YALE
PEABODY
MUSEUM

P.O. BOX 208118 | NEW HAVEN CT 06520-8118 USA | PEABODY.YALE. EDU

JOURNAL OF MARINE RESEARCH

The *Journal of Marine Research,* one of the oldest journals in American marine science, published important peer-reviewed original research on a broad array of topics in physical, biological, and chemical oceanography vital to the academic oceanographic community in the long and rich tradition of the Sears Foundation for Marine Research at Yale University.

An archive of all issues from 1937 to 2021 (Volume 1–79) are available through EliScholar, a digital platform for scholarly publishing provided by Yale University Library at https://elischolar.library.yale.edu/.

Requests for permission to clear rights for use of this content should be directed to the authors, their estates, or other representatives. The *Journal of Marine Research* has no contact information beyond the affiliations listed in the published articles. We ask that you provide attribution to the *Journal of Marine Research*.

Yale University provides access to these materials for educational and research purposes only. Copyright or other proprietary rights to content contained in this document may be held by individuals or entities other than, or in addition to, Yale University. You are solely responsible for determining the ownership of the copyright, and for obtaining permission for your intended use. Yale University makes no warranty that your distribution, reproduction, or other use of these materials will not infringe the rights of third parties.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
https://creativecommons.org/licenses/by-nc-sa/4.0/

THE SEA: THE SCIENCE OF OCEAN PREDICTION
Journal of Marine Research, 75, 683–725, 2017

Numerical modelling in a multiscale ocean

by Dale B. Haidvogel[1,2], Enrique N. Curchitser[1], Sergey Danilov[3], and Baylor Fox-Kemper[4]

ABSTRACT

Systematic improvement in ocean modelling and prediction systems over the past several decades has resulted from several concurrent factors. The first of these has been a sustained increase in computational power, as summarized in Moore's Law, without which much of this recent progress would not have been possible. Despite the limits imposed by existing computer *hardware*, however, significant accruals in system performance over the years have been achieved through novel innovations in system *software*, specifically the equations used to represent the temporal evolution of the oceanic state as well as the numerical solution procedures employed to solve them.

Here, we review several recent approaches to system design that extend our capability to deal accurately with the multiple time and space scales characteristic of oceanic motion. The first two are methods designed to allow flexible and affordable enhancement in spatial resolution within targeted regions, relying on either a set of nested structured grids or, alternatively, a single unstructured grid. Finally, spatial discretization of the continuous equations necessarily omits finer, subgrid-scale processes whose effects on the resolved scales of motion cannot be neglected. We conclude with a discussion of the possibility of introducing subgrid-scale parameterizations to reflect the influences of unresolved processes.

Keywords: Ocean modelling, nested grids, finite element and volume methods, subgrid-scale parameterization

1. Introduction

The equations of motion for the ocean cannot be solved exactly except under highly idealized conditions. This is so for several reasons. First, the equations governing the dynamics of the open ocean in their exact form are nonlinear partial differential equations (PDEs) for which analytic, closed-form solutions are unobtainable. Even were this not the case, many if not all of the environmental fields needed to specify a given problem—bathymetry, coastline geometry, surface forcing, initial and boundary conditions, etc.—are nonanalytic fields determined from observations; these in turn are uncertain to within observational error

1. Rutgers University, New Brunswick, New Jersey 08901, USA
2. Corresponding author: ***e-mail: dale@marine.rutgers.edu***
3. Alfred Wegener Institute, 27570 Bremerhaven, Germany
4. Brown University, Providence, Rhode Island 02912, USA

© 2017 Dale B. Haidvogel, Enrique N. Curchitser, Sergey Danilov, and Baylor Fox-Kemper.

