

Tracer Conservation with an Explicit Free Surface Method for z -Coordinate Ocean Models

STEPHEN M. GRIFFIES AND RONALD C. PACANOWSKI

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

MARTIN SCHMIDT

Baltic Sea Research Institute Warnemünde, Rostock Warnemunde, Germany

V. BALAJI

Cray/SGI, NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 2 December 1999, in final form 24 March 2000)

ABSTRACT

This paper details a free surface method using an explicit time stepping scheme for use in z -coordinate ocean models. One key property that makes the method especially suitable for climate simulations is its very stable numerical time stepping scheme, which allows for the use of a long density time step, as commonly employed with coarse-resolution rigid-lid models. Additionally, the effects of the undulating free surface height are directly incorporated into the baroclinic momentum and tracer equations. The novel issues related to local and global tracer conservation when allowing for the top cell to undulate are the focus of this work. The method presented here is quasi-conservative locally and globally of tracer when the baroclinic and tracer time steps are equal. Important issues relevant for using this method in regional as well as large-scale climate models are discussed and illustrated, and examples of scaling achieved on parallel computers provided.

1. Introduction

An explicit free surface method has been developed for the z -coordinate Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model (MOM) (for documentation of the model, see Pacanowski and Griffies 2000). The scheme allows for the use of a free surface with substantially improved stability properties over the older approach of Killworth et al. (1991, referred to as KSWP in the following), as well as for greatly improved conservation of tracers, such as heat and salt, relative to either the KSWP or Dukowicz and Smith (1994) approach. The result is an algorithm suitable for climate as well as regional ocean models.

The purpose of this paper is to present the method and to discuss its physical and numerical properties. Solution examples are provided of the Goldsbrough–Stommel circulation (see Huang 1993) driven by surface freshwater forcing, a baroclinic adjustment problem

similar to that considered by Marsigli in the seventeenth century (see Gill 1982), and a global ocean simulation.

Before detailing the algorithm and showing examples, it is useful to provide some context for the present work.

a. Tracer conservation and freshwater forcing

For a Boussinesq fluid, volume conservation over a vertical column yields the free surface height equation

$$\partial_t \eta = -\nabla \cdot \mathbf{U} + q_w, \quad (1)$$

where η is the deviation of the fluid surface elevation from its resting position at $z = 0$; $\mathbf{U} = \int_{-H}^{\eta} \mathbf{u} \, dz$ is the vertically integrated horizontal velocity, also known as the transport; and $q_w = P - E + R$ represents a volume flux per unit area (with units of a velocity) of freshwater passing across the air–sea interface due to precipitation, evaporation, and/or runoff. It is through this balance of volume that dynamical assumptions, which affect η and \mathbf{U} , determine the ability of a model to faithfully represent freshwater forcing.

Many ocean models employ a constant volume. For example, the rigid-lid approximation sets $\partial_t \eta = 0$, thence eliminating fast barotropic waves. With this assumption, volume conservation via Eq. (1) leads to the

Corresponding author address: Dr. Stephen M. Griffies, NOAA/Geophysical Fluid Dynamics Laboratory, P.O. Box 308, Forrestal Campus, Princeton, NJ 08542.
E-mail: smg@gfdl.gov

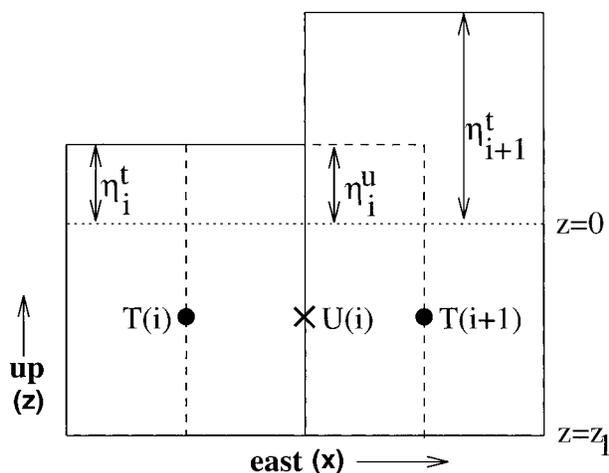


FIG. 1. Schematic of rectangular surface tracer and velocity cells with the free surface. The tracer points are denoted by a solid dot and the tracer cells are enclosed by solid lines; a velocity point is denoted by an \times and is enclosed by dashed lines. These points have fixed vertical position $z = z_1/2 < 0$, regardless of the value for the free surface height. The thickness of a tracer cell is $h' = -z_1 + \eta'$, where $z_{k=0} = \eta'$ is the prognostic surface height determined through volume conservation [i.e., the vertically integrated continuity equation (1)]. The thickness of a velocity cell is $h'' = -z_1 + \eta''$, where η'' is determined by taking the minimum height of the four surrounding tracer cells, as prescribed by the partial cell approach of Pacanowski and Gnanadesikan (1998). Thicknesses of the cells are assumed to be positive, so that $-z_1 + \eta' > 0$. In models without an explicit representation of tides, this constraint in practice means that $|z_1|$ must be greater than roughly 2 m. For models with tides, $|z_1|$ may need to be somewhat larger, depending on the tidal range considered.

balance $q_w = \nabla \cdot \mathbf{U}$, which forms the basis for the work of Huang (1993). The corresponding barotropic dynamics require both a streamfunction and velocity potential, thus introducing two elliptic problems. Elliptic problems are generally inefficient to solve, especially in the presence of realistic geometry, topography, and surface forcing, and moreso on parallel computers (see section 5). Bryan (1969) made the further assumption that the vertically integrated velocity has zero divergence, resulting in a zero vertical velocity at the ocean surface: $w(z=0) = -\nabla \cdot \mathbf{U} = 0$. This assumption eliminates the velocity potential, yet at the cost of precluding a direct freshwater forcing.

A free surface model, in principle, should trivially allow for the introduction of freshwater, since it only requires the addition of q_w to the free surface height equation (1). However, doing so consistently and conservatively requires some changes to the baroclinic and tracer equations that are not commonly made. To illustrate this point, consider what conservation means for total salt in a single grid cell in the special case of zero salt fluxes, but generally nonzero fresh water fluxes. With rectangular model cells (Fig. 1), salt is conserved if

$$\partial_t(h's) = 0, \quad (2)$$

where h' is the thickness of a model tracer cell. Salt conservation with a rigid lid, in which $\partial_t \eta = 0$, means

salinity is constant. For a free surface model, $h' = |z_1| + \eta$, where $z = z_1 < 0$ is the bottom of the surface tracer cell, and so salt conservation means $(|z_1| + \eta)\partial_t s = -s\partial_t \eta$. When $|z_1| \gg |\eta|$, this equation is often approximated by $|z_1|\partial_t s = -s\partial_t \eta$ (e.g., KSWP; Dukowicz and Smith 1994; Wolff et al. 1997; Marshall et al. 1997), which generally precludes salt conservation whenever the surface height changes, as through the addition of freshwater [see Barnier (1998) or Gordon et al. (2000) for a discussion of the virtual salt fluxes applied to constant volume ocean models]. Such limitations are also relevant for heat and other tracers. Our aim here is to relax this approximation and to provide a conservative scheme valid over a wide range of vertical resolutions, including high resolutions where $|z_1| \sim |\eta|$.

In principle, implementation of tracer forcing in flux form is sufficient to ensure that tracer is globally conserved. However, with time-dependent cell thicknesses, details related to the discrete time stepping scheme, including the use of time filters necessary with the three-time-level leapfrog scheme, serve to preclude strict conservation. Additionally, global conservation is neither necessary nor sufficient to ensure locally conservative behavior. For example, consider an ocean at rest with globally constant tracer concentration. Then allow a wind to blow so to initiate velocity and surface height fluctuations, yet do not introduce any tracer surface fluxes. In addition to global tracer content remaining fixed, the tracer concentration locally at every point in the ocean will maintain the same constant value. This behavior provides an example of “local” tracer conservation, which is not a necessary result of global conservation. Instead, it arises so long as there is *compatibility* between tracer and volume budgets.

b. Splitting between fast and slow dynamics

For computational efficiency, primitive equation ocean models aim to exploit the timescale split between the fast barotropic gravity waves, which can propagate at some 200 m s^{-1} , and the much slower (some 100 times slower) remaining dynamics. Such models approximate the barotropic mode by the depth-averaged motion, and the baroclinic mode is approximated by the deviation from the depth average. In a rigid-lid ocean model, vertical averaging provides a clean split between the fast and slow modes. In contrast, vertical averaging does not completely separate out the fast dynamics in a free surface model, since in this case the barotropic mode is weakly depth dependent. The papers by KSWP, Dukowicz and Smith (1994), Higdon and Bennett (1996), Higdon and de Szoeke (1997), and Hallberg (1997) provide discussions of these issues, which can be quite subtle yet important for the purpose of providing a stable split between the fast and slow modes in realistic ocean models. Additionally, details of this split have crucial implications for the tracer conservation properties of the model.

c. Modeling context and summary of key results

MOM has traditionally been a rigid-lid ocean model, where the streamfunction algorithm of Bryan (1969) was used to solve for the barotropic component of the dynamics. In the 1990s, important alternatives have been added, including the explicit free surface of KSWP, the rigid lid–surface pressure of Smith et al. (1992) and Dukowicz et al. (1993), and the implicit free surface of Dukowicz and Smith (1994). Version 3 of MOM (MOM 3) includes *all* of these algorithms within a single model framework. There are numerous additional developments using a free surface in z -coordinate ocean models, such as that used in the Ocean Circulation and Climate Advanced Modelling project (OCCAM; Webb et al. 1998), where the KSWP scheme is refined, as well as the Hamburg Ocean Primitive Equation (Wolff et al. 1997), Massachusetts Institute of Technology (Marshall et al. 1997), and Océan Parallélisé (OPA; Roullet and Madec 2000) ocean models, each of which employ an implicit time stepping algorithm. The present paper is the result of an effort to assess each of these free surface schemes, and possible alternatives such as those of Blumberg and Mellor (1987) and Mellor (1996), who use a terrain-following coordinate model, and Bleck and Smith (1990) and Hallberg (1997) who employ isopycnal layered models.

When testing the various free surface methods, we kept the following goals in mind: 1) MOM's dynamical core should provide a faithful representation of large-scale ocean dynamics in an efficient manner for use in both coarse ($>1^\circ$) and fine ($<1^\circ$) resolution global and regional experiments. 2) Because of the increasing importance of parallel computation, the model should scale well as the number of computer processors increases. 3) The input of freshwater should occur naturally and allow tracer conservation. As described in this paper, our preferred approach is to use an explicit free surface method, which is shown here to satisfy these three goals.

There are important differences between the free surface method as described in this paper and the other free surface methods currently in use with z -coordinate models. First, the KSWP, Dukowicz and Smith (1993), Wolff et al. (1997), and Marshall et al. (1997) approaches assume the top model box to have fixed volume for purposes of the tracer and baroclinic momentum budgets, and so do not satisfy the third goal above. Second, the OCCAM model (Webb et al. 1998) relaxes this constraint, although details of their method are not documented. As shown in this paper, allowing the top box to change in time is necessary, but not sufficient, to globally and locally conserve tracer when using a free surface method. Details of the time stepping scheme are crucial and warrant careful analysis. Third, the OPA model also relaxes the fixed top cell volume constraint, yet it employs an implicit approach that is arguably less straightforward than an explicit approach, in addition does not tend to perform as well on parallel machines

[though advances are available, as documented by Guyon et al. (2000)].

d. Contents of this paper

This paper consists of the following sections. Section 2 describes the model's discrete tracer equations. It highlights points concerning tracer conservation in the presence of a time-dependent surface cell. Section 3 describes the model's discrete equations of motion and details the time stepping algorithm for the barotropic and baroclinic systems. Section 4 exhibits results from idealized numerical examples that provide some practical illustrations of the method. Section 5 discusses issues related to using the method in a parallel computational environment. Section 6 finishes the paper with conclusions.

