ADVANCING EARTH AND SPACE SCIENCES

JAMES | Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2024MS004506

Key Points:
- We derive and discuss the arbitrarily averaged hydrostatic Boussinesq ocean equations in generalized vertical coordinates
- Known results for Eulerian- and isopycnal-mean equations in depth and isopycnal coordinates are recovered as special cases
- The properties of eddy terms depend on the averaging choice, and are not always consistent with current parameterization implementations

Correspondence to:
M. F. Jansen,
mfj@uchicago.edu

Citation:
Jansen, M. F., Adcroft, A., Griffies, S. M., & Grooms, I. (2024). The averaged hydrostatic Boussinesq ocean equations in generalized vertical coordinates. *Journal of Advances in Modeling Earth Systems*, *16*, e2024MS004506. https://doi.org/10.1029/2024MS004506

Received 12 JUN 2024
Accepted 14 NOV 2024

The Averaged Hydrostatic Boussinesq Ocean Equations in Generalized Vertical Coordinates

Malte F. Jansen[1], Alistair Adcroft[2], Stephen M. Griffies[2,3], and Ian Grooms[4]

[1]Department of the Geophysical Sciences, The University of Chicago, Chicago, IL, USA, [2]Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA, [3]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, [4]Department of Applied Mathematics, University of Colorado, Boulder, CO, USA

Abstract Due to their limited resolution, numerical ocean models need to be interpreted as representing filtered or averaged equations. How to interpret models in terms of formally averaged equations, however, is not always clear, particularly in the case of hybrid or generalized vertical coordinate models, which limits our ability to interpret the model results and to develop parameterizations for the unresolved eddy contributions. We here derive the averaged hydrostatic Boussinesq equations in generalized vertical coordinates for an arbitrary thickness-weighted average. We then consider various special cases and discuss the extent to which the averaged equations are consistent with existing ocean model formulations. As previously discussed, the momentum equations in existing depth-coordinate models are best interpreted as representing Eulerian averages (i.e., averages taken at fixed depth), while the tracer equations can be interpreted as either Eulerian or thickness-weighted isopycnal averages. Instead we find that no averaging is fully consistent with existing formulations of the parameterizations in semi-Lagrangian discretizations of generalized vertical coordinate ocean models such as MOM6. A coordinate-following average would require "coordinate-aware" parameterizations that can account for the changing nature of the eddy terms as the coordinate changes. Alternatively, the model variables can be interpreted as representing either Eulerian or (thickness-weighted) isopycnal averages, independent of the model coordinate that is being used for the numerical discretization. Existing parameterizations in generalized vertical coordinate models, however, are not always consistent with either of these interpretations, which, respectively, would require a three-dimensional divergence-free eddy tracer advection or a form-stress parameterization in the momentum equations.

Plain Language Summary Numerical ocean models represent continuous three-dimensional physical fields using discrete data points and hence cannot adequately represent variability at all scales. Instead model variables need to be interpreted as filtered versions of the continuous physical fields. We here derive the evolution equations for horizontally filtered fields, where the horizontal filtering follows arbitrary surfaces, with the only requirement being that the surfaces do not fold over (i.e., they are iso-surfaces of a field that is monotonic in depth). The equations for the filtered variables are formulated using an arbitrary vertical coordinate system, thus making them applicable to a wide range of different numerical ocean models. We then consider different physically motivated choices for the averaging surfaces and express the equations using different common choices for the vertical coordinate system. We conclude by discussing which equation sets are consistent with the equations solved by existing models and/or which modifications are needed to achieve consistency. The results have important implications for ocean model development, because the models need to include "parameterizations" to represent the effect of motions that have been removed by the filter. The formulation of these parameterizations needs to be consistent with the filtering operation that was assumed to obtain the model equations.

1. Introduction

Due to their limited resolution, numerical ocean models cannot resolve the full spectrum of motions. This limitation is widely recognized and motivates the need for so-called parameterizations (or closures), that is, additional terms added to the model's equations that are meant to capture the effect of dynamics that cannot be explicitly resolved by the models.

We can formalize the separation into resolvable and sub-grid-scale motions by applying a suitable averaging to the equations of motion, such that the averaged variables are sufficiently smooth to be explicitly resolvable. Due to the nonlinearity of the equations of motion the equations for the averaged quantities retain terms that involve deviations from the average (i.e., contributions from the unresolved flow), which then need to be parameterized. Applying a formal averaging to the equations of motion allows us to identify (a) what the model variables are meant to represent and (b) what needs to be captured by the parameterizations. An explicitly defined averaging procedure therefore allows us to cleanly compare the model results against observations and to devise and test parameterizations for the unresolvable eddy terms.

1.1. Types of Averages

In practice it is often not clear what kind of average is suitable, and many averaging procedures have been suggested. Existing averaging procedures may be broadly organized into three categories: (a) Reynolds averaging; (b) (non-Reynolds) spatial filtering; (c) grid-cell averaging as part of the numerical discretization. The third category applies most immediately to finite-volume methods, where the discrete model variables by construction represent grid-box or grid-face averages (e.g., Griffies et al., 2020). If this is the only averaging considered, the nature of the "sub-grid" terms that need to be parameterized depends directly on choices of the numerical discretization, which complicates the development of adequate parameterizations. Moreover, full resolution of variability on all scales that can be represented by the model grid is an unrealistic goal, given imperfect numerical methods. We will therefore here focus on the first two approaches, where the continuous equations are averaged explicitly before the discretization. The goal of the averaging is to obtain fields that are relatively smooth at the grid-scale, such that the results are less sensitive to the details of the numerical discretization, although we notice that a second grid-cell averaging (of the filtered equations) may still be performed as part of a finite-volume discretization.

Eddy parameterizations are often motivated by the eddy terms arising in the Reynolds-averaged equations. Reynolds averages are defined by the convenient property that $\overline{\bar{a}b} = \bar{a}\bar{b}$, where the overbar denotes the averaging operator, which is commonly taken to be a time-mean (e.g., Gent et al., 1995; McDougall & McIntosh, 2001), a zonal-mean in a zonally re-entrant domain (e.g., Bachman & Fox-Kemper, 2013) or an ensemble mean over different possible realizations of the turbulent flow (e.g., Maddison & Marshall, 2013; Uchida et al., 2022; Young, 2012). However, it has also long been acknowledged that models are fundamentally limited in their ability to represent small spatial scales, such that a spatial filtering with a finite-size filter-stencil would be the more appropriate averaging operator (Fox-Kemper & Menemenlis, 2008; McDougall & McIntosh, 2001). Spatial filters are generally not Reynolds operators (e.g., Buzzicotti et al., 2023) and can only be approximated as such in the presence of a clear scale separation, which does not generally exist in the ocean. Fortunately, many of the results obtained using Reynolds averaging readily carry over to spatial filters, as will be made explicit for the results of this manuscript.

1.2. Averaging Coordinate Surfaces

Independent of whether the average is in space, time, or over an ensemble, different choices can be made for the averaging coordinate, with the most widely used approaches being "Eulerian" averages, where the average is taken at fixed height (e.g., Bachman & Fox-Kemper, 2013), and isopycnal averages, where the average is taken along a surface of constant potential or neutral density (e.g., Young, 2012). For isopycnal averages we can further employ thickness-weighted and non-weighted averages.

Although it is natural to interpret the model variables as representing averages along the model coordinate, this is not the only possible interpretation. Indeed, McDougall and McIntosh (2001) argued that the formulation of the Gent and McWilliams (1990) (here after GM) parameterization used in many z-coordinate ocean models is not fully consistent with the form of the eddy terms in the Eulerian averaged tracer equations. For better consistency, they argued that the model's tracer variables should instead be interpreted as thickness-weighted isopycnal averages (with the average taken along the isopycnal whose mean depth equals the depth of the model level), while the model velocities represent Eulerian averages. However, this interpretation has not been consistently adopted (e.g., Bachman & Fox-Kemper, 2013; Zanna & Bolton, 2020).

Similarly, Loose et al. (2023) pointed out that the common formulation of isopycnal ocean models is not fully consistent with an isopycnally averaged interpretation, neither with nor without thickness weighting. Specifically, the eddy terms in the momentum and tracer equations become non-conservative when using a non-thickness weighted isopycnal average, which is both undesirable and inconsistent with the existing parameterization approaches. A thickness-weighted average in turn yields no eddy term in the continuity equation, which is inconsistent with the appearance of a parameterized eddy volume flux in isopycnal models. Instead, thickness-weighting introduces an eddy pressure gradient contribution (related to the form stress) in the momentum equations. Loose et al. (2023) propose that this term can be parameterized in the form of a vertical viscosity, following the ideas of P. Rhines and Young (1982) and Greatbatch and Lamb (1990), which in practice yields similar results as the existing GM-like parameterization. However, this approach is not presently used in any operational ocean models. In summary, the question of how to formally interpret the numerical model equations in terms of averaged equations remains not fully settled in both z-coordinate and isopycnal ocean models.

Matters are further complicated by the rise of hybrid or "generalized" vertical coordinate models, such as the Hybrid Coordinate Model (HYCOM, Bleck, 2002; Chassignet et al., 2007), the General Estuarine Transport Model (GETM, Hofmeister et al., 2010), the ocean component of the Model for Prediction Across Scales (MPAS-Ocean, Petersen et al., 2015), or the Modular Ocean Model 6 (MOM6, Adcroft et al., 2019). In MOM6 and HYCOM the time discretization is semi-Lagrangian, which allows for wide flexibility in the choice of target vertical coordinate (Griffies et al., 2020). The MOM6 configuration employed in the Geophysical Fluid Dynamics Laboratory's current global climate and Earth System models uses a hybrid target coordinate, following potential density surfaces over most of the ocean's interior while transitioning to depth coordinates near the surface and in unstratified regions (Adcroft et al., 2019). A similar hybrid vertical coordinate is used in HYCOM, where the vertical coordinate further transitions to being terrain-following over shallow coastal seas (Chassignet et al., 2007). How to interpret these models in terms of formally averaged equations remains largely unexplored.

1.3. Manuscript Overview

In this manuscript we derive the averaged dynamical equations in arbitrary vertical coordinates and discuss implications for the interpretation of numerical model variables and the need for parameterizations. We start in Section 2 by deriving the averaged hydrostatic Boussinesq equations in arbitrary ("generalized") vertical coordinates, with the average following the vertical coordinate. We focus on a generalized thickness-weighted average (TWA), which is the only conservative average in a non-Eulerian coordinate system (cf. Loose et al., 2023). In the special case where the vertical coordinate is height, z, the generalized thickness-weighting has no effect, such that the traditional Eulerian averaged equations are retained as a special case of the generalized TWA equations. We notice that the generalized TWA equations have previously been derived by Maddison and Marshall (2013) using a more formal mathematical approach, and are largely identical to the isopycnal thickness-weighted-average (TWA) equations as discussed in Young (2012).

In Section 3 we consider the case where the averaging surfaces do not line up with the vertical coordinate used to express the final averaged equations. This most general case is derived by transforming the averaged equations obtained in Section 2 into a second arbitrary vertical coordinate system. Armed with this general result we consider various special cases in Section 4, some of which have been discussed before. For example, the exact residual mean equations are obtained as a special case of the general equations if the averaging coordinate is chosen as potential density and the averaged equations are expressed using a vertical coordinate that represents the averaged height of isopycnals. In Section 5 we compare the derived equation sets to those solved in numerical models, to determine which interpretations, if any, are most consistent with the existing model formulations, and to identify which changes to the models may improve consistency with the equations. We conclude in Section 6 with a discussion of the main results and conclusions.

2. The Generalized Thickness-Weighted Averaged Equations

2.1. The Hydrostatic Boussinesq Equations in Generalized Vertical Coordinates

In this paper we focus on the hydrostatic Boussinesq equations, which are used in most global ocean models. However, we note that all results are generalizable to the compressible equations since the hydrostatic non-Boussinesq equations with pressure-based vertical coordinates are isomorphic to the hydrostatic Boussinesq equations in z-coordinates (Marshall et al., 2004). A summary of symbols used throughout this manuscript is provided in Table 1.

10.1029/2024MS004506

Table 1
Summary of Select Symbols Used in the Text

Symbol	Meaning
c	Arbitrary scalar tracer (including temperature and salinity)
$\mathbf{v}$	3D velocity vector: $\mathbf{v} = u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$
$\mathbf{u}$	Horizontal velocity vector: $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$
∇	3D nabla operator: $\nabla = \mathbf{i}\partial_x\|_z + \mathbf{j}\partial_y\|_z + \mathbf{k}\partial_z$
∇_r	"Horizontal" nabla operator with derivatives at fixed r: $\nabla = \mathbf{i}\partial_x\|_r + \mathbf{j}\partial_y\|_r$
z_r	Generalized "thickness" w.r.t. coordinate r: $z_r \equiv \partial_r z$
$\dot{r}$	Lagrangian rate of change of "r": $\dot{r} \equiv Dr/Dt$
$\dot{c}$	Tracer tendencies due to diffusion and sources or sinks
$\mathcal{F}_{x/y}$	Zonal/meridional accelerations due to frictional and external forces
$\overline{(\cdot)}^r$	Average at fixed r
$\widehat{(\cdot)}^r$	Generalized thickness-weighted average at fixed r: $\widehat{(\cdot)}^r \equiv \overline{z_r(\cdot)}^r/\overline{z_r}^r$
$\overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{c}/\mathbf{u}/\mathbf{v}}}^r$	Eddy tracer/momentum flux divergence (Equations 13 and 14)
$\overline{\nabla}^r \cdot \widehat{\mathbf{E}^{\mathbf{u}/\mathbf{v}}}^r$	Eliassen-Palm (EP) flux divergence (Equations 16 and 17)
$\dot{r}^{\#a}$	Lagrangian rate of change of r following a-averaged flow (Equation 22)
$w^{\#b}$	"Residual" vertical velocity for isopycnal TWA (Equation 52)
$b^{\#}(x,y,z,t)$	Buoyancy surface whose mean height is z, that is, $\overline{z}^b\left(x,y,b^{\#}(x,y,z,t),t\right) = z$
$\widehat{(\cdot)}^{b^{\#}}$	Thickness-weighted average along buoyancy surface $b = b^{\#}$ (Equation 53)
$\widetilde{(\cdot)}$	Denotes a model variable
$\tau^{u/v}$	Zonal/merional component of (parameterized) stress tensor
$\mathbf{D}$	(parameterized) Diffusivity tensor
$\mathbf{v}_{GM}$	(parameterized) 3D Eddy advective velocity (following Gent and McWilliams, 1990)
$\mathbf{A}$	(parameterized) Skew flux tensor (Equation 72)
$\mathbf{u}_{GM}$	(parameterized) 2D eddy advective velocity

The hydrostatic Boussinesq equations in generalized vertical coordinates can be written in flux form as

$$\partial_t|_r z_r + \nabla_r \cdot (z_r \mathbf{u}) + \partial_r\left(z_r \dot{r}\right) = 0 \tag{1}$$

$$\partial_t|_r (z_r c) + \nabla_r \cdot (z_r \mathbf{u} c) + \partial_r\left(z_r \dot{r} c\right) = z_r \dot{c}. \tag{2}$$

$$\partial_t|_r (z_r u) + \nabla_r \cdot (z_r \mathbf{u} u) + \partial_r\left(z_r \dot{r} u\right) - f z_r v = -z_r \rho_0^{-1} \partial_x|_z p + z_r \mathcal{F}_x \tag{3}$$

$$\partial_t|_r (z_r v) + \nabla_r \cdot (z_r \mathbf{u} v) + \partial_r\left(z_r \dot{r} v\right) + f z_r u = -z_r \rho_0^{-1} \partial_y|_z p + z_r \mathcal{F}_y \tag{4}$$

$$\partial_r p = -\rho g\, \partial_r z, \tag{5}$$

where r is a generalized vertical coordinate (assumed to be monotonic in z), $z_r \equiv \partial_r z$ is the generalized thickness (also the Jacobian of the transformation between z-coordinates and r-coordinates), $\mathbf{u} \equiv u\,\mathbf{i} + v\,\mathbf{j}$ is the horizontal velocity, and ∇_r denotes a horizontal nabla operator with derivatives taken at fixed r: $\nabla_r \equiv \mathbf{i}\partial_x|_r + \mathbf{j}\partial_y|_r$. $\mathcal{F}_{x/y}$ represents accelerations by friction and external forces. (Notice that the unit vectors $\mathbf{i}$ and $\mathbf{j}$ are here parallel to surfaces of constant height and independent of the choice of r; (e.g., see Figure 2 in McDougall et al., 2014).) The horizontal pressure gradient at fixed z can be written as

$$\partial_x|_z p = \partial_x|_r p + \rho\, g \partial_x|_r z \tag{6}$$

and

$$\partial_y|_z p = \partial_y|_r p + \rho\, g \partial_y|_r z \tag{7}$$

in generalized vertical coordinates. The identities in Equations 6 and 7 will be used in the following derivations. However, we note that in numerical models the pressure gradient acceleration is often not evaluated in the specific form of the two terms on the RHS of Equations 6 and 7, which often cancel at leading order (e.g., Adcroft et al., 2008; Bleck, 2002; Lin, 1997). Equation 2 is a placeholder for multiple tracer equations, with temperature and salinity being the dynamically active tracers, which can be used to compute density via a suitable equation of state (i.e., $\rho = \tilde{\rho}(S, T, z)$). We will here only provide limited discussion of how to approximate the equation of state in terms of averaged quantities (see McDougall and McIntosh (1996), Brankart (2013), and Stanley et al. (2020)).