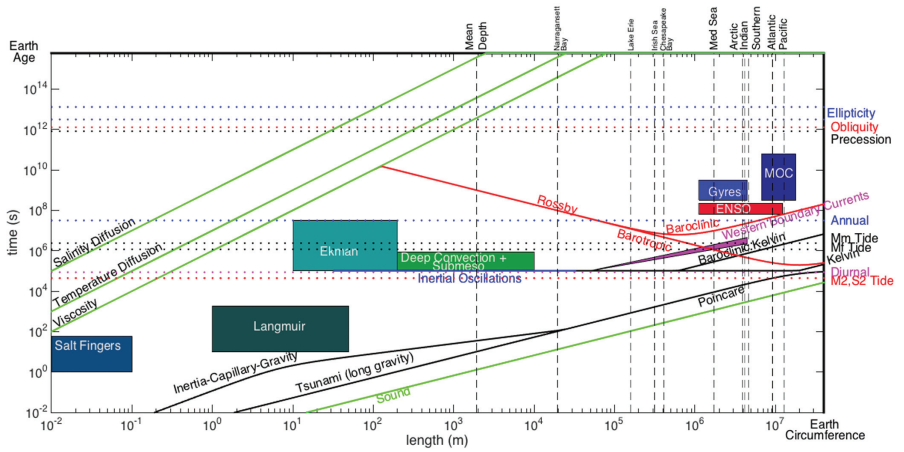

Figure 1. Space-time scales of some oceanic processes. The latitude has been set to 25° N. The characteristic length scale of water bodies is taken to be sqrt(area).

and often nonuniformly distributed in space and time. As a consequence of these restrictions, solutions to the equations of motion, unless greatly simplified, must be determined numerically via approximate solution procedures.

Obtaining predictions of future states in the ocean therefore raises several related, and often competing, issues. These include the following: first, how to discretize the continuous equations of motion and their component variables (and what potential sources of error may be introduced thereby). Of particular recent interest, given the broad range of scales relevant to ocean prediction (e.g., global, regional, coastal, estuarine, etc.), are multiscale discretization techniques that allow timely and affordable predictions. Second, because discretization necessarily omits some smaller-scale processes, the inclusion of the effects of subgrid-scale influences via appropriate parameterizations must be considered.

In what follows, we first discuss some general properties and consequences of the numerical solution of the oceanic equations of motion. We next briefly review two methodologies used to extend the range of spatial scales covered by such solutions, namely the use of nested structured grids and unstructured grids, respectively. Finally, we close with a discussion of the "scale-aware" parameterization of subgrid-scale processes.

2. A brief introduction to ocean numerical modelling: Issues, techniques, and consequences

Ocean processes cover an enormous range of space and time scales; some examples are shown schematically in Figure 1. Notwithstanding this complexity, the equations governing the time evolution of oceanic processes are for the most part well understood. The primitive equations that form the basis for many ocean circulation prediction systems are derived

from the conservation laws for mass, momentum (Newton's Law, F = ma), heat (the First Law of Thermodynamics), and tracers (e.g., salinity), supplemented with an equation of state relating the density of the ocean to in situ properties (temperature, salinity, pressure). The primary assumptions made along the way are the Boussinesq and hydrostatic approximations. The derivation and implications of the primitive equations are reviewed in the companion contribution by Jacobs and Fox-Kemper (2017).

In their "simplest" form, the primitive equation set incorporates seven nonlinear PDEs describing the space/time evolution of the seven primary oceanic variables: velocity components in three space dimensions (u, v, w), temperature and salinity, density, and pressure. In addition to conservation of mass, momentum, and heat upon which they are based, the primitive equation set has many additional desirable properties, among them conservation of higher-order quantities such as energy, vorticity, and enstrophy (vorticity squared).

Although we will not dwell overly long on the (very important) issue of computational efficiency, it is to be noted that the hydrostatic primitive equations encompass processes of both hyperbolic (wave) and parabolic (diffusion) types, but avoid the elliptic character of the equation set in the absence of the hydrostatic approximation. This yields great computational savings in that solutions of elliptic equations are not required. However, in exchange for this economy, the solution of the primitive equations requires some "regularization" (smoothing) if convergence under grid refinement is to be guaranteed (e.g., Vitousek and Fringer 2011).

