

Inclusion of Thermobaricity in Isopycnic-Coordinate Ocean Models

SHAN SUN

NASA/Goddard Institute for Space Studies, New York, New York

RAINER BLECK AND CLAES Rooth

University of Miami, Miami, Florida

JOHN DUKOWICZ

Los Alamos National Laboratory, Los Alamos, New Mexico

ERIC CHASSIGNET

University of Miami, Miami, Florida

PETER KILLWORTH

Southampton Oceanography Centre, Southampton, United Kingdom

15 January 1998 and 2 March 1999

ABSTRACT

Buoyancy anomalies caused by thermobaricity, that is, the modulation of seawater compressibility by potential temperature anomalies, underlie a long-standing argument against the use of potential-density-framed numerical models for realistic circulation studies. The authors show that this problem can be overcome by relaxing the strict correspondence between buoyancy and potential density in isopycnic-coordinate models. A parametric representation of the difference between the two variables is introduced in the form of a “virtual potential density,” which can be viewed as the potential density that would be computed from the in situ conditions using the compressibility coefficient for seawater of a fixed (but representative) salinity and potential temperature. This variable is used as a basis for effective dynamic height computations in the dynamic equations, while the traditionally defined potential density may be retained as model coordinate. The conservation properties of the latter assure that adiabatic transport processes in a compressibility-compliant model can still be represented as exactly two-dimensional. Consistent with its dynamic significance, the distribution of virtual potential density is found to determine both the local static stability and, to a lesser degree, the orientation of neutrally buoyant mixing surfaces. The paper closes with a brief discussion of the pros and cons of replacing potential density by virtual potential density as vertical model coordinate.

1. Introduction

The shallow-water equations have proven to be a versatile tool for elucidating barotropic aspects of ocean circulation dynamics. In “stacked” or multilayer form they have also been used to investigate baroclinic dynamics, albeit under circumstances where neglect of compressibility effects and diapycnal mixing processes can be defended based on scaling arguments (e.g., Pedlosky 1982; Gill 1982). Numerical solutions of these equations have been studied in the oceanic context since

the 1960s (e.g., Welander 1966; Holland 1967). Over the years, these efforts have led to the development of full-fledged oceanic general circulation models, which are referred to as layer models when conceived as exact representations of a set of incompressible and immiscible fluid layers, but otherwise are referred to as “isopycnic coordinate” or simply “isopycnic” models (e.g., Bleck et al. 1992; Oberhuber 1993).

Isopycnic-coordinate models are traditionally being formulated in terms of potential (as opposed to in situ) density. There are two reasons for this: Potential density surfaces are material under adiabatic conditions, rendering transport processes two-dimensional in the oceanic interior, and they approximately correspond to neutrally buoyant surfaces along which most of the interior

Corresponding author address: Dr. Shan Sun, NASA/Goddard Institution for Space Studies, 2880 Broadway, New York, NY 10025.
E-mail: sun@giss.nasa.gov

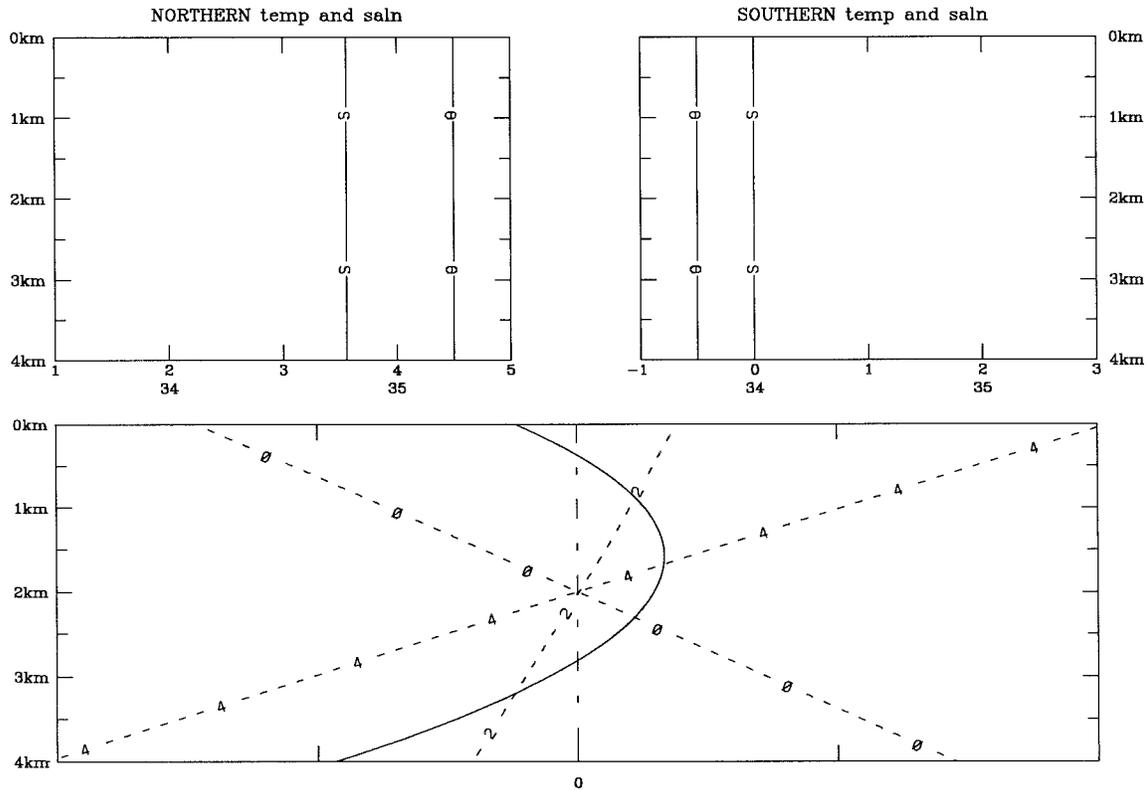


FIG. 1. Lower panel: interhemispheric pressure difference (in relative units) based on the θ/S profiles shown in the upper panels. Solid: in situ density; short-dashed: potential densities ρ_0 , ρ_2 , ρ_4 .

oceanic mixing takes place. This “tradition” manifests itself in two ways, which we wish to distinguish here sharply: (i) coordinate layers are defined as layers of constant potential density and (ii) the dynamic equations treat seawater as incompressible; that is, they do not differentiate between the in situ density of the water in each coordinate layer and its potential density.

Modeling conventions (i) and (ii) give rise to two problems which, though seemingly disparate, are both related to the fact that seawater is compressible—specifically, that its compressibility has a *thermobaric* component, that is, depends on temperature and salinity.

- In haloclinic regions, the slope of a potential density surface is a function of the reference pressure used in defining potential density. The sensitivity of this slope to salinity variations is such that a reference pressure yielding monotonic variation of potential density with depth everywhere in the World Ocean does not exist. In other words, no matter which reference pressure one chooses, a geographic location can always be found where potential isopycnals will fold. This obviously interferes with the use of potential density as a vertical coordinate in global ocean models.
- Variations of in situ density along potential isopycnals caused by the thermobaric effect give rise to pressure forces and circulation systems, which, though weak,

cannot safely be ignored in a general circulation model. A demonstration of the existence and likely relevance of such pressure forces is given below.