2.2. Averaged Equations

Taking an average at fixed r of Equations 1–5 yields

$$\partial_t|_r \overline{z_r}^r + \nabla_r \cdot (\overline{z_r}^r \hat{\mathbf{u}}^r) + \partial_r\left(\overline{z_r}^r \widehat{\dot{r}}^r\right) = 0 \tag{8}$$

$$\partial_t|_r (\overline{z_r}^r \hat{c}^r) + \nabla_r \cdot (\overline{z_r}^r \hat{\mathbf{u}}^r \hat{c}^r) + \partial_r\left(\overline{z_r}^r \widehat{\dot{r}}^r \hat{c}^r\right) = -\overline{z_r}^r \overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{c}}}^r + \overline{z_r}^r \widehat{\dot{c}}^r \tag{9}$$

$$\partial_t|_r (\overline{z_r}^r \hat{u}^r) + \nabla_r \cdot (\overline{z_r}^r \hat{\mathbf{u}}^r \hat{u}^r) + \partial_r\left(\overline{z_r}^r \widehat{\dot{r}}^r \hat{u}^r\right) - f \overline{z}^r \hat{v}^r = -\overline{z_r}^r \overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{u}}}^r - \rho_0^{-1} \overline{z_r} \widehat{\partial_x|_z p}^r + \overline{z_r}^r \widehat{\mathcal{F}_x}^r \tag{10}$$

$$\partial_t|_r (\overline{z_r}^r \hat{v}^r) + \nabla_r \cdot (\overline{z_r}^r \hat{\mathbf{u}}^r \hat{v}^r) + \partial_r\left(\overline{z_r}^r \widehat{\dot{r}}^r \hat{v}^r\right) + f \overline{z}^r \hat{u}^r = -\overline{z_r}^r \overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{v}}}^r - \rho_0^{-1} \overline{z_r} \widehat{\partial_y|_z p}^r + \overline{z_r}^r \widehat{\mathcal{F}_y}^r \tag{11}$$

$$\partial_r \overline{p}^r = -\hat{\rho}^r g \overline{z_r}^r \tag{12}$$

where $\overline{(\cdot)}^r$ denotes an average at fixed r, which we here assume to commute with the partial derivatives. This assumption holds for a wide range of typical averaging operations, including ensemble averages and spatial filters with a constant filtering kernel. (Although we note that the same assumption would not generally hold if the average was taken on a surface that does not align with the coordinate system.) $\widehat{(\cdot)}^r \equiv \overline{z_r(\cdot)}^r / \overline{z_r}^r$ denotes the corresponding generalized thickness weighted average. This generalized thickness weighted averaging has previously been introduced by Maddison and Marshall (2013) and Klingbeil et al. (2019), and it reduces to the more traditional isopycnal thickness weighted average if the vertical coordinate r represents buoyancy (or potential density). For simplicity we have here assumed that the Coriolis parameter, f, is constant over the filter stencil, such that $\widehat{f\mathbf{u}}^r = f\hat{\mathbf{u}}^r$. If that assumption is relaxed, we obtain an additional eddy Coriolis term in the momentum equations.

In general, the advective eddy flux divergence takes the form

$$\overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{c}}}^r \equiv (\overline{z_r}^r)^{-1} \nabla_r \cdot (\overline{z_r}^r (\widehat{\mathbf{u}c}^r - \hat{\mathbf{u}}^r \hat{c}^r)) + (\overline{z_r}^r)^{-1} \partial_r\left(\overline{z_r}^r\left(\widehat{\dot{r}c}^r - \widehat{\dot{r}}^r \hat{c}^r\right)\right), \tag{13}$$

where $\overline{\nabla}^r \cdot$ denotes the divergence in a metric space where a volume element is defined by $dx \times dy \times d\overline{z}^r$, which arises as the natural space for the averaged equations. (We will return to this point in Section 4.2.) For a Reynolds average, the eddy flux divergence can be written as

$$\overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{c}}}^r \equiv (\overline{z_r}^r)^{-1} \nabla_r \cdot (\overline{z_r}^r \widehat{\mathbf{u}'' c''}^r) + (\overline{z_r}^r)^{-1} \partial_r\left(\overline{z_r}^r \widehat{\dot{r}'' c''}^r\right) \tag{14}$$

where double primes denote deviations from the thickness-weighted average.

For Equations 8–12 to form a closed set of equations, we need closures for the eddy tracer and momentum flux divergences, $\overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{c}}}^r, \overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{u}}}^r, \overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{v}}}^r$, an equation of state for the mean quantities, $\hat{\rho}^r = \tilde{\rho}^r(\hat{S}^r, \hat{T}^r, \overline{z})$ (e.g., Brankart, 2013; McDougall & McIntosh, 1996; Stanley et al., 2020), and a closure for the eddy components of the pressure gradient term. For a Reynolds average, the pressure gradient acceleration can be expressed as (see Appendix A)

$$\overline{z_r}^r \widehat{\partial_x|_z p}^r = \overline{z_r}^r \partial_x|_r \overline{p}^r + \overline{z_r}^r \hat{\rho}^r g \partial_x|_r \overline{z}^r + \partial_x|_r \left(\overline{z_r' p'}^r\right) - \partial_r \left(\overline{p' \partial_x|_r z'}^r\right) \tag{15a}$$

$$= \overline{z_r}^r \partial_x|_{\overline{z}^r} \overline{p}^r + \partial_x|_r \left(\overline{z_r' p'}^r\right) - \partial_r \left(\overline{p' \partial_x|_r z'}^r\right) \tag{15b}$$

and similarly for the y-component. The general (non-Reynolds average) form is readily obtained by substituting $\overline{z_r' p'}^r \rightarrow \overline{z_r p}^r - \overline{z_r}^r \overline{p}^r$ and $\overline{p' \partial_x|_r z'}^r \rightarrow \overline{p \partial_x|_r z}^r - \overline{p}^r \partial_x|_r \overline{z}^r$. The formulation of the eddy terms in Equation 15a or Equation 15b clarifies that the eddy pressure gradient force only redistributes momentum, with the vertical momentum flux $\left(\overline{p' \partial_x|_r z'}^r\right)$ readily identifiable as the form stress acting on a generalized coordinate surface.

We can absorb the eddy contributions to the pressure gradient force into the eddy momentum flux divergence by defining (cf. Young, 2012; Maddison & Marshall, 2013).

$$\overline{\nabla}^r \cdot \widehat{\mathbf{E}^{\mathbf{u}}}^r \equiv \overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{u}}}^r + \left(\rho_0 \overline{z_r}^r\right)^{-1} \partial_x|_r \left(\overline{z_r' p'}^r\right) - \left(\rho_0 \overline{z_r}^r\right)^{-1} \partial_r \left(\overline{p' \partial_x|_r z'}^r\right) \tag{16}$$

$$\overline{\nabla}^r \cdot \widehat{\mathbf{E}^{\mathbf{v}}}^r \equiv \overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{v}}}^r + \left(\rho_0 \overline{z_r}^r\right)^{-1} \partial_y|_r \left(\overline{z_r' p'}^r\right) - \left(\rho_0 \overline{z_r}^r\right)^{-1} \partial_r \left(\overline{p' \partial_y|_r z'}^r\right), \tag{17}$$

where $\overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{u}}}^r$ and $\overline{\nabla}^r \cdot \widehat{\mathbf{J}^{\mathbf{v}}}^r$ are defined as in Equation 14. -$\widehat{\mathbf{E}^{\mathbf{u}}}^r$ and -$\widehat{\mathbf{E}^{\mathbf{v}}}^r$ can be interpreted as generalized Eliassen-Palm (EP) flux vectors (Eliassen & Palm, 1961) and describe the total eddy momentum flux via both advective and pressure contributions. Using Equations 16 and 17 we can write the generalized thickness-weighted average (TWA) horizontal momentum equations as

$$\partial_t|_r \left(\overline{z_r}^r \hat{u}^r\right) + \nabla_r \cdot \left(\overline{z_r}^r \hat{\mathbf{u}}^r \hat{u}^r\right) + \partial_r \left(\overline{z_r}^r \hat{\dot{r}}^r \hat{u}^r\right) - f \overline{z}^r \hat{v}^r = -\overline{z_r}^r \overline{\nabla}^r \cdot \widehat{\mathbf{E}^{\mathbf{u}}}^r - \rho_0^{-1} \overline{z_r} \partial_x|_{\overline{z}^r} \overline{p}^r + \overline{z_r}^r \widehat{\mathcal{F}_x}^r \tag{18}$$

$$\partial_t|_r \left(\overline{z_r}^r \hat{v}^r\right) + \nabla_r \cdot \left(\overline{z_r}^r \hat{\mathbf{u}}^r \hat{v}^r\right) + \partial_r \left(\overline{z_r}^r \hat{\dot{r}}^r \hat{v}^r\right) + f \overline{z}^r \hat{u}^r = -\overline{z_r}^r \overline{\nabla}^r \cdot \widehat{\mathbf{E}^{\mathbf{v}}}^r - \rho_0^{-1} \overline{z_r} \partial_y|_{\overline{z}^r} \overline{p}^r + \overline{z_r}^r \widehat{\mathcal{F}_y}^r \tag{19}$$

Except for the additional eddy terms, Equations 8, 9, 12, 18 and 19 have the same form as the unaveraged equations in Section 2.1, with the substitutions $z \rightarrow \overline{z}^r$, $p \rightarrow \overline{p}^r$, $\mathbf{u} \rightarrow \hat{\mathbf{u}}^r$, $\dot{r} \rightarrow \hat{\dot{r}}^r$, $c \rightarrow \hat{c}^r$ and $\rho \rightarrow \hat{\rho}^r$. Except for the specific form of the pressure gradient term (and associated EP flux contributions), the equations are also identical to the isopycnal TWA equations as discussed in Young (2012), with the buoyancy, b, here replaced by the generalized vertical coordinate, r.

The equations, moreover, reduce to the Eulerian mean equations for $r = z$, in which case $z_r \rightarrow z_z = 1$ and hence the generalized "thickness weighted" average simply reduces to the Eulerian average. Because the Eulerian-averaged equations are already a subset of the generalized TWA equations, and because non-thickness weighted averaging with any vertical coordinate that has non-constant thickness is non-conservative and therefore undesirable (see Loose et al., 2023), we will only consider generalized "thickness weighted" averaging in the main part of this manuscript. A brief discussion of the non-thickness-weighted averaged momentum equations is included in Appendix C.

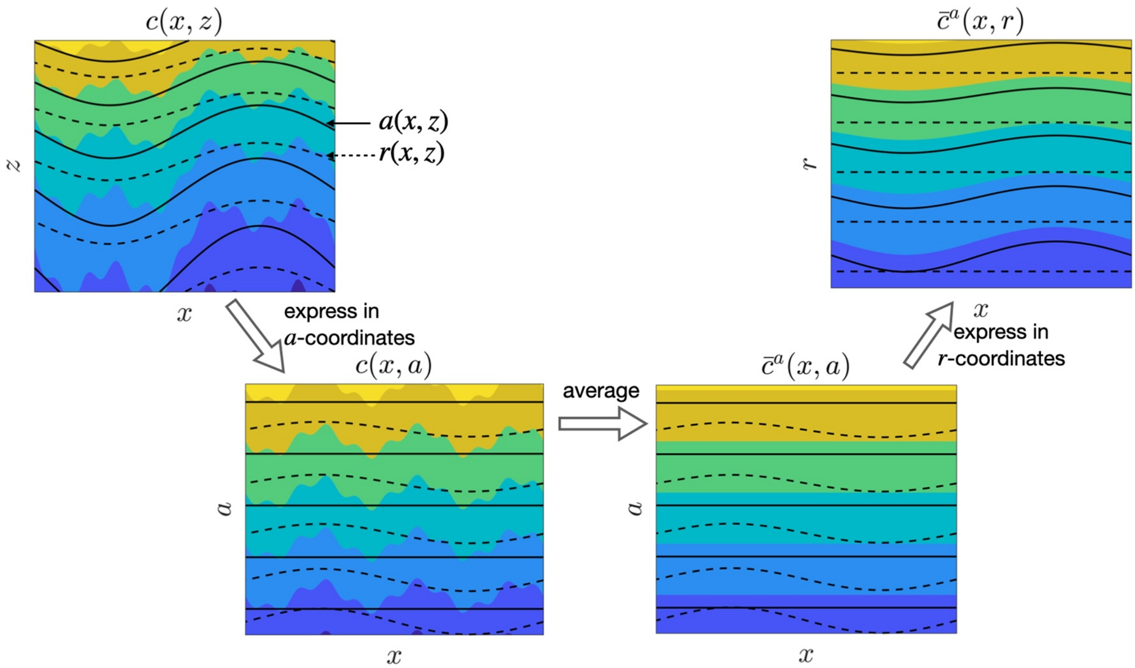

Figure 1. Illustration of an averaging along an arbitrary vertical coordinate, a, and expression of the averaged field using another arbitrary vertical coordinate, r. The figure on the top left shows an example field $c(x,z)$ (shading) in (x,z)-space, together with isolines of a (solid) and r (dashed). We can first express this field in (x,a)-space to obtain $c(x,a)$, which is shown on the bottom left. Taking the average (at fixed a) we obtain $\overline{c}^a(x,a)$ (bottom right) which can finally be transformed into (x,r)-space to obtain $\overline{c}^a(x,r)$ (top right). Notice that $\overline{c}^a(x,r)$ is not generally constant in the averaging dimension (here x), unless r is itself constant along a surfaces in the averaging dimension (as is the case if r is itself an averaged quantity). We will return to this issue below in the context of the residual mean equations (cf. Figure 2).

3. Coordinate Transformation of Averaged Equations

In general, the averaging may not need to follow the same vertical coordinate as the model. For example, McDougall and McIntosh (2001) argue that the tracers in z-coordinate models that use the GM parameterization should be interpreted as thickness-weighted isopycnal averages. To see how such interpretations can be justified more generally, we here derive the arbitrarily averaged Boussinesq equations expressed in generalized vertical coordinates. We denote the averaging coordinate with "a" (i.e., averages are assumed to be taken along surfaces of constant a) while the generalized vertical coordinate used to express the final equations remains denoted with "r." To derive the equations we start from the equations expressed in a-coordinates, take a (thickness-weighted) average (at fixed a) and then transform the resulting equations into r-coordinates (as sketched in Figure 1). Various specific choices for the averaging coordinate (a) and the model coordinate (r) will be discussed in Section 4. For example, for a thickness-weighted isopycnal average (i.e., $a \rightarrow b$) and using the mean isopycnal height as the vertical coordinate (i.e., $r \rightarrow \overline{z}^b$) we recover the "residual mean" equations discussed in Young (2012) and advocated as an interpretation for the z-coordinate model tracer equations by McDougall and McIntosh (2001).

We start from the averaged tracer Equation 9, expressed using the averaging coordinate "a" (i.e., we simply replace $r \rightarrow a$ in Equation 9). Multiplying both sides with $\overline{z}_r^a/\overline{z_a}^a$, where $\overline{z}_r^a \equiv \partial_r \overline{z}^a$, we get

$$\frac{\overline{z}_r^a}{\overline{z_a}^a}\left[\partial_t|_a(\overline{z_a}^a\,\hat{c}^a) + \nabla_a\cdot(\overline{z_a}^a\hat{\mathbf{u}}^a\,\hat{c}^a) + \partial_a\left(\overline{z_a}^a\,\hat{\dot{a}}^a\,\hat{c}^a\right)\right] = -\overline{z}_r^a\overline{\nabla}^a\cdot\hat{\mathbf{J}^{\mathbf{c}}}^a + \overline{z}_r^a\,\hat{\dot{c}}^a. \tag{20}$$

As shown in Appendix B, the L.H.S. of Equation 20 can be expressed in r-coordinates as

$$\frac{\overline{z}_r^a}{\overline{z_a}^a}\left[\partial_t|_a(\overline{z_a}^a\,\hat{c}^a) + \nabla_a\cdot(\overline{z_a}^a\hat{\mathbf{u}}^a\,\hat{c}^a) + \partial_a\left(\overline{z_a}^a\,\hat{\dot{a}}^a\,\hat{c}^a\right)\right] = \partial_t|_r(\overline{z}_r^a\,\hat{c}^a) + \nabla_r\cdot(\overline{z}_r^a\hat{\mathbf{u}}^a\,\hat{c}^a) + \partial_r\left(\overline{z}_r^a\,\dot{r}^{\#a}\,\hat{c}^a\right) \tag{21}$$

where we introduced the Lagrangian rate of change of r following the a-averaged flow

$$\dot{r}^{\#a} \equiv \frac{\widehat{D}^a}{Dt} r \tag{22a}$$

$$\equiv \partial_t|_a r + \hat{\mathbf{u}}^a \cdot \nabla_a r + \widehat{\dot{a}}^a \partial_a r \tag{22b}$$

$$= -(\partial_r a)^{-1} \partial_t|_r a - (\partial_r a)^{-1} \hat{\mathbf{u}}^a \cdot \nabla_r a + \widehat{\dot{a}}^a (\partial_r a)^{-1}. \tag{22c}$$

Note that in the last step we used the coordinate transformation identities in Equation B1. From Equations 20 and 21, it follows that the a-averaged tracer equation can be expressed using an arbitrary vertical coordinate, r, as

$$\partial_t|_r \left(\overline{z}_r^a \hat{c}^a\right) + \nabla_r \cdot \left(\overline{z}_r^a \hat{\mathbf{u}}^a \hat{c}^a\right) + \partial_r \left(\overline{z}_r^a \, \dot{r}^{\#a} \hat{c}^a\right) = -\overline{z}_r^a \overline{\nabla}^a \cdot \widehat{\mathbf{J}^{\mathbf{c}}}^a + \overline{z}_r^a \widehat{\dot{c}}^a. \tag{23}$$

The continuity equation follows immediately from Equation 23 with $c = 1$, and the a-averaged momentum equations in r-coordinates can be derived analogously to the tracer equation (using Equation 21 with $c \to u, v$). Hydrostatic balance can be written as

$$\partial_r \overline{p}^a = \partial_a \overline{p}^a \partial_r a \tag{24a}$$

$$= -g \overline{z_a}^a \hat{\rho}^a \partial_r a \tag{24b}$$

$$= -g \partial_r \overline{z}^a \hat{\rho}^a, \tag{24c}$$

where in the second step we used the averaged hydrostatic balance in Equation 12 with $r \to a$.