The essence of determining approximate solutions to the equations of motion is the representation of each of the four-dimensional dependent variables—velocity for instance—as a discrete function described by a finite set of numbers, rather than as a continuous function of space and time, i.e., u(x, y, z, t). As is clear, this representation requires specific methodologies for discretization of the four independent variables. After discretization in space and time, the resulting approximate difference equations (ADEs), supplemented with an appropriate set of initial and boundary conditions, are solved to obtain solutions representing future ocean states.

The primitive equations are essentially "exact" for processes for which the hydrostatic and Boussinesq approximations are appropriate (i.e., the upper-right quadrant of Fig. 1) as well as conservative of the properties noted previously. However, solutions to the ADE are clearly approximate only, and several questions therefore arise. These include the following: in what way(s) do errors arise, how big are they, and how may they be minimized and/or avoided; do solutions to the ADE converge to the exact solution as discretization intervals in space/time are reduced; and can the conservation of properties displayed in the original exact equations be maintained in the ADE despite inevitable error?

We discuss these questions, by example, next. To do so, we will use a convenient, reduced form of the equations of motion, namely the linearized shallow water equations that apply to a single, shallow layer of constant density. Treatment of ocean stratification will be discussed below. In the absence of density variation, and with the inclusion of a conserved tracer field, the reduced equation set is

$$\frac{\partial u}{\partial t} + U\frac{\partial u}{\partial x} - fv = -g\frac{\partial \eta}{\partial x} + K\nabla^2 u$$
$$\frac{\partial v}{\partial t} + U\frac{\partial v}{\partial x} + fu = -g\frac{\partial \eta}{\partial y} + K\nabla^2 v$$
$$\frac{\partial \eta}{\partial t} + H\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0$$
$$\frac{\partial T}{\partial t} + U\frac{\partial T}{\partial x} = K\nabla^2 T$$

where the dependent variables are as follows: (u,v), the horizontal velocity components in the x and y coordinate directions, respectively; η, the elevation of the sea surface above its resting position; and T, a tracer field. The parameters appearing in the equations, assumed constant, include the following: U, an advecting zonal velocity; f, the Coriolis parameter; g, the acceleration of gravity; H, the resting layer depth; and K, a horizontal mixing coefficient.

a. Spatial differencing: Taylor series and a simple geostrophic example

Spatial discretization of the dependent variables and spatial operators in the equations of motion can be approached in several ways. The methods most commonly employed in ocean modelling are the finite difference (FD), finite volume (FV), and finite element (FE) methods.[5] We will use the first of these techniques here to illustrate some general principles and results. The implementation of the latter two methods is also discussed.

The FD method assumes that each dependent variable, say u(x,y,z,t), is defined by its values at a finite number of spatial points (x_j, y_k, z_l) where j, k, and l are integer subscripts. Typically, these "grid points" form a structured grid and are uniformly spaced, although some degree of nonuniform spatial resolution can be accommodated with appropriate coordinate transformations (Thompson, Warsi, and Mastin 1982).

For simplicity, consider the function u(x,t) in one space dimension, and suppose that we seek approximate forms for the partial derivatives ∂_x, ∂_x^2, etc., along with an estimate of the error incurred by the approximation. Consider a set of equally spaced grid points ($x_j = j\Delta x$) and an associated set of values $u(x_j) \equiv u_j$. FD approximations to differential equations can be obtained in several ways. The most often used is the Taylor series method. The Taylor series expansions for the behavior of a smooth function about a central point x_j are

$$u(x_j + \Delta x) = u(x_j) + (\Delta x)u' + \left(\frac{\Delta x^2}{2}\right)u'' + \left(\frac{\Delta x^3}{3!}\right)u''' + O(\Delta x^4)$$
$$u(x_j - \Delta x) = u(x_j) - (\Delta x)u' + \left(\frac{\Delta x^2}{2}\right)u'' - \left(\frac{\Delta x^3}{3!}\right)u''' + O(\Delta x^4)$$