The subsurface circulation in the Atlantic is characterized, in descending order, by northward-flowing upper Antarctic Intermediate Water, southward-flowing North Atlantic Deep Water, and northward-flowing Antarctic Bottom Water [Fig. 6.15 in Open University (1989)]. This double reversal of flow direction corresponds to—and most likely is dynamically controlled by—a double reversal of the interhemispheric pressure force with depth, created by the higher compressibility of the colder southern water.

To illustrate the existence of this pressure force, we consider two idealized water columns, one salty and relatively warm and the other fresh and cold, representing wintertime subpolar North and South Atlantic water columns, respectively. We prescribe θ and S in the two water columns as $(4.5^\circ\text{C}, 34.8 \text{ psu})$ and $(-0.5^\circ\text{C}, 34.0 \text{ psu})$, respectively; this is illustrated in the top two panels of Fig. 1. The solid curve in the lower panel of Fig. 1 represents the interhemispheric pressure difference (in arbitrary units) based on the in situ density ρ , with the barotropic component removed; it indeed shows the expected double reversal with depth. The three dashed curves in the same panel show the inter-

hemispheric pressure difference obtained by substituting potential densities ρ_0 , ρ_2 , ρ_4 (potential densities referenced to pressure levels 0, 20, and 40 MPa, respectively) for ρ at all depths. Owing to the fact that $\nabla\rho_r$ by definition is independent of depth, these curves are linear and hence show only a single reversal of the pressure difference with depth. This suggests that a traditional isopycnic model, irrespective of which reference level its vertical coordinate is based on, would be unable to simulate the interleaving of Atlantic water masses described above.

The obvious solution to the first problem mentioned—folding of coordinate surfaces—is to generalize the procedure for placing layer interfaces in the fluid, that is, to deviate from the notion that interfaces *must* everywhere coincide with potential isopycnals. Experiments with a suitably generalized vertical coordinate, largely based on atmospheric modeling know-how (Bleck and Benjamin 1993), are under way and will be reported on at some future date. The purpose of the present paper is to suggest ways in which the second problem mentioned, the inability of isopycnic models to simulate thermobarically driven circulation systems, can be overcome.

2. Density representation in isopycnic models

In the following, variables without sub- or superscripts will be used to denote in situ or local conditions (pressure p , specific volume α , . . .), while subscript r will denote *potential* values achieved after adiabatic (de)compression to reference pressure p_r .

If the incompressibility assumption (ii) is what prevents thermobaric effects from being included in isopycnic models, then the remedy is obvious: one must liberate oneself from the tradition of letting the definition of the vertical coordinate dictate the water density within each layer. In other words, water density in the model must be allowed to vary not only from layer to layer, but also within coordinate layers.

The question then becomes: which density-related, x , y , t -dependent variable should be chosen to represent the mass field in isopycnic model layers?

One obvious and seemingly safe choice is to set the density within coordinate layers to the in situ water density. Unfortunately, this has severe consequences for the accuracy with which the horizontal pressure gradient force can be computed in steeply inclined coordinate layers. The argument is as follows.

Transformation from Cartesian (x, y, z) to isopycnic (x, y, α_r) coordinates¹ under hydrostatic conditions changes the pressure gradient term in the horizontal momentum equation (loosely referred to as pressure gradient “force,” or PGF) into

$$-\alpha\nabla_z p = -\nabla_{\alpha_r} M + p\nabla_{\alpha_r} \alpha, \quad (1)$$

where $M = p\alpha + gz$ is the Montgomery potential (Montgomery 1937) and the subscripted ∇ operator denotes differentiation in x , y direction with the subscript held constant. A form of the hydrostatic equation convenient for evaluating M is

$$\frac{\partial M}{\partial \alpha} = p.$$

In steeply inclined α_r layers, the two terms $\nabla_{\alpha_r} M$ and $p\nabla_{\alpha_r} \alpha$ in (1) can individually become orders of magnitude larger than their residual $\alpha\nabla_z p$, thereby affecting numerical precision in evaluating the latter. This problem is related to the one discovered years ago in models using terrain-following (so-called σ) coordinates (e.g., Janjic 1977), even though it seems to be more serious in the present context. Practical tests conducted with the Miami Isopycnic Coordinate Ocean Model indeed have shown that the approach based on (1) not only leads to accuracy problems, but to insurmountable numerical stability problems (Sun 1997).

Defining the Montgomery potential in terms of α_r instead of α , as in $M = p\alpha_r + gz$, does not alleviate this problem. In this case, (1) is replaced by

$$-\alpha\nabla_z p = -\nabla_{\alpha_r} M + (\alpha_r - \alpha)\nabla_{\alpha_r} p, \quad (2)$$

which again is a two-term expression where both terms have been found to become unacceptably large in case of steeply inclined α_r surfaces. In fact, the term involving $\nabla_{\alpha_r} p$ serves as a particularly good indicator of the degeneracy in the PGF calculation created by steeply inclined α_r surfaces.

The excessive magnitude of the right-hand side terms in (2) can be attributed to the combination of an α_r -based Montgomery potential with an α -based hydrostatic equation. In traditional isopycnic models, a two-term representation of the PGF is avoided by evaluating the PGF in the form $-\nabla_{\alpha_r} M$ where M is defined as in (2) but satisfies the modified hydrostatic equation $\partial M/\partial \alpha_r = p$. This PGF expression is exact only if the in situ water density of the ocean being modeled is identical to its potential density [referred to earlier as modeling convention (ii)]. It would seem, therefore, that a choice must be made in isopycnic modeling between the “ σ coordinate error” and the thermobaricity error. Fortunately, this is not so.

We will demonstrate that a modified potential density, which we choose to name *virtual potential density*, matches the in situ density in its ability to capture the dynamic aspects of the oceanic mass field in isopycnic models, while avoiding the pitfalls in the pressure gradient computation just mentioned. After establishing this result, which has possible implications for nonisopycnic models as well, we will explore the applicability of virtual potential density as vertical coordinate in layer models. Of particular interest in this context are folding tendencies and the buoyancy neutrality, or lack thereof,

¹ We use the adjective “isopycnic” throughout this paper to denote constant *potential* density.

of potential isopycnals versus surfaces of virtual potential density.

3. Basic definitions

The equation of state for seawater expresses its density (ρ) (or its reciprocal, α) in terms of the salinity (S), temperature (T), and pressure (p). Adiabatic (de)compression to a reference pressure p_r causes α and T to change to values denoted respectively as potential specific volume (α_r) and potential temperature ($T_r \equiv \theta$); these will henceforth be considered as primary thermodynamic variables in place of α and T . We can, without loss of generality, use different reference pressures for these variables, and choose for conformity with oceanographic practice 1 bar = 0.1 MPa (fn 2) as the reference for θ , while leaving the choice of reference pressure for α_r to be determined as appropriate for specific applications. This convention is advantageous in model applications, since θ under this definition is identical in the surface layer to the in situ temperature, which is a controlling parameter in air–sea interaction and radiative transfer processes.