The full a-averaged hydrostatic Boussinesq equations can then be written in r-coordinates as:

$$\partial_t|_r \left(\overline{z}_r^a\right) + \nabla_r \cdot \left(\overline{z}_r^a \hat{\mathbf{u}}^a\right) + \partial_r \left(\overline{z}_r^a \, \dot{r}^{\#a}\right) = 0 \tag{25}$$

$$\partial_t|_r \left(\overline{z}_r^a \hat{c}^a\right) + \nabla_r \cdot \left(\overline{z}_r^a \hat{\mathbf{u}}^a \hat{c}^a\right) + \partial_r \left(\overline{z}_r^a \, \dot{r}^{\#a} \hat{c}^a\right) = -\overline{z}_r^a \overline{\nabla}^a \cdot \widehat{\mathbf{J}^{\mathbf{c}}}^a + \overline{z}_r^a \widehat{\dot{c}}^a \tag{26}$$

$$\partial_t|_r \left(\overline{z}_r^a \hat{u}^a\right) + \nabla_r \cdot \left(\overline{z}_r^a \hat{\mathbf{u}}^a \hat{u}^a\right) + \partial_r \left(\overline{z}_r^a \, \dot{r}^{\#a} \hat{u}^a\right) - f \overline{z}_r^a \hat{v}^a = -\overline{z}_r^a \overline{\nabla}^a \cdot \widehat{\mathbf{E}^{\mathbf{u}}}^a - \rho_0^{-1} \overline{z}_r^a \partial_x|_{\overline{z}^a} \overline{p}^a + \overline{z}_r^a \widehat{\mathcal{F}}_x^a \tag{27}$$

$$\partial_t|_r \left(\overline{z}_r^a \hat{v}^a\right) + \nabla_r \cdot \left(\overline{z}_r^a \hat{\mathbf{u}}^a \hat{v}^a\right) + \partial_r \left(\overline{z}_r^a \, \dot{r}^{\#a} \hat{v}^a\right) + f \overline{z}_r^a \hat{u}^a = -\overline{z}_r^a \overline{\nabla}^a \cdot \widehat{\mathbf{E}^{\mathbf{v}}}^a - \rho_0^{-1} \overline{z}_r^a \partial_y|_{\overline{z}^a} \overline{p}^a + \overline{z}_r^a \widehat{\mathcal{F}}_y^a \tag{28}$$

$$\partial_r \overline{p}^a = -\hat{\rho}^a g \partial_r \overline{z}^a. \tag{29}$$

Using that $\partial_x|_{\overline{z}^a} = \partial_x|_r - \left(\partial_x|_r \overline{z}^a\right)\left(\overline{z}_r^a\right)^{-1} \partial_r$ and hydrostatic balance (29), the mean pressure gradient term can be expressed in r-coordinates as

$$\partial_x|_{\overline{z}^a} \overline{p}^a = \partial_x|_r \overline{p}^a + \hat{\rho}^a g \partial_x|_r \overline{z}^a. \tag{30}$$

Numerical models do not usually evaluate the pressure gradient term via the two terms on the R.H.S. of Equation 30, which to first order cancel when the coordinate system does not follow height or pressure surfaces, but the expression is here included for the sake of completeness of the analytical equations.

Notice that the eddy fluxes in the generalized average Equations 25–29 depend on the averaging choice but not on the coordinate system ultimately chosen to express the equations. Coordinate-system independent eddy flux terms are desirable for hybrid coordinate models as it allows parameterizations to be independent of the model coordinate, which may vary across the model domain.

Except for the additional eddy contributions, Equations 25–29 are again formally identical to the unaveraged equations in Section 2.1 with the substitutions $\mathbf{u} \to \hat{\mathbf{u}}^a$, $c \to \hat{c}^a$, $\dot{c} \to \widehat{\dot{c}}^a$, $\rho \to \hat{\rho}^a$, $\mathcal{F}_{x/y} \to \widehat{\mathcal{F}_{x/y}}^a$, $p \to \overline{p}^a$,

Table 2
Summary of Special Cases Discussed in Section 4, Together With Previous Publications Where the Respective Equations Have Been Discussed

Averaging coordinate $a =$	Vertical (model) coordinate $r =$	Section	Previously discussed in
z (height)	r (general)	4.1	N/A
z (height)	z (height)	4.1.1	for example, Ch. 4 of Pope (2000)
z (height)	b (isopycnal)	4.1.2	N/A
b (isopycnal)	r (general)	4.2	N/A
b (isopycnal)	b (isopycnal)	4.2.1	Young (2012)
b (isopycnal)	$\overline{z}^b$ (mean height)	4.2.2	Young (2012)
b/z (mixed)	r (general)	4.3.1	N/A
b/z (mixed)	z (height)	4.3.2	McDougall and McIntosh (2001)

$z_r \rightarrow \overline{z}_r^a$ and $\dot{r} \rightarrow \dot{r}^{\#a}$. The last two substitutions warrant some further discussion. Notice that $\overline{z}_r^a \equiv \partial_r \overline{z}^a \neq \overline{z_r}^a$, that is, $\overline{z}_r^a$ is not necessarily equal to the a-averaged generalized thickness but is instead defined directly as the r-derivative of the a-averaged height. Moreover, $\dot{r}^{\#a} \neq \widehat{\dot{r}}^a$, that is, $\dot{r}^{\#a}$ is not simply the generalized TWA of $\dot{r}$. Instead, $\dot{r}^{\#a}$ is the rate of change of r following the generalized TWA flow. This point was previously emphasized in Section 3 of Young (2012) for the special case of the buoyancy-averaged equations expressed in depth coordinates.

4. Special Cases

In this section we consider a suite of special cases for averaging coordinate and model vertical coordinate. We touch base with the literature, where available, with Table 2 summarizing the cases considered here.

4.1. The Eulerian Mean Equations in Generalized Vertical Coordinates: $a \rightarrow z$

With $a = z$ (in which case $\widehat{()}^a \rightarrow \widehat{()}^z = \overline{()}^z$) we obtain the Eulerian-averaged equations in generalized vertical coordinates as

$$\partial_t|_r z_r + \nabla_r \cdot (z_r \overline{\mathbf{u}}^z) + \partial_r \left(z_r \dot{r}^{\#z} \right) = 0 \tag{31}$$

$$\partial_t|_r (z_r \overline{c}^z) + \nabla_r \cdot (z_r \overline{\mathbf{u}}^z \overline{c}^z) + \partial_r \left(z_r \dot{r}^{\#z} \overline{c}^z \right) = -z_r \nabla \cdot \overline{\mathbf{J}^{\mathbf{c}}}^z + z_r \overline{\dot{c}}^z \tag{32}$$

$$\partial_t|_r (z_r \overline{u}^z) + \nabla_r \cdot (z_r \overline{\mathbf{u}}^z \overline{u}^z) + \partial_r \left(z_r \dot{r}^{\#z} \overline{u}^z \right) - f z_r \overline{v}^z = -z_r \nabla \cdot \overline{\mathbf{J}^{\mathbf{u}}}^z - \rho_0^{-1} z_r \partial_x|_z \overline{p}^z + z_r \overline{\mathcal{F}_x}^z \tag{33}$$

$$\partial_t|_r (z_r \overline{v}^z) + \nabla_r \cdot (z_r \overline{\mathbf{u}}^z \overline{v}^z) + \partial_r \left(z_r \dot{r}^{\#z} \overline{v}^z \right) + f z_r \overline{u}^z = -z_r \nabla \cdot \overline{\mathbf{J}^{\mathbf{v}}}^z - \rho_0^{-1} z_r \partial_y|_z \overline{p}^z + z_r \overline{\mathcal{F}_y}^z \tag{34}$$

$$\partial_r \overline{p}^z = -\overline{\rho}^z g z_r \tag{35}$$

with $\overline{\mathbf{J}^{\mathbf{c}}}^z = \overline{\mathbf{v}' c'}^z$, $\overline{\mathbf{J}^{\mathbf{u}}}^z = \overline{\mathbf{v}' u'}^z$, $\overline{\mathbf{J}^{\mathbf{v}}}^z = \overline{\mathbf{v}' v'}^z$ (where $\mathbf{v}$ denotes the 3-dimensional velocity vector),

$$\dot{r}^{\#z} = \frac{\overline{D}^z}{Dt} r = \partial_t|_z r + \overline{\mathbf{v}}^z \cdot \nabla r \tag{36}$$

and ∇ (without subscript) represents the 3-dimensional divergence and gradient operators. Notice that the generalized EP-flux reduces to the advective eddy momentum flux for Eulerian averaging (i.e., $\overline{\mathbf{E}^{\mathbf{u/v}}}^z = \overline{\mathbf{J}^{\mathbf{u/v}}}^z$) as $z_z = 1$ (and hence $z_z' = 0$) and $\partial_x|_z z = 0$ (cf. Equations 16 and 17). The averaged horizontal pressure gradient can be expressed in r-coordinates as

$$\partial_x|_z \overline{p}^z = \partial_x|_r \overline{p}^z + \overline{\rho}^z g \, \partial_x|_r z. \tag{37}$$

Aside from the additional eddy contributions, Equations 31–35 are again the same as the unaveraged equations in Section 2.1 with $\mathbf{u}$, c, p, ρ, and $\mathcal{F}_{x/y}$ replaced by their respective Eulerian averages, and $\dot{r} \rightarrow \dot{r}^{\#z}$, which gives the Lagrangian rate of change of r following the Eulerian-mean flow. Noticeably, the thickness remains z_r, that is, it is not affected by the Eulerian averaging, unless r itself is chosen to depend on the averaged quantities (which may be desirable or even necessary in practice). The eddy contributions appear as the Eulerian eddy flux convergences of tracer and momentum.

4.1.1. The Eulerian Mean Equations in Z-Coordinates: $a \rightarrow z, r \rightarrow z$

With $r \rightarrow z$, Equations 31–35 reduce to the well-known z-coordinate Eulerian mean equations:

$$\nabla \cdot \overline{\mathbf{v}}^z = 0 \tag{38}$$

$$\partial_t|_z \overline{c}^z + \nabla_z \cdot (\overline{\mathbf{u}}^z \overline{c}^z) + \partial_z (\overline{w}^z \overline{c}^z) = -\nabla \cdot \overline{\mathbf{J}^{\mathbf{c}}}^z + \overline{\dot{c}}^z \tag{39}$$

$$\partial_t|_z (\overline{u}^z) + \nabla_z \cdot (\overline{\mathbf{u}}^z \overline{u}^z) + \partial_z (\overline{w}^z \overline{u}^z) - f \overline{v}^z = -\nabla \cdot \overline{\mathbf{J}^{\mathbf{u}}}^z - \rho_0^{-1} \partial_x|_z \overline{p}^z + \overline{\mathcal{F}_x}^z \tag{40}$$

$$\partial_t|_z (\overline{v}^z) + \nabla_z \cdot (\overline{\mathbf{u}}^z \overline{v}^z) + \partial_r (\overline{w}^z \overline{v}^z) + f \overline{u}^z = -\nabla \cdot \overline{\mathbf{J}^{\mathbf{v}}}^z - \rho_0^{-1} \partial_y|_z \overline{p}^z + \overline{\mathcal{F}_y}^z \tag{41}$$

$$\partial_z \overline{p}^z = -\overline{\rho}^z g. \tag{42}$$

4.1.2. The Eulerian Mean Equations in Isopycnal Coordinates: $a \rightarrow z, r \rightarrow b$

The Eulerian mean equations in isopycnal coordinates are simply given by Equations 31–35 with r replaced by a suitable buoyancy coordinate, b.

4.2. The Isopycnal TWA Equations in Generalized Vertical Coordinates: $a \rightarrow b$

The isopycnal TWA equations in generalized vertical coordinates are given by Equations 25–29 with $a \rightarrow b$, where b is a suitably defined buoyancy. For the most general case of a nonlinear equation of state and with buoyancy, b, defined with respect to potential density, the equations do not obviously simplify (in an exact manner) and are hence not repeated here for the special case of $a \rightarrow b$.

For a linear equation of state with a materially conserved (for adiabatic flow) buoyancy variable in the form

$$b = -g\,(\rho - \rho_0)/\rho_0, \tag{43}$$

we can use that $\hat{\rho}^b = \overline{\rho}^b = \rho$ and we can define a Montgomery potential $M = p + \rho g z$, which, for a Reynolds average, allows us to express the pressure contribution to the EP flux in Equation 16 as

$$\partial_x|_b \left(\overline{z_b' p'}^b \right) - \partial_b \left(\overline{p' \partial_x|_b z'}^b \right) = \overline{z_b' \partial_x|_b p'}^b - \overline{\partial_b p' \partial_x|_b z'}^b \tag{44a}$$

$$= \overline{z_b' \partial_x|_b p'}^b + \overline{\partial_b z' \rho g \partial_x|_b z'}^b \tag{44b}$$

$$= \overline{z_b' \partial_x|_b M'}^b \tag{44c}$$

$$= \partial_b \left(\overline{z' \partial_x|_b M'}^b \right) + \partial_x|_b \left(\rho_0 \frac{\overline{z'^2}^b}{2} \right). \tag{44d}$$

where in the last step we used that $\partial_b M = -\rho_0 z$, and thus

$$\overline{\nabla}^b \cdot \widehat{\mathbf{E}^{\mathbf{u}}}^b = \overline{\nabla}^b \cdot \widehat{\mathbf{J}^{\mathbf{u}}}^b + \left(\rho_0 \overline{z_r}^r\right)^{-1} \partial_x|_b \left(\rho_0 \frac{\overline{z'^2}^b}{2} \right) + \left(\rho_0 \overline{z_r}^r\right)^{-1} \partial_b \left(\overline{z' \partial_x|_b M'}^b \right) \tag{45}$$

and similarly for the meridional component (cf. Section 6 in Young, 2012). Notice, however, that the ocean components of climate models generally use a nonlinear equation of state, in which case there is no materially conserved (for adiabatic flow) buoyancy variable that allows for these simplifications.

4.2.1. The Isopycnal TWA Equations in Isopycnal Coordinates: $a \rightarrow b$ and $r \rightarrow b$

The isopycnal TWA equations in isopycnal coordinates are given directly by the generalized TWA equations (8,9,10/18, 11/19 and 12) with $r \rightarrow b$. Again, no obvious exact simplification is obtained for a buoyancy variable based on potential density with a realistic equation of state.

For a linear equation of state, where buoyancy is defined as in Equation 43, we can again express the EP flux divergences as in Equation 45 and we can moreover express the mean pressure gradient in terms of the Montgomery potential gradient as

$$\partial_x|_{\overline{z}^b}\overline{p}^b = \partial_x|_b\overline{p}^b - \left(\overline{z_b}^b\right)^{-1}\partial_b\overline{p}^b\partial_x|_b\overline{z}^b \tag{46a}$$

$$= \partial_x|_b\overline{M}^b \tag{46b}$$

where in the last step we used that $\partial_b\overline{p}^b = -\rho g\overline{z}_b^b$

4.2.2. The "Residual Mean" Equations: $a \rightarrow b$ and $r \rightarrow \overline{z}^b$

With $a \rightarrow b$ and $r \rightarrow \overline{z}^b$ we recover the exact "residual mean" equations as discussed in Young (2012). This choice gives $\overline{z}_r^a \rightarrow \partial_{\overline{z}^b}\overline{z}^b = 1$, and Equations 25–29 simplify to

$$\nabla_{\overline{z}^b}\cdot\hat{\mathbf{u}}^b + \partial_{\overline{z}^b}w^{\#b} = 0 \tag{47}$$

$$\partial_t|_{\overline{z}^b}\hat{c}^b + \nabla_{\overline{z}^b}\cdot\left(\hat{\mathbf{u}}^b\hat{c}^b\right) + \partial_{\overline{z}^b}\left(w^{\#b}\hat{c}^b\right) = -\overline{\nabla}^b\cdot\widehat{\mathbf{J}^{\mathbf{c}}}^b + \widehat{\dot{c}}^b \tag{48}$$

$$\partial_t|_{\overline{z}^b}\hat{u}^b + \nabla_{\overline{z}^b}\cdot\left(\hat{\mathbf{u}}^b\hat{u}^b\right) + \partial_{\overline{z}^b}\left(w^{\#b}\hat{u}^b\right) - f\hat{v}^b = -\overline{\nabla}^b\cdot\widehat{\mathbf{E}^{\mathbf{u}}}^b - \rho_0^{-1}\partial_x|_{\overline{z}^b}\overline{p}^b + \widehat{\mathcal{F}_x}^b \tag{49}$$

$$\partial_t|_{\overline{z}^b}\hat{v}^b + \nabla_{\overline{z}^b}\cdot\left(\hat{\mathbf{u}}^b\hat{v}^b\right) + \partial_{\overline{z}^b}\left(w^{\#b}\hat{v}^b\right) + f\hat{u}^b = -\overline{\nabla}^b\cdot\widehat{\mathbf{E}^{\mathbf{v}}}^b - \rho_0^{-1}\partial_y|_{\overline{z}^b}\overline{p}^b + \widehat{\mathcal{F}_y}^b \tag{50}$$

$$\partial_{\overline{z}^b}\overline{p}^b = -\hat{\rho}^b g \tag{51}$$

where

$$w^{\#b} = \partial_t|_b\overline{z}^b + \hat{\mathbf{u}}^b\cdot\nabla_b\overline{z}^b + \widehat{\dot{b}}^b\partial_b\overline{z}^b. \tag{52}$$

Save for the eddy terms, these equations have the same form as the unaveraged equations in z-coordinates. Eddy terms appear as isopycnal TWA eddy flux convergences in the tracer and momentum equations, with the eddy momentum fluxes including advective as well as pressure contributions (Equations 16 and 17).