2. Discrete tracer budget

The purpose of this section is to derive the discrete tracer budgets within the free surface ocean model with a time-dependent top cell volume. For orientation and notation, Fig. 1 details a zonal-depth cross section of the *rectangular* control volumes over which the model's surface equations are discretized. Notably, the position of a grid point within a cell is assumed to be fixed in time, hence maintaining the Eulerian nature of MOM. However, the corresponding grid cell volume generally changes according to temporal undulations of the surface height. We do *not* make the approximation that the surface height is small relative to any other length scale in the model, including the ocean depth.

a. Continuous time tracer budget

Omitting meridional gradients for brevity, the continuous time budget for the total amount of tracer within a discrete tracer cell (Fig. 1) is given by

$$\partial_t(Ah'T) = -dy(\tilde{F}_{i+1}^x - \tilde{F}_i^x) - A(F_{k-1}^z - F_k^z), \quad (3)$$

where T is the tracer per unit volume, that is, a tracer concentration; $A = dx dy$ is the horizontal area of the tracer cell; $h' = -z_i + \eta'$ is the thickness of the tracer cell; i denotes the zonal grid point and k the vertical; and these labels are exposed only when needed. Also, \tilde{F}^x is the *thickness weighted* horizontal advective and diffusive fluxes, which are computed as in Pacanowski and Gnanadesikan (1998) to account for the generally space–time-dependent grid cell thicknesses.

In Eq. (3), F_k^z is the vertical turbulent and advective tracer flux that passes through the cell interface k , with $F_k^z > 0$ representing an upward flux of tracer. In particular, $F_{k=0}^z$ represents the tracer flux through the time-dependent sea surface $z_{k=0} = \eta'$. This flux involves vertical tracer advection relative to the undulating sea surface, which is induced by freshwater flux through the sea surface, as well as parameterized turbulent flux:

$$F_{k=0}^z = F_{k=0}^{z,\text{urb}} - q_w T_{k=0}. \quad (4)$$

Here, $F_{k=0}^z$ must be calculated from a boundary condition that equates this flux with the total tracer flux Q_T entering the ocean from other component models, such as atmosphere, river, and/or ice models. Generally Q_T has a contribution from parameterized turbulence as well as a tracer flux with freshwater,

$$Q_T = Q_T^{\text{urb}} - q_w T_w, \quad (5)$$

where T_w is the tracer concentration in the freshwater. Although T_w and $T_{k=0}$ may be of the same order of magnitude, the terms $q_w T_w$ and $q_w T_{k=0}$ stand for different physical processes. The term $q_w T_w$ represents the amount of tracer passing through the air–sea interface with freshwater, whereas $q_w T_{k=0}$ represents the advection of tracer in the ocean relative to the sea surface.

Specification of the surface tracer flux involves details of how the air–sea interface is modeled. A simple “closure” for the turbulence term, which is often used for ocean-only simulations, is a restoring condition $Q_T^{\text{urb}} = \gamma h'(T_1 - T^*)$, where T^* is the reference tracer, γ is an inverse damping time, and T_1 is the time-lagged surface tracer concentration. For a more refined coupling of ocean and atmosphere, the flux Q_T^{urb} and the tracer T_w have to be prescribed, for example, by a boundary layer model. For salt, which does not pass the sea surface and is conserved exactly, $Q_T^{\text{urb}} = 0$ and $T_w = 0$ is appropriate (salt is considered in detail in section 2g). For other tracers, approximations depend on the details of the boundary layer model [a simple approximation is given in section 2f; see also Pacanowski and Griffies (2000) or Large and Pond (1981) for more discussion].

b. Time discretization

The question arises whether it is preferable to time step the thickness-weighted tracer concentration $h'T$ or the tracer concentration T . Tests with both approaches were performed and showed negligible differences. Notably, as shown below, both approaches have the *same* temporal order of accuracy, and so choosing one is a matter of convenience. As time stepping T involved the least changes to the existing MOM code, we detail that approach in the following. A similar reasoning applies to the momentum equation time discretization discussed in section 3.

Dividing the tracer budget (3) by the tracer cell cross-sectional area, and performing the product rule on $\partial_i(h'T)$, yields the tracer concentration budget

$$h' \partial_t T = -\partial_x \tilde{F}^x - (\delta_{k1} T_k \partial_t \eta + F_{k-1}^z - F_k^z), \quad (6)$$

where δ_{k1} is the Kronecker delta, which is nonzero only for the surface cell when $k = 1$. This budget is time stepped with a leapfrog scheme along with a Robert time filter (e.g., Haltiner and Williams 1980) to suppress time splitting. When needing the cell thickness, the thickness is set to that value at the current time step $h'(\tau)$.

c. Compatibility between volume and tracer budgets

Recall the example considered in section 1a, in which an ocean at rest with an initially constant tracer concentration is perturbed via a wind stress yet without tracer sources. Again, the solution should maintain the same constant tracer concentration locally at each grid cell, even as the surface cells undulate. To do so, it is necessary for there to be compatibility between the volume and tracer budgets. In particular, upon setting the tracer concentration to a constant, the tracer budget in the surface grid cells must reduce to the discretized surface height equation (1). Compatibility between the budgets thus imposes a constraint on how to time step the surface height, and such is detailed in section 3c.

d. Total tracer content in the discrete model

What is desired is the temporal combination of thickness and tracer that measures the evolution of total tracer content in the model. Details of the time stepping scheme for both the tracer and surface height determine the form of such a measure; indeed, whether or not it exists. To illustrate these points, consider, as in the introduction, a single tracer cell in the absence of tracer fluxes. Doing so leads to the balance

$$h(\tau) \partial_\tau T(\tau) + T(\tau) \partial_\tau h(\tau) = 0, \quad (7)$$

where the superscript on the tracer cell thickness is suppressed ($h = h'$). Use of a leapfrog scheme for *both* tracer and surface height, with equal time step τ , renders the discrete conservation law

$$\begin{aligned} h(\tau)T(\tau + \Delta\tau) + T(\tau)h(\tau + \Delta\tau) \\ = h(\tau - \Delta\tau)T(\tau) + T(\tau - \Delta\tau)h(\tau). \end{aligned} \quad (8)$$

Alternative time stepping schemes for either tracer and/or thickness will necessarily lead to distinct discrete conservation laws, or even to the inability to identify such at all. We have explored a number of possibilities, with the choice of leapfrog for both tracer and thickness being the most natural way to provide compatibility between the tracer and volume budgets.

As previously mentioned, an alternative is to leap frog the thickness-weighted tracer hT , which in the absence of boundary tracer fluxes leads to the discrete conservation law:

$$h(\tau + \Delta\tau)T(\tau + \Delta\tau) = T(\tau - \Delta\tau)h(\tau - \Delta\tau). \quad (9)$$

Importantly, both of the conservation laws (8) and (9) are *equivalent* to second-order accuracy in $\Delta\tau$. That is, they both reduce to $\partial_i(hT) = 0$ at the same rate as $\Delta\tau$ is reduced. Hence, although nonstandard, the conservation law (8) resulting from separately time stepping tracer concentration and thickness is just as valid as the alternative (9). Again, our choice for (8) is based on convenience.

More generally, manipulation of the tracer concentration budget (6), again assuming that both the tracer

concentration and surface height are time stepped via a leapfrog, leads, in the absence of boundary fluxes, to conservation of the total tracer content as measured by

$$\sum_{ijk} A_{ij} [h_k(\tau) T_k(\tau + \Delta\tau) + \delta_{ki} T_k(\tau) h_k(\tau + \Delta\tau)], \quad (10)$$

where A_{ij} is the horizontal area of a tracer cell. Again, this measure of total tracer content is the same, to second-order accuracy in $\Delta\tau$, as the analogous measure realized via time stepping hT instead of T .

e. Local compatibility versus global conservation

The compatibility condition, which leads to a locally conservative scheme, and global tracer conservation cannot be simultaneously maintained with a stable three-time-level discretization. For example, use of a time filter on either the tracer (via the Robert filter) or the surface height (via a time averaging procedure described in section 3), preclude strict conservation of the total tracer content given by Eq. (10). However, these filters do not preclude local compatibility between volume and tracer budgets. Alternatively, omitting the time filters on the surface height and tracer in the tracer budget (11) allows for exact global conservation, yet such breaks the local compatibility so long as a time filter is used in the surface height equation (11). We provide two examples in section 4 that aim to quantify these points. Given the inability to provide exact discrete local and global conservation properties with an undulating free surface and the leapfrog scheme, we generally believe it prudent that conservation properties should be assessed on a case by case basis.

Alternative time stepping schemes exist, such as two-time-level schemes (e.g., Hallberg 1995, 1997), which provide potential remedies to the issues raised here. However, pursuit of such alternatives goes beyond our present scope.

f. Comments on surface fluxes

It is useful to further elucidate the different terms describing the surface tracer fluxes. For this purpose, consider use of the free surface equation (1) and the surface tracer flux expression (5), which provides the relation

$$\begin{aligned} T_1 \partial_t \eta + F_{k=0}^z &= T_1 \partial_t \eta + Q_T \\ &= Q_T^{\text{trb}} - T_1 \nabla \cdot \mathbf{U} + q_w (T_1 - T_w). \end{aligned} \quad (11)$$

Notably, neither the tracer $T_{k=0}$ nor the turbulent flux $F_{k=0}^z$ are explicitly needed. Relatedly, the distinction should be made between $q_w T_1$, which naturally appears here in the relation (11), and $q_w T_{k=0}$, which appears in the expression for $F_{k=0}^z$ in Eq. (4).

Assuming that the tracer concentration in the freshwater is the same as in the surface cell, $T_w \approx T_1$, makes the tracer time tendency independent of any explicit freshwater forcing. Instead, the dependence is restricted

to the affects that freshwater has on the convergence $-\nabla \cdot \mathbf{U}$. This approximation simplifies the setup of boundary conditions for the tracer flux. However, in general T_w and T_1 are distinct and so T_w may need to be specified explicitly from data or another component model.

g. The special case of salt

The air–sea interface effectively acts as an impenetrable barrier for salts due to their large hydration energy. Therefore, neglecting the formation of sea spray and the interchange of salt with sea ice, and in the absence of salt entering through the solid earth boundaries, total ocean salt is constant. In turn, the salt concentration $\rho_o s$ within a grid cell, where s is salinity, changes only through advective and turbulent fluxes from the interior ocean, as well as time tendencies in the volume of the grid cell. From the flux expression (5), a zero salt flux through the air–sea interface formally represents a balance

$$F_{k=0}^z = F_{k=0}^{z,\text{trb}} - q_w s_{k=0} = Q_s = 0 \quad (12)$$

between the turbulent flux and advective flux. Discussion of this balance has been given by Huang (1993) and Dewar and Huang (1996), where the turbulent flux was called an *antiadvective flux*. On the atmosphere side, both the advective and turbulent flux components vanish, $s_w = 0$ and $Q_s^{\text{trb}} = 0$. In relation (11), this result indicates that surface salinity is affected only through the $s_1 \partial_t \eta$ term. A more thorough treatment of the salinity surface boundary condition is given by Beron-Vera et al. (1999).

With large salinity deviations, as may occur in the presence of realistic freshwater fluxes, a more general approach than that discussed by Bryan and Cox (1972) appears necessary to compute ocean density. Bryan and Cox approximate the full United Nations Educational, Scientific and Cultural Organization (UNESCO) equation of state by different cubic polynomials for each model depth level. This approach is warranted for models in which salinity is close to 35 psu and for which the model's vertical cell thickness is independent of horizontal position. The latter assumption is broken by the Pacanowski and Gnanadesikan (1998) partial bottom cells, and the former is broken by realistic freshwater fluxes. Hence, for the purpose of computing the model's hydrostatic pressure, we employ the full UNESCO equation of state as formulated by Jackett and McDougall (1995). In this approach, in situ density is computed as

$$\rho(\tau) = \rho[\theta(\tau), s(\tau), p(\tau - \Delta\tau)], \quad (13)$$

where θ is the model's potential temperature and s is salinity. The pressure p is lagged one time step instead of performing the iterative approach described by Dewar et al. (1998). The added cost of computing density via Jackett and McDougall's UNESCO formulation has been

found to be a modest 5%–10% depending on the model configuration.