A general functional relation between in situ specific volume and potential specific volume $\alpha_r(\theta, S) = \alpha(\theta, S, p_r)$ is

$$\alpha(\theta, S, p) = \alpha_r(\theta, S) \exp \int_p^{p_r} \kappa(\theta, S, p') dp', \quad (3)$$

which results from integrating the expression for the adiabatic compressibility

$$\kappa = \frac{1}{\rho} \left(\frac{\partial \rho}{\partial p} \right)_{\theta, S} = - \frac{1}{\alpha} \left(\frac{\partial \alpha}{\partial p} \right)_{\theta, S} \quad (4)$$

over the pressure interval (p, p_r).

We evaluate κ using the equation of state developed by Jackett and McDougall (1995, hereafter JMCD), an approximation of the UNESCO formula written in terms of θ instead of T . To give a general impression of the κ variability between the surface and the sea floor, we show in Fig. 2 plots of κ_0 and κ_5 (compressibility at 0 and 50 MPa) over the range $-2^\circ \leq \theta \leq 32^\circ\text{C}$ and $30 \leq S \leq 38$ psu. The figure indicates that κ varies by $\sim 15\%$ over both the 5-km depth range and the 32-K temperature range. The variation with salinity is somewhat lower, $\sim 3\%$ over the 0.8% salinity range. The variability in p is seen to be largely uncorrelated with that in θ and S .

The Bjerknes circulation theorem emphasizes the distinction between *barotropic* fluids, whose density depends only on pressure, and *baroclinic* fluids, where this is not the case. According to the theorem, buoyancy forces can drive circulation systems only in baroclinic

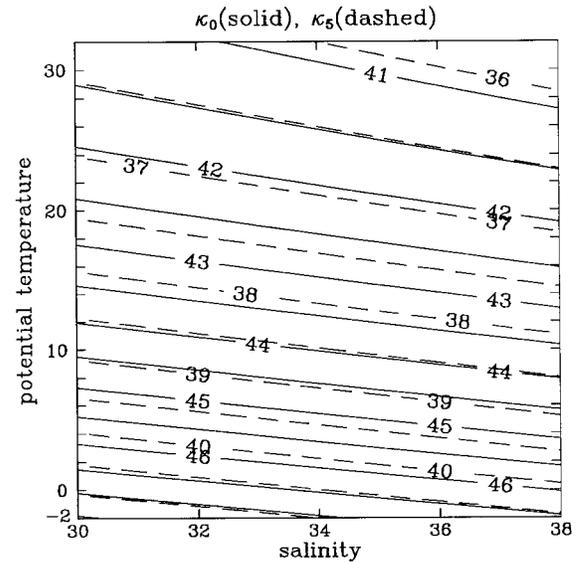


FIG. 2. Adiabatic compressibility (10^{-5} MPa^{-1}) at 0 MPa (κ_0 , solid) and 50 MPa (κ_5 , dashed) as function of θ and S .

fluids; that is, fluids whose density is affected by variables other than pressure—temperature and salinity in the oceanic case. In order to isolate the components of density variability that are capable of creating circulation-inducing buoyancy contrasts from those that are dynamically passive, we proceed now to split κ into a pressure-dependent component $\kappa^{(p)}$ and a residual thermobaric component, which we will abbreviate here by $\kappa^{(\theta)}$, even though it may exhibit some salinity and pressure dependence as well. Judging from Fig. 2, $\kappa^{(p)}$ is at least one order of magnitude larger than $\kappa^{(\theta)}$. A density field from which the effect of the $\kappa^{(p)}$ -related compressibility has been removed can therefore be expected to be much closer to the (genuine) potential than to the in situ density field. Thus, if a way can be found to remove the effect of $\kappa^{(p)}$ from the oceanic density field, we will be able to substantially reduce the size of the term $p \nabla_{\alpha_r} \alpha$ in (1) (and of the term $\nabla_{\alpha_r} M$ at the same time).

Substituting $\kappa^{(p)} + \kappa^{(\theta)}$ for κ in (3) leads to

$$\alpha(\theta, S, p) = \alpha_r(\theta, S) A_r(p) \exp \int_p^{p_r} \kappa^{(\theta)}(\theta, S, p') dp', \quad (5)$$

where

$$A_r(p) = \exp \int_p^{p_r} \kappa^{(p)}(p') dp' \quad (6)$$

describes the (de)compression effect attributable to the p -dependent component of compressibility alone: $A_r(p)$ is the term whose effect on the dynamic equations we are trying to remove.

The next step is to combine the other two factors on the right-hand side of (5) into

² 1 MPa corresponds approximately to a 100-m depth interval.

TABLE 1. Coefficients $c_{1,\dots,11}$ for κ in (8) and (9).

c_1	c_2	c_3	c_4	c_5	c_6
1.84412×10^{-4}	-0.128026	46.3424	-3.07406×10^{-5}	0.00460241	-0.289649
c_7	c_8	c_9	c_{10}	c_{11}	
-0.109102	7.95530×10^{-4}	0.00110342	1.42401×10^{-5}	-1.31958×10^{-5}	

$$\alpha^*(\theta, S, p) \equiv \alpha_r(\theta, S) \exp \int_p^{p_r} \kappa^{(\theta)}(\theta, S, p') dp' \\ = \frac{\alpha(\theta, S, p)}{A_r(p)}. \quad (7)$$

The variable α^* so defined would be identical to the conventional potential specific volume $\alpha_r(\theta, S)$ if compressibility were a function of p alone, that is, if there were no thermobaricity. We can also say that α^* is the potential specific volume a water parcel assumes during adiabatic (de)compression to pressure p_r , if its θ, S, p -dependent compressibility coefficient is replaced by one for standard salinity and potential temperature. Borrowing from the practice of expressing atmospheric buoyancy variations caused by water vapor in terms of a “virtual” temperature, we will refer to α^* as *virtual potential specific volume*, the reciprocal of *virtual potential density*.

Even though α^* depends on both p and p_r , we choose for simplicity not to refer to it as α_r^* , given that virtual potential densities referenced to different pressures p_r , p_s only differ by a constant factor $A_r(p_s)$. In other words, α^* surfaces associated with different reference pressures are congruent.

It is worth pointing out that the problems addressed in this paper do not appear in the atmospheric context. The compressibility of an ideal gas is $\kappa = c_v/(c_p p)$. Thus, to the extent that the atmospheric constituents can be treated as ideal gases and the dependence of c_p and c_v on the water vapor content can be neglected, the atmosphere is an example of a fluid where $\kappa = \kappa^{(p)}$.