Notice that the choice of $\overline{z}^b$ (rather than the unaveraged z) as the vertical coordinate is necessary to eliminate the thickness from the equations, as $\partial_z\overline{z}^b = \partial_z b\,\partial_b\overline{z}^b \neq 1$. The choice further guarantees that the averaged fields remain constant along the averaging dimension in the case of a Reynolds average—as sketched in Figure 2. Similarly, for a spatial filter, the choice guarantees that the filtered fields remain horizontally smooth along the model coordinate. The notation of Young (2012) drops the overbar on the z-coordinate in the averaged equation. As the residual mean Equations 47–51 do not depend on z but only $\overline{z}^b$, this difference may be viewed simply as a notational re-definition: $\overline{z}^b \rightarrow z$. This modified definition of the z-coordinate, however, is important to keep in mind when interpreting model variables, parameters, and boundary conditions. For example, boundary fluxes can affect model quantities at all $\overline{z}^b$-levels for which the corresponding isopycnal outcrops into the boundary anywhere within the averaging region (i.e., boundary conditions cannot technically be applied strictly at $\overline{z}^b = z_B$, where z_B is the height of the boundary). A simple re-definition of the vertical coordinate, moreover, cannot be

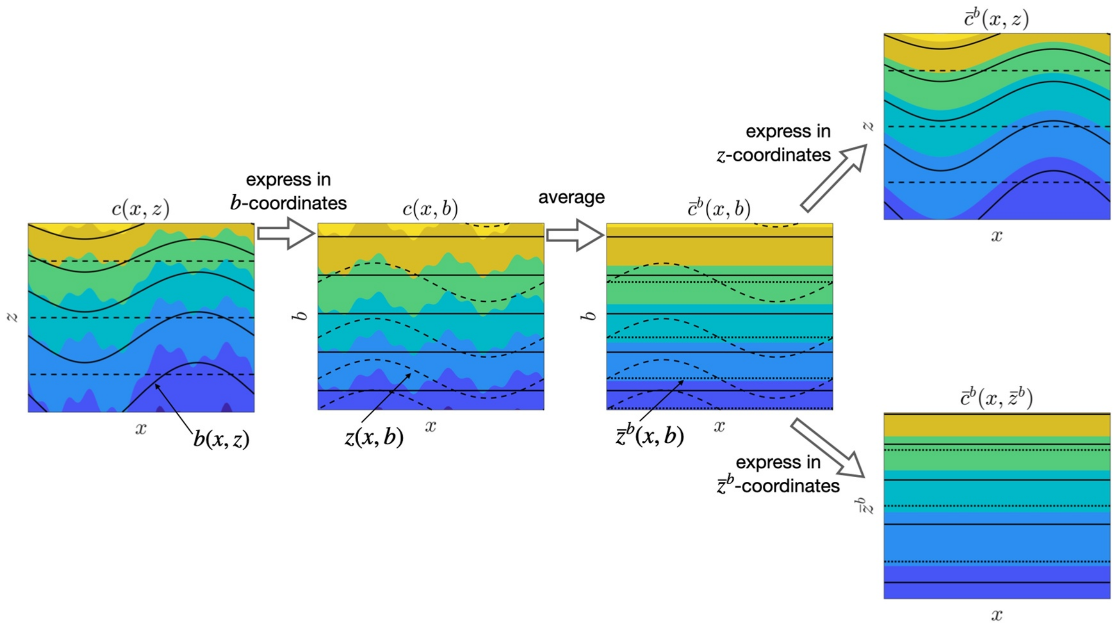

Figure 2. Illustration of the averaging in the "residual" mean equations, where the average is taken along surfaces of constant buoyancy, b, and the averaged equations are expressed using the averaged height $\bar{z}^b$ as the vertical coordinate. The figure on the left shows an example field $c(x,z)$ (shading) in physical space, together with isolines of b (solid) and z (dashed). The second panel from the left shows the same fields expressed in (x,b)-space. Taking an average in the x-direction (along surfaces of constant b) yields $\bar{c}^b(x,b)$, which is shown in the third panel together with isolines of b(solid), $z(x,b)$ (dashed) and $\bar{z}^b(x,b)$ (dotted). Transforming $\bar{c}^b(x,b)$ into $\left(x,\bar{z}^b\right)$ -space yields the residual mean field, $\bar{c}^b\left(x,\bar{z}^b\right)$, shown on the bottom right. This field is constant along the averaging dimension (here x), and we notice that this result holds generally if the averaged field is expressed using a vertical coordinate that is itself an averaged quantity. For comparison, the top right panel shows $\bar{c}^b(x,z)$, which is not constant in x.

applied in the mixed Eulerian/TWA interpretation of the z-coordinate equations advocated by McDougall and McIntosh (2001), as will be discussed in the following section.

4.3. Mixed Eulerian/TWA Equations

McDougall and McIntosh (2001) argue that the tracers in z-coordinate models employing the GM parameterization should be interpreted as representing isopycnal thickness-weighted averages ($a \to b$), while the model velocities represent Eulerian averages ($a \to z$), with GM parameterizing the difference between the Eulerian and isopycnal TWA velocities. The corresponding model equations are obtained by combining the Eulerian mean momentum Equations 40–42 and residual mean tracer Equation 48. However, since the model coordinate, r, can only represent either $\bar{z}^b$ or z, this requires a new definition of the residual mean quantities in actual z-space. Specifically, one can define

$$\hat{c}^{b^\#}(x,y,z,t) \equiv \hat{c}^b\left(x,y,b^\#(x,y,z,t),t\right) \tag{53}$$

where $b^\#(x,y,z,t)$ is defined as in De Szoeke and Bennett (1993), McDougall and McIntosh (2001) and Young (2012) such that $\bar{z}^b\left(x,y,b^\#(x,y,z,t),t\right) = z$; that is, $b^\#(x,y,z,t)$ describes the buoyancy surface whose mean height is z. Notice that $\hat{c}^{b^\#}$ is here not an average along surfaces of constant $b^\#$, but a thickness-weighted average along the b-surface with $b = b^\#(x,y,z,t)$—as sketched in Figure 3.

4.3.1. Mixed Eulerian/TWA Equations in Generalized Vertical Coordinates

We can also express $b^\#$ and $\hat{c}^{b^\#}$ in an arbitrary vertical coordinate, as $b^\#(x,y,r,t) = b^\#(x,y,z(x,y,r,t),t)$ and $\hat{c}^{b^\#}(x,y,r,t) = \hat{c}^{b^\#}(x,y,z(x,y,r,t),t) = \hat{c}^b\left(x,y,b^\#(x,y,r,t),t\right)$. With analog definitions for $\hat{u}^{b^\#}$, $\hat{v}^{b^\#}$ and $\hat{\dot{r}}^{\#b^\#}$, Equations 25 and 26 with $a \to b$ and using that $\bar{z}^b\left(x,y,b^\#,t\right) = z$ imply

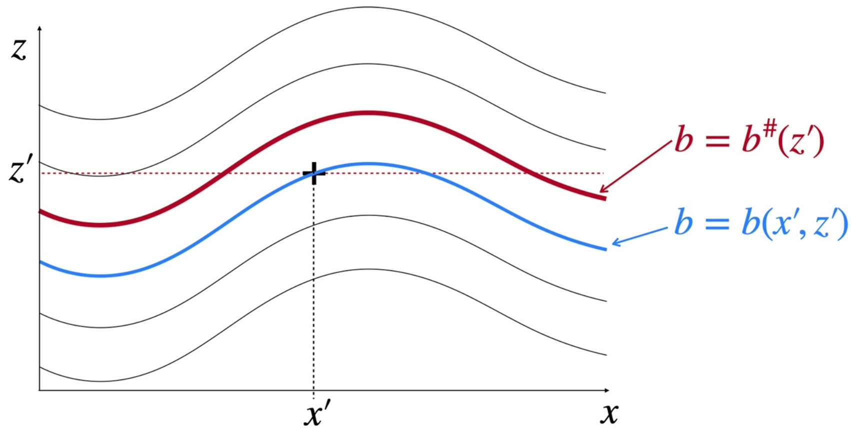

Figure 3. Illustration of $b^\#$ and the corresponding average, here sketched for the case of a domain-wide average in x. The solid lines represent an example of buoyancy surfaces (i.e., surfaces of constant b) in $(x,z)-$space. The blue line highlights the specific buoyancy surface $b = b(x',z')$, where the point (x',z') is marked by the plus sign. The red line marks the specific buoyancy surface $b = b^\#(z')$, that is the buoyancy surface whose averaged height is equal to z'. The quantity $\hat{c}^{b^\#}(z') \equiv \hat{c}^b\left(b^\#(z')\right)$, represents an average of c along the buoyancy surface marked in red. For comparison, $\hat{c}^b(x',z') = \hat{c}^b(b(x',z'))$ is the average taken along the buoyancy surface marked in blue. Notice that, for a domain-wide average in x, as considered in this example, $b^\#$ and $\hat{c}^{b^\#}$ are no longer functions of the averaging coordinate, x (or may be viewed as constant in x). However, for a spatial filtering, which may be a more appropriate for numerical models, $b^\#$ and $\hat{c}^{b^\#}$ retain all dimensions of the original field.

$$\partial_t|_r (z_r) + \nabla_r \cdot \left(z_r \hat{\mathbf{u}}^{b^\#}\right) + \partial_r\left(z_r\, \dot{r}^{\#b^\#}\right) = 0 \tag{54}$$

$$\partial_t|_r\left(z_r \hat{c}^{b^\#}\right) + \nabla_r\cdot\left(z_r \hat{\mathbf{u}}^{b^\#}\hat{c}^{b^\#}\right) + \partial_r\left(z_r\, \dot{r}^{\#b^\#}\hat{c}^{b^\#}\right) = z_r \widehat{\dot{c}}^{b^\#} - z_r \nabla\cdot\widehat{\mathbf{J}^{\mathbf{c}}}^{b^\#}. \tag{55}$$

Defining the "quasi-Stokes" velocities as $\mathbf{u}^* \equiv \hat{\mathbf{u}}^{b^\#} - \overline{\mathbf{u}}^z$ and $\dot{r}^* = \dot{r}^{\#b^\#} - \dot{r}^{\#z}$, Equations 54 and 55 together with the Eulerian mean momentum Equations 33–35 form the following set of equations:

$$\partial_t|_r (z_r) + \nabla_r\cdot\left(z_r\left(\overline{\mathbf{u}}^z + \mathbf{u}^*\right)\right) + \partial_r\left(z_r\left(\dot{r}^{\#z} + \dot{r}^*\right)\right) = 0 \tag{56}$$

$$\partial_t|_r\left(z_r \hat{c}^{b^\#}\right) + \nabla_r\cdot\left(z_r\left(\overline{\mathbf{u}}^z + \mathbf{u}^*\right)\hat{c}^{b^\#}\right) + \partial_r\left(z_r\left(\dot{r}^{\#z} + \dot{r}^*\right)\hat{c}^{b^\#}\right) = z_r \widehat{\dot{c}}^{b^\#} - z_r\nabla\cdot\widehat{\mathbf{J}^{\mathbf{c}}}^{b^\#} \tag{57}$$

$$\partial_t|_r\left(z_r\overline{u}^z\right) + \nabla_r\cdot\left(z_r\overline{\mathbf{u}}^z\overline{u}^z\right) + \partial_r\left(z_r\, \dot{r}^{\#z}\overline{u}^z\right) - f z_r\overline{v}^z = -z_r\nabla\cdot\overline{\mathbf{J}^{\mathbf{u}}}^z - \rho_0^{-1} z_r\partial_x|_z\overline{p}^z + z_r\overline{\mathcal{F}_x}^z \tag{58}$$

$$\partial_t|_r\left(z_r\overline{v}^z\right) + \nabla_r\cdot\left(z_r\overline{\mathbf{u}}^z\overline{v}^z\right) + \partial_r\left(z_r\, \dot{r}^{\#z}\overline{v}^z\right) + f z_r\overline{u}^z = -z_r\nabla\cdot\overline{\mathbf{J}^{\mathbf{v}}}^z - \rho_0^{-1} z_r\partial_y|_z\overline{p}^z + z_r\overline{\mathcal{F}_y}^z \tag{59}$$

$$\partial_r\overline{p}^z = -\overline{\rho}^z g z_r, \tag{60}$$

where $\partial_x|_z\overline{p}^z$ can be expressed in r-coordinates as given in Equation 37. For Equations 56–60 to be a closed set of equations, we need closures for the quasi-Stokes velocities and the eddy tracer and momentum flux divergences, as well as an equation of state that provides the Eulerian-averaged density $(\overline{\rho}^z)$ as a function of the isopycnal TWA temperature and salinity ($\hat{T}^{b^\#}$ and $\hat{S}^{b^\#}$). The latter issue has previously been addressed by McDougall and McIntosh (2001), who argue that the required approximation is no more inaccurate than computing the Eulerian-averaged density, $\overline{\rho}^z$, from the Eulerian-averaged temperature, $\overline{T}^z$, and salinity, $\overline{S}^z$ (see their Appendix B).

Notice also that the continuity Equation 56 can alternatively be replaced with the Eulerian mean continuity Equation 31, and the combination moreover yields that

Table 3
Overview of Different Interpretations of z-Coordinate Models Using the Equation Set in Section 5.1.1 in Terms of Averaged Equations

Interpretation	Equations	Discussion	$\nabla\cdot(\mathbf{v_{GM}}\tilde{c}) =$	$\nabla\cdot(\mathbf{D}\nabla\tilde{c}) =$	$\nabla\cdot\tau^{u,v} =$
Mixed average	§4.3.2	§5.1.2	$\nabla_z\cdot\left(\mathbf{u}^*\hat{c}^{b^\#}\right) + \partial_z\left(w^*\hat{c}^{b^\#}\right)$	$-\nabla\cdot\widehat{\mathbf{J}^{\mathbf{c}}}^{b^\#}$	$-\nabla\cdot\overline{\mathbf{J}^{\mathbf{u,v}}}^z$
Eulerian average	§4.1.1	§5.1.3	skew contrib. to $\nabla\cdot\overline{\mathbf{J}^{\mathbf{c}}}^z$	diffusive contrib. to $-\nabla\cdot\overline{\mathbf{J}^{\mathbf{c}}}^z$	$-\nabla\cdot\overline{\mathbf{J}^{\mathbf{u,v}}}^z$
Isopycnal average	§4.2.2	§5.1.4	0	$-\overline{\nabla}^b\cdot\widehat{\mathbf{J}^{\mathbf{c}}}^b$	$-\overline{\nabla}^b\cdot\widehat{\mathbf{E}^{\mathbf{u,v}}}^b$

Note. The second and third columns give the sections where the equations and implications are discussed, and the third to fifth columns summarize what the respective parameterizations need to represent in each interpretation.

$$\nabla_r\cdot(\mathbf{u}^* z_r) + \partial_r\left(z_r\,\dot{r}^*\right) = 0, \tag{61}$$

that is, the three-dimensional quasi-Stokes velocity is divergence free (as previously pointed out by McDougall and McIntosh (2001)). Notice that the quasi-Stokes velocity differs from the 2D divergent "bolus" velocity $\left(\mathbf{u}_{bolus} = \hat{\mathbf{u}}^b - \overline{\mathbf{u}}^b\right)$.

4.3.2. Mixed Eulerian/TWA Equations in z Coordinates

In the z-coordinate representation, that is, with $r \to z$, Equations 56–60 reduce to the equations proposed by McDougall and McIntosh (2001) as an interpretation for z-coordinate ocean models that employ the GM parameterization (although the full set of equations is never explicitly written out in McDougall and McIntosh (2001)):

$$\nabla_z\cdot\overline{\mathbf{u}}^z + \partial_z\overline{w}^z = 0 \tag{62}$$

$$\partial_t|_z\hat{c}^{b^\#} + \nabla_z\cdot\left(\left(\overline{\mathbf{u}}^z + \mathbf{u}^*\right)\hat{c}^{b^\#}\right) + \partial_z\left(\left(\overline{w}^z + w^*\right)\hat{c}^{b^\#}\right) = \widehat{\dot{c}}^{b^\#} - \nabla\cdot\widehat{\mathbf{J}^{\mathbf{c}}}^{b^\#} \tag{63}$$

$$\partial_t|_z\overline{u}^z + \nabla_z\cdot(\overline{\mathbf{u}}^z\overline{u}^z) + \partial_z(\overline{w}^z\overline{u}^z) - f\overline{v}^z = -\nabla\cdot\overline{\mathbf{J}^{\mathbf{u}}}^z - \rho_0^{-1}\partial_x|_z\overline{p}^z + \overline{\mathcal{F}_x}^z \tag{64}$$

$$\partial_t|_z\overline{v}^z + \nabla_z\cdot(\overline{\mathbf{u}}^z\overline{v}^z) + \partial_z(\overline{w}^z\overline{v}^z) + f\overline{u}^z = -\nabla\cdot\overline{\mathbf{J}^{\mathbf{v}}}^z - \rho_0^{-1}\partial_y|_z\overline{p}^z + \overline{\mathcal{F}_y}^z \tag{65}$$

$$\partial_z\overline{p}^z = -g\overline{\rho}^z, \tag{66}$$

where we used that $\dot{z}^{\#z} = \overline{w}^z$ and $\dot{z}^* = \dot{z}^{\#b^\#} - \dot{z}^{\#z} = w^{\#b^\#} - \overline{w}^z \equiv w^*$.

5. Comparison to Existing Model Formulations

We now compare the averaged equations derived in the previous section to the equations solved in numerical ocean models, to infer which interpretations (if any) are consistent with the formulation of the equations and parameterizations in existing models, or which changes are necessary to achieve consistency. Although the main focus here is on generalized vertical coordinate models, we start with a brief review of z-coordinate models. We will not explicitly consider isopycnal coordinate models, as purely isopycnal coordinates are rarely used in realistic global ocean models, although we note that the generalized vertical coordinate results carry over directly to the isopycnal coordinate case, as discussed in the previous subsection. Overviews of the different interpretations are also provided in Table 3 (for $z-$coordinate models) and Table 4 (for generalized vertical coordinate models).

5.1. Z-Coordinate Models

5.1.1. Model Equations

The equations solved by z-coordinate numerical ocean models can usually be written in the following form:

Table 4
Overview of Different Interpretations of Generalized Vertical Coordinate Models Using the Equation Set in Section 5.2.1 in Terms of Averaged Equations

Interpretation	Equations	Discussion	$\mathbf{u_{GM}} =$	$\nabla \cdot (\mathbf{D}\,\nabla\,\tilde{c}) =$	$\nabla \cdot \tau^{u,v} =$
Coordinate-following average	§2.2	§5.2.2	0	$-\overline{\nabla}^r \cdot \widehat{\mathbf{J^c}}^r$	$-\overline{\nabla}^r \cdot \widehat{\mathbf{E^{u,v}}}^r$
Eulerian average	§4.1	§5.2.3	0	$-\nabla \cdot \overline{\mathbf{J^c}}^z$	$-\nabla \cdot \overline{\mathbf{J^{u,v}}}^z$
Isopycnal TWA	§4.2	§5.2.4	0	$-\overline{\nabla}^b \cdot \widehat{\mathbf{J^c}}^b$	$-\overline{\nabla}^b \cdot \widehat{\mathbf{E^{u,v}}}^b$
Mixed Average	§4.3.1	§5.2.5	inconsistent	$-\nabla \cdot \widehat{\mathbf{J^c}}^{b^\#}$	$-\nabla \cdot \overline{\mathbf{J^{u,v}}}^z$

Note. The second and third columns give the sections where the equations and implications are discussed, and the third to fifth columns summarize what the respective parameterizations need to represent in each interpretation. Notice that no interpretation predicts the appearance of a 2D divergent eddy advection in the continuity equation, such that all interpretations require $\mathbf{u}_{GM} = 0$ to achieve consistency with the model equations. In the mixed average interpretation we instead need to represent tracer advection via the 3D divergent-free "quasi-Stokes" advection, which is inconsistent with the GM implementation via a 2D divergent velocity.

$$\nabla \cdot \tilde{\mathbf{v}} = 0 \tag{67}$$

$$\partial_t|_z \tilde{c} + \nabla \cdot ((\tilde{\mathbf{v}} + \mathbf{v}_{GM})\,\tilde{c}) = \tilde{\dot{c}} + \nabla \cdot (\mathbf{D}\,\nabla \tilde{c}) \tag{68}$$

$$\partial_t|_z \tilde{u} + \nabla \cdot (\tilde{\mathbf{v}}\tilde{u}) - f\tilde{v} = \nabla \cdot \tau^u - \rho_0^{-1}\,\partial_x|_z \tilde{p} + \tilde{\mathcal{F}}_x \tag{69}$$

$$\partial_t|_z \tilde{v} + \nabla \cdot (\tilde{\mathbf{v}}\tilde{v}) + f\tilde{u} = \nabla \cdot \tau^v - \rho_0^{-1}\,\partial_y|_z \tilde{p} + \tilde{\mathcal{F}}_y \tag{70}$$

$$\partial_z \tilde{p} = -g\tilde{\rho}, \tag{71}$$

where the tilde denotes a model variable (which may be interpreted in terms of a suitable average), $\mathbf{D}$ represents a diffusivity tensor, and $\tau^{u/v}$ represents the zonal/meridional component of a viscous stress tensor. $\mathbf{v}_{GM}$ represents a divergence-free eddy advection, which, if included, is typically parameterized following Gent and McWilliams (1990). The GM tracer tendency can alternatively be expressed in terms of a "skew flux" as

$$\nabla \cdot (\mathbf{v}_{GM}\tilde{c}) = -\nabla \cdot (\mathbf{A}\nabla \tilde{c}) \tag{72}$$

where $\mathbf{A}$ is an antisymmetric tensor (see Griffies, 1998, for details). Although numerical implementations may use either the "advective" or "skew flux" representation of the GM parameterization, the two formulations are analytically equivalent. The non-parameterized terms in the tracer and momentum equations may also be expressed in multiple analytically equivalent forms, but this does not affect our conclusions here.

