgradient eddy closures in ocean models. Journal of Geophysical Research 105 (C12), 28,613 - 28,627.

Scott, R. B., Wang, F., 2005. Direct evidence of an oceanic inverse kinetic energy cascade from satellite altimetry. Journal of Physical Oceanography 35, 1650–1666.

Stammer, D., 1997. Global characteristics of ocean variability estimated from regional TOPEX/Poseidon altimeter measurements. Journal of Physical Oceanography 27, 1743–1769.

Stevenson, S., Fox-Kemper, B., Jochum, M., Rajagopalan, B., Yeager, S. G., 2010. ENSO model validation using wavelet probability analysis. Journal of Climate 23, 5540–5547.
URL http://journals.ametsoc.org/doi/abs/10.1175/2010JCLI3609.1

Stone, P. H., 1972. On non-geostrophic baroclinic stability: Part III. The momentum and heat transports. Journal of the Atmospheric Sciences 29, 419–426.

Taylor, G. I., 1921. Diffusion by continuous movements. Proceedings of the London Mathematical Society 20, 196–212.

Thompson, A. F., Young, W. R., 2007. Two-layer baroclinic eddy heat fluxes: Zonal flows and energy balance. Journal of the Atmospheric Sciences 64, 3214–3231.

Tulloch, R., Marshall, J., Hill, C., Smith, K. S., Jun. 2011. Scales, growth rates, and spectral fluxes of baroclinic instability in the ocean. Journal of Physical Oceanography 41 (6), 1057–1076.

Veronis, G., 1975. The role of models in tracer studies. In: Numerical Models of Ocean Circulation: Proceedings of a Symposium Held at Durham, New Hampshire, October 17-20, 1972. National Academy of Sciences, pp. 133–146.

Wittenberg, A. T., 2009. Are historical records sufficient to constrain ENSO simulations? Geophysical Research Letters 36, L12702. doi: 10.1029/2009GL038710.

Zika, J. D., McDougall, T. J., Sloyan, B. M., Jan. 2010. A tracer-contour inverse method for estimating ocean circulation and mixing. Journal of Physical Oceanography 40 (1), 26–47.