4. Practical implications

Analysis in terms of thermobaric effects provides a unified perspective on several of the generally perceived limitations of potential density-based reference frames that arise when water mass variability in θ and S is involved. The key problem areas addressed here are

- 1) impact of thermobaricity on the dynamic balance, specifically the horizontal and vertical components of the pressure gradient force;
- 2) the representation of static stability, and the problem of isopycnic surface folding;
- 3) the deviations from buoyancy neutrality of isopycnic lateral displacements.

a. Polynomial approximation of κ

For practical model use, we generated a least squares polynomial approximation to (4), based on the state equation of JMcD, in the ranges of $-2^\circ \leq \theta \leq 32^\circ\text{C}$, $30 \leq S \leq 38$ psu, and $0 \leq p \leq 55$ MPa. After considerable experimentation, we settled on a polynomial of third degree in θ , quadratic in p , linear in S , and including the composite terms θS , θp , $\theta^2 p$, and $\theta S p$.

Setting $\kappa = \kappa^{(p)} + \kappa^{(\theta)}$, the 11-term polynomial so defined splits into

$$\kappa^{(p)} \times 10^5 = c_1 p^2 + c_2 p + c_3 \quad (8)$$

and

$$\kappa^{(\theta)} \times 10^5 = c_4 \theta^3 + c_5 \theta^2 + c_6 \theta + c_7 S' + c_8 \theta S' \\ + c_9 \theta p + c_{10} \theta S' p + c_{11} \theta^2 p, \quad (9)$$

where $S' = S - 35$ psu, κ is in MPa^{-1} , and p is in MPa. The coefficients c_1, \dots, c_{11} are shown in Table 1. The accuracy of the polynomial fit of κ [Eqs. (8) and (9) combined] is illustrated in Fig. 3. The thermobaric component $\kappa^{(\theta)}$ alone [Eq. (9)] is shown in Fig. 4.

b. Pressure gradient force representation in the dynamic equations

The PGF in a hydrostatic fluid can be expressed as the isobaric gradient of the geopotential $\phi \equiv gz$, that is, $\alpha \nabla_z p = \nabla_p \phi$. Integrating the hydrostatic equation $\partial \phi / \partial p = -\alpha$ from the sea surface (zero pressure) to pressure p then allows us to write the PGF as

$$-\alpha \nabla_z p = -\nabla_p \phi = -\nabla \phi_{srf} + \int_0^p \nabla_p \alpha dp', \quad (10)$$

where ϕ_{srf}/g is the sea surface height. By virtue of

$$\nabla_p \ln \alpha = \nabla_p \ln \alpha^*, \quad (11)$$

which follows from (7), the integrand in the last term of (10) can be expressed as

$$\nabla_p \alpha = \frac{\alpha}{\alpha^*} \nabla_p \alpha^* = A_r(p) \nabla_p \alpha^*,$$

allowing (10) to be written in the form

$$-\alpha \nabla_z p = -\nabla_p \phi \\ = -\nabla \phi_{srf} + \int_0^p A_r(p') \nabla_p \alpha^* dp'. \quad (12)$$

Now if a modified geopotential ϕ^* is defined by

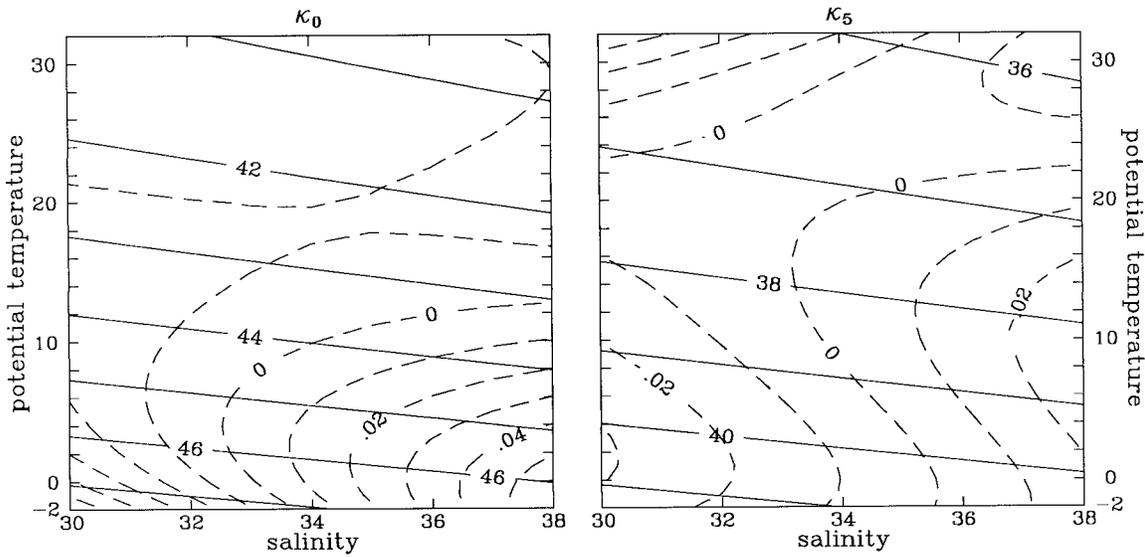


FIG. 3. Adiabatic compressibility (10^{-5} MPa^{-1}) at 0 and 50 MPa (κ_0 and κ_5) as function of θ and S . Solid: from JMcD (1995); dashed: [(8) + (9)] minus JMcD (1995).

$$\frac{\partial \phi^*}{\partial p} \equiv -\alpha^*, \quad (13)$$

then $\phi^* = \phi_{srf}^* - \int_0^p \alpha^* dp'$ and

$$-\nabla_p \phi^* = -\nabla_p \phi_{srf}^* + \int_0^p \nabla_p \alpha^* dp'. \quad (14)$$

Setting $\phi_{srf}^* = \phi_{srf}$ allows us to combine the last expression with (12) into

$$-\nabla_p \phi^* = -\nabla_p \phi - \int_0^p [A_r(p') - 1] \nabla_p \alpha^* dp'. \quad (15)$$

Viewing (10) as a variant of the thermal wind equation, we can argue that $\nabla_p \phi$ and $\int \nabla_p \alpha dp'$ are of comparable magnitude, and that this also holds for $\int \nabla_p \alpha^* dp'$ by virtue of (12). According to (6) and Fig. 4, the factor $[A_r(p) - 1]$ in an ocean 5 km deep is at most of order 10^{-2} . Hence we may conclude that

$$-\nabla_p \phi \approx -\nabla_p \phi^*. \quad (16)$$

This is an important result. It implies that estimates of the horizontal PGF are insensitive to whether the hydrostatic equation has been integrated using the in situ or the virtual potential density, that is, whether or not the oceanic density field is subject to compression by the pressure-related part $\kappa^{(p)}$ of the compressibility coefficient.