5.1.2. Interpretation as Mixed Eulerian/Isopycnal TWA Equations

McDougall and McIntosh (2001) argued that z-coordinate models using the GM parameterization should be interpreted in terms of the mixed Eulerian (momentum) and isopycnal TWA (tracer) interpretation, as given by Equations 62–66. In this interpretation $\tilde{\mathbf{v}} = \overline{\mathbf{v}}^z$, $\tilde{c} = \hat{c}^{b^\#}$, $\tilde{\dot{c}} = \hat{\dot{c}}^{b^\#}$, $\tilde{p} = \overline{p}^z$, $\tilde{\rho} = \overline{\rho}^z$, $\tilde{\mathcal{F}}_{x/y} = \overline{\mathcal{F}_{x,y}}^z$, the viscous stress tensor parameterizes the Eulerian mean eddy momentum fluxes (i.e., $\nabla \cdot \tau^{u/v} = -\nabla \cdot \overline{\mathbf{J^{u,v}}}^z$), the GM advection represents the "quasi-Stokes drift" ($\mathbf{v}_{GM} = \mathbf{v}^*$) (which, by construction, is not tracer dependent) and the diffusive tracer flux represents the isopycnal TWA eddy flux (i.e., $\nabla \cdot (\mathbf{D}\,\nabla\,\overline{c}^z) = -\nabla \cdot \widehat{\mathbf{J^c}}^{b^\#}$) (with any possible skew-flux component to the isopycnal TWA eddy flux neglected).

5.1.3. Interpretation as Eulerian Mean Equations

As discussed in the introduction, it is, however, still common for z-coordinate models to be interpreted in terms of the Eulerian mean Equations 38–42. In this case the tilde simply denotes an average at fixed depth (i.e., $\widetilde{()} = \overline{()}^z$), the eddy momentum fluxes are parameterized via a viscous stress (i.e., $\nabla \cdot \tau^{u/v} = -\nabla \cdot \overline{\mathbf{J}^{\mathbf{u},\mathbf{v}}}^{z}$), and the eddy tracer flux divergence is parameterized as

$$\nabla \cdot \overline{\mathbf{J}^{\mathbf{c}}}^{z} = -\nabla \cdot [(\mathbf{A} + \mathbf{D})\nabla \overline{c}^{z}] \tag{73}$$

where $\mathbf{A}$ is the antisymmetric GM skew flux tensor (Equation 72), and $\mathbf{D}$ is a symmetric diffusivity tensor, representing along-isopycnal mixing associated with mesoscale eddies and (the much smaller) diapycnal mixing by small-scale turbulence (Redi, 1982; Solomon, 1971).

The main caveat of the Eulerian mean interpretation is that, even for adiabatic flow with a conserved buoyancy variable (b) the Eulerian mean eddy buoyancy flux $\left(\overline{\mathbf{J}^{\mathbf{b}}}^{z}\right)$ is not strictly along buoyancy surfaces, unlike what is assumed in the formulation of the parameterizations. In addition, both the symmetric diffusivity tensor and the antisymmetric skew flux tensor (or, equivalently, the eddy advective velocity) are not necessarily tracer independent (e.g., Kamenkovich et al., 2021; Sun et al., 2021). However, the variance budget does require that the eddy buoyancy flux is *on-average* along isopycnals for adiabatic flow (e.g., Eden et al., 2007). The simplifying assumption to represent the eddy flux as locally along isopycnals is therefore not fundamentally different from the assumption that along-isopycnal eddy fluxes can be represented as locally down-gradient (which also can only be justified in a mean sense). Similarly, the possible tracer dependence also applies to diffusive closures, which are widely used (e.g., Adcroft et al., 2019; Gnanadesikan et al., 2015; Petersen et al., 2015). Given the much greater simplicity of the Eulerian mean interpretation, we therefore consider this interpretation to remain a reasonable alternative to the mixed Eulerian/residual-mean interpretation suggested by McDougall and McIntosh (2001).

5.1.4. Interpretation as Residual Mean Equations

Alternatively, Equations 47–51 provide a full "residual mean" interpretation of the z-coordinate model equations, although this requires a re-interpretation of the model's vertical coordinate as $\overline{z}$. With this interpretation, the model variables in Equation 67–71 are $\widetilde{\mathbf{u}} = \hat{\mathbf{u}}^{b}$ $\tilde{w} = w^{\#b}, \tilde{c} = \hat{c}^{b}, \widetilde{\dot{c}} = \widehat{\dot{c}}^{b}, \tilde{p} = \overline{p}^{b}, \tilde{\rho} = \hat{\rho}^{b}$ and $\tilde{\mathcal{F}}_{x,y} = \widehat{\mathcal{F}_{xy}}^{b}$. However, this interpretation is only consistent with Equations 67–71 if the GM parameterization is not used (i.e., $\mathbf{v}_{GM} = 0$). Moreover, the stress tensor then needs to represent the full EP flux $\left(\nabla \cdot \tau^{u,v} = -\overline{\nabla}^{b} \cdot \widehat{\mathbf{E}^{\mathbf{u},\mathbf{v}}}^{b}\right)$, including both the eddy momentum advection and the eddy form stress (cf. Equations 16 and 17). In the currently used model formulations, where the eddy stress is represented via a viscous stress tensor with relatively weak vertical viscosity, this interpretation is at best justified if the vertical component of the eddy form stress (i.e., the last term in Equations 16 and 17) is assumed negligible. However, the vertical eddy form stress can readily be incorporated into the viscous stress tensor following P. Rhines and Young (1982) and Greatbatch and Lamb (1990), as has been done by Ferreira and Marshall (2006), Zhao and Vallis (2008) and Saenz et al. (2015).

5.2. Hybrid and Generalized Vertical Coordinate Models

5.2.1. Model Equations

We here focus primarily on the model equations in MOM6, which we are most familiar with, and then briefly compare to other models. The MOM6 model equations are given by Adcroft et al. (2019):

$$\partial_t|_r \tilde{z}_r + \nabla_r \cdot (\tilde{z}_r(\widetilde{\mathbf{u}} + \mathbf{u}_{GM})) + \partial_r\left(\tilde{z}_r \widetilde{\dot{r}}\right) = 0 \tag{74}$$

$$\partial_t|_r(\tilde{z}_r \tilde{c}) + \nabla_r \cdot (\tilde{z}_r(\widetilde{\mathbf{u}} + \mathbf{u}_{GM})\,\tilde{c}) + \partial_r\left(\tilde{z}_r \widetilde{\dot{r}}\,\tilde{c}\right) = \tilde{z}_r \nabla \cdot (\mathbf{D}\,\nabla \tilde{c}) + \tilde{z}_r \widetilde{\dot{c}} \tag{75}$$

$$\partial_t|_r \tilde{u} + \widetilde{\dot{r}}\,\partial_r \tilde{u} - \left(f + \tilde{\zeta}\right)\tilde{v} + \partial_x|_r\left(|\widetilde{\mathbf{u}}|^2/2\right) = \nabla \cdot \tau^{u} - \rho_0^{-1}\,\partial_x|_{\tilde{z}}\tilde{p} + \tilde{\mathcal{F}}_x \tag{76}$$

$$\partial_t|_r \tilde{v} + \widetilde{\dot{r}}\,\partial_r \tilde{v} + \left(f + \tilde{\zeta}\right)\tilde{u} + \partial_y|_r\left(|\widetilde{\mathbf{u}}|^2/2\right) = \nabla \cdot \tau^{v} - \rho_0^{-1}\,\partial_y|_{\tilde{z}}\tilde{p} + \tilde{\mathcal{F}}_y \tag{77}$$

$$\partial_r \tilde{p} = -\tilde{\rho} g\,\partial_r \tilde{z}, \tag{78}$$

where the tilde denotes a model variable (which may be interpreted in terms of a suitable average), ∇ (without subscript) represents the 3-dimensional (coordinate-system independent) divergence and gradient operators, $\mathbf{D}$ represents a symmetric anisotropic diffusivity tensor, with a large isopycnal and much smaller diapycnal component, and $\tau^{u/v}$ represents the zonal/meridional component of the viscous stress tensor. $\mathbf{u}_{GM}$ represents an eddy advection, which, if included, is typically parameterized broadly following Gent and McWilliams (1990), although notice that, unlike in the original GM parameterization for z-coordinate models, $\mathbf{u}_{GM}$ here is a two-dimensional divergent velocity field rather than a non-divergent three-dimensional velocity.

For comparison with the equations discussed in this manuscript we can use the continuity equation to re-write the model's momentum equation in flux form as

$$\partial_t|_r(\tilde{z}_r\tilde{u}) + \nabla_r\cdot(\tilde{z}_r\tilde{\mathbf{u}}\,\tilde{u}) + \partial_r\left(\tilde{z}_r\tilde{\dot{r}}\,\tilde{u}\right) + \tilde{u}\,\nabla_r\cdot(\tilde{z}_r\mathbf{u}_{GM}) - f\tilde{z}_r\tilde{v} \;=\; \tilde{z}_r\nabla\cdot\tau^u - \rho_0^{-1}\tilde{z}_r\partial_x|_{\tilde{z}}\tilde{p} + \tilde{z}_r\tilde{\mathcal{F}}_x \tag{79}$$

$$\partial_t|_r(\tilde{z}_r\tilde{v}) + \nabla_r\cdot(\tilde{z}_r\tilde{\mathbf{u}}\,\tilde{v}) + \partial_r\left(\tilde{z}_r\tilde{\dot{r}}\,\tilde{v}\right) + \tilde{v}\,\nabla_r\cdot(\tilde{z}_r\mathbf{u}_{GM}) + f\tilde{z}_r\tilde{u} \;=\; \tilde{z}_r\nabla\cdot\tau^v - \rho_0^{-1}\tilde{z}_r\partial_y|_{\tilde{z}}\tilde{p} + \tilde{z}_r\tilde{\mathcal{F}}_y. \tag{80}$$

Notice that the divergence of the GM advection appears here as a result of its appearance in the continuity equation. This formulation illustrates that the inclusion of a divergent GM advection in the continuity equation but not the momentum equations violates the conservation of momentum.

The HYCOM model solves principally the same equations as MOM6 (74–78), except that the GM advective velocity is computed based on a (biharmonic) interface height diffusion (Bleck, 2002). The MPAS ocean model (Ringler et al., 2013), instead uses a 3-dimensional divergence-free eddy advective velocity more directly following Gent and McWilliams (1990) and Gent et al. (1995).

Notice that MOM6 and HYCOM employ a semi-Lagrangian temporal discretization (e.g. Durran, 2010, Chapter 7), where the vertical coordinate follows the flow during time-stepping (i.e., $\tilde{\dot{r}} = 0$ in Equations 74–77) and the model state is then re-mapped onto a target grid between time steps (Bleck, 2002; Griffies et al., 2020).

5.2.2. Interpretation in Terms of Coordinate-Following Generalized TWA

One plausible interpretation of generalized vertical coordinate models is in terms of the generalized TWA equations with the average following the model coordinate, as sketched in Figure 4 (Equations 8–12). In this case $\tilde{z} = \overline{z}^r$, $\tilde{\mathbf{u}} = \hat{\mathbf{u}}^r$, $\tilde{\dot{r}} = \hat{\dot{r}}^r$, $\tilde{c} = \hat{c}^r$, $\tilde{\dot{c}} = \hat{\dot{c}}^r$, $\tilde{p} = \overline{p}^r$, $\tilde{\rho} = \hat{\rho}^r$, $\tilde{\mathcal{F}}_{x/y} = \widehat{\mathcal{F}_{xy}}^r$, and the diffusive tracer flux parameterizes the generalized TWA eddy tracer transport (i.e., $\nabla\cdot(\mathbf{D}\,\nabla\,\tilde{c}) = -\overline{\nabla}^r\cdot\widehat{\mathbf{J}^{\mathbf{c}}}^r$). However, as the generalized TWA equations do not include an eddy contribution in the continuity equation, this interpretation is only consistent with the model Equations 74–78 if no GM parameterization is used (i.e., $\mathbf{u}_{GM} = 0$). Moreover, the stress tensor in this interpretation needs to represent both the eddy momentum advection and the generalized eddy form stress $\left(\nabla\cdot\tau^{u/v} = -\overline{\nabla}^r\cdot\widehat{\mathbf{E}^{\mathbf{u},\mathbf{v}}}^r\right)$ — cf. Equations 18 and 19. Finally, we note that $\widehat{\mathbf{J}^{\mathbf{c}}}^r$ may in general have a skew-flux component in addition to a diffusive component, which would need to be assumed negligible when using a symmetric diffusivity tensor.

A challenge with the generalized TWA interpretation of a hybrid coordinate model is that the physical interpretation of the variables, and, perhaps more importantly, of the eddy terms that need to be parameterized, changes throughout the domain (e.g., as the coordinate transitions from being isopycnal to Eulerian), thus requiring parameterizations to be "coordinate system aware." This situation poses a challenge for parameterization development, although it is not necessarily an insurmountable problem. Indeed the interface height diffusion (which determines $\mathbf{u}_{GM}$) in HYCOM is coordinate aware as it acts on the layer interface height, and is thus automatically turned off where the coordinate follows z-surfaces, although the appearance of this closure in the continuity equation is not consistent with the equations derived here. At present, we are not aware of any existing hybrid vertical coordinate model that employs coordinate-aware parameterizations consistent with the coordinate-following generalized TWA equations in Equations 8–12.

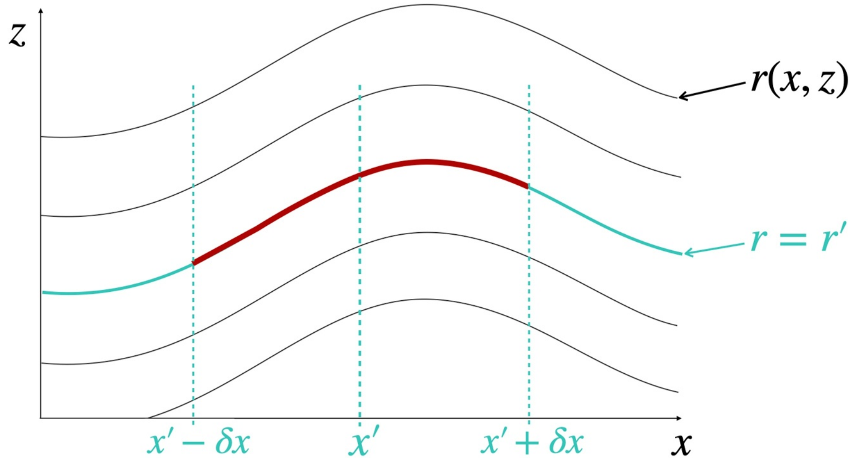

Figure 4. Sketch of coordinate-following average interpretation, where the average is here assumed to be a box filter along the x-dimension with width $2\delta x$. The thin black lines show iso-surfaces of the vertical coordinate "r" in (x-z) space, with the cyan line indicating a specific r-surface for which $r = r'$. In the coordinate-following average interpretation, a model variable at the point (x', r') then represents an average of the original field along the r-surface $r = r'$ between $x' - \delta x$ and $x' + \delta x$—that is the surface highlighted by the thick red line.

5.2.3. Interpretation in Terms of Eulerian Mean Equations

If we regard the choice of model coordinate system as independent of the averaging, the Eulerian average (i.e., Equations 31–35) may provide a simple interpretation also for generalized vertical coordinate models, although this interpretation is again consistent with the model equations in Equations 74–78 only if no GM parameterization is used (i.e., $\mathbf{u}_{GM} = 0$).