5. Other spatial representation methods based upon higher-order expansions in continuous Fourier or polynomial basis functions, although enjoying a lengthy history in atmospheric prediction, have seen more limited use in ocean modelling. See, e.g., Wunsch et al. (1997) and Choi et al. (2004).

where the prime notation designates a partial derivative (i.e., $u' \equiv \partial_x u$). By combining these expressions, a variety of FD approximations may be derived, including

$$u'(x_j) = \frac{u(x_j + \Delta x) - u(x)}{\Delta x} + O(\Delta x)$$

$$u'(x_j) = \frac{u(x_j + \Delta x) - u(x_j - \Delta x)}{2\Delta x} + O(\Delta x^2)$$

and

$$u''(x_j) = \frac{u(x_j + \Delta x) - 2u(x_j) + u(x_j - \Delta x)}{\Delta x^2} + O(\Delta x^2).$$

Here $O(\Delta x)$ and $O(\Delta x^2)$ represent residual (error) terms proportional to Δx and Δx^2, respectively, and indicate the rate at which the error made by the approximation may be reduced as the grid spacing Δx is refined. An approximation whose leading-order error term is proportional to Δx is referred to as "first-order" in space; for an error term proportional to Δx^2 the approximation is said to be of "second-order." Higher-order approximations may be obtained by using additional gridpoint values (e.g., $u(x_j \pm 2\Delta x)$).

As the first two expressions illustrate, FD approximations can be constructed in alternate ways, each having their own distinct error properties. Note that centered FD approximations (those using gridpoint values set equally to either side of the central point) are generally more accurate than one-sided approximations (those using sets of values asymmetrically distributed to one side or the other).

As a first example, consider geostrophic currents in the ocean. At lowest order in a Rossby number expansion, the open ocean is in quasi-steady hydrostatic and geostrophic balance. From the first of our prior equations, north-south geostrophic flow near the ocean surface is given by $fv = g\frac{\partial \eta}{\partial x}$ (e.g., Stewart 2008). Let us suppose that the sea surface perturbation is oscillatory in space $\eta(x) = \eta_0 e^{ikx}$, where η_0 is the amplitude and k is the wavenumber (2π/wavelength) of the sea surface disturbance. From this, the north-south component of geostrophic flow is found to be exactly $v = (\eta_0 g/f)ike^{ikx}$.

It is instructive to ask what geostrophic flow we would obtain if we were to adopt a centered, second-order approximation to the term $\frac{\partial \eta}{\partial x}$ on an FD grid of spacing Δx. In this case, we assume

$$\frac{\partial \eta}{\partial x} = \eta_0 \left(e^{ik(x+\Delta x)} - e^{ik(x-\Delta x)}\right)/2\Delta x$$

from which, after substitution and a bit of manipulation, we obtain

$$v = (\eta_0 g/f)ike^{ikx}(\sin(k\Delta x)/k\Delta x).$$

Our FD approximation matches the exact result except for an additional factor of $(\sin(k\Delta x)/k\Delta x)$. The error in utilizing our centered, second-order approximation is therefore determined by the departure of this additional multiplicative factor from a value of unity. Note that the FD result may be written $v = (\eta_0 g/f)ik'e^{ikx}$ where the modified