3. Discrete momentum budget

In this section, we formulate the discrete momentum budget, which shares much in common with the tracer budget just discussed, and then present a solution algorithm for the model's approximated barotropic and baroclinic modes.

a. Continuous time momentum budget

Ignoring meridional gradients for brevity, the continuous time budget for the zonal momentum of a Boussinesq fluid occupying a rectangular velocity cell is given by

$$\begin{aligned} \partial_t(\rho_o A h^u u)_i &= (\rho_o A h^u f v)_i - h^u dy(p_{i+1} - p_i) \\ &\quad - \rho_o dy(\tilde{F}_{i+1}^x - \tilde{F}_i^x) \\ &\quad - \rho_o A (F_{k-1}^z - F_k^z). \end{aligned} \quad (14)$$

In this equation, A is the horizontal area of the velocity cell, h^u is its thickness (see Fig. 1 for the relation between h^u and h'), and p_i is the hydrostatic pressure acting on the vertical face of the velocity cell. As with the zonal tracer flux, the momentum flux \tilde{F}^x is weighted by the generally time-dependent thickness of the cell. Metric terms arising from momentum advection and friction on a sphere (e.g., Griffies and Hallberg 2000) have been omitted for brevity, but they are present in the numerical code.¹

Here, F^z represents the vertical advective and turbulent momentum flux passing across the horizontal cell faces. In particular, $F_{k=0}^z$ represents the momentum flux passing through the ocean surface at $z_{k=0} = \eta^u$. It consists of the usual vertical turbulent momentum flux and a contribution from the vertical advection of momentum relative to the moving sea surface,

$$F_{k=0}^z = F_{k=0}^{z,\text{turb}} - q_w u_{k=0}. \quad (15)$$

The surface flux $F_{k=0}^z$ must be prescribed by the boundary condition that the total vertical momentum flux through the air–sea interface is continuous at $z_{k=0} = \eta^u$. That is, $F_{k=0}^z = \mathcal{M}$, where the momentum flux \mathcal{M} , which arises from another component model such as an atmospheric boundary layer, takes the form

$$\mathcal{M} = -(\rho_o^{-1} \tau_{\text{winds}}^x + q_w u_w), \quad (16)$$

which has contributions from wind stress as well as

momentum flux with freshwater. Although the total momentum flux is continuous at $z_{k=0} = \eta^u$, the two components need not be. For example, the freshwater velocity u_w may be different from $u_{k=0}$, although most climate models assume they are equal and take them equal to the horizontal velocity $u_{k=1}$ in the top model grid cell. This point is further discussed below.

On the B grid used in MOM, the offset in the horizontal between velocity and tracer/density points means that the hydrostatic pressure at the western face of a surface velocity cell is given by

$$p_i = g \overline{\rho_i (|z_1|/2 + \eta_i)} + p_i^{\text{atm}}, \quad (17)$$

where $\overline{\rho_i}$ is a meridional grid average and η_i is on the tracer point. The first piece of this pressure, $p_b \equiv (g|z_1|/2) \overline{\rho_i}$, represents a discretization of the baroclinic pressure in the surface cell, whose horizontal gradients arise from baroclinicity in the density field. Note that z_1 has been assumed independent of horizontal position, hence its removal from the meridional average operator. The second contribution, $p_s \equiv g \overline{\rho_i \eta_i}$, represents a discretization of the surface pressure, whose gradients arise from those in the density-weighted free surface height. Here, p^{atm} represents an atmospheric pressure contribution, which is dropped in the following for brevity, yet is maintained in the numerical code when coupling to an atmospheric model.

For Boussinesq models, it is common to approximate the surface pressure as $p_s \approx g \rho_o \overline{\eta_i}$, rather than the hydrostatic form $g \overline{\rho_i \eta_i}$. With this approximation, surface pressure gradients arise *solely* from gradients in the free surface height. To maintain self-consistency with the hydrostatic baroclinic pressure field, while incurring only trivial computational expense, we maintain the hydrostatic form $p_s = g \overline{\rho_i \eta_i}$ in the model.

b. General strategy for the time stepping

Manipulations of the budget (14) lead to the zonal velocity equation

$$\partial_t u_k = f v_k - \partial_x p_s / \rho_o + \tilde{G}_k, \quad (18)$$

where \tilde{G}_k represents the forcing

$$\begin{aligned} h_k^u \tilde{G}_k &= h_k^u G_k - \delta_{k1} u_k \partial_t \eta \\ &= -h_k^u \partial_x p_b / \rho_o - \partial_x \tilde{F}^x \\ &\quad - (F_{k-1}^z - F_k^z + \delta_{k1} u_k \partial_t \eta). \end{aligned} \quad (19)$$

Notably, the cell thicknesses h^u used for computing the terms in G_k are taken at baroclinic time τ , as are the other inviscid contributions such as pressure and advection.

The freshwater velocity u_w in the surface momentum flux (16) requires additional efforts to specify the complete boundary conditions. However, the momentum flux with freshwater is in many cases smaller than the uncertainty in the parameterized wind stress. So, simple approximations for u_w may be exploited to reduce the

¹ Additionally, the constant Boussinesq density, ρ_o , is not set to 1 g cm⁻³, as previously done in the Bryan–Cox–Semtner model (Bryan 1969; Cox 1984; Semtner 1974) or previous versions of MOM (Pacanowski et al. 1991). Instead, $\rho_o = 1.035$ g cm⁻³, from which density in the World Ocean generally deviates by less than 2% (Gill 1982, p. 47), whereas using $\rho_o = 1.0$ g cm⁻³ is less accurate.

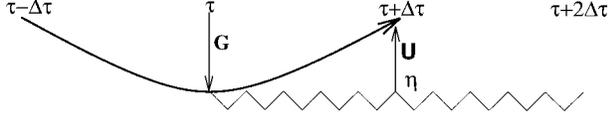


FIG. 2. Schematic of the split-explicit time stepping scheme. Time increases to the right. The baroclinic time steps are denoted by $\tau - \Delta\tau$, τ , $\tau + \Delta\tau$, and $\tau + 2\Delta\tau$. The curved line represents a baroclinic leapfrog time step, and the smaller barotropic time steps $N\Delta t = 2\Delta\tau$ are denoted by the zigzag line. First, a baroclinic leapfrog time step updates the baroclinic mode to $\tau + \Delta\tau$. Then, using the vertically integrated forcing $G(\tau)$ [Eq. (21)] computed at baroclinic time step τ , a forward-backward time stepping scheme integrates the surface height and vertically integrated velocity from τ to $\tau + 2\Delta\tau$ using N barotropic time steps of length Δt , while keeping $G(\tau)$ and the ocean depth $D(\tau)$ fixed. Time averaging the barotropic fields over the $N + 1$ time steps (endpoints included) centers the vertically integrated velocity and free surface height at baroclinic time step $\tau + \Delta\tau$. Note that for accurate centering, N is set to an even integer.

model complexity. Using the free surface height equation (1), and the surface flux $F_{k=0}^z = \mathcal{M}$, with \mathcal{M} given by Eq. (16), leads to

$$u_1 \partial_t \eta + F_{k=0}^z = -\rho_o^{-1} \tau_{winds}^x - u_1 \nabla \cdot \mathbf{U} + q_w(u_1 - u_w). \quad (20)$$

Hence, the approximation $u_w \approx u_1$ removes the explicit dependence of baroclinic momentum on the freshwater flux. Freshwater does influence baroclinic momentum indirectly, however, through its effects on the convergence $-\nabla \cdot \mathbf{U}$ of the vertically integrated velocity. The approximation $u_w \approx u_1$ should be well justified for many cases, and generalizations for special cases such as heavy rainfall are straightforward.

For the barotropic mode, MOM time steps the vertically integrated transport $U = \sum_k h_k u_k$, as is common in shallow water models. Its evolution takes the form

$$\partial_t U = u_1 \partial_t \eta + \sum_k h_k \partial_t u_k = fV - D \partial_x p_s / \rho_o + G, \quad (21)$$

where Eq. (18) was used for $\partial_t u_k$, the vertically integrated forcing is given by $G = \sum_k h_k G_k$ with G_k defined in Eq. (19), and $D = \sum_k h_k$ is the time-dependent ocean depth. Once the transport and ocean depth are updated, the updated barotropic velocity can be diagnosed through $\bar{u}(\tau + \Delta\tau) = U(\tau + \Delta\tau)/D(\tau + \Delta\tau)$.

c. Time stepping algorithm

The focus in this section is on time and depth discretization, with Fig. 2 summarizing the following algorithm. For purposes of brevity, the horizontal spatial discretization discussed in the previous subsection will not be exposed. Discrete baroclinic times and time steps will be denoted by the Greek τ and $\Delta\tau$, respectively, whereas the barotropic analogs will use the Latin t and Δt .

1) ALGORITHM BASICS

Allowing the grid cell thicknesses to evolve introduces a fundamentally new element to the traditional algorithms relevant for constant cells in z -coordinate models (e.g., Bryan 1969; Semtner 1974; Cox 1984; Killworth et al. 1991; Dukowicz and Smith 1994). However, since we impose positive cell thicknesses for all cells, including the surface, modifications to the constant cell approach should be relatively modest. That is a goal of the proposed algorithm.

The basic idea is directly analogous to the approach of Bryan (1969), Semtner (1974), and Cox (1984). That is, the velocity at an arbitrary depth level k and baroclinic time $\tau' = \tau + \Delta\tau$ is split into two components:

$$u_k(\tau') = B_{km}(\tau) u_m(\tau') + [\delta_{km} - B_{km}(\tau)] u_m(\tau') \equiv \hat{u}_k(\tau, \tau') + \bar{u}(\tau, \tau'). \quad (22)$$

This equation is an identity valid for any baroclinic times τ' and τ , with $\tau' = \tau + \Delta\tau$ chosen in the following. The utility of the identity relies on the form of the ‘‘baroclinicity operator’’ used to split the velocity field

$$B_{km}(\tau) = \delta_{km} - D(\tau)^{-1} h_m^u(\tau), \quad (23)$$

where δ_{km} is the Kronecker delta, summation over the repeated vertical level index m is implied, and

$$D(\tau) = \sum_k h_k^u(\tau) = D_o + \eta^u(\tau) \quad (24)$$

is the ocean depth at baroclinic time τ over a column of velocity points, with D_o the resting ocean depth.

This baroclinicity operator projects out the depth-independent part of a field and keeps its approximate baroclinic portion, where the projection is based on the distribution of cell thicknesses at time τ . Introduction of two baroclinic time labels τ and τ' to Eq. (22) is necessitated by the freedom afforded the ocean depth to change in time. That is, τ' represents the baroclinic time of the full velocity field $u_k(\tau')$, whereas τ represents the time used to define the baroclinicity operator $B_{km}(\tau)$ and which defines the split.

If the split introduced in Eq. (22) is successful, $\hat{u}_k(\tau, \tau')$ will evolve on a slow timescale, $\Delta\tau$, and $\bar{u}(\tau, \tau')$ will evolve on the fast timescale, $\Delta t = 2N^{-1}\Delta\tau$, with N determined by the ratio of external to internal gravity wave speeds. The method therefore proceeds by separately updating $\hat{u}_k(\tau, \tau')$ and $\bar{u}(\tau, \tau')$ by exploiting the timescale split. Upon doing so, the right-hand side of the identity (22) will be specified, hence allowing for an update of the full velocity field via

$$u_k(\tau + \Delta\tau) = \hat{u}_k(\tau, \tau + \Delta\tau) + \bar{u}(\tau, \tau + \Delta\tau). \quad (25)$$

Upon updating the velocity cell thickness to $h^u(\tau + \Delta\tau)$, the baroclinic and barotropic velocities $\hat{u}_k(\tau + \Delta\tau)$ and $\bar{u}(\tau + \Delta\tau)$ can then be diagnosed. The following subsections detail this approach.