In the Eulerian mean interpretation, sketched in Figure 5, the model variables are interpreted as $\tilde{z} = z$, $\widetilde{\mathbf{u}}, = \overline{\mathbf{u}}^z$, $\widetilde{\dot{r}} = \dot{r}^{\#z}$, $\tilde{c} = \overline{c}^z$, $\widetilde{\dot{c}} = \overline{\dot{c}}^z$, $\tilde{p} = \overline{p}^z$, $\tilde{\rho} = \overline{\rho}^z$ and $\tilde{\mathcal{F}}_{x/y} = \overline{\mathcal{F}_{x,y}}^z$. The viscous stress tensor needs to represent the effect of the Eulerian eddy momentum fluxes (i.e., $\nabla \cdot \tau^{\mathbf{u}/\mathbf{v}} = -\nabla \cdot \overline{\mathbf{J}^{\mathbf{u}/\mathbf{v}}}^z$) and the tracer diffusion needs to capture the Eulerian eddy tracer flux (i.e., $\nabla \cdot (\mathbf{D} \nabla \tilde{c}) = -\nabla \cdot \overline{\mathbf{J}^{\mathbf{c}}}^z$). Notice that in the semi-Lagrangian time discretization we

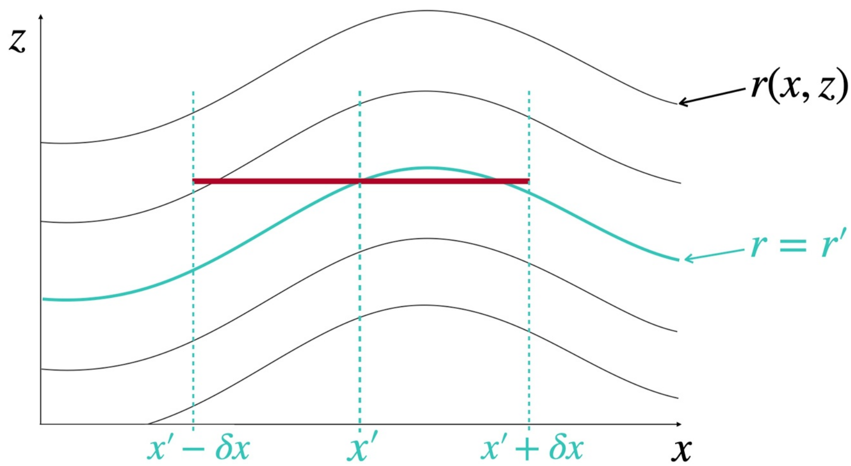

Figure 5. Sketch of the Eulerian average interpretation, where the average is again assumed to be a box filter along the x-dimension with width $2\delta x$. The thin black lines show iso-surfaces of the vertical coordinate "r" in (x-z) space, with the cyan line highlighting a specific r-surface with $r = r'$. In the Eulerian-mean interpretation, a model variable at the point (x', r') represents an average of the original field taken at fixed height between $x' - \delta x$ and $x' + \delta x$—that is the surface highlighted by the thick red line. An exception is the model's cross-coordinate velocity, $\widetilde{\dot{r}}$, which is defined as in Equation 36 as the rate of change of r following the Eulerian-mean flow.

10.1029/2024MS004506

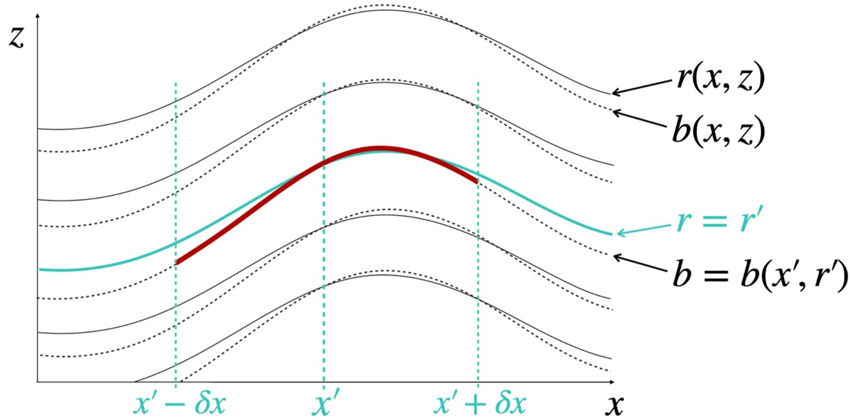

Figure 6. Sketch of isopycnal average interpretation, where the average is again assumed to be a box filter along the x-dimension with width $2\delta x$. Isopycnal surfaces (i.e., surfaces of constant b) are sketched as dashed lines. The solid black lines show iso-surfaces of the vertical coordinate "r," with the cyan line highlighting the specific r-surface with $r = r'$. In the isopycnal mean interpretation, a model variable at the point (x', r') represents an average of the original field taken along a buoyancy surface between $x' - \delta x$ and $x' + \delta x$— that is the surface highlighted by the thick red line. An exception is again the model's cross-coordinate velocity, $\widetilde{\dot{r}}$, which is defined as the rate of change of r following the isopycnal thickness weighted mean flow (see Equation 22).

set $\widetilde{\dot{r}} = \dot{r}^{\#z} = 0$, which in this interpretation implies that the vertical coordinate follows the Eulerian mean flow (i.e., between remapping steps the coordinate is Lagrangian with respect to the Eulerian mean flow rather than the full unaveraged flow). Notice also that in hybrid coordinate models where the coordinate follows isopycnals in the interior, we set $r = b(\tilde{\theta}, \tilde{S}, z) = b(\overline{\theta}^z, \overline{S}^z, z)$, that is the vertical coordinate is itself defined in terms of averaged quantities (which guarantees that model variables will be smooth along the model coordinate surfaces).

However, the implementation of GM in MOM6 and HYCOM is inconsistent with the Eulerian mean continuity Equation 31, which does not contain an eddy term. Instead, the "advective" effect of eddies appears as a skew flux contribution to the eddy term in the tracer equations in this interpretation, which could be parameterized as in z-coordinate models via either an antisymmetric component to the eddy diffusivity tensor or a 3D eddy advection, as done in the MPAS ocean model.

The interpretation of the model variables as Eulerian averages is arguably desirable for the justification of boundary conditions and boundary layer parameterizations, which are generally formulated assuming the Eulerian mean flow as given. In models, such as MPAS, that implement GM via a divergence-free advection in the continuity equation, the Eulerian mean interpretation therefore is a convenient choice. However, for numerical reasons, the same implementation of GM would be disadvantageous in semi-Lagrangian models with isopycnal target coordinates in the ocean interior, such and MOM6 and HyCOM. We will elaborate on this issue in Section 6.

5.2.4. Interpretation in Terms of Isopycnal TWA Equations

If the GM parameterization is not used (i.e., $\mathbf{u}_{GM} = 0$), Equations 74–78 can also be interpreted in terms of the isopycnal TWA (Equations 25–29 with $a \rightarrow b$), as sketched in Figure 6. In this interpretation, $\tilde{z} = \overline{z}^b$, $\widetilde{\mathbf{u}}, = \hat{\mathbf{u}}^b$, $\widetilde{\dot{r}} = \dot{r}^{\#b}$, $\tilde{c} = \hat{c}^b$, $\widetilde{\dot{c}} = \widehat{\dot{c}}^b$, $\tilde{p} = \overline{p}^b$, $\tilde{\rho} = \hat{\rho}^b$ and $\tilde{\mathcal{F}}_{x/y} = \widehat{\mathcal{F}_{xy}}^b$. The tracer diffusion needs to capture the isopycnal TWA eddy tracer flux, that is $\nabla \cdot (\mathbf{D} \nabla \tilde{c}) = -\overline{\nabla}^b \cdot \widehat{\mathbf{J}^{\mathbf{c}}}^b$. Treating the diffusivity tensor $\mathbf{D}$ as symmetric then implies that we are ignoring any potential skew-flux contribution to $\widehat{\mathbf{J}^{\mathbf{c}}}^b$. Finally, the viscous stress tensor needs to generally represent both the advective isopycnal TWA eddy momentum fluxes and the eddy form stress (i.e., $\nabla \cdot \tau^{\mathbf{u/v}} = -\overline{\nabla}^b \cdot \widehat{\mathbf{E}^{\mathbf{u/v}}}^b$). When using semi-Lagrangian time-stepping, where we set $\widetilde{\dot{r}} = 0$, the implication of the isopycnal-TWA interpretation is that the vertical coordinate is Lagrangian with respect to the isopycnal TWA flow.

A potential caveat of the isopycnal TWA interpretation is that bulk formulas and boundary layer parameterizations tend to be formulated in terms of Eulerian mean quantities, which are not available in this formulation.

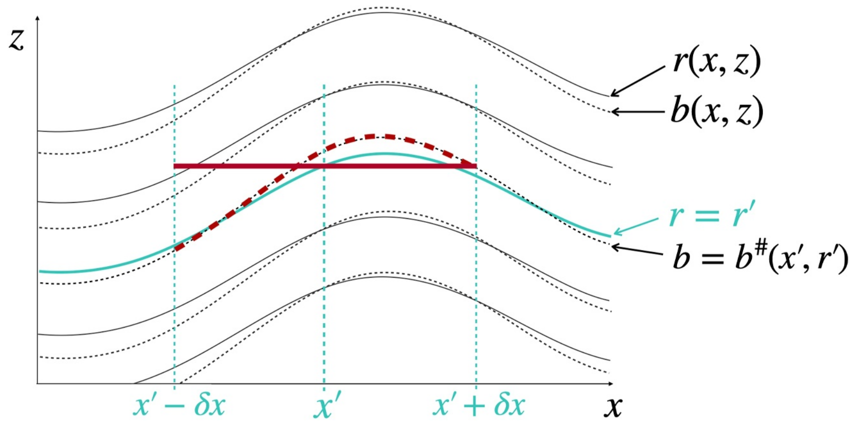

Figure 7. Sketch of the mixed Eulerian/isopycnal average interpretation, where the average is again assumed to be a box filter along the x-dimension with width $2\delta x$. Isopycnal surfaces (i.e., surfaces of constant b) are sketched as dashed lines. The solid black lines show iso-surfaces of the vertical coordinate "r," with the cyan line highlighting the specific r-surface with $r = r'$. In the mixed interpretation, the tracer variables at the point (x', r') represent a thickness-weighted average of the original fields taken along a buoyancy surface between $x' - \delta x$ and $x' + \delta x$— that is the surface highlighted by the dashed red line, while the velocities, pressure and density are averaged at fixed height —that is along the solid red line. An exception is again the model's cross-coordinate velocity, $\widetilde{\dot{r}}$, which is defined as the rate of change of r following the Eulerian mean flow (see Equation 36).

Perhaps more importantly, however, the implementation of the "GM" parameterization in Equations 74–78 is again generally inconsistent with the isopycnal TWA interpretation. Specifically the appearance of a "GM" eddy advection term in the continuity equation is again inconsistent with the isopycnal TWA equations which do not include an eddy contribution in the continuity equation (Equation 25 with $a \to b$). Instead, the effect of mesoscale eddies in reducing isopycnal slopes would need to be parameterized via the eddy form stress in the momentum equations (which in turn drives an ageostrophic circulation that tends to flatten isopycnals). A form-stress parameterization (that achieves a similar effect as the GM parameterization) can be implemented via an enhanced vertical viscosity (see P. Rhines & Young, 1982; Greatbatch & Lamb, 1990; Ferreira & Marshall, 2006; Zhao & Vallis, 2008; Loose et al., 2023).

5.2.5. Interpretation in Terms of Mixed Eulerian/Isopycnal TWA

Without the GM parameterization, and assuming the quasi-Stokes advection in Equations 56 and 57 to be negligible (i.e., $\mathbf{u}^* \approx 0$ and $\dot{r}^* \approx 0$), the model Equations 74–78 can also be interpreted in terms of the mixed Eulerian/isopycnal average in Equations 56–60. In this interpretation, which is sketched in Figure 7, $\tilde{z} = z$, $\widetilde{\mathbf{u}} = \overline{\mathbf{u}}^z, \widetilde{\dot{r}} = \dot{r}^{\#z}, \tilde{c} = \hat{c}^{b^\#}, \widetilde{\dot{c}} = \widehat{\dot{c}}^{b^\#}, \tilde{p} = \overline{p}^z, \tilde{\rho} = \overline{\rho}^z$ and $\tilde{\mathcal{F}}_{x/y} = \overline{\mathcal{F}_{x,y}}^z$. The tracer diffusion then needs to capture the isopycnal TWA eddy tracer flux (i.e., $\nabla \cdot (\mathbf{D} \nabla \tilde{c}) = -\nabla \cdot \widehat{\mathbf{J}^{\mathbf{c}}}^{b^\#}$) (where a symmetric diffusivity tensor again amounts to neglecting any skew-flux contribution to $\widehat{\mathbf{J}^{\mathbf{c}}}^{b^\#}$), and the viscous stress tensor needs to represent the Eulerian eddy momentum fluxes (i.e., $\nabla \cdot \tau^{\mathbf{u}/\mathbf{v}} = -\nabla \cdot \overline{\mathbf{J}^{\mathbf{u}/\mathbf{v}}}^z$). In semi-Lagrangian time-stepping, where we set $\widetilde{\dot{r}} = 0$, the implication is that the coordinate follows the Eulerian mean flow.

Unfortunately, the mixed interpretation is still not consistent with the model Equations 74–78 when the GM advection is included. While the appearance of the *horizontal* eddy advection in the continuity and tracer Equations 56 and 57 is consistent with the GM implementation in MOM6, the continuity and tracer equations in the mixed interpretation also include the vertical eddy advection, while the momentum equations do not. That is, in Equations 74 and 75 we would need to assume $\widetilde{\dot{r}} = \dot{r}^{\#z} + \dot{r}^*$ to achieve consistency with the mixed interpretation, but in Equations 76 and 77 we would need to assume $\widetilde{\dot{r}} = \dot{r}^{\#z}$, so there is no choice of $\widetilde{\dot{r}}$ that leads to a complete consistent set of equations.

Similar to the Eulerian mean interpretation, the model equations could be made consistent with the mixed Eulerian/isopycnal TWA formulation by including a 3-dimensional divergence-free eddy advection (representing

$\mathbf{u}^*$ and $\dot{r}^*$ in Equation 57) in the tracer equations, and either including the same 3-dimensional eddy advection in the continuity equation, or removing the eddy advection in the continuity equation altogether (A divergence-free eddy advection has no effect in the continuity equation.)

6. Conclusions

We derived and discussed the arbitrarily averaged equations in generalized vertical coordinates (Equations 25–29). The equations can be written in a form that mirrors the unaveraged equations, with a relatively straightforward interpretation of the "resolved" variables (which, however, are not all directly equal to the average of the respective variable) plus additional eddy forcing terms that can be written in the form of eddy flux divergences. These eddy flux divergences are fundamentally coordinate-system independent, but instead depend on the averaging coordinate. The implication is that eddy parameterizations need to be developed specific to the choice of average but not the choice of the model coordinate (Although certain averaging choices may be more natural and/or numerically advantageous for different model coordinate systems.)

We also considered special cases for common averages (Eulerian and isopycnal) and formulated the resulting equations in generalized and specific coordinate systems, which allows us to (a) recover known results for isopycnal and Eulerian mean equations and (b) consider candidates for the interpretation of existing generalized vertical coordinate models.

Various interpretations (Eulerian mean, isopycnal TWA, or a mixed Eulerian/isopycnal interpretation) are consistent with the existing generalized vertical coordinate model formulations if no GM parameterization (or interface-height diffusion) is used, and the eddy form stress and/or skew fluxes are assumed to be negligible (as may be appropriate for "eddy resolving" models).

However, no interpretation has been found that is consistent with the implementation of the GM parameterization (or interface height diffusion) as a 2D divergent eddy advection in the continuity and tracer equations (as used in MOM6 and HyCOM). This lack of a physical interpretation severely limits our ability to improve or test the parameterizations. We therefore suggest that implementation of the "GM" parameterization (or a dynamically similar parameterization of the form stress—e.g., Greatbatch & Lamb, 1990; Loose et al., 2023) in generalized vertical coordinate models should be modified to match one of the interpretations discussed in this paper. The candidates are essentially the same as for z-coordinate models, although numerical considerations may make an isopycnal TWA interpretation particularly desirable for semi-Lagrangian discretizations (as used in MOM6 and HYCOM) where the vertical coordinate approximately follows isopycnals in the interior.

Consistency with both the Eulerian mean and the mixed Eulerian/TWA interpretation can be achieved by implementing the GM parameterization via a 3D divergence-free advection (or, equivalently, a skew flux) in the tracer equation only, as done in the MPAS ocean model. In semi-Lagrangian models, this implementation would imply that the vertical coordinate follows the Eulerian mean flow.

For models that use isopycnal target-coordinates in the interior, it is, however, numerically advantageous for the Lagrangian coordinate to follow the residual flow (as this requires no re-interpolation in an adiabatic isopycnal limit). This goal can be achieved by using the isopycnal TWA interpretation, in which case the GM parameterization needs to be replaced with a closure in the momentum equations, for example, following Greatbatch and Lamb (1990). Such a closure has recently been implemented in MOM6 and tested successfully in a purely isopycnal configuration (Loose et al., 2023). Since the eddy form stress depends on the choice of averaging, and not the model coordinate system, the same closure should readily be applicable in generalized vertical coordinates.

The interpretation of generalized vertical coordinate models in terms of a coordinate-following average is also interesting to entertain, as it most naturally conforms to the model's numerics. Indeed, the finite volume discretization naturally implies a volume-weighted grid-box average following the model coordinate (Griffies et al., 2020). For a hybrid isopycnal-z-coordinate model, the coordinate-following average also has the advantage that it naturally reduces to an Eulerian average near the surface and where stratification vanishes (which is useful for parameterizations in those regions). The major hurdle, however, is that the parameterizations will have to be "coordinate system aware," which is expected to significantly complicate parameterization development.

We focus in this paper on conservative "thickness-weighted" averages with the generalized thickness defined as $\partial_a z$ for any averaging coordinate, a. As discussed in Loose et al. (2023) for the specific case of isopycnal averaging, any non-thickness weighted average following a coordinate surface with non-constant thickness (i.e., $\partial_a z \neq$ const.) is non-conservative and hence leads to non-conservative equations for the mean quantities. Accepting this major limitation, one could interpret the velocities in isopycnal and perhaps generalized vertical coordinate models as non-thickness-weighted isopycnal averages, which then introduces the bolus transport into the thickness weighted continuity and tracer equations. As discussed in Appendix C, this interpretation is arguably the most consistent with the implementation of the "GM" parameterization in existing isopycnal coordinate models, although we consider the non-conservative nature of the equations to be highly undesirable. For generalized vertical coordinate models this interpretation moreover leads to an inconsistency in the treatment of the vertical velocity between the continuity and momentum equations.

We end by noting that consistency of the model equations with the averaged equations is not just desirable for theoretical reasons but is fundamental for parameterization development, which relies on a clear definition of the eddy terms that need to be parameterized. The need for a consistent and agreed upon definition is particularly urgent given the recent rise in data-driven parameterization development, where parameterizations are "trained" offline based on filtered high-resolution data sets (e.g., Bachman et al., 2015, 2020; Guillaumin & Zanna, 2021; Perezhogin et al., 2023; Zanna & Bolton, 2020; Zhang et al., 2023). When applied online, these parameterizations can only be expected to be successful if the numerical model formulation is consistent with the filtering operation assumed in the training of the parameterizations.