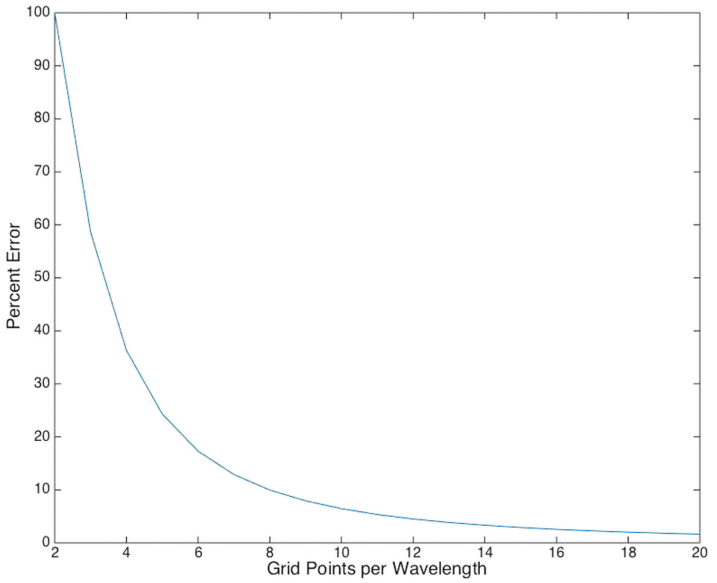

Figure 2. Percent error in the value of surface geostrophic current when a second-order finite difference (FD) approximation is used to evaluate the surface pressure gradient on a uniform grid.

wavenumber $k' = k\sin(k\Delta x)/k\Delta x$ (e.g., Moin 2010). Because k' is real, the central differencing operator is seen in this case to produce no phase error, as it would if k' had an imaginary component.

On a uniform FD grid, the admissible wavenumbers k range from approximately 0 (very fine resolution relative to the wavelength of η) to a value of (π/Δx). The latter limit applies because the finest perturbation resolvable on a grid of spacing Δx has a wavelength of 2Δx. As (2π/kΔx) is equal to the number of grid points per wavelength of η, we can compute the percent error in our approximation as a function of grid point resolution. The result is shown in Figure 2.

Several features in Figure 2 are of interest. Suppose that we wanted to guarantee an error less than 5%. For this discretization and the assumed form of the sea surface perturbation, we would require roughly 12 grid points per wavelength of the significant sea surface undulations. At reduced resolution, say 6 points per wavelength, we are already up to roughly 20% error. Finally, for the finest resolvable sea surface perturbation, the 2Δx wave, the error is 100%; there is no geostrophic flow at all at this scale despite the fact that the 2Δx wave has, in the continuum, the most rapid spatial variation.

b. Horizontal discretization in two dimensions: Propagating waves on staggered grids

Thus far we have assumed that all variables are available on a common set of horizontal grid points. Such an arrangement is not necessary however. An alternative is the use of

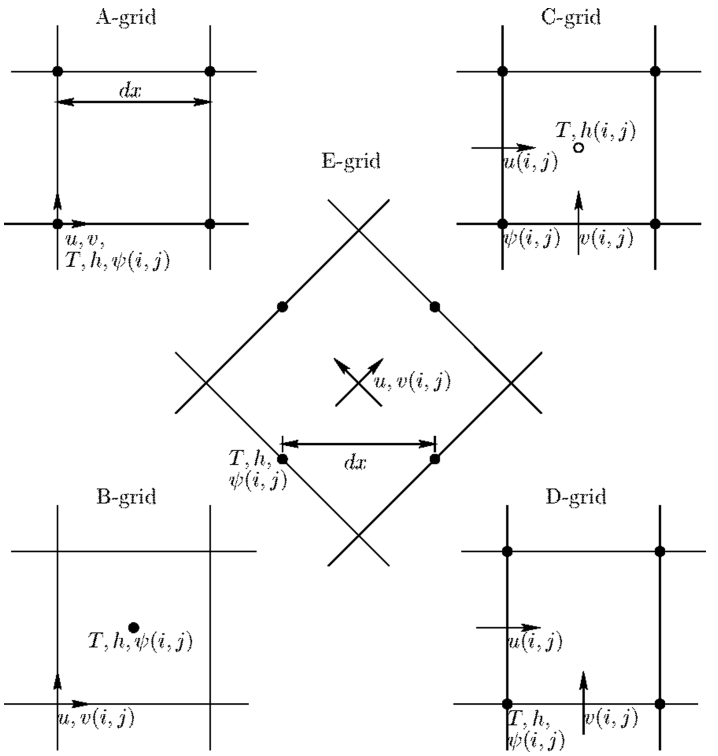

Figure 3. The arrangement of variables on the Arakawa grids.