It is instructive to verify that replacing α^* by α_r in situations where thermobaric effects are important yields a PGF approximation inferior to (16). Retracing the steps leading from (10) to (16), but with α_r taking the place of α^* , one immediately realizes that the relationship between $\nabla_p \ln \alpha$ and $\nabla_p \ln \alpha_r$, analogous to (11) is complicated by the fact that the isobaric gradient of $\kappa^{(\theta)}$ does not vanish. The equation taking the place of (11) can be inferred from (5) to be

$$\nabla_p \ln \alpha = \nabla_p \ln \alpha_r + \int_p^{p_r} \nabla_p \kappa^{(\theta)}(\theta, S, p') dp'. \quad (17)$$

At pressure levels far from p_r , the last term on the right-hand side of (17) is not necessarily small. In fact, it can overshadow the term on the left, thereby rendering $|\nabla_p \ln \alpha_r|$ much larger than $|\nabla_p \ln \alpha|$.

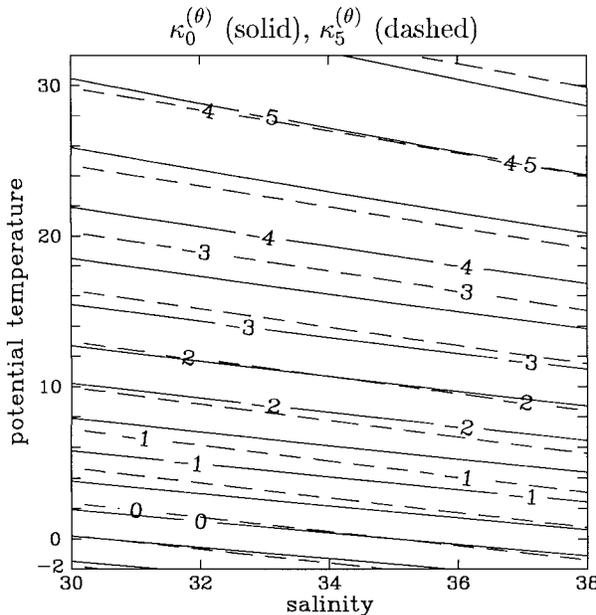


FIG. 4. Thermobaric compressibility (10^{-5} MPa^{-1}) at 0 MPa ($\kappa_0^{(\theta)}$, solid) and 50 MPa ($\kappa_5^{(\theta)}$, dashed) as function of θ and S .

Starting from (17) and substituting α_r for α^* in (13) and (14) leads to the following analog of (15):

$$-\nabla_p \phi_r = -\nabla_p \phi - \int_0^p \left(\frac{\alpha}{\alpha_r} - 1 \right) \nabla_p \alpha_r dp' - \int_0^p \alpha \int_{p'}^{p_r} \nabla_p \kappa^{(\theta)}(\theta, S, p'') dp'' dp', \quad (18)$$

where ϕ_r is the geopotential associated with α_r .

The single-integral term in (18) is the analog of the last term in (15); both are similar in magnitude and thus small compared to $\nabla_p \phi$. The double-integral term, however, which is spawned by the last term in (17) and thus represents the error in the α_r -based PGF calculation stemming from the neglect of thermobaricity, cannot be counted on as being small. This obviously will prevent us from declaring $-\nabla_p \phi_r$ a good approximation to $-\nabla_p \phi$. The effect of the last term in (18) is particularly strong if $\nabla_p \kappa^{(\theta)}$ —or $\nabla_p \theta$, for that matter, given that $\kappa^{(\theta)}$ is predominately a function of θ —remains constant over a large depth range.

Switching from isobaric to isopycnic differentiation in (16) yields an approximation to the PGF in the form

$$[-\alpha \nabla_z p \approx] -\nabla_p \phi^* = -\nabla_{\alpha_r} M^* + p \nabla_{\alpha_r} \alpha^*, \quad (19)$$

where

$$M^* = p \alpha^* + \phi^*$$

is the Montgomery potential based on α^* . Note the formal similarity between (19) and (1).

A hydrostatic equation stated in terms of α^* can be derived by differentiating the modified Montgomery potential in the vertical, $\partial M^* = p \partial \alpha^* + \alpha^* \partial p + \partial \phi^*$. By virtue of (13), this expression reduces to

$$\frac{\partial M^*}{\partial \alpha^*} = p.$$

In conclusion, the addition of the last term in (19) and the use of α^* in the Montgomery potential and the hydrostatic equation comprise the changes needed to account for thermobaric effects in isopycnic coordinate models. The three components of the PGF (we interpret here the hydrostatic equation as a statement about the vertical PGF) retain their original isopycnic appearance, with the gradient of virtual potential specific volume along isopycnic surfaces as a correction term.

In order to gain an understanding of the magnitude of the last term in (19) in relation to the PGF itself, we write this term as

$$p \alpha^* \nabla_{\alpha_r} \ln \alpha^* = p \alpha^* \nabla_{\alpha_r} \int_p^{p_r} \kappa^{(\theta)} dp' \quad (20)$$

and consider the case of a potential isopycnic descending from 0 to 5 km (50 MPa) over a horizontal distance L . Assuming a representative value of 10^{-5} MPa^{-1} for $|\kappa^{(\theta)}|$ and a middepth value of $2 \times 10^4 \text{ m}^2 \text{ s}^{-2}$ for $p \alpha^*$,

we obtain for (20) a magnitude of roughly $(10 \text{ m}^2 \text{ s}^{-2})/L$. Assuming that the sloping isopycnal is associated with a 10 cm s^{-1} current, the left-hand side of (19) has a magnitude of 10^{-5} m s^{-2} at midlatitudes. Thus, (20) and the PGF are of comparable magnitude if $L = 1000 \text{ km}$, that is, if the isopycnal slope is 1:200.

Isopycnals occasionally are more steeply sloped than this, implying that (20) is not always small compared to the PGF. Practical tests conducted on isopycnal fields derived from Levitus (1982) data suggest, however, that the above estimate is overly pessimistic. At $\sim 117\,000$ grid points where the geostrophic velocity exceeded 1 cm s^{-1} , the term (20) was found to exceed 10% of the PGF at roughly 1 in 1000 points, while no point was found where (20) reached the same magnitude as the PGF.

Generally speaking, it is clear that the term (20) is much smaller than the corresponding term

$$p \nabla_{\alpha_r} \alpha = p \alpha \nabla_{\alpha_r} \int_p^{p_r} \kappa dp$$

in (1), given that κ is 40 times larger than the value for $|\kappa^{(\theta)}|$ assumed above.

The example discussed in the introduction can be used to illustrate that substituting virtual potential specific volume for in situ specific volume in the three-dimensional PGF calculation indeed produces satisfactory results. The pressure difference calculated from the virtual potential density field referenced to 0, 20, and 40 MPa, respectively, is plotted in Fig. 5 using dashes of different length. The solid curve from Fig. 1 has been added as reference. Even though the abscissa scale has been expanded to accentuate possible differences, the three dashed curves are found to be virtually indistinguishable from the reference curve.

c. Static stability and folding of potential isopycnals

For the purpose of the following discussion we find it useful to define a “mean ambient pressure” p_l whose relationship to the local and to the reference pressure is given by $p = p_r + \delta p + \bar{p}$ where $\delta p = p_l - p_r$ is the “ambient pressure bias” and \bar{p} is a small pressure deviation.