2) COMPUTING $\hat{u}_k(\tau, \tau + \Delta\tau)$

By construction, evolution of the velocity field $\hat{u}_k(\tau, \tau')$ is unaffected by vertically independent forces, such as those from surface pressure gradients. Therefore, it is sufficient to update the “primed” velocity:

$$u'_k(\tau + \Delta\tau) = u_k^r(\tau - \Delta\tau) + 2\Delta\tau[fv_k(\tau) + \tilde{G}_k(\tau)], \quad (26)$$

which represents a temporal discretization of the full momentum equation (18), yet without the surface pressure gradient, thus allowing for use of the baroclinic time step. The lagged velocity $u_k^r(\tau - \Delta\tau) = u_k(\tau - \Delta\tau) + (\alpha/2)[u_k(\tau) - 2u_k(\tau - \Delta\tau) + u_k^r(\tau - 2\Delta\tau)]$ is a Robert time-filtered version of the full velocity field. A weak form of such filtering, with $\alpha = 0.01$, has been found sufficient to suppress splitting between the two leapfrog branches (e.g., Haltiner and Williams 1980). It is also useful to dampen fast dynamics that may partially leak through the baroclinicity operator due to the generally imperfect separation between the slow and fast dynamics.

The baroclinic piece of the primed velocity $u'_k(\tau + \Delta\tau)$ yields

$$\hat{u}_k(\tau, \tau + \Delta\tau) = B_{km}(\tau)u'_m(\tau + \Delta\tau), \quad (27)$$

as can be shown by adding the missing surface pressure gradient to Eq. (26), where it is annihilated by application of $B_{km}(\tau)$. At this point, we have specified one-half of the updated full velocity given by Eq. (25).

3) COMPUTING $\bar{u}(\tau, \tau + \Delta\tau)$

The computation of $\bar{u}(\tau, \tau + \Delta\tau)$ constitutes a most crucial part of any free surface algorithm, since details largely determine the stability and conservative aspects of the scheme. The following approach, which uses a time averaging over barotropic time steps integrated from τ to $\tau + 2\Delta\tau$, is commonly used in explicit free surface methods. However, its details differ somewhat from other approaches, and so prompts a reasonably thorough presentation.

The barotropic “subcycling” uses a forward-backward time stepping scheme, also used by KSWP (see also, e.g., Haltiner and Williams 1980). In this approach, surface height is first time stepped using the small barotropic time step Δt :

$$\begin{aligned} \eta^*(\tau, t_{n+1}) &= \eta^*(\tau, t_n) \\ &\quad - \Delta t[\nabla \cdot \mathbf{U}^*(\tau, t_n) - q_w(\tau)]. \end{aligned} \quad (28)$$

The asterisk is used to denote intermediate values of the barotropic fields, each of which are updated on the barotropic time steps. The time label τ on η^* and \mathbf{U}^* denotes the baroclinic time at which the vertically integrated forcing $G(\tau)$, the tracer fields, and the total depth of the ocean $D(\tau)$ are held for the duration of the barotropic time stepping from τ to $\tau + 2\Delta\tau$. This is also the time that sets the barotropic time steps via

$$t_n = \tau + n\Delta t, \quad (29)$$

with $n \in [0, N]$. For stability purposes, we found it important to take the initial condition $\eta^*(\tau, t_{n=0})$ as the time average $\bar{\eta}^*(\tau)$ computed from the previous barotropic integration taking place over $\tau - \Delta\tau$ to $\tau + \Delta\tau$ [time averaging is defined by Eq. (32) discussed below]. Note that it is assumed that the freshwater flux q_w is constant over the small barotropic time steps, since the hydrological fluxes are typically updated on each baroclinic time step.

Having the surface height η^* updated to the new barotropic time step allows for an update of the transport

$$\begin{aligned} U^*(\tau, t_{n+1}) &= U^*(\tau, t_n) + f\Delta t \left[\frac{V^*(\tau, t_n) + V^*(\tau, t_{n+1})}{2} \right] \\ &\quad + \Delta t[G(\tau) - D(\tau)\partial_x \tilde{p}_s^*(\tau, t_{n+1})], \end{aligned} \quad (30)$$

where $\tilde{p}_s^*(\tau, t_{n+1}) = g\eta^*(\tau, t_{n+1})\rho(\tau)/\rho_o$ is the surface pressure normalized by the Boussinesq density. The Coriolis force is computed using a Crank–Nicholson semi-implicit time stepping scheme, which was also used by KSWP (see also Haltiner and Williams 1980).

After N barotropic time steps, the vertical transport and surface height are time averaged to produce

$$\bar{U}^*(\tau + \Delta\tau) = \frac{1}{N+1} \sum_{n=0}^N U^*(\tau, t_n) \quad (31)$$

$$\bar{\eta}^*(\tau + \Delta\tau) = \frac{1}{N+1} \sum_{n=0}^N \eta^*(\tau, t_n). \quad (32)$$

The time-averaged fields are centered on the baroclinic time step $\tau + \Delta\tau$, so long as N is an even integer. The time-averaged surface height $\bar{\eta}^*(\tau + \Delta\tau)$ is used to initialize the next suite of barotropic integrations from $\tau + \Delta\tau$ to $\tau + 3\Delta\tau$, and $\bar{U}^*(\tau + \Delta\tau)$ is used to update $\bar{u}(\tau, \tau + \Delta\tau)$ via

$$\bar{u}(\tau, \tau + \Delta\tau) = \bar{U}^*(\tau + \Delta\tau)/D(\tau). \quad (33)$$

Having $\bar{u}(\tau, \tau + \Delta\tau)$ then allows for the full velocity field $u_k(\tau + \Delta\tau)$ to be updated according to Eq. (25).

4) COMPUTING $\eta(\tau + \Delta\tau)$ AND $\bar{u}(\tau + \Delta\tau)$

Although the updated full velocity field $u_k(\tau + \Delta\tau)$ is now known, the updated surface height $\eta(\tau + \Delta\tau)$ and barotropic velocity $\bar{u}(\tau + \Delta\tau) = U(\tau + \Delta\tau)/D(\tau + \Delta\tau)$ remain to be determined. A number of options were tried, with the following two approaches initially attempted. (a) Stop the barotropic subcycling at $\tau + \Delta\tau$ and use the unaveraged barotropic fields:

$$\eta(\tau + \Delta\tau) = \eta^*(\tau, \tau + \Delta\tau) \quad \text{and} \quad (34)$$

$$U(\tau + \Delta\tau) = U^*(\tau, \tau + \Delta\tau). \quad (35)$$

This choice was originally used by KSWP, yet we, and others such as the OCCAM group (see Webb et al. 1998), have found it to result in a very unstable algo-

rithm and so it is not suitable. (b) Use of the time-averaged barotropic fields given by Eqs. (31) and (32),

$$\eta(\tau + \Delta\tau) = \overline{\eta}^*(\tau + \Delta\tau) \quad \text{and} \quad (36)$$

$$U(\tau + \Delta\tau) = \overline{U}^*(\tau + \Delta\tau), \quad (37)$$

allows for a substantial improvement in stability, and it is for this reason that OCCAM chooses this approach. However, this choice is not self-consistent in the sense that $U(\tau + \Delta\tau)$ so defined will not generally be equivalent to $\sum_k h_k^u(\tau + \Delta\tau)u_k(\tau + \Delta\tau)$ due to time-dependent thicknesses. Additionally, use of the time average for the surface height precludes compatibility between the tracer and volume budgets as discussed in section 2c, and it also does not allow for a specification of a discretely conserved total tracer as described in section 2d. Therefore, we propose the following alternative approach that satisfies our goals.

We first update the thicknesses, and then diagnose the integrated transport. To do so, compatibility as discussed in section 2c indicates that since the tracer concentration is time stepped with a leapfrog scheme, so too must the surface height budget:

$$\eta(\tau + \Delta\tau) = \overline{\eta}^*(\tau - \Delta\tau) - 2\Delta\tau[\nabla \cdot \mathbf{U}(\tau) - q_w(\tau)]. \quad (38)$$

As the surface height generally evolves on a fast time-scale, use of this proposed baroclinic leapfrog time step is generally unstable. To maintain stability, as well as general smoothness of the free surface field, it has been found necessary to use the time-averaged surface height $\overline{\eta}^*(\tau - \Delta\tau)$ rather than a more traditional Robert filtered height. Doing so does not effect the compatibility condition, and so has been found suitable. Using this “big leapfrog” step to define $\eta(\tau + \Delta\tau)$ then allows for an update of the surface tracer and velocity cell thicknesses $h_k^t(\tau + \Delta\tau)$ and $h_k^u(\tau + \Delta\tau)$.

Given the updated cell thicknesses, the updated transport $U(\tau + \Delta\tau)$ can be diagnosed

$$U(\tau + \Delta\tau) = \sum_k h_k^u(\tau + \Delta\tau)u_k(\tau + \Delta\tau), \quad (39)$$

thus ensuring self-consistency between the updated full velocity $u_k(\tau + \Delta\tau)$ and the updated vertically integrated velocity. Finally, to complete the time stepping, the updated barotropic velocity is diagnosed through

$$\overline{u}(\tau + \Delta\tau) = U(\tau + \Delta\tau)/D(\tau + \Delta\tau), \quad (40)$$

where $D(\tau + \Delta\tau) = \sum h_k^u(\tau + \Delta\tau)$ is the new depth of a velocity cell column. The updated baroclinic velocity follows from its definition $\hat{u}_k(\tau + \Delta\tau) = u_k(\tau + \Delta\tau) - \overline{u}(\tau + \Delta\tau)$. Note that a similar reconciliation or readjustment of the velocities is discussed by Hallberg (1995, p. 221) for use in his isopycnal model.

d. Comments on the algorithm

1) GENERALITY

Although the immediate application of the algorithm is for the free surface method, the algorithm allows any thickness within a vertical column to vary in time, so long as it remains positive. In particular, it is thought that this approach will find use for models with a time-varying bottom boundary layer thickness, such as that proposed by Killworth and Edwards (1999).

2) STABILITY

As expected, the baroclinic and barotropic time steps are determined by the usual Courant–Freidrichs–Lewy constraints set by waves and advection. Notably, time averaging has not been found to adversely affect the propagation of barotropic waves relative to the results using the KSWP approach. We comment further on this point when presenting a version of the Marsigli problem in section 4c.

Time averaging over the barotropic steps has been found to stabilize the model so that stretching of tracer time steps to values larger than the baroclinic time step is readily available. The stretching of tracer time steps is ubiquitous when spinning up to a thermodynamic equilibrium coarse-resolution, rigid-lid, z -coordinate ocean models (e.g., Bryan 1984; Killworth et al. 1984; Danabasoglu et al. 1996). Therefore, we consider the added stability of the present scheme to be of great practical value.

3) COMPUTATIONAL TIMING

Because of the ability to match the baroclinic and tracer time steps to those commonly used in rigid-lid models, the present scheme has been found to be computationally comparable to the rigid lid on a single computer processor. Furthermore, due to the absence of the topographic instability present in rigid-lid models, which was described by Killworth (1987) and Dukowicz et al. (1993), the free surface is more economical with nontrivial topography. Section 4a provides an example to support these comments, and section 5 presents a discussion of computational aspects relevant for parallel machines.