horizontally staggered grids in which the dependent variables are offset from each other in various ways. Five examples of horizontal staggering, the so-called Arakawa "A," "B," "C," "D," and "E" grids (Arakawa 1966), are shown in Figure 3. Historically, the most widely used arrangements of variables for FD primitive equation ocean models have been the B and C grids; however, the majority of models in widespread current use have adopted the C grid. The arrangement of variables on the C grid places the discrete values of the sea surface height and tracer fields at the center of each grid cell, and the values of the u and v velocity components on the cell edges in the sense of normally directed flow (Fig. 3).

Consider the impact on processes resulting from the discretization of their governing equations on a C grid. We first return briefly to the previous geostrophic example. Note that the arrangement of variables on the C grid facilitates the discretization of the surface pressure gradient terms in the equations for geostrophic balance. Because of the positioning of the values of η, the sea surface gradient terms needed at u and v grid points can both be obtained from central differences across a single Δx or Δy. (On our nonstaggered example previously, the central difference was obtained across 2Δx). For second-order differencing, an error reduction of a factor of four is realized on the staggered C grid.

A broad space-time spectrum of propagating waves exists within the ocean and can be associated with gravitational, planetary, and/or topographic restoring forces (e.g., Pedlosky 2003). Two classes of wave motions that are intimately involved in basin-scale adjustment are inertia-gravity waves, which mediate gravitational adjustment, and planetary or Rossby waves, which are the primary basin-wide agents of geostrophic adjustment. Systematic errors in the representation of these wave processes have consequences for the manner in which numerical ocean circulation models respond to time-varying forcing.

As a second example, consider our shallow water equation set under the assumption of no large-scale advection (U = 0) or diffusion (K = 0). Defining the horizontal divergence $\delta = \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)$, the equations for (u, v, η) can be combined to form the hyperbolic equation for the propagation of inertia-gravity waves

$$\frac{\partial^2 \delta}{\partial t^2} + (f^2 - gH\nabla^2)\delta = 0.$$

The insertion of a trial wave solution proportional to $e^{i(kx+ly-\omega t)}$ into this equation yields the exact dispersion relation for inertia-gravity waves:

$$\left(\frac{\omega}{f}\right)^2 = 1 + R_d^2(k^2 + l^2)$$

where $R_d = \sqrt{gH}/f$ is the Rossby deformation radius.

As we have noted, on a C grid the discretized form of the surface pressure gradient terms are easily obtained by a central difference across a single grid interval (Δx or Δy). However, the discretized forms of the Coriolis terms (requiring fv at a u point and fu at a v location) each require a four-way average of v and u values, respectively. This is in contrast to a nonstaggered grid or the Arakawa B grid, neither of which would require spatial averaging to specify a centered estimate of the Coriolis terms.

For a propagating, wave-like solution of the form $e^{i(kx+ly-\omega t)}$, the effect of *spatial differencing* on the estimates of the surface pressure gradient terms follows from the previous geostrophic example, yielding multiplicative terms proportional to $\sin\left(\frac{k\Delta x}{2}\right)$ and $\sin\left(\frac{l\Delta y}{2}\right)$ in the respective momentum equations. In contrast, the effects of the *spatial averaging* on the Coriolis terms can be shown to result in factors of $\cos\left(\frac{k\Delta x}{2}\right)$ and $\cos\left(\frac{l\Delta y}{2}\right)$ for averaging in x and y.