The most easily understood consequence of thermobaricity is its impact on the relation between isopycnal layer thickness and static stability.

The Brunt–Väisälä frequency N is the conventional measure of static stability. It is defined as

$$N^2 = g \frac{\partial}{\partial z} \ln \alpha_l, \quad (21)$$

where α_l is the potential specific volume with respect to the mean ambient pressure $p_l = p_r + \delta p$ (held constant during vertical differentiation), and z is directed upward.

By setting $p = p_l$ in (7) and differentiating the outside

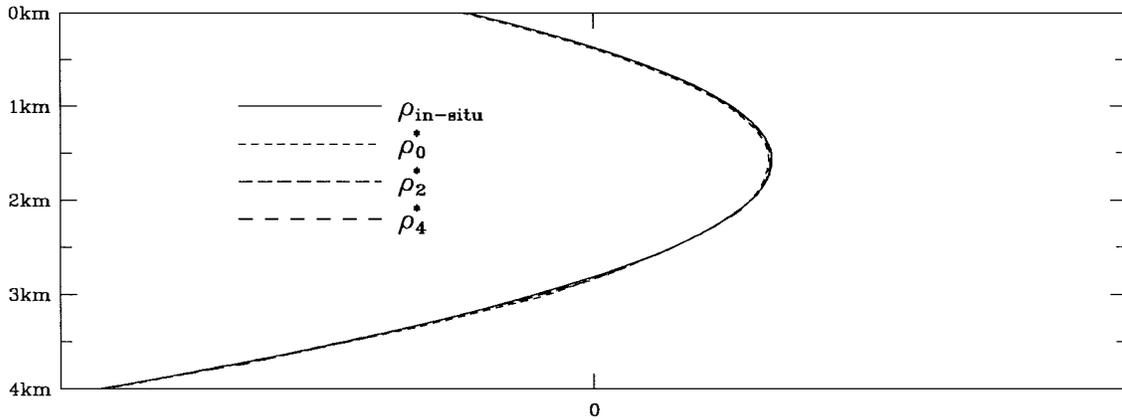


FIG. 5. Interhemispheric pressure difference (in relative units) based on θ/S profiles shown in Fig. 1. Solid: using in situ density (repeated from Fig. 1); dashed: using virtual potential densities ρ_0^* , ρ_2^* , ρ_4^* .

members of that equation with respect to z while holding p_l constant, one finds that (21) can also be written in terms of α^* :

$$N^2 = g \frac{\partial}{\partial z} \ln \alpha^*(\theta, S, p_l). \quad (22)$$

Hence, the frequency of buoyancy oscillations remains unchanged if the specific volume of seawater in the model is changed from its in situ value to α^* .

Taking the logarithm in the first part of (7) gives, after setting $p = p_l$,

$$\ln \alpha^*(\theta, S, p_l) = \ln \alpha_r(\theta, S) + \int_{p_l}^{p_r} \kappa^{(0)}(\theta, S, p') dp'.$$

Combining the z derivative of the above and (22) yields

$$g \frac{\partial \ln \alpha_r}{\partial z} \simeq N^2 + g \delta p \left(\frac{\partial \kappa^{(0)}}{\partial \theta} \frac{\partial \theta}{\partial z} + \frac{\partial \kappa^{(0)}}{\partial S} \frac{\partial S}{\partial z} \right). \quad (23)$$

This equation provides a diagnostic criterion for the occurrence of isopycnic-coordinate surface folding due to excessive pressure bias. The conditions under which potential isopycnals fold have previously been discussed by You and McDougall (1990). Our intent in deriving the above formula is to highlight the roles played by the ambient pressure bias and the coefficient $\kappa^{(0)}$ encompassing the essence of thermobaricity.

The critical pressure bias δp_{crit} beyond which folding occurs is found by setting (23) to zero. The potential for folding, indicated by a small absolute value of δp_{crit} , is large if the bracketed term in (23) is large in absolute terms. Given that $\kappa^{(0)}$ is primarily a function of θ (see Fig. 4) and $\partial \kappa^{(0)} / \partial \theta < 0$, the threat of folding in situations where $p_r > p_l$ is thus greatest if θ decreases upward, while in situations where $p_r < p_l$, the threat is greatest if θ increases upward. Since $\partial \kappa^{(0)} / \partial S < 0$, a vertical salinity gradient of the same sign as the potential temperature gradient increases the potential for folding.

Reviewing existing global hydrographic datasets, one

finds that degeneracy of the depth-to-potential density coordinate transformation near the surface and near the bottom in subpolar regions cannot simultaneously be avoided by any choice of reference pressure. Because of the wide range of variations in temperature and salinity in the subpolar surface layer, a practical choice for global modeling is to use a high enough reference pressure to avoid the problem in the abyssal waters, while accommodating near-surface conditions by a modified coordinate mapping procedure there.

Folding of surfaces $\alpha^* = \text{const}$ (an issue of interest if α^* is to be used as vertical model coordinate) can be investigated by deriving an equation analogous to (23) but with $\partial \ln \alpha^* / \partial z$ replacing $\partial \ln \alpha_r / \partial z$ on the right-hand side. Note that α^* in this context is to be taken as a function of in situ pressure, not of p_l as in (22). A derivation paralleling that of (22), but without setting $p = p_l$, yields

$$g \frac{\partial}{\partial z} \ln \alpha^*(\theta, S, p) = N^2 + \rho g^2 \kappa^{(0)}. \quad (24)$$

Assuming, as we did in discussing (20), a representative value of 10^{-5} MPa^{-1} for $|\kappa^{(0)}|$, the term $\rho g^2 \kappa^{(0)}$ in (24) is of order 10^{-6} s^{-2} . This value is small compared to N^2 in the thermocline, though definitely not in the deep ocean. However, by adding a constant to (9) [and modifying c_3 in (8) accordingly], $\kappa^{(0)}$ can be reduced to near zero at cold temperatures, thereby rendering the term $\rho g^2 \kappa^{(0)}$ small compared to N^2 in the deep ocean as well. With this in mind, we can state that the depth variation of virtual potential density closely approximates the depth variation of *locally referenced* potential density and that multivaluedness of α^* in the vertical is therefore likely to be uncommon.

d. Neutrality of potential and virtual potential isopycnals

Aside from the folding of isopycnals, the deviation from neutrality of constant potential density surfaces is

a concern in isopycnic modeling. Lateral stirring predominantly takes place along neutral surfaces. Thus, alignment of coordinate surfaces with neutral surfaces not only results in a clean separation of isopycnal and diapycnal mixing in the model, but also allows naturally occurring mixing to at least partially mask numerical dispersion errors associated with along-coordinate advection.