Relatedly, when using partial bottom cells of Pacanowski and Gnanadesikan (1998), as commonly used now in MOM for representing the bottom topography, use of the time-dependent top model cells engenders only a trivial added cost relative to the case of constant top model cell volumes. This experience is in contrast to the 10% added cost of allowing the surface cell to vary in the implicit method of Roullet and Madec (2000).

4) GRID SPLITTING

As pointed out in KSWP, the surface height on a B grid is prone to grid splitting, which manifests as a checkerboard pattern. The present scheme is less prone to this splitting than other algorithms we tried, and numerous tests indicate that the splitting is easily suppressed with mild filtering as described by KSWP or other filters described in Pacanowski and Griffies (2000). Notably, as the nonlinear free surface described here incorporates the surface height undulations into the tracer and momentum equations, suppression of the grid splitting is more important than in the linearized free surface methods.

5) VOLUME CONSERVATION

To conserve the model volume, it is sufficient to ensure that the surface height η evolves in a conservative manner. Both explicit time stepping discretizations (28) and (38) trivially satisfy such conservation. Therefore, all model tests indicate that volume is conserved to within computer roundoff.

6) ENERGETIC CONSISTENCY

Energetic consistency of the methods used here have been shown elsewhere. First, the arguments given by Bryan (1969) account for the momentum advection terms, which are discretized using second-order centered differences. The result is a redistribution of local kinetic energy, yet a preservation of its global integral. Second, the arguments in the appendix of Pacanowski and Gnanadesikan (1998) show that the partial cell methods ensure that the change in energy due to horizontal pressure forces balances potential energy change when density is linearly dependent on temperature and salinity. Importantly, this consistency is realized only when incorporating the undulating surface grid cell thickness into *both* the tracer and baroclinic velocity equations, as done here. Third, appendix B of Dukowicz and Smith (1994) provides a complementary analysis of the energetic consistency within their implicit free surface method, much of which is relevant for the present considerations.

7) VERTICAL VELOCITIES

Vertical velocities are diagnosed at baroclinic time steps using volume conservation within a grid cell. In particular, volume conservation over a surface cell indicates that the vertical velocity at the bottom face of this cell arises from the horizontal convergence of volume in this cell, time tendencies in the thickness of this cell, and volume passing across the top face from freshwater fluxes.

It is useful to see precisely how these velocities are determined. For this purpose, consider a vertical integral

of the continuity equation $\nabla \cdot \mathbf{u} + \partial_z w = 0$ over a rectangular top model grid cell:

$$w_{k=1} = w(\eta) + \int_{z_1}^{\eta} dz \nabla \cdot \mathbf{u}. \quad (41)$$

Use of the surface kinematic boundary condition

$$[\partial_t + \mathbf{u}(\eta) \cdot \nabla] \eta = w(\eta) + q_w, \quad (42)$$

and Leibnitz's rule with $\nabla_{z_1} = 0$, renders

$$\begin{aligned} w_{k=1} &= \partial_t \eta - q_w + \nabla \cdot \int_{z_1}^{\eta} dz \mathbf{u} \\ &= -\nabla \cdot \mathbf{U} + \nabla \cdot \int_{z_1}^{\eta} dz \mathbf{u}, \end{aligned} \quad (43)$$

where the last step used the vertically integrated continuity equation (1). This exact expression for $w_{k=1}$ is approximated in the model with a discretized version of

$$w_{k=1} \approx -\nabla \cdot \mathbf{U} + \nabla \cdot (h\mathbf{u}), \quad (44)$$

where $h = \eta + |z_1|$ is the surface cell thickness and the horizontal velocity \mathbf{u} is that in the surface cell $k = 1$. Notably, unlike most z -coordinate free surface models, there has been no linearization of the surface kinematic boundary condition made by dropping the surface height advection $\mathbf{u}(\eta) \cdot \nabla \eta$.

Given an expression for the convergence $-\nabla \cdot \mathbf{U}$, diagnosing $w_{k=1}$ in this manner allows for the remaining interior vertical velocities to be successively found through further integration of the continuity equation downward through a vertical column. On a B grid, $-\nabla \cdot \mathbf{U}$ is centered on a tracer point. Hence, Eq. (44) yields the vertical velocity on the bottom face of the surface tracer cell. The vertical velocity on the surrounding velocity cells is constructed as a volume conserving average of the surrounding tracer cell vertical velocities.

In MOM, the bottom of the ocean on tracer cells is a flat surface, representing the "lopped off" surfaces of topography.² Hence, the vertical velocity must vanish at this location. A self-consistency check on how accurately the model's numerics conserve volume amounts to testing how well this property is satisfied when integrating downward from the ocean surface, starting from the vertical velocity given by Eq. (44). Adding freshwater to the model provides a nontrivial test of these properties. The present scheme produces zero vertical velocities at the ocean bottom on tracer cells, to within computer roundoff, regardless of the topography or surface forcing.

² The bottom velocity cell generally does not sit on the ocean bottom, and so can support a vertical velocity due to sloping topography. Details of how MOM handles this velocity are given in Webb (1995) as well as in Pacanowski and Griffies (2000).

8) ORDER OF THE SOLUTION METHOD

The method presented above solves for the updated “quasi-baroclinic” velocity $\hat{u}_k(\tau, \tau + \Delta\tau)$ prior to the updated quasi-barotropic velocity $\bar{u}(\tau, \tau + \Delta\tau)$. The prefix quasi is used since these velocities are defined with respect to the ocean depth at time τ instead of $\tau + \Delta\tau$. This approach is motivated largely because it allows for minimal fundamental changes to the MOM 3 structure, in which the baroclinic and tracer equations are solved prior to the barotropic equations. However, a more straightforward approach is to solve for the updated surface height $\eta(\tau + \Delta\tau)$ and barotropic velocity $\bar{u}(\tau + \Delta\tau)$, and then to update the baroclinic and tracer equations. This latter approach is being pursued in MOM 4.

4. Numerical examples

For the purpose of numerically illustrating the algorithm, we present selected results from four experiments. The first and second are from an idealized, flat-bottom, coarse-resolution, sector model driven by time-independent buoyancy and momentum forcing. Details of the configuration are provided in the caption to Fig. 3. The third example is that of a higher-resolution regional model with two basins connected by a channel with a shallow sill. This configuration is shown in detail later (Fig. 5). The final example is a coarse-resolution global model that quantifies the issues of local versus global conservation.

a. Free surface compared to rigid lid

We start by providing a direct comparison between the explicit free surface and the rigid-lid streamfunction method in the sector model. Each experiment used the same tracer time step of 1 day and the same baroclinic time step of 1 h. The rigid-lid experiment used a 1 h time step for the streamfunction, whereas the free surface used 300 s for the barotropic time step. With the time averaging used with the free surface method, there are a total of 24 barotropic time steps integrated for each baroclinic time step. Both the free surface and rigid-lid methods used a virtual salt flux and zero freshwater forcing.

The experiments were run for 4000 tracer years, at which point both reached a quasi-equilibrium. When run on a single processor on GFDL’s Cray T90, the free surface method took a few percent longer than the rigid-lid method. Although this comparison is a function of model configuration and computer details, as well as elliptic solver algorithm, the key point is that such a comparison is impressive for the free surface model, since the rigid-lid model is ideally suited for such a simple configuration. That is, the flat-bottom, time-independent forcing, and absence of mesoscale eddies means that the number of elliptic solver scans can be

quite low, often requiring only a single scan to update the streamfunction.

More realistic bottom topography, surface forcing, and mesoscale eddies will greatly increase the number of elliptic solver scans required by the rigid-lid model. In practice, more realistic situations often prompt one to greatly reduce or loosen the criteria used for determining a solution to the elliptic problem. Indeed, in some cases it is difficult to argue that a “solution” has even been found. In contrast, the ratio of barotropic to baroclinic wave speeds is independent of model resolution and surface forcing. Rather, it is determined by the bottom topography and stratification. Consequently, for the explicit free surface method, there is no issue regarding convergence to a solution, as convergence is guaranteed by construction.

As with KSWP and Dukowicz and Smith (1994), we check numerical integrity of the free surface solution by comparing with the more highly tested rigid-lid results. For example, Fig. 3 provides a sample of the free surface solution; the rigid-lid results are nearly identical. Since the vertical velocity at the ocean surface vanishes at equilibrium in a constant forced, coarse-resolution free surface model without freshwater forcing, the two methods should indeed provide nearly identical answers. A detailed comparison (not shown) reveals negligible differences for all aspects of the simulations. Other model configurations also have been run, with similar comparisons.

We conclude from these tests that the explicit free surface method maintains the convenience of the rigid lid for purposes of idealized coarse-resolution modeling. The solutions are likewise nearly the same when run under the same forcing. Both of these conclusions have remained valid with more realistic climate model experiments now routinely run at GFDL.

b. Goldsbrough–Stommel circulation

As discussed by Huang (1993), the Goldsbrough–Stommel circulation arises from freshwater forcing, and this circulation is absent in traditional rigid-lid models that use the virtual salt flux. Figure 4 illustrates this circulation when run with the free surface for 4000 yr. The freshwater forcing, as used by Huang, is given by

$$q_w = \frac{q_w^0}{\cos\phi} \left(1 - 2 \frac{\phi - \phi_S}{\phi_N - \phi_S} \right), \quad (45)$$

where $q_w^0 = 1 \text{ m yr}^{-1}$ and ϕ_N and ϕ_S are the northern and southern latitudinal boundaries, respectively. The domain integral of this water flux vanishes. There is no wind forcing nor temperature forcing, and temperature is uniform. The barotropic circulation is quite weak and is comparable to that obtained by Huang. The baroclinic circulation, as documented through the overturning streamfunction, is strong and saline direct, yet somewhat weaker than that of Huang, perhaps because of our use