Taking a uniform grid (Δx = Δy) for convenience, and for the moment assuming exact time differencing, the corresponding form for the dispersion relation on a C grid becomes

$$\left(\frac{\omega}{f}\right)^2 = \cos^2\left(\frac{k\Delta x}{2}\right)\cos^2\left(\frac{l\Delta x}{2}\right) + 4\left(\frac{R_d}{\Delta x}\right)^2\left[\sin^2\left(\frac{k\Delta x}{2}\right) + \sin^2\left(\frac{l\Delta x}{2}\right)\right].$$

The departure of the approximate dispersion relation (ADR) from its exact form is found to be dependent on the two nondimensional parameters $\left(\frac{k\Delta x}{2}\right)$ and $\left(\frac{l\Delta x}{2}\right)$; these are (π times) the ratio of the grid spacing to the wavelength of the wave in the x and y directions. Recalling

that k and l are, respectively, 2π divided by the x and y wavelengths of the wave, and that the finest resolvable wavelength is 2Δx, these "error parameters" range in value from near zero (the limit of very fine resolution of the wave) to $\left(\frac{\pi}{2}\right)$. From this it is clear that the ADR approaches the exact form in the limit (Δx$\rightarrow$0).[6]

The maximum error incurred on the C grid occurs for the most poorly resolved wave $\left(k = l = \frac{\pi}{\Delta x}\right)$, for which

$$\left(\frac{\omega}{f}\right)^2 = 0 + \left[\left(\frac{4}{\pi^2}\right) R_d^2 (k^2 + l^2)\right]_{k=l=\pi/\Delta x}$$

where the "0" has been included to emphasize that the contribution to the ADR from the Coriolis parameter has disappeared entirely. In addition, the contribution to the ADR from the gravity waves has been reduced by a factor of π^2. In other words, the phase speed $\left(c_{ph} = \frac{\omega}{k}\right)$ of the gravity waves has been reduced by a factor of π. Note that, in the continuum, pure gravity waves are nondispersive; that is, their phase speed is constant, independent of wavenumber. However, when produced on a C grid, they are dispersive.

In the preceding semi-discrete example (discrete in space, exact in time) the bi-directional wave (hyperbolic) character of the solutions is maintained, albeit with errors in the phase speed dependent on wavenumber. This fortuitous outcome is a consequence of the symmetric differencing and averaging operations defined on the C grid, which do not change the character of the resulting ADE. Further discussion of the influence of staggered grids on the propagation of geophysical waves can be found in Wajsowicz (1986) and Dukowicz (1995).

c. Time marching of the wave equation: Mix and match

So far, so good. However, numerical solution of the equations of motion requires that they be discretized in time as well as in space. That is, approximate forms for the time derivative must be specified that relate the future values of the dependent variables to their past values. As with approximations in space, there are many approaches to time differencing. A thorough discussion of time differencing is beyond the scope here. Instead, we will use several simple time-marching schemes to support our general remarks. The reader is directed to more comprehensive discussions (e.g., Durran 1999).

Consider the generic tracer equation $\frac{\partial T}{\partial t} = F(T)$ where F is a linear function describing the time-evolution of T, e.g., the advection, diffusion, or combined advection–diffusion equations. As in space, we divide the continuous time interval into discrete increments, or time steps, Δt. Then integration of the tracer equation in time yields a relationship between past and future values of the prognostic variable, i.e., a time-marching scheme. Time integration may in principle be performed over any number of past time levels. In

6. Note that $\sin(k\Delta x/2) \rightarrow (k\Delta x/2)$ and $\cos(k\Delta x/2) \rightarrow 1$ as $\Delta x \rightarrow 0$.

practice, integration from t_0 to $t_0 + \Delta t$ (a "one-step" scheme) or from $t_0 - \Delta t$ to $t_0 + \Delta t$ (a "two-step" scheme) are typically employed.