McDougall (1988) expresses the slope difference between α_r and neutral surfaces as a function of the isoneutral θ gradient, the vertical gradients of θ and S , and the pressure variation of the saline contraction and thermal expansion coefficients. Our aim here is to cast the nonneutrality condition explicitly in terms of $\kappa^{(\theta)}$.

It follows from the definition of neutral surfaces that they are tangential to (strictly speaking, osculating with respect to) a constant potential density surface referenced to the mean ambient pressure p_l (McDougall 1987). The nonneutrality of a potential isopycnal referenced to a "distant" pressure p_r can be expressed in terms of the gradient of potential density referenced to p_l along this surface. Setting $p = p_l$ and keeping both p_l and p_r constant, we obtain from (5)

$$\begin{aligned} \nabla_{\alpha_r} \ln \alpha_l &= \nabla_{\alpha_r} \int_{p_l}^{p_r} \kappa^{(\theta)}(\theta, S, p') dp' \\ &\approx -\delta p \nabla_{\alpha_r} \kappa^{(\theta)}. \end{aligned} \quad (25)$$

The deviation of neutral surfaces from the coordinate isopycnals, measured in terms of potential specific volume anomaly, is thus proportional to the product of the ambient pressure bias δp and the isopycnal gradient of the thermobaric compressibility.

As pointed out before and shown in Fig. 4, $\kappa^{(\theta)}$ is primarily a function of θ . Hence, (25) can be further reduced to

$$\nabla_{\alpha_r} \ln \alpha_l \approx -\delta p \frac{\partial \kappa^{(\theta)}}{\partial \theta} \nabla_{\alpha_r} \theta. \quad (26)$$

Within the limits of the approximations made in deriving this expression [which is similar to one in section 7.2 of McDougall (1988)], we can state that in oceanic regions where water mass properties are tightly coupled in θ - S space (meaning that both θ and S , and by implication $\kappa^{(\theta)}$, are functions of α_r alone), potential isopycnals are neutral surfaces regardless of the magnitude of the pressure bias δp . Stated differently, nonneutrality is only an issue where water masses of different θ , S properties meet at pressures markedly different from the reference pressure p_r .

It is of interest to explore whether the nonneutrality problem associated with α_r surfaces can be reduced by switching to an α^* -based coordinate system. A measure of the deviation of α^* surfaces from neutrality can be found by evaluating the gradient of α_l at constant α^* . Setting $p = p_l$ in (5) and combining it with (7) gives

$$\alpha_l = \alpha^* A_r(p_l) \exp\left(-\int_p^{p_l} \kappa^{(\theta)}(\theta, S, p') dp'\right).$$

This yields

$$\begin{aligned} \nabla_{\alpha^*} \ln \alpha_l &= -\nabla_{\alpha^*} \int_p^{p_l} \kappa^{(\theta)}(\theta, S, p') dp' \\ &= -\int_p^{p_l} \nabla_{\alpha^*} \kappa^{(\theta)}(\theta, S, p') dp' + \kappa^{(\theta)} \nabla_{\alpha^*} p. \end{aligned}$$

Evaluated at $p = p_l$, this becomes

$$\nabla_{\alpha^*} \ln \alpha_l = \kappa^{(\theta)} \nabla_{\alpha^*} p. \quad (27)$$

Comparison of (26) and (27) shows that the nature of the approximation error in the α_r and α^* is very different.

As implied by the discussion of (26), α_r surfaces become neutral surfaces as $\delta p \rightarrow 0$. Incidentally, this provides an analytical basis for the practice in hydrographic diagnostics where patching isopycnic-coordinate maps with reference pressures separated by $O(10)$ MPa provides good global continuity in water-mass distribution characteristics (Lynn and Reid 1968). The equivalence also works for conditions of weak variability of $\kappa^{(\theta)}$ on isopycnals, that is, under conditions of approximate lateral water mass homogeneity.

An α^* surface, on the other hand, is exactly neutral only if it is isobaric. Thus, if a finite pressure bias δp combined with an isopycnal water mass gradient creates a significant neutral surface slope relative to an α_r surface, the associated α^* surface will be more neutral as long as its slope is small. Note that δp does not appear in (27).

The quantitative importance of the right-hand side of (27) is hard to judge. Like the term $\rho g^2 \kappa^{(\theta)}$ in (24), it contains $\kappa^{(\theta)}$ as an undifferentiated factor; hence, its magnitude in different regions of (θ, S, p) space can be manipulated by an additive constant. While minimizing $\kappa^{(\theta)}$ in the abyssal ocean appeared to be the best strategy for minimizing the importance of the term $\rho g^2 \kappa^{(\theta)}$ in (24), (27) suggests that buoyancy neutrality of α^* surfaces benefits most from reducing the overall magnitude of $\kappa^{(\theta)}$, that is, from removing a nonzero mean. If this "mean" is evaluated in a mass-weighted sense, the fact that most of the water in the ocean is close to freezing implies that the two requirements actually can be satisfied simultaneously.

Judging from the above discussion, α^* does not appear to offer a clear advantage over α_r as far as buoyancy neutrality is concerned. In both cases there will be oceanic regions where numerical diffusion along a coordinate surface has a nonzero diapycnal component. This problem can be ameliorated to some extent by mixing tensor rotation (Griffies et al. 1998), but one should note that improving the accuracy of explicit mixing processes in the model leaves unaffected the diapycnal mixing

resulting from numerical dispersion errors in the horizontal transport calculation in situations where coordinate surfaces and neutral surfaces are not aligned.

5. Discussion

We have introduced in this paper a new variable called “virtual potential density” (and its reciprocal, virtual potential specific volume α^*), which can be viewed as the potential density that would be computed from in situ conditions using a compressibility coefficient reflecting the pressure dependence of seawater compressibility alone.

Our main result, which is of potential interest not only in the context of isopycnic models but in any ocean model featuring steeply inclined coordinate layers, is that α^* is a near-perfect substitute for in situ specific volume in representing the oceanic mass field in dynamic model computations, including the full complement of vertical and horizontal pressure forces.

Specifically, α^* defines the layerwise increments of a modified Montgomery potential, $\partial M^* = \alpha^* \partial p$, which defines a major portion of the horizontal PGF in the momentum equations. This provides the basis for including the thermobaricity effect in isopycnic coordinate models, and ensures that in their absence the model dynamics reduce to the conventional incompressible case where one sets $\partial M = \alpha_r \partial p$.

Consistent with the notion that the spatial distribution of α^* defines much of the stratification impact on the isopycnic system dynamics, we have shown that the Brunt–Väisälä frequency is represented correctly. This implies that substituting α^* for α in a numerical model does not affect internal wave speeds.