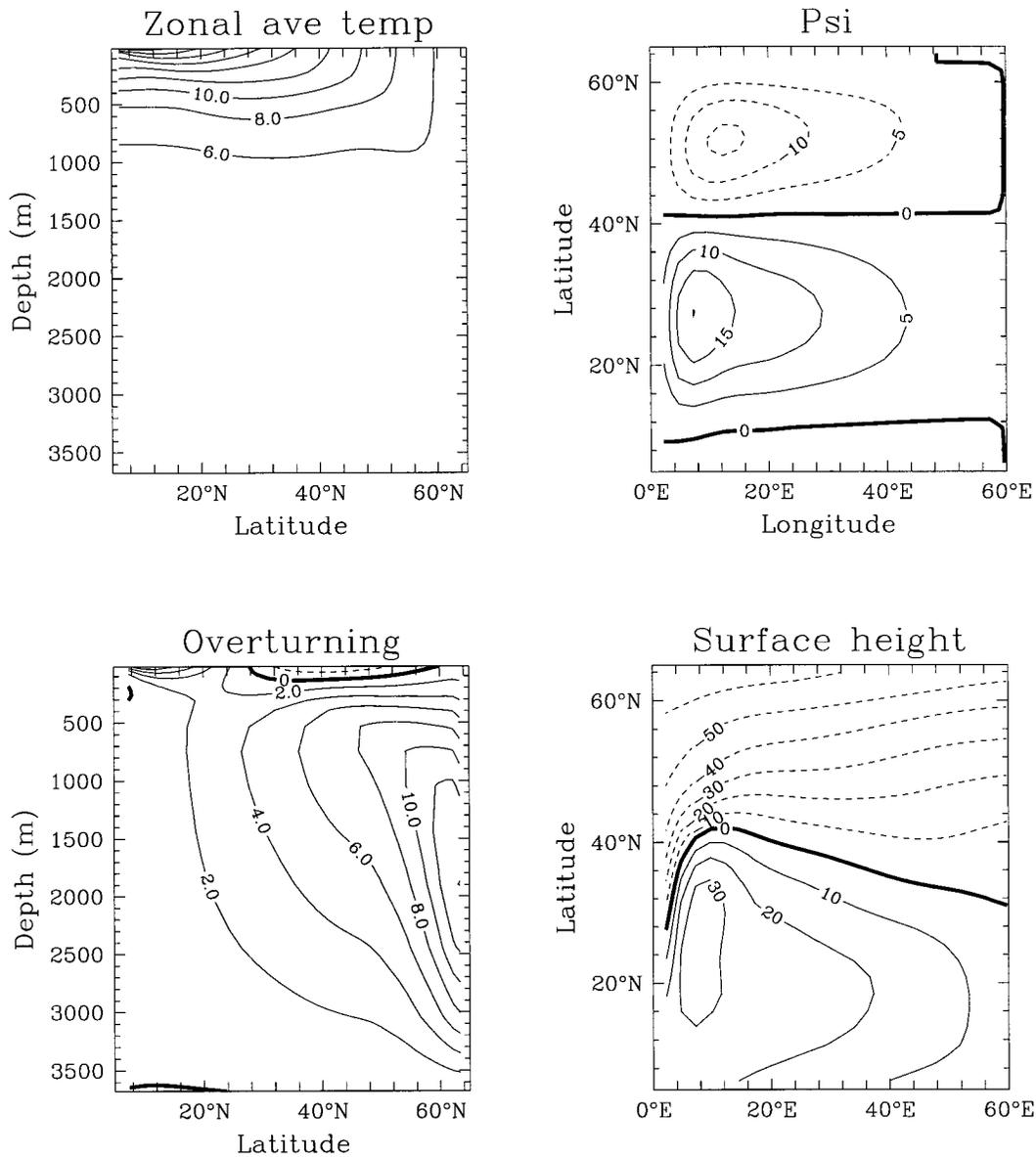


FIG. 3. The sector model configuration employed here consists of a flat-bottom $60^\circ \times 60^\circ$ sector from 5° to 65° N. The grid is locally square, so that $\Delta\phi = \Delta\lambda \cos\phi$, where $\Delta\lambda = 2.5^\circ$ is the longitudinal resolution. There are 18 unevenly spaced grid points in the vertical. The vertical tracer diffusivity is $0.5 \text{ cm}^2 \text{ s}^{-1}$, the vertical momentum viscosity is $10 \text{ cm}^2 \text{ s}^{-1}$, the horizontal tracer diffusivity is $10^7 \text{ cm}^2 \text{ s}^{-1}$, and horizontal momentum viscosity is $2 \times 10^9 \text{ cm}^2 \text{ s}^{-1}$. The Quick scheme of Leonard (1979), as detailed in Holland et al. (1998), is used for tracer advection. Temperature in the surface level of 35-m resting thickness is restored to the linear temperature profile given by Cox and Bryan (1984). A 35-day timescale is used, which delivers roughly a $47 \text{ W m}^{-2} (\text{C}^\circ)^{-1}$ flux sensitivity. Salinity is uniform. Winds are zonal as given by Bryan (1987). The rigid-lid model used a conjugate gradient solver. Shown here are results from the free surface model after 4000 tracer years; the rigid-lid results are virtually identical. (upper left) Zonally averaged temperature. (lower left) Overturning streamfunction ($S_v = 10^6 \text{ m}^3 \text{ s}^{-1}$). (upper right) Barotropic streamfunction (S_v). (lower right) Free surface height (cm). Note that the surface height diagnosed from the rigid-lid solution is quite close to that shown here from the free surface method.

of half the vertical diffusivity. It remains unclear how important this circulation is for realistic climate simulations, since it is often easily masked in the presence of wind and heat forcing. Regardless, its presence is a natural consequence of the use of freshwater forcing with the free surface method.

When integrating the model to reach the solution in Fig. 4, we used 1-day tracer and 1-h baroclinic velocity time steps as in the previous experiment. Although neither salt nor water were added to the model, the total salt does not remain precisely constant during the integration. The reason, as discussed in section 2, is that

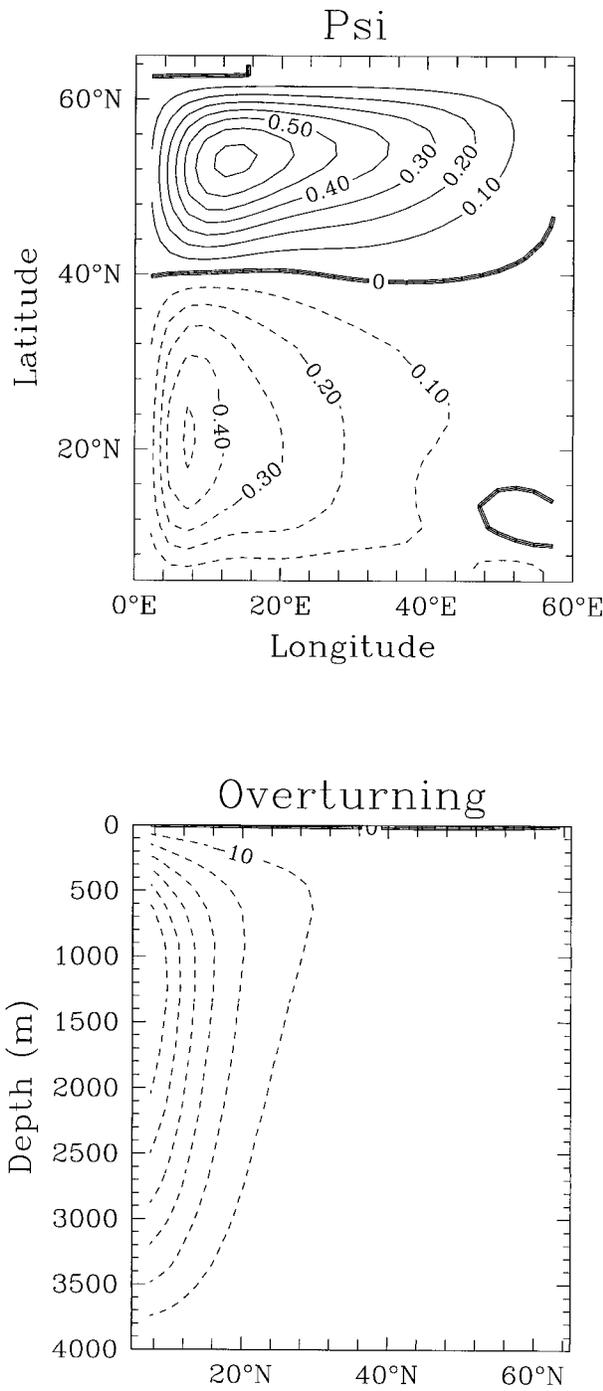


FIG. 4. Illustrating the Goldsbrough–Stommel circulation. (top) The barotropic streamfunction (Sv). (bottom) Overturning streamfunction, with each contour representing 10 Sv. These results should be compared to Figs. 6a and 7a in Huang (1993).

the tracer and baroclinic momentum time steps were unequal. The consequences on the equilibrium solution reached at 4000 yr have not been assessed. Doing so requires rerunning the experiment with equal tracer and baroclinic time steps.

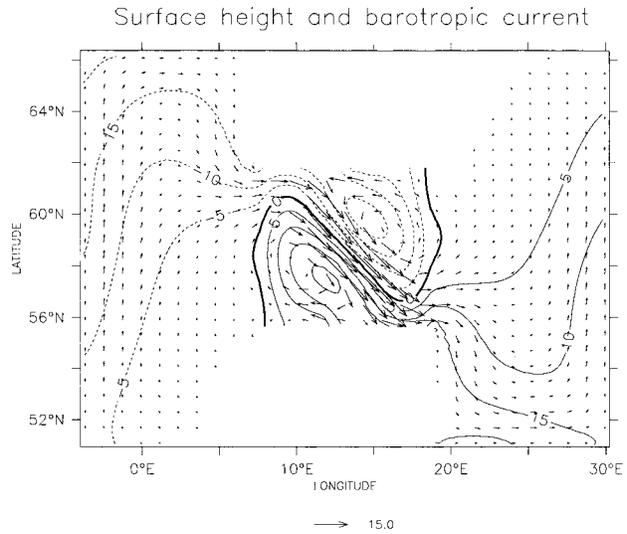


FIG. 5. The model configuration consists of a domain from roughly 51° to 66.5°N, 35° in long, and 100-m depth. The resolution is 0.3° long, 0.15° lat, with uniform 4-m-thick boxes. A sill is placed in the middle of the domain, running from south to north and reaching to within 20 m of the surface. Partial bottom cells of Pacanowski and Gnanadesikan (1998) are used to represent the sill. Tracer and baroclinic momentum time steps are 240 s, and the barotropic time step is 60 s. The vertical tracer diffusivity is 1.0 cm² s⁻¹, the vertical momentum viscosity is 10 cm² s⁻¹, the horizontal tracer diffusivity is 10⁷ cm² s⁻¹, and the horizontal momentum viscosity is 5 × 10⁸ cm² s⁻¹. The model is initialized from rest with zero sea level elevation and with a uniform linear temperature profile from 10°C at the surface and 5°C at the bottom. Salinity varies vertically from 33 to 35 psu west of the sill, and from 10 to 12 psu east of the sill. Shown here is the surface height and barotropic velocity fields (only every third velocity point is shown) after 1 day of adjustment from the initial conditions. Note the strong dipole sitting on top of the sill, and the predominantly northward barotropic currents next to the outer walls as remnants of barotropic Kelvin waves. A vector representing a current speed of 15 cm s⁻¹ is indicated at the bottom of the figure.

c. Marsigli problem

The purpose of this section is to directly compare the new free surface method to that of KSWP. For this purpose, we consider an unforced baroclinic adjustment problem motivated from the system first discussed for the Bosphorus by Marsigli in 1681 (see Gill 1982, p. 96).

The experimental details are given in the caption to Fig. 5. A strong baroclinic pressure gradient arising from a salinity front rapidly piles a sea level difference of about 40 cm between the two basins. After about two model days, barotropic and average baroclinic pressure gradients balance each other, and the sea level difference becomes roughly time independent. A weak cross-channel geostrophic circulation persists that prevents a salinity exchange between both basins. Although the setup appears artificial, a similar balance can be observed in the Baltic Sea during periods of weak wind forcing. Because the sea level changes rapidly and there is a strong interbasin salinity gradient, this adjustment problem provides a severe test for how well the model conserves total salt.

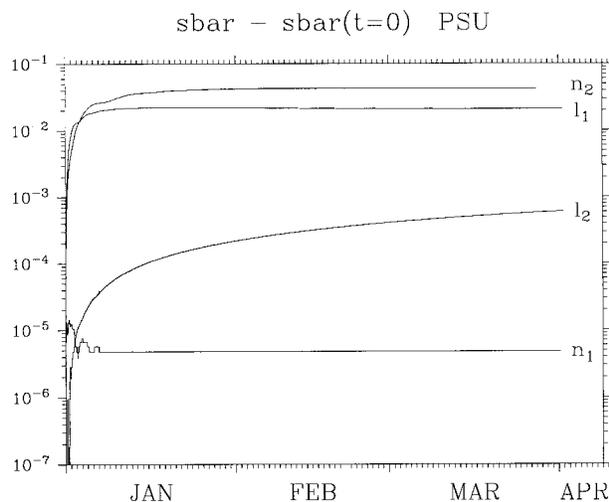


FIG. 6. Time series of the deviation of domain-averaged salinity from its initial condition, where each model time step is plotted (note the logarithmic vertical axis). All model versions start with the same initial conditions. The time series l_1 results from running a linearized free surface, with a constant cell thickness in the tracer and baroclinic velocity equations. The tracer and baroclinic time steps are both 240 s. The time series l_2 employs exactly the same linearized free surface scheme but the salt between the upper model boundary at $z = 0$ and the sea level $z = \eta$ is included in the domain tracer integral. The time series n_1 , which shows negligible deviation from zero, uses the nonlinear free surface discussed in this paper with tracer and baroclinic time steps both equal to 240 s. The time series n_2 uses the same nonlinear free surface, yet with a tracer time step of 720 s, and baroclinic time step of 240 s. All cases used 60 s for the barotropic time step.