Appendix A: The Averaged Pressure Gradient

Using that $\partial_x|_z = \partial_x|_r - \left(\partial_x|_r z\right)(\partial_r z)^{-1}\partial_r$ and assuming a Reynolds average, the thickness weighted pressure gradient acceleration and be written as

$$\overline{z_r}^r \widehat{\partial_x|_z p}^r = \overline{z_r \partial_x|_z p}^r \tag{A1a}$$

$$= \overline{z_r \partial_x|_r p}^r - \overline{\partial_r p\, \partial_x|_r z}^r \tag{A1b}$$

$$= \partial_x|_r\left(\overline{z_r p}^r\right) - \partial_r\left(\overline{p \partial_x|_r z}^r\right) \tag{A1c}$$

$$= \partial_x|_r\left(\overline{z_r}^r \overline{p}^r\right) - \partial_r\left(\overline{p}^r \partial_x|_r \overline{z}^r\right) + \partial_x|_r\left(\overline{z_r' p'}^r\right) - \partial_r\left(\overline{p' \partial_x|_r z'}^r\right) \tag{A1d}$$

$$= \overline{z_r}^r \partial_x|_r \overline{p}^r - \partial_r \overline{p}^r \partial_x|_r \overline{z}^r + \partial_x|_r\left(\overline{z_r' p'}^r\right) - \partial_r\left(\overline{p' \partial_x|_r z'}^r\right) \tag{A1e}$$

$$= \overline{z_r}^r \partial_x|_{\overline{z}^r} \overline{p}^r + \partial_x|_r\left(\overline{z_r' p'}^r\right) - \partial_r\left(\overline{p' \partial_x|_r z'}^r\right) \tag{A1f}$$

and similarly for the y-component. The general form (not assuming a Reynolds average) is readily obtained by substituting $\overline{z_r' p'}^r \rightarrow \overline{z_r p}^r - \overline{z_r}^r \overline{p}^r$ and $\overline{p' \partial_x|_r z'}^r \rightarrow \overline{p \partial_x|_r z}^r - \overline{p}^r \partial_x|_r \overline{z}^r$.

The pressure gradient acceleration can alternatively be expressed (again, assuming a Reynolds average) as

$$\widehat{\partial_x|_z p}^r = \overline{\partial_x|_z p}^r + \frac{1}{\overline{z_r}^r} \overline{z_r' \partial_x|_z p'}^r \tag{A2a}$$

$$= \partial_x|_r \overline{p}^r + \overline{\rho g \partial_x|_r z}^r + \frac{1}{\overline{z_r}^r} \overline{z_r' \partial_x|_z p'}^r \tag{A2b}$$

$$= \partial_x|_r \overline{p}^r + \hat{\rho}^r g \partial_x|_r \overline{z}^r + \overline{\rho'' g \partial_x|_r z}^r + \frac{1}{\overline{z_r}^r} \overline{z_r' \partial_x|_z p'}^r \tag{A2c}$$

$$= \partial_x|_{\overline{z}^r} \overline{p}^r + \overline{\rho'' g \partial_x|_r z}^r + \frac{1}{\overline{z_r}^r} \overline{z_r' \partial_x|_z p'}^r \tag{A2d}$$

(and analogously for $\widehat{\partial_y|_z p}^{r}$). Notice that the mean part of the pressure gradient force is given by the gradient of the mean pressure at fixed $\overline{z}^r$. The second term on the RHS of Equation A2d is equal to the difference between the mean pressure gradient at fixed z and the gradient of the mean pressure at fixed $\overline{z}^r$: $\overline{\partial_x|_z p}^{r} - \partial_x|_{\overline{z}^r}\overline{p}^r$. The specific form of this term was here derived assuming that $\overline{\hat{\rho}^r \partial_x|_r z}^{r} = \hat{\rho}^r \overline{\partial_x|_r z}^{r}$ (in the step leading to A2c), which, in addition to the usual Reynolds properties, requires that $\overline{1/\overline{z}^r}^{r} = 1/\overline{z}^r$. Although not formally a Reynolds property, this assumption holds for commonly used Reynolds averages (such as ensemble or zonal averages). More generally, the term simply becomes $\overline{\rho g \partial_x|_r z}^{r} - \hat{\rho}^r g \partial_x|_r \overline{z}^r$. The last term on the RHS of Equation A2d can be directly related to the geostrophic eddy thickness flux: $\overline{z_r' \partial_x|_z p'}^{r} = f\rho_0 \overline{z_r' v_g'}^{r}$, where we assumed that $f \neq 0$. For a spatial filter that is not a Reynolds average, this term becomes $\overline{z_r \partial_x|_z p}^{r} - \overline{z_r}^{r}\overline{\partial_x|_z p}^{r} = f\rho_0\left(\overline{z_r v_g}^{r} - \overline{z_r}^{r}\overline{v_g}^{r}\right)$.

Appendix B: Derivation of Equation 21

Using the following identities

$$\partial_a = (\partial_r a)^{-1}\partial_r, \quad \nabla_a = \nabla_r - (\nabla_r a)(\partial_r a)^{-1}\partial_r, \quad \partial_t|_a = \partial_t|_r - \left(\partial_t|_r a\right)(\partial_r a)^{-1}\partial_r, \tag{B1}$$

we find

$$\begin{aligned}\partial_t|_a\left(\overline{z_a}^a \hat{c}^a\right) + \nabla_a\cdot\left(\overline{z_a}^a \hat{\mathbf{u}}^a \hat{c}^a\right) + \partial_a\left(\overline{z_a}^a \hat{\dot{a}}^a \hat{c}^a\right) &= \partial_t|_r\left(\overline{z_a}^a \hat{c}^a\right) - (\partial_r a)^{-1}\left(\partial_t|_r a\right)\partial_r\left(\overline{z_a}^a \hat{c}^a\right)\\ &\quad + \nabla_r\cdot\left(\overline{z_a}^a \hat{\mathbf{u}}^a \hat{c}^a\right) - (\partial_r a)^{-1}\nabla_r a\cdot\partial_r\left(\overline{z_a}^a \hat{\mathbf{u}}^a \hat{c}^a\right)\\ &\quad + (\partial_r a)^{-1}\partial_r\left(\overline{z_a}^a \hat{\dot{a}}^a \hat{c}^a\right).\end{aligned} \tag{B2}$$

With $\overline{z}^a_r \equiv \partial_r \overline{z}^a = \partial_a \overline{z}^a \partial_r a = \overline{z_a}^a \partial_r a$ we then get

$$\begin{aligned}\frac{\overline{z}^a_r}{\overline{z_a}^a}\left[\partial_t|_a\left(\overline{z_a}^a \hat{c}^a\right) + \nabla_a\cdot\left(\overline{z_a}^a \hat{\mathbf{u}}^a \hat{c}^a\right) + \partial_a\left(\overline{z_a}^a \hat{\dot{a}}^a \hat{c}^a\right)\right] &= \partial_r a\, \partial_t|_r\left(\overline{z_a}^a \hat{c}^a\right) - \left(\partial_t|_r a\right)\partial_r\left(\overline{z_a}^a \hat{c}^a\right)\\ &\quad + \partial_r a\, \nabla_r\cdot\left(\overline{z_a}^a \hat{\mathbf{u}}^a \hat{c}^a\right) - \nabla_r a\cdot\partial_r\left(\overline{z_a}^a \hat{\mathbf{u}}^a \hat{c}^a\right)\\ &\quad + \partial_r\left(\overline{z_a}^a \hat{\dot{a}}^a \hat{c}^a\right) \qquad \text{(B3a)}\\ &= \partial_t|_r\left(\overline{z}^a_r \hat{c}^a\right) - \left(\overline{z_a}^a \hat{c}^a \partial_t|_r(\partial_r a) + \partial_t|_r a\, \partial_r\left(\overline{z_a}^a \hat{c}^a\right)\right)\\ &\quad + \nabla_r\cdot\left(\overline{z}^a_r \hat{\mathbf{u}}^a \hat{c}^a\right) - \left(\overline{z_a}^a \hat{\mathbf{u}}^a \hat{c}^a\cdot\nabla_r(\partial_r a) + \nabla_r a\cdot\partial_r\left(\overline{z_a}^a \hat{\mathbf{u}}^a \hat{c}^a\right)\right)\\ &\quad + \partial_r\left(\overline{z_a}^a \hat{\dot{a}}^a \hat{c}^a\right) \qquad \text{(B3b)}\\ &= \partial_t|_r\left(\overline{z}^a_r \hat{c}^a\right) + \nabla_r\cdot\left(\overline{z}^a_r \hat{\mathbf{u}}^a \hat{c}^a\right)\\ &\quad + \partial_r\left(-\overline{z_a}^a \hat{c}^a \partial_t|_r a - \overline{z_a}^a \hat{\mathbf{u}}^a \hat{c}^a\cdot\nabla_r a + \overline{z_a}^a \hat{\dot{a}}^a \hat{c}^a\right) \qquad \text{(B3c)}\\ &= \partial_t|_r\left(\overline{z}^a_r \hat{c}^a\right) + \nabla_r\cdot\left(\overline{z}^a_r \hat{\mathbf{u}}^a \hat{c}^a\right) + \partial_r\left(\overline{z}^a_r\, \dot{r}^{\#a} \hat{c}^a\right) \qquad \text{(B3d)}\end{aligned}$$

with $\dot{r}^{\#a}$ as defined in Equation 22.

10.1029/2024MS004506

Appendix C: Semi-Thickness-Weighted Averaging

Adcroft et al. (2019) argue that the GM advective velocity in MOM6 is to be interpreted in terms of the "bolus" velocity, which in turn is usually defined as the difference between the thickness-weighted and non-thickness-weighted isopycnally averaged velocity, that is, $\mathbf{u}_b = \hat{\mathbf{u}}^b - \overline{\mathbf{u}}^b = \overline{\mathbf{u}' z_b'}^b / \overline{z}_b^b$, where the second equality assumes a Reynolds average (P. B. Rhines, 1982; McDougall & McIntosh, 2001; Vallis, 2006). The bolus velocity appears in the continuity and tracer equations if (and only if) they are averaged along isopycnals with thickness-weighting, but the velocities solved for in the momentum equations represent non-thickness-weighted isopycnal averages. These semi-thickness-weighted isopycnal average equations are derived in this Appendix. The derivation follows the same approach as taken in the main part of this manuscript, that is we first take the average of the equations in isopycnal (b) coordinates, and then transform the averaged equations into generalized (r) coordinates. For simplicity we here assume a Reynolds average. The generalization to a non-Reynolds average works analogously to the results in the main manuscript and does not affect any of our conclusions.

Combining the thickness-weighted average continuity and tracer (Equations 8 and 9 with $r \rightarrow b$), with a non-thickness weighted average of the momentum (Equations 3 and 4 with $r \rightarrow b$), we can obtain a set of b-averaged equations of the following form (where for discussion purposes we will here assume b to represent a suitably defined buoyancy variable, but formally it can be an arbitrary field that is monotonic in depth):

$$\partial_t|_b \overline{z_b}^b + \nabla_b \cdot \left(\overline{z_b}^b \left(\overline{\mathbf{u}}^b + \mathbf{u}_b \right) \right) + \partial_b \left(\overline{z_b}^b \left(\overline{\dot{b}}^b + \dot{b}_b \right) \right) = 0 \tag{C1}$$

$$\partial_t|_b \left(\overline{z_b}^b \, \hat{c}^b \right) + \nabla_b \cdot \left(\overline{z_b}^b \left(\overline{\mathbf{u}}^b + \mathbf{u}_b \right) \hat{c}^b \right) + \partial_b \left(\overline{z_b}^b \left(\overline{\dot{b}}^b + \dot{b}_b \right) \hat{c}^b \right) = -\overline{z_b}^b \overline{\nabla}^b \cdot \widehat{\mathbf{J}^{\mathbf{C}}}^b + \overline{z_b \dot{c}}^b \tag{C2}$$

$$\partial_t|_b \overline{u}^b + \overline{\mathbf{u}}^b \cdot \nabla_b \overline{u}^b + \overline{\dot{b}}^b \partial_b \overline{u}^b - f \overline{v}^b = -\overline{\mathbf{u}' \cdot \nabla_b u'}^b - \overline{\dot{b}' \partial_b u'}^b - \rho_0^{-1} \overline{\partial_x|_z p}^b + \overline{\mathcal{F}_x}^b \tag{C3}$$

$$\partial_t|_b \overline{v}^b + \overline{\mathbf{u}}^b \cdot \nabla_b \overline{v}^b + \overline{\dot{b}}^b \partial_b \overline{v}^b + f \overline{u}^b = -\overline{\mathbf{u}' \cdot \nabla_b v'}^b - \overline{\dot{b}' \partial_b v'}^b - \rho_0^{-1} \overline{\partial_y|_z p}^b + \overline{\mathcal{F}_y}^b \tag{C4}$$

$$\partial_b \overline{p}^b = -\hat{\rho}^b g \overline{z_b}^b \tag{C5}$$

where $\dot{b}_b = \overline{z_b' \dot{b}'}^b / \overline{z_b}^b$, and the pressure gradient term can be expressed in b-coordinates as

$$\overline{\partial_x|_z p}^b = \partial_x|_b \overline{p}^b + \overline{\rho g \partial_x|_b z}^b \tag{C6a}$$

$$= \partial_x|_b \overline{p}^b + \hat{\rho}^b g \partial_x|_b \overline{z}^b + \overline{\rho'' g \partial_x|_b z}^b \tag{C6b}$$

$$= \partial_x|_{\overline{z}^b} \overline{p}^b + \overline{\rho'' g \partial_x|_b z}^b, \tag{C6c}$$

and similarly for $\overline{\partial_y|_z p}^b$. The eddy term in Equation C6c vanishes if ρ is constant along b-surfaces (as is the case for a linear equation of state with a buoyancy variable defined via Equation 43).

Notice that the eddy momentum flux contributions in Equations C3 and C4 are not in the form of a flux-divergence and are hence not conservative, as pointed out by Loose et al. (2023). This is in disagreement with the usual implementation of the eddy stress in numerical ocean models. Moreover, the appearance of the bolus transport in the continuity equation leads to additional non-conservative terms in the momentum budget, associated with the divergence of the bolus transport, which becomes apparent when formulating the momentum equations in flux form:

$$\partial_t|_b\left(\overline{z_b}^b\overline{u}^b\right)+\nabla_b\cdot\left(\overline{z_b}^b\overline{\mathbf{u}}^b\overline{u}^b\right)+\partial_b\left(\overline{z_b}^b\overline{\dot{b}}^b\overline{u}^b\right)+\overline{u}^b\nabla_b\cdot\left(\overline{z_b}^b\mathbf{u}_b\right)+\overline{u}^b\partial_b\left(\overline{z_b}^b\dot{b}_b\right)-f\overline{z_b}^b\overline{v}^b$$
$$=-\overline{z_b}^b\overline{\mathbf{u}'\cdot\nabla_b u'}^b-\overline{z_b}^b\overline{\dot{b}'\partial_b u'}^b-\frac{\overline{z_b}^b}{\rho_0}\overline{\partial_x|_z p}^b+\overline{z_b}^b\overline{\mathcal{F}_x}^b \tag{C7}$$

$$\partial_t|_b\left(\overline{z_b}^b\overline{v}^b\right)+\nabla_b\cdot\left(\overline{z_b}^b\overline{\mathbf{u}}^b\overline{v}^b\right)+\partial_b\left(\overline{z_b}^b\overline{\dot{b}}^b\overline{v}^b\right)+\overline{v}^b\nabla_b\cdot\left(\overline{z_b}^b\mathbf{u}_b\right)+\overline{v}^b\partial_b\left(\overline{z_b}^b\dot{b}_b\right)+f\overline{z_b}^b\overline{u}^b$$
$$=-\overline{z_b}^b\overline{\mathbf{u}'\cdot\nabla_b v'}^b-\overline{z_b}^b\overline{\dot{b}'\partial_b v'}^b-\frac{\overline{z_b}^b}{\rho_0}\overline{\partial_y|_z p}^b+\overline{z_b}^b\overline{\mathcal{F}_y}^b \tag{C8}$$

Despite these shortcomings, Equations C1–C5 offer arguably the most obvious interpretation of existing isopycnal coordinate models that solve Equations 74–78 with $r\rightarrow b$. In this interpretation, $\tilde{z}=\overline{z}^b$, $\tilde{\mathbf{u}}=\overline{\mathbf{u}}^b$, $\tilde{\dot{b}}=\overline{\dot{b}}^b$, $\tilde{c}=\hat{c}^b$, $\tilde{\dot{c}}=\hat{\dot{c}}^b$, $\tilde{p}=\overline{p}^b$, $\tilde{\rho}=\hat{\rho}^b$, $\tilde{\mathcal{F}}_{x/y}=\overline{\mathcal{F}_{x,y}}^b$, $\mathbf{u}_{GM}=\mathbf{u}_b$ and we need to assume that $\dot{b}_b\approx 0$, which is a reasonable assumption in the ocean interior, where mesoscale eddies are assumed to be largely adiabatic. As pointed out above and in Section 5.2, the eddy bolus/GM advection does neither conserve momentum in the averaged equations nor in the model formulation, which is at least consistent. The tracer diffusion needs to capture the isopycnal TWA eddy tracer flux (i.e., $\nabla\cdot(\mathbf{D}\nabla\tilde{c})=-\nabla\cdot\hat{\mathbf{J}^c}^b$), and the divergence of the viscous stress tensor needs to represent the advective eddy momentum tendencies and any eddy contributions to the pressure gradient term (i.e., $\nabla\cdot\tau^{\mathbf{u}}=-\overline{\mathbf{u}'\cdot\nabla_b u'}^b-\overline{\dot{b}'\partial_b u'}^b-\rho_0^{-1}\overline{\rho''g\partial_x|_b z}^b$ and $\nabla\cdot\tau^{\mathbf{v}}=-\overline{\mathbf{u}'\cdot\nabla_b v'}^b-\overline{\dot{b}'\partial_b v'}^b-\rho_0^{-1}\overline{\rho''g\partial_y|_b z}^b$)—the non-conservative parts of the eddy momentum tendencies hence need to be assumed negligible.