By comparing the vertical gradient of α_r to that of potential specific volume α_r , we arrived at a criterion for the single-valuedness of the latter with respect to z . Sign changes in the vertical α_r gradient signal a degeneracy in the depth-to-isopycnic coordinate transformation. The single-valuedness requirement constrains the reference pressure choice whenever θ and/or S increases upward in the abyssal ocean, or decreases upward near the surface. In today’s ocean, this constraint can be satisfied in the abyssal waters by choosing a middepth reference pressure such as 20 MPa. The cold halocline near-surface layers in subpolar regions cannot be accommodated in this way, though, without either introducing a layer patching scheme with different reference pressures, or some form of hybrid coordinate transformation in the near-surface layers.

The gradient of α_r along a potential isopycnal is a useful measure of the degree to which the latter deviates from the local neutral surface. Since the slope difference is mainly a function of the isopycnic potential temperature gradient, weighted by the ambient pressure bias, it vanishes asymptotically wherever the temperature–salinity correlation is tight, or where the ambient pressure bias is small. This implies that potential isopycnals

are piecewise nearly neutral for diagnostic purposes. Significant deviations occur only in regions of pronounced water mass contrasts, and even there only for large pressure biases.

The agreement between the vertical gradients of α_r and α^* expressed by (24) suggests that α^* is single-valued in z even in situations where α_r is not. It is therefore of interest to explore the pros and cons of α^* as vertical model coordinate. The ideal coordinate would be one that is single-valued in z , constant along neutral surfaces, material under adiabatic conditions, and endowed with a simple PGF formula. Attempts at finding a variable that satisfies several of these conditions have been undertaken by Jackett and McDougall (1997) and, more recently, by Eden and Willebrand (1998).

Our analysis suggests that surfaces $\alpha^* = \text{const}$ do not systematically track neutral surfaces (locally referenced potential isopycnals) any better than do surfaces $\alpha_r = \text{const}$. From this perspective, α^* offers no apparent advantage over potential density as a vertical model coordinate. However, α^* is advantageous in that the two-term expression, $-\nabla M^* + p \nabla \alpha^*$, for the horizontal PGF given in (19) reduces to $-\nabla M^*$ if evaluated in (x, y, α^*) coordinates. That the PGF can be expressed as the gradient of a potential function ($M^* = p \alpha^* + \phi^*$ in this case) makes it easy for a model using α^* as vertical coordinate to emulate dynamic constraints governing spinup of horizontal circulation systems.

There appear to be two reasons for choosing α^* over α_r as vertical coordinate, namely, simplicity of the PGF formulation and less likelihood of its being multivalued in z . The main disadvantage of α^* is its nonmaterial character in adiabatic flow. As far as buoyancy neutrality is concerned, neither variable appears to have a clear advantage. The relative weights of the four factors considered undoubtedly depends on the particular model application and will have to be determined in practical tests.

An important reason for retaining α_r as a vertical coordinate would be that it is conserved in adiabatic flow. This greatly facilitates the dynamic interpretation of model results because it unambiguously links inter-layer mass exchange to diapycnal processes acting on the fluid. More importantly, it greatly reduces the diffusion associated with the numerical evaluation of vertical transport terms in the model equations. Virtual potential density does not qualify in this regard, as can be seen from the equation for the time rate of change of α^* obtainable from (7). Under adiabatic conditions this “generalized vertical velocity” in α^* coordinates reduces to

$$\frac{d}{dt} \ln \alpha^* = -\kappa^{(\theta)} \frac{dp}{dt},$$

whereas $d\alpha_r/dt$ reduces to zero.

Acknowledgments. This work was sponsored by the

U.S. Department of Energy under the Computer Hardware, Advanced Mathematics, and Model Physics (CHAMMP) program, Grant DEFG05-94ER61943, and by the National Science Foundation, Grant OCE92-06643.

REFERENCES

- Bleck, R., and S. G. Benjamin, 1993: Regional weather prediction with a model combining terrain-following and isentropic coordinates. Part I: Model description. *Mon. Wea. Rev.*, **121**, 1770–1785.
- , C. Rooth, D. Hu, and L. Smith, 1992: Salinity-driven thermocline transients in a wind- and thermohaline-forced isopycnic coordinate model of the North Atlantic. *J. Phys. Oceanogr.*, **22**, 1486–1505.
- Eden, C., and J. Willebrand, 1998: Neutral density revisited. *Deep-Sea Res.*, in press.
- Gill, A. E., 1982: *Atmosphere–Ocean Dynamics*. Academic Press, 662 pp.
- Griffies, S. M., A. Gnanadesikan, R. C. Pacanowski, V. T. Larichev, J. K. Dukowicz, and R. D. Smith, 1998: Isoneutral diffusion in a z-coordinate ocean model. *J. Phys. Oceanogr.*, **28**, 805–830.
- Holland, W. R., 1967: On the wind-driven circulation in an ocean with bottom topography. *Tellus*, **19**, 582–599.
- Jackett, D. R., and T. J. McDougall, 1995: Minimal adjustment of hydrographic profiles to achieve static stability. *J. Atmos. Oceanic Technol.*, **12**, 381–389.
- , and —, 1997: A neutral density variable for the world's oceans. *J. Phys. Oceanogr.*, **27**, 237–263.
- Janjic, Z., 1977: Pressure gradient force and advection scheme used for forecasting with steep and small scale topography. *Contrib. Atmos. Phys.*, **50**, 186–199.
- Levitus, S., 1982: *Climatological Atlas of the World Ocean*. NOAA Prof. Paper No. 13, U.S. Govt. Printing Office, 173 pp.
- Lynn, R. J., and J. L. Reid, 1968: Characteristics and circulation of deep and abyssal waters. *Deep-Sea Res.*, **15**, 577–598.
- McDougall, T. J., 1987: Neutral surfaces. *J. Phys. Oceanogr.*, **17**, 1950–1964.
- , 1988: Neutral-surface potential vorticity. *Progress in Oceanography*, Vol. 20, Pergamon, 185–221.
- Montgomery, R. B., 1937: A suggested method for representing gradient flow in isentropic surfaces. *Bull. Amer. Meteor. Soc.*, **18**, 210–212.
- Oberhuber, J. M., 1993: Simulation of the Atlantic circulation with a coupled sea ice–mixed layer–isopycnal general circulation model. Part I: Model description. *J. Phys. Oceanogr.*, **23**, 808–829.
- Open University, 1989: *Ocean Circulation*. Pergamon Press, 238 pp.
- Pedlosky, J., 1982: *Geophysical Fluid Dynamics*. Springer-Verlag, 624 pp.
- Sun, S., 1997: Compressibility effects in the Miami Isopycnic Coordinate Ocean Model. Ph.D. dissertation, University of Miami, 138 pp. [Available from Dr. Shan Sun, NASA/Goddard Institution for Space Studies, 2880 Broadway, New York, NY 10025.]
- Welander, P., 1966: A two-layer frictional model of wind-driven motion in a rectangular oceanic basin. *Tellus*, **18**, 54–62.
- You, Y., and T. J. McDougall, 1990: Neutral surfaces and potential vorticity in the world's oceans. *J. Geophys. Res.*, **95**, 13 235–13 261.