Figure 5 shows the sea level and barotropic current after 1 day using the nonlinear free surface method. The current between the basins is deflected to the right by the Coriolis force. There is a strong dipole sitting on top of the sill, with a high to the southwest and a low to the northeast. The barotropic flow at the outer edges of the subbasins is directed northward in both basins, which indicates that this current is the result of barotropic Kelvin waves encircling the subbasins. Such waves, and the remaining currents, cannot be represented in a rigid-lid model.

In general, the solution for the sea level, velocity, temperature, and salinity differ only slightly between the KSWP and the new approach during the adjustment. The close agreement of these features provides an important positive assessment of the new method. In particular, it indicates that the time averaging approach used here does not overly damp the barotropic waves, at least as compared to the KSWP algorithm.

As there are no surface water fluxes, the total volume of the ocean domain should remain constant, as indeed it does to within computer roundoff (not shown). Likewise, total salt and domain-averaged salinity (total salt divided by total volume) should remain constant, where total salt is computed via Eq. (10). An accounting of the domain-averaged salinity is shown in Fig. 6. Shown

are four time series over a 3-month period that track how salinity deviates from its initial value.

The time series l_1 results from running the free surface of KSWP, with a constant cell thickness in the tracer and baroclinic velocity equations. The tracer and baroclinic time steps are both 240 s and the barotropic time step is 60 s. During the first two model days the sea level in the western basin is decreasing rapidly but the sea level in the eastern basin is increasing. Since the upper model boundary with the KSWP scheme is at $z = 0$, saline water is entering the model area in the western basin but brackish water leaves the model area through the level $z = 0$ in the eastern basin. Consequently, the domain-averaged salinity in the model area is increasing. As long as the sea surface salinity is uniform, the salt gain in the western basin and the salt loss in the eastern basin would be canceled exactly by a sea level variation in the opposite direction. However, such cancellation is not general, and so nonconservation of total salt is the norm.

Since the sea surface height is known, the salt between the upper model boundary at $z = 0$ and the sea level $z = \eta$ can be included in the diagnostic for domain tracer content. As an example, the time series l_2 employs exactly the same linearized free surface scheme as l_1 but uses this improved tracer diagnostic. Compared with l_1 , the conservation of salt is better; however, there is a clear trend in the domain-averaged salinity. This gain of salt stems from undulations of the sea level in correspondence with a varying sea surface salinity.

The time series n_1 , which shows minor deviation from zero, uses the new free surface discussed in this paper with tracer and baroclinic time steps both equal to 240 s, with a 60-s barotropic time step. The conservation of tracer is improved substantially when compared with the linearized free surface results. The deviation from zero is less than 10^{-5} psu.

The time series n_2 uses the same free surface, yet with unequal baroclinic and tracer time steps, with a tracer time step of 720 s, and baroclinic time step of 240 s. As discussed in section 2, tracer conservation is not ensured in this case, as is indeed shown by the roughly 0.04 psu salinity drift.

d. Quantifying local versus global conservation

The Marsigli example provided an illustration of the global conservation properties of the new algorithm when using a discretization that maintains local compatibility between the volume and tracer budgets. As mentioned in section 2e, there is an alternative approach that is available for maintaining more exact global conservation properties, yet at the cost of sacrificing the local compatibility. We present here an example that compares the two approaches and concludes that the scheme that performs superior for global conservation is preferable for climate modeling purposes.

We consider a coarse-resolution global model run

with two idealized conditions and using two different forms for the discretization of the $T\partial_t\eta$ term appearing in the tracer concentration budget (6). Discretization (a) uses $T(\tau)[\eta(\tau + \Delta\tau) - \bar{\eta}^*(\tau - \Delta\tau)]$, as used in the Marsigli example just presented. This form provides for local self-consistency between the tracer and volume budgets as discussed when deriving Eq. (38). However, as discussed in section 2d, it will not provide for exact global conservation. Discretization (b) drops the time average on the lagged surface height and instead uses $T(\tau)[\eta(\tau + \Delta\tau) - \eta(\tau - \Delta\tau)]$. This expression leads to exact global conservation in the case where there is no Robert filtering on the tracer, yet it is not locally self-consistent with a volume budget that uses a time-filtered surface height $\bar{\eta}^*(\tau - \Delta\tau)$ for stability.

The global model has roughly three degrees of horizontal resolution, uses 15 vertical levels down to 5000-m depth, and is run with a 30-min time step for the tracer and baroclinic momentum and 30-s barotropic time step. The first experiment initialized the ocean at rest with constant 25°C water and then added a climatological wind stress yet no heat fluxes. As presented in section 2c, the solution should maintain constant 25°C water everywhere. For an indefinite integration period, discretization a indeed remains precisely at 25°C. In contrast, discretization b shows water after 30 days with a temperature of 24.999 95°C. Although not exact, this is quite a small deviation even when extrapolated to integrations of centuries to millennia.

The second experiment initialized the temperature field with a zonally averaged version of the Levitus (1982) analysis. The surface flux consisted of the same climatological wind stress as the previous example as well as a uniform and constant 10 W m⁻² surface heat flux. We assess here whether the heat input through the ocean surface is equivalent to the heat absorbed by the ocean as deduced by measuring the ocean's total heat content via Eq. (10). We focus on a single model day since this is a typical time period over which winds and surface tracer flux are held fixed in coupled climate models.

After 1 day for discretization a, the heat input through the ocean surface minus the change in ocean heat content was equivalent to a spatially averaged error over the globe of -0.5 W m⁻². Note that over the first half of the day, the error was -1.6 W m⁻², which diminished to -0.37 W m⁻² after 2 days. These are nontrivial errors, upward of 10%–20% of the imposed heating from the atmosphere. In contrast, discretization b, along with Robert filtering on the tracer, reduced the error from -0.5 W m⁻² to a negligible -0.003 W m⁻². Removing the Robert filter brings the error to 10^{-9} W m⁻², which arises from computer roundoff.

Both solution methods showed negligible qualitative differences for the structure of the surface height over the 30-day integration period. However, given the small errors incurred by discretization b for both the first and second set of experiments, we conclude that it is a pref-

erable approach for longer integrations where conservation is key.

5. Computational aspects

The purpose of this section is to provide an overview of computational issues related to the use of a particular barotropic algorithm in a parallel computational environment. The papers by Dukowicz et al. (1993), Webb (1996), Webb et al. (1997), Marshall et al. (1997), and Guyon et al. (2000) are complementary to the following.

a. Fundamental considerations

The dominant feature of current supercomputing technology is the large and rapidly growing disparity between processor clock speeds and the memory bandwidth and latency. The rate at which processors are able to execute their instructions upon data far outstrips the ability of memory to deliver data to the processor. All current research into supercomputing architecture (see, e.g., Culler and Singh 1998) is directed at resolving this problem.

Traditional supercomputers, such as the Cray vector machines, are largely dependent on prohibitively expensive custom memory to deliver data at a rate sufficient to keep vector pipelines busy. The microprocessor-based computers currently being produced rely instead on commodity memory, but implement a deep memory hierarchy, where multiple levels of smaller and faster data caches are placed between the main memory and the processor. For performance comparable to vector systems, scalable cache-based microprocessor supercomputers achieve speedup through massive parallel partitioning of the problem, where the problem is distributed among many processors working concurrently and exchanging data on a network. This architectural model leads to measures of algorithm performance substantially different from the vector model. Notably, the following issues become central.

- 1) *Temporal locality*. Given the relatively high cost of memory access, algorithms that have a high rate of data reuse are preferable. The *computational intensity*, defined as the ratio of floating-point operations per memory request, is required to be as high as possible. The match between the size of the data subset allotted to a processor, and the size of its cache, is a key element in this aspect of performance.
- 2) *Spatial locality*. In the interests of minimizing communication among processors, algorithms that maintain spatial locality of data references are most useful.
- 3) *Communication profile*. The communication performance of the interprocessor hardware interconnect is measured in terms of its *latency* (startup time per message) and *bandwidth* (the inverse of the transmission rate). Communication latency can be a cause

TABLE 1. MOM 3 performance on a Cray T3E-900 for a coarse-resolution model. The model has one-dimensional domain decomposition along latitude rows, runs with the MOM 3 memory window, and uses Cray-SHMEM communication. Results here are from a Southern Hemisphere configuration using a 4° Mercator grid extending from the equator to 74° S with $90 \times 28 \times 40$ grid points. The model used a tracer time step of 43 200 s, baroclinic time step of 4320 s, and explicit free surface time step of 216 s. Model run times are given in s. Shown are the number of T3E computer processors (N), number of computed latitude rows per processor (j/N), time for the main model loop (main), time for the free surface portion of the model (fs), total model run time (total), scaling (scale), and the megaflops per second (Mflop/s). Scaling is defined to be the total run time taken for an experiment to complete on one processor, divided by the time taken for N processors as scaled by the number of processors: i.e., $\text{scale} = T_1/(NT_N)$.

N	j/N	Main	fs	Total	Scale (%)	Mflop/s
1	28	27.10	1.10	28.22	100	32.86
2	14	15.18	0.40	15.58	90	59.51
4	7	10.94	0.27	11.22	63	82.65
7	4	7.48	0.21	7.70	53	120.43
14	2	1.45	0.19	1.65	122	562.00

for concern on most commercially available machines today, barring a few exceptions such as the Cray T3E. The amount of data that needs to be transferred between processors is a function of the spatial locality of the algorithm and the problem size. Algorithms that can accomplish the maximum data transfer in the fewest possible messages are preferable.

b. Algorithms available in MOM 3

As mentioned in the introduction, there are three basic algorithms in MOM 3 for solving the barotropic mode: the rigid-lid streamfunction of Bryan (1969), with modifications to the rigid-lid surface pressure method of Smith et al. (1992) and Dukowicz et al. (1993); the implicit free surface method of Dukowicz and Smith (1994); and the explicit free surface method of KSWP as well as that discussed here.

The rigid-lid model, although parallelized (Redler et al. 1998), has not been the focus of recent algorithm research at GFDL, largely because of the physical issues previously raised in this paper as well as problems with global data dependencies related to island integrals [see Bryan (1969) as well as Smith et al. 1992 for discussion].

The implicit free surface method, in contrast, has a relatively long history of use on parallel machines. This algorithm solves the barotropic system implicitly in time, hence allowing for the barotropic mode time step to be lengthened to that of the baroclinic mode. As such, there is no subcycling of the barotropic mode as required with the explicit approach. In so doing, however, the algorithm requires the solution of an elliptic problem.

The elliptic solvers used in the implicit free surface method are variants of the basic iterative Jacobi solver, where methods are used to accelerate the rate of con-

TABLE 2. Same as Table 1 but for a more refined grid of 1° Mercator resolution with $360 \times 112 \times 40$ grid points, tracer time step of 4320 s, baroclinic time step of 4320 s, and explicit free surface time step of 60 s.

N	j/N	Main	fs	Total	Scale (%)	Mflop/s
1	112	1793.41	142.29	1935.72	100	39.77
2	56	1088.09	144.38	1232.48	98	77.96
4	28	545.62	78.79	624.43	97	153.87
7	16	316.80	50.85	367.42	94	261.51
8	14	277.57	45.07	322.66	94	297.78
14	8	162.72	31.53	194.27	89	494.53
16	7	143.55	27.96	172.34	88	557.52
28	4	86.04	22.99	109.05	79	881.10

vergence. For example, the conjugate gradient (CG) solver computes the optimal vector along which to converge. Computation of this vector involves a global sum across the entire domain; hence, the data dependencies are global, and involve global communication at each step in the solver iteration. It is observed in MOM that the number of iterations for the CG method is ~ 100 for a 1° model with static boundary conditions, and goes up with increasing resolution. Additionally, the iteration count is highly dependent on the model's representation of bottom topography and surface forcing, with more realistic topography and time-dependent surface forcing generally requiring much larger iteration counts. Parallel implementation of the CG requires communication at each iteration. Other, faster methods, with better parallel performance, such as multigrid methods, are known to exist, but also require increased steps as the resolution goes up.