To express the momentum Equations C3 and C4 in an arbitrary vertical coordinate r, we use that

$$\frac{\overline{D}^b\overline{u}^b}{Dt}\equiv\partial_t|_b\overline{u}^b+\overline{\mathbf{u}}^b\cdot\nabla_b\overline{u}^b+\overline{\dot{b}}^b\partial_b\overline{u}^b \tag{C9a}$$

$$=\partial_t|_r\overline{u}^b-\partial_t|_r b(\partial_r b)^{-1}\partial_r\overline{u}^b+\overline{\mathbf{u}}^b\cdot\nabla_r\overline{u}^b-\left(\overline{\mathbf{u}}^b\cdot\nabla_r b\right)(\partial_r b)^{-1}\partial_r\overline{u}^b+\overline{\dot{b}}^b(\partial_r b)^{-1}\partial_r\overline{u}^b \tag{C9b}$$

$$=\partial_t|_r\overline{u}^b+\overline{\mathbf{u}}^b\cdot\nabla_r\overline{u}^b-(\partial_r b)^{-1}\left[\partial_t|_r b+\overline{\mathbf{u}}^b\cdot\nabla_r b-\overline{\dot{b}}^b\right]\partial_r\overline{u}^b \tag{C9c}$$

$$=\partial_t|_r\overline{u}^b+\overline{\mathbf{u}}^b\cdot\nabla_r\overline{u}^b+\dot{r}^{\dagger b}\partial_r\overline{u}^b \tag{C9d}$$

where

$$\dot{r}^{\dagger b}\equiv\frac{\overline{D}^b r}{Dt} \tag{C10a}$$

$$=\partial_t|_b r+\overline{\mathbf{u}}^b\cdot\nabla_b r+\overline{\dot{b}}^b\partial_b r \tag{C10b}$$

$$=-(\partial_r b)^{-1}\left[\partial_t|_r b+\overline{\mathbf{u}}^b\cdot\nabla_r b-\overline{\dot{b}}^b\right] \tag{C10c}$$

is the Lagrangian rate of change of r following the non-thickness-weighted b-averaged flow.

Using Equation C9 to express the non-thickness-weighted momentum (Equations C3 and C4) in generalized vertical coordinates, and combining with the thickness-weighted continuity equation, tracer equation, and hydrostatic balance in arbitrary vertical coordinates (Equations 25, 26 and 29), we can write the semi-thickness-weighted b-averaged equations in generalized vertical coordinates as

$$\partial_t|_r\left(\overline{z}_r^b\right) + \nabla_r \cdot \left(\left(\overline{\mathbf{u}}^b + \mathbf{u}_b\right)\overline{z}_r^b\right) + \partial_r\left(\overline{z}_r^b\, \dot{r}^{\#b}\right) = 0 \tag{C11}$$

$$\partial_t|_r\left(\overline{z}_r^b\,\hat{c}^b\right) + \nabla_r \cdot \left(\overline{z}_r^b\left(\overline{\mathbf{u}}^b + \mathbf{u}_b\right)\hat{c}^a\right) + \partial_r\left(\overline{z}_r^b\, \dot{r}^{\#b}\,\hat{c}^b\right) = -\overline{z}_r^b\overline{\nabla}^b \cdot \widehat{\mathbf{J}^{\mathbf{c}}}^b + \overline{z}_r^b\widehat{\dot{c}}^b \tag{C12}$$

$$\partial_t|_r\overline{u}^b + \overline{\mathbf{u}}^b \cdot \nabla_r\overline{u}^b + \dot{r}^{\dagger b}\partial_r\overline{u}^b - f\overline{v}^b = -\overline{\mathbf{u}' \cdot \nabla_b u'}^b - \overline{\dot{b}'\partial_b u'}^b - \rho_0^{-1}\overline{\partial_x|_z p}^b + \overline{\mathcal{F}_x}^b \tag{C13}$$

$$\partial_t|_r\overline{v}^b + \overline{\mathbf{u}}^b \cdot \nabla_r\overline{v}^b + \dot{r}^{\dagger b}\partial_r\overline{v}^b + f\overline{u}^b = -\overline{\mathbf{u}' \cdot \nabla_b v'}^b - \overline{\dot{b}'\partial_b v'}^b - \rho_0^{-1}\overline{\partial_y|_z p}^b + \overline{\mathcal{F}_y}^b \tag{C14}$$

$$\partial_r\overline{p}^b = -\hat{\rho}^b g\partial_r\overline{z}^b, \tag{C15}$$

where the pressure gradient term can be written in r-coordinates as

$$\overline{\partial_x|_z p}^b = \partial_x|_r\overline{p}^b + \hat{\rho}^b g\partial_x|_r\overline{z}^b + \overline{\rho'' g\partial_x|_b z}^b. \tag{C16}$$

We again keep the eddy terms in b-coordinates as their effects need to be parameterized.

The semi-thickness-weighted b-averaged Equations C11–C15 closely resemble the equations solved by MOM6 and HYCOM (Equations 74–78). However, the effective "vertical" velocities appearing in the continuity and tracer versus momentum equations again differ. To interpret the existing model equations in terms of the semi-thickness weighted isopycnal average, we would need to assume that $\widetilde{\dot{r}} = \dot{r}^{\#b} = \dot{r}^{\dagger b}$, but generally $\dot{r}^{\#b} \neq \dot{r}^{\dagger b}$. Notice that the formulation of HYCOM and MOM6 evolved from isopycnal coordinate models (i.e., $r = b$) in which case $\dot{b}^{\#b} = \dot{b}^{\dagger b}$ as long as the eddies are adiabatic — this is the case discussed above. For a general vertical coordinate, however, $\dot{r}^{\#b} \neq \dot{r}^{\dagger b}$, even for adiabatic flow. When using semi-Lagrangian time-stepping, setting $\dot{r}^{\#b} = 0$ in the continuity equation implies that the coordinate follows the thickness-weighted average flow. We then cannot also set $\dot{r}^{\dagger b} = 0$ in the momentum equation (which in turn would require the coordinate to follow the non-thickness-weighted flow), unless we assume the two flows to be the same (i.e., the bolus transport is not parameterized at all). This inconsistency, together with the non-conservative form of the momentum equations and the eddy momentum flux, leads us to conclude that an interpretation of the generalized vertical coordinate model equations in terms of the semi-thickness-weighted isopycnal average equations is neither desirable nor fully consistent with the existing model implementations.

Data Availability Statement

Scripts to create the plots in Figures 1 and 2 can be found at Jansen (2024).

Acknowledgments

MFJ acknowledges support from the National Science Foundation (NSF) through award OCE-1912163. IG acknowledges support from the National Science Foundation (NSF) through award OCE-1912332. We thank Robert Hallberg, Noora Loose and Bill Young for interesting and useful discussions (although we note that the views expressed here have not been endorsed by them or anyone else). We also thank Rainer Bleck and Mark Petersen for providing details on the formulation of the HYCOM and MPAS ocean models. Comments from three anonymous reviewers helped with clarity in the presentation.

References

Adcroft, A., Anderson, W., Blanton, C., Bushuk, M., Dufour, C. O., Dunne, J. P., et al. (2019). The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *Journal of Advances in Modeling Earth Systems*, *11*(10), 3167–3211. https://doi.org/10.1029/2019ms001726

Adcroft, A., Hallberg, R., & Harrison, M. (2008). A finite volume discretization of the pressure gradient force using analytic integration. *Ocean Modelling*, *22*(3–4), 106–113. https://doi.org/10.1016/j.ocemod.2008.02.001

Bachman, S. D., & Fox-Kemper, B. (2013). Eddy parameterization challenge suite i: Eady spindown. *Ocean Modelling*, *64*, 12–28. https://doi.org/10.1016/j.ocemod.2012.12.003

Bachman, S. D., Fox-Kemper, B., & Bryan, F. O. (2015). A tracer-based inversion method for diagnosing eddy-induced diffusivity and advection. *Ocean Modelling*, *86*, 1–14. https://doi.org/10.1016/j.ocemod.2014.11.006

Bachman, S. D., Fox-Kemper, B., & Bryan, F. O. (2020). A diagnosis of anisotropic eddy diffusion from a high-resolution global ocean model. *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001904. https://doi.org/10.1029/2019ms001904

Bleck, R. (2002). An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates. *Ocean Modelling*, *4*(1), 55–88. https://doi.org/10.1016/s1463-5003(01)00012-9

Brankart, J.-M. (2013). Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modelling. *Ocean Modelling*, *66*, 64–76. https://doi.org/10.1016/j.ocemod.2013.02.004

Buzzicotti, M., Storer, B. A., Khatri, H., Griffies, S., & Aluie, H. (2023). Spatio-temporal coarse-graining decomposition of the global ocean geostrophic kinetic energy. *Journal of Advances in Modeling Earth Systems*, *15*(6), e2023MS003693. https://doi.org/10.1029/2023MS003693

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Hogan, P. J., Wallcraft, A. J., et al. (2007). The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. *Journal of Marine Systems*, *65*(1–4), 60–83. https://doi.org/10.1016/j.jmarsys.2005.09.016

De Szoeke, R. A., & Bennett, A. F. (1993). Microstructure fluxes across density surfaces. *Journal of Physical Oceanography*, *23*(10), 2254–2264. https://doi.org/10.1175/1520-0485(1993)023<2254:mfads>2.0.co;2

Durran, D. R. (2010). *Numerical methods for fluid dynamics with applications to geophysics* (2nd ed.). Springer.

Eden, C., Greatbatch, R. J., & Olbers, D. (2007). Interpreting eddy fluxes. *Journal of Physical Oceanography*, *37*(5), 1282–1296. https://doi.org/10.1175/jpo3050.1

Eliassen, A., & Palm, E. (1961). On the transfer of energy in stationary mountain waves. *Geofysiske Publikasjoner*, *22*(3).

Ferreira, D., & Marshall, J. (2006). Formulation and implementation of a "residual-mean" ocean circulation model. *Ocean Modelling*, *13*(1), 86–107. https://doi.org/10.1016/j.ocemod.2005.12.001

Fox-Kemper, B., & Menemenlis, D. (2008). Can large eddy simulation techniques improve mesoscale rich ocean models? In M. W. Hecht & H. Hasumi (Eds.), *Ocean Modeling in an Eddying Regime* (Vol. 177, pp. 319–337). American Geophysical Union. https://doi.org/10.1029/177GM19

Gent, P. R., & McWilliams, J. C. (1990). Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography*, *20*(1), 150–155. https://doi.org/10.1175/1520-0485(1990)020<0150:imiocm>2.0.co;2

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*(4), 463–474. https://doi.org/10.1175/1520-0485(1995)025<0463:peitti>2.0.co;2

Gnanadesikan, A., Pradal, M.-A., & Abernathey, R. (2015). Isopycnal mixing by mesoscale eddies significantly impacts oceanic anthropogenic carbon uptake. *Geophysical Research Letters*, *42*(11), 4249–4255. https://doi.org/10.1002/2015gl064100

Greatbatch, R. J., & Lamb, K. G. (1990). On parameterizing vertical mixing of momentum in non-eddy resolving ocean models. *Journal of Physical Oceanography*, *20*(10), 1634–1637. https://doi.org/10.1175/1520-0485(1990)020<1634:opvmom>2.0.co;2

Griffies, S. M. (1998). The Gent-McWilliams skew flux. *Journal of Physical Oceanography*, *28*(5), 831–841. https://doi.org/10.1175/1520-0485(1998)028<0831:tgmsf>2.0.co;2

Griffies, S. M., Adcroft, A., & Hallberg, R. W. (2020). A primer on the vertical Lagrangian-remap method in ocean models based on finite volume generalized vertical coordinates. *Journal of Advances in Modeling Earth Systems*, *12*(10), e2019MS001954. https://doi.org/10.1029/2019ms001954

Guillaumin, A. P., & Zanna, L. (2021). Stochastic-deep learning parameterization of ocean momentum forcing. *Journal of Advances in Modeling Earth Systems*, *13*(9), e2021MS002534. https://doi.org/10.1029/2021ms002534

Hofmeister, R., Burchard, H., & Beckers, J.-M. (2010). Non-uniform adaptive vertical grids for 3d numerical ocean models. *Ocean Modelling*, *33*(1–2), 70–86. https://doi.org/10.1016/j.ocemod.2009.12.003

Jansen, M. F. (2024). Scripts for plots in Figures 1 and 2 in manuscript "The averaged hydrostatic boussinesq equations in generalized vertical coordinates". *Zenodo*. https://doi.org/10.5281/zenodo.11509777

Kamenkovich, I., Berloff, P., Haigh, M., Sun, L., & Lu, Y. (2021). Complexity of mesoscale eddy diffusivity in the ocean. *Geophysical Research Letters*, *48*(5), e2020GL091719. https://doi.org/10.1029/2020gl091719

Klingbeil, K., Becherer, J., Schulz, E., de Swart, H. E., Schuttelaars, H. M., Valle-Levinson, A., & Burchard, H. (2019). Thickness-weighted averaging in tidal estuaries and the vertical distribution of the Eulerian residual transport. *Journal of Physical Oceanography*, *49*(7), 1809–1826. https://doi.org/10.1175/jpo-d-18-0083.1

Lin, S. J. (1997). A finite volume integration method for computing pressure gradient force in general vertical coordinates. *Quarterly Journal of the Royal Meteorological Society*, *123*(542), 1749–1762. https://doi.org/10.1002/qj.49712354214

Loose, N., Marques, G. M., Adcroft, A., Bachman, S., Griffies, S. M., Grooms, I., et al. (2023). Comparing two parameterizations for the restratification effect of mesoscale eddies in an isopycnal ocean model. *Journal of Advances in Modeling Earth Systems*, *15*(12), e2022MS003518. https://doi.org/10.1029/2022MS003518

Maddison, J. R., & Marshall, D. P. (2013). The Eliassen–Palm flux tensor. *Journal of Fluid Mechanics*, *729*, 69–102. https://doi.org/10.1017/jfm.2013.259

Marshall, J., Adcroft, A., Campin, J.-M., Hill, C., & White, A. (2004). Atmosphere–ocean modeling exploiting fluid isomorphisms. *Monthly Weather Review*, *132*(12), 2882–2894. https://doi.org/10.1175/mwr2835.1

McDougall, T. J., Groeskamp, S., & Griffies, S. M. (2014). On geometric aspects of interior ocean mixing. *Journal of Physical Oceanography*, *44*(8), 2164–2175. https://doi.org/10.1175/JPO-D-13-0270.1

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*(12), 2653–2665. https://doi.org/10.1175/1520-0485(1996)026⟨2653:TTRMVP⟩2.0.CO;2

McDougall, T. J., & McIntosh, P. C. (2001). The temporal-residual-mean velocity. Part II: Isopycnal interpretation and the tracer and momentum equations. *Journal of Physical Oceanography*, *31*(5), 1222–1246. https://doi.org/10.1175/1520-0485(2001)031<1222:ttrmvp>2.0.co;2

Perezhogin, P., Zanna, L., & Fernandez-Granda, C. (2023). Generative data-driven approaches for stochastic subgrid parameterizations in an idealized ocean model. *Journal of Advances in Modeling Earth Systems*, *15*(10), e2023MS003681. https://doi.org/10.1029/2023ms003681

Petersen, M. R., Jacobsen, D. W., Ringler, T. D., Hecht, M. W., & Maltrud, M. E. (2015). Evaluation of the arbitrary Lagrangian–Eulerian vertical coordinate method in the MPAS-ocean model. *Ocean Modelling*, *86*, 93–113. https://doi.org/10.1016/j.ocemod.2014.12.004

Pope, S. G. (2000). *Turbulent flows*. Cambridge.

Redi, M. (1982). Oceanic isopycnal mixing by coordinate rotation. *Journal of Physical Oceanography*, *12*(10), 1154–1158. https://doi.org/10.1175/1520-0485(1982)012<1154:oimbcr>2.0.co;2

Rhines, P., & Young, W. (1982). Homogenization of potential vorticity in planetary gyres. *Journal of Fluid Mechanics*, *122*(-1), 347–367. https://doi.org/10.1017/s0022112082002250

Rhines, P. B. (1982). Basic dynamics of the large-scale geostrophic circulation. In *Summer Study Program in Geophysical Fluid Dynamics* (Vol. 1, p. 47). Woods Hole Oceanographic Institution.

Ringler, T., Petersen, M., Higdon, R. L., Jacobsen, D., Jones, P. W., & Maltrud, M. (2013). A multi-resolution approach to global ocean modeling. *Ocean Modelling*, *69*, 211–232. https://doi.org/10.1016/j.ocemod.2013.04.010

Saenz, J. A., Chen, Q., & Ringler, T. (2015). Prognostic residual mean flow in an ocean general circulation model and its relation to prognostic Eulerian mean flow. *Journal of Physical Oceanography*, *45*(9), 2247–2260. https://doi.org/10.1175/jpo-d-15-0024.1

Solomon, H. (1971). On the representation of isentropic mixing in ocean models. *Journal of Physical Oceanography*, *1*(3), 233–234. https://doi.org/10.1175/1520-0485(1971)001⟨0233:OTROIM⟩2.0.CO;2

Stanley, Z., Grooms, I., Kleiber, W., Bachman, S. D., Castruccio, F., & Adcroft, A. (2020). Parameterizing the impact of unresolved temperature variability on the large-scale density field: Part 1. Theory. *Journal of Advances in Modeling Earth Systems*, *12*(12), e2020MS002185. https://doi.org/10.1029/2020MS002185

Sun, L., Haigh, M., Shevchenko, I., Berloff, P., & Kamenkovich, I. (2021). On non-uniqueness of the mesoscale eddy diffusivity. *Journal of Fluid Mechanics*, *920*, A32. https://doi.org/10.1017/jfm.2021.472

Uchida, T., Jamet, Q., Dewar, W. K., Le Sommer, J., Penduff, T., & Balwada, D. (2022). Diagnosing the thickness-weighted averaged eddy-mean flow interaction from an eddying North Atlantic ensemble: The Eliassen-Palm flux. *Journal of Advances in Modeling Earth Systems*, *14*(5), e2021MS002866. https://doi.org/10.1029/2021ms002866

Vallis, G. K. (2006). *Atmospheric and oceanic fluid dynamics*. Cambridge University Press.

Young, W. R. (2012). An exact thickness-weighted average formulation of the boussinesq equations. *Journal of Physical Oceanography*, *42*(5), 692–707. https://doi.org/10.1175/jpo-d-11-0102.1

Zanna, L., & Bolton, T. (2020). Data-driven equation discovery of ocean mesoscale closures. *Geophysical Research Letters*, *47*(17), e2020GL088376. https://doi.org/10.1029/2020gl088376

Zhang, C., Perezhogin, P., Gultekin, C., Adcroft, A., Fernandez-Granda, C., & Zanna, L. (2023). Implementation and evaluation of a machine learned mesoscale eddy parameterization into a numerical ocean circulation model. *Journal of Advances in Modeling Earth Systems*, *15*(10), e2023MS003697. https://doi.org/10.1029/2023ms003697

Zhao, R., & Vallis, G. K. (2008). Parameterizing mesoscale eddies with residual and Eulerian schemes, and a comparison with eddy-permitting models. *Ocean Modelling*, *23*(1–2), 1–12. https://doi.org/10.1016/j.ocemod.2008.02.005