The explicit free surface has improved data locality, since the computations involve only local derivatives and no elliptic problem. The ratio of the baroclinic to the barotropic time step is given by the ratio of the phase speeds of the barotropic gravity wave to the first baroclinic wave. This ratio is determined by ocean depth and stratification and, so, is independent of grid resolution. Hence, the method's computational performance is generally independent of grid resolution.

c. Examples of performance on scalable systems

MOM 3 is parallelized in one dimension, along latitude rows only. Scalability is thus measured in terms of latitude rows per processor. We show scaling results from a Cray T3E at two different resolutions using the explicit free surface algorithm documented here. Table 1 shows scaling results for a Southern Hemisphere configuration using a 4° Mercator grid. The superlinear scaling above seven rows per processing element (PE) is a result of arrays becoming small enough to fit in cache. Table 2 shows results for the same domain at 1° resolution.

The results for the 1° model are more appropriate for current and future needs of parallel ocean modeling. They indicate that one can increase the number of pro-

processors in MOM 3 until 8 rows/PE without significant loss of scalability. Note that with 4 rows/PE, scalability is 80%, whereas it has been found to be about 60% for MOM's version of the implicit free surface method. Furthermore, as noted previously, absolute times for the implicit free surface method are highly dependent on the conjugate gradient iteration count needed to achieve convergence to the subjectively chosen tolerance level.

6. Conclusions

Some key conclusions of this paper were noted in the introduction, with the central ones highlighted here.

By adapting the partial cell framework of Pacanowski and Gnanadesikan (1998) to the undulating top model grid cell, the explicit free surface scheme described here fully incorporates the effects of the surface height into the tracer and baroclinic momentum equations. Doing so allows for the surface boundary conditions to be formulated in a physically consistent and meaningful manner, to treat tracer and freshwater fluxes consistently and quasi-conservatively, and to maintain energetic consistency in which buoyancy effects in the density equation balance the work done by pressure from the momentum equation.

The explicit free surface method is stable and allows long tracer time steps. This result represents a key practical feature of the algorithm for use in coarse-resolution climate studies. Relatedly, when run on parallel computers, the absence of an elliptic problem, present in the rigid-lid and implicit free surface methods, enhances the processor scaling of the explicit free surface method.

Acknowledgments. We thank members of the Ocean, Climate, and Prediction Groups at GFDL for testing this scheme, and earlier versions, in various model configurations. In particular, we thank Jeff Anderson, Bob Hallberg, Matt Harrison, George Mellor, Thomas Neumann, Young-Gyu Park, Igor Polyakov, Tony Rosati, Torsten Seifert, Mike Spelman, Eli Tziperman, David Webb, Mike Winton, and Bruce Wyman for enjoyable conversations and useful suggestions. Bob Hallberg and Igor Polyakov deserve special thanks for extensive suggestions and useful critiques. Comments from the anonymous reviewers are also greatly appreciated. We thank Jerry Mahlman, the director of GFDL, for his support and encouragement.

REFERENCES

- Barnier, B., 1998: Forcing the oceans. *Ocean Modeling and Parameterization*, E. P. Chassignet and J. Verron, Eds., NATO Advanced Study Institute, Kluwer Academic, 45–80.
- Beron-Vera, F. J., J. Ochoa, and P. Ripa, 1999: A note on boundary conditions for salt and freshwater balances. *Ocean Modelling*, **1**, 111–118.
- Bleck, R., and L. T. Smith, 1990: A wind-driven isopycnic coordinate model of the north and equatorial Atlantic Ocean. 1. Model development and supporting experiments. *J. Geophys. Res.*, **95** (C3), 3273–3285.
- Blumberg, A. F., and G. L. Mellor, 1987: A description of a three-dimensional coastal ocean circulation model. *Three-Dimensional Coastal Ocean Models*, N. Heaps, Ed., Coastal and Estuarine Sciences, Vol. 4, Amer. Geophys. Union, 1–16.
- Bryan, F. O., 1987: Parameter sensitivity of primitive equation ocean general circulation models. *J. Phys. Oceanogr.*, **17**, 970–985.
- Bryan, K., 1969: A numerical method for the study of the circulation of the world ocean. *J. Comput. Phys.*, **4**, 347–376.
- , 1984: Accelerating the convergence to equilibrium of ocean-climate models. *J. Phys. Oceanogr.*, **14**, 666–673.
- , and M. D. Cox, 1972: An approximate equation of state for numerical models of the ocean circulation. *J. Phys. Oceanogr.*, **2**, 510–514.
- Cox, M. D., 1984: A primitive equation, 3-dimensional model of the ocean. GFDL Ocean Group Tech. Rep. 1, 143 pp.
- , and K. Bryan, 1984: A numerical model of the ventilated thermocline. *J. Phys. Oceanogr.*, **14**, 674–687.
- Culler, D. E., and J. P. Singh, 1998: *Parallel Computer Architecture: A Hardware/Software Approach*. Morgan Kaufmann, 1100 pp.
- Danabasoglu, G., J. C. McWilliams, and W. G. Large, 1996: Approach to equilibrium in accelerated global oceanic models. *J. Climate*, **9**, 1092–1110.
- Dewar, W. K., and R. X. Huang, 1996: On the forced flow of salty water in a loop. *Phys. Fluids*, **8**, 954–970.
- , Y. Hsueh, T. J. McDougall, and D. Yuan, 1998: Calculation of pressure in ocean simulations. *J. Phys. Oceanogr.*, **28**, 577–588.
- Dukowicz, J. K., and R. D. Smith, 1994: Implicit free-surface method for the Bryan–Cox–Semtner ocean model. *J. Geophys. Res.*, **99**, 7991–8014.
- , —, and R. C. Malone, 1993: A reformulation and implementation of the Bryan–Cox–Semtner ocean model on the connection machine. *J. Atmos. Oceanic Technol.*, **10**, 195–208.
- Gill, A. E., 1982: *Atmosphere–Ocean Dynamics*. Academic Press, 662 pp.
- Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood, 2000: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dyn.*, **16**, 147–168.
- Griffies, S. M., and R. W. Hallberg, 2000: Biharmonic friction with a Smagorinsky viscosity for use in large-scale eddy-permitting ocean models. *Mon. Wea. Rev.*, **128**, 2935–2946.
- Guyon, M., G. Madec, F. X. Roux, M. Imbard, C. Herbaut, and P. Fronier, 1999: Parallelization of the OPA ocean model. *Calculateurs Paralleles*, **11**, 499–517.
- Hallberg, R., 1995: Some aspects of the circulation in ocean basins with isopycnals intersecting the sloping boundaries. Ph.D. thesis, University of Washington, Seattle, WA, 244 pp.
- , 1997: Stable split time stepping schemes for large-scale ocean modeling. *J. Comput. Phys.*, **135**, 54–65.
- Haltiner, G. J., and R. T. Williams, 1980: *Numerical Prediction and Dynamic Meteorology*. John Wiley, 477 pp.
- Higdon, R. L., and A. F. Bennett, 1996: Stability analysis of operator splitting for large-scale ocean modeling. *J. Comput. Phys.*, **123**, 311–329.
- , and R. A. de Szoeke, 1997: Barotropic–baroclinic time splitting for ocean circulation modeling. *J. Comput. Phys.*, **135**, 30–53.
- Holland, W. R., J. C. Chow, and F. O. Bryan, 1998: Application of a third-order upwind scheme in the NCAR ocean model. *J. Climate*, **11**, 1487–1493.
- Huang, R. X., 1993: Real freshwater flux as a natural boundary condition for the salinity balance and thermohaline circulation forced by evaporation and precipitation. *J. Phys. Oceanogr.*, **23**, 2428–2446.
- Jackett, D. R., and T. J. McDougall, 1995: Minimal adjustment of hydrographic profiles to achieve static stability. *J. Atmos. Oceanic Technol.*, **12**, 381–389.
- Killworth, P. D., 1987: Topographic instabilities in level model OGCM's. *Ocean Modelling* (unpublished manuscript), **75**, 9–12.
- , and N. R. Edwards, 1999: A turbulent bottom boundary layer

- code for use in numerical models. *J. Phys. Oceanogr.*, **29**, 1221–1238.
- , J. M. Smith, and A. E. Gill, 1984: Speeding up ocean circulation models. *Ocean Modelling* (unpublished manuscript), **56**, 1–5.
- , D. Stainforth, D. J. Webb, and S. M. Paterson, 1991: The development of a free-surface Bryan–Cox–Semtner ocean model. *J. Phys. Oceanogr.*, **21**, 1333–1348.
- Large, W. G., and S. Pond, 1981: Open ocean flux measurements in moderate to strong winds. *J. Phys. Oceanogr.*, **11**, 324–336.
- Leonard, B. P., 1979: A stable and accurate convective modelling procedure based on quadratic upstream interpolation. *Comput. Methods Appl. Mech. Eng.*, **19**, 59–98.
- Levitus, S., 1982: *Climatological Atlas of the World Ocean*. NOAA Prof. Paper 13, U.S. Government Printing Office, Washington, DC, 173 pp.
- Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997: A finite-volume, incompressible Navier–Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.*, **102**, 5753–5766.
- Mellor, G. L., 1996: User's guide for a three-dimensional, primitive equation, numerical ocean model. Program in Atmospheric and Oceanic Studies, Princeton University, Princeton, NJ, 40 pp. [Available from Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ 08542.]
- Pacanowski, R. C., and A. Gnanadesikan, 1998: Transient response in a z -level ocean model that resolves topography with partial cells. *Mon. Wea. Rev.*, **126**, 3248–3270.
- , and S. M. Griffies, 2000: The MOM 3.1 manual. NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, 680 pp.
- , K. Dixon, and A. Rosati, 1991: The GFDL modular ocean model user guide. GFDL Ocean Group Tech. Rep. 2, Geophysical Fluid Dynamics Laboratory, Princeton, NJ, 16 pp.
- Redler, R., K. Ketelsen, J. Dengg, and C. W. Böning, 1998: A high-resolution numerical model for the circulation of the Atlantic Ocean. *Proceedings of the Fourth European CRAY-SGI MPP Workshop*, H. Lederer and F. Hertweck, Eds., Max-Planck-Institut für Plasmaphysik, 95–108.
- Roulet, G., and G. Magec, 2000: Salt conservation, free surface and varying volume. A new formulation for Ocean GCMs. *J. Geophys. Res.*, **105**, 23 927–23 947.
- Semtner, A. J., Jr., 1974: An oceanic general circulation model with bottom topography. Numerical Simulation of Weather and Climate, Tech. Rep. 9, Department of Meteorology, University of California, Los Angeles.
- Smith, R. D., J. K. Dukowicz, and R. C. Malone, 1992: Parallel ocean general circulation modeling. *Physica D*, **60**, 38–61.
- Webb, D. J., 1995: The vertical advection of momentum in Bryan–Cox–Semtner ocean general circulation models. *J. Phys. Oceanogr.*, **25**, 3186–3195.
- , 1996: An ocean model code for array processor computers. *Comput. Geophys.*, **22**, 569–578.
- , A. C. Coward, B. A. de Cuevas, and C. S. Gwilliam, 1997: A multiprocessor ocean general circulation model using message passing. *J. Atmos. Oceanic Technol.*, **14**, 175–183.
- , B. A. de Cuevas, and A. C. Coward, 1998: The first main run of the OCCAM global ocean model. Southampton Oceanography Centre Internal Doc. 34, 44 pp.
- Wolff, J.-O., E. Maier-Reimer, and S. Legutke, 1997: The Hamburg Ocean Primitive Equation Model HOPE. DKRZ Tech. Rep. 13, 98 pp.