Several simple one- and two-step time-marching schemes can be written in the following form: $T^{(n+1)} = \alpha T^{(n)} + \beta T^{(n-1)} + \Delta t \{\gamma F^{(n+1)} + \kappa F^{(n)} + \varepsilon F^{(n-1)}\}$. Here, α–ε are constants, and the superscript notation denotes the time step, i.e., $T^{(n)} \equiv T(x, n\Delta t)$. It has been assumed that values of T are available at a minimum of two prior time levels. Simple time-marching schemes that fit this description include the following: Euler forward ($\alpha = 1$, $\kappa = 1$; $\beta = \gamma = \varepsilon = 0$), Euler backward ($\alpha = 1, \gamma = 1$), trapezoidal $\left(\alpha = 1, \gamma = \kappa = \frac{1}{2}\right)$, leapfrog ($\beta = 1, \kappa = 2$), and Adams–Bashforth ($\alpha = 1, \kappa = 3/2, \varepsilon = -1/2$).

From our discussion of differencing in space, it may be anticipated that time differences that are "centered"—that is, with the estimate of F centered within the interval spanned by the time steps of T—will be more accurate than those that are not. This is indeed the case. Of our five simple time-marching schemes, the two Euler methods are of first order in time. The remaining three, including the well-known leapfrog scheme, are centered in this sense and consequently of second order. Note that higher-order accuracy can be obtained with either a one- or two-step scheme.

In general, two types of error may arise in solutions to the ADE. The first, dispersive error, we have already encountered. The second is *amplitude error*, in which the magnitude of the approximate solution differs from its exact counterpart. A particularly unfortunate situation arises when the time-marching scheme allows the approximate solution to grow in amplitude without limit, rather than behave stably as in the previous propagating wave example wherein the amplitude of the wave is preserved in time. This is termed "numerical instability" and caution is always needed to avoid an approximate solution that "blows up."

In previous reference to the semi-discrete, bi-directional wave equation, we have noted that it is helpful for the ADE to preserve the character—in this instance, hyperbolic—of the exact equation set. As we will see, this is not the case for all time-marching schemes. However, here is one that does so:

$$\left(\frac{\eta^{(n+1)} - \eta^{(n)}}{\Delta t}\right) = -H\left(\frac{\partial u^{(n)}}{\partial x} + \frac{\partial v^{(n)}}{\partial y}\right)$$

$$\left(\frac{u^{(n+1)} - u^{(n)}}{\Delta t}\right) = f\left(\frac{v^{(n+1)} + v^{(n)}}{2}\right) - g\left(\frac{\partial \eta}{\partial x}\right)^{(n+1)}$$

$$\left(\frac{v^{(n+1)} - v^{(n)}}{\Delta t}\right) = -f\left(\frac{u^{(n+1)} + n^{(n)}}{2}\right) - g\left(\frac{\partial \eta}{\partial y}\right)^{(n+1)}$$

wherein a mixture of Euler forward, Euler backward, and trapezoidal treatments are all used in combination. This treatment seems strange at first sight; however, manipulation of the ADE for (u, v, η) yields the equivalent ADE for horizontal divergence, namely

$$\left(\frac{\delta^{(n+1)} - 2\delta^{(n)} + \delta^{(n-1)}}{\Delta t^2}\right) = -f^2\left(\frac{\delta^{(n+1)} + 2\delta^{(n)} + \delta^{(n-1)}}{4}\right) + gH\nabla^2\delta^{(n)}.$$

Here, we have left the spatial derivatives in their continuous form for convenience only; in the fully discrete ADE they would need to be approximated using symmetric averages and differences as in the C grid examples shown previously.

The result of our time-marching treatments in this case is a second-order (in time and space) bi-directional wave equation in which there arises both a centered time-differencing term on the left-hand side of the equation and a centered time-averaging term on the right. Insertion of our trial wave equation, as shown previously, shows that the former yields a multiplicative factor of $\left[-4\sin^2\left(\frac{\omega\Delta t}{2}\right)/\Delta t^2\right]$ and the latter a factor of $\cos^2\left(\frac{\omega\Delta t}{2}\right)$. It is easily verified that the solution to the ADE approaches that of the exact equation as the time step is refined. A more complete analysis of the propagation of inertia–gravity waves for forward–backward time-differencing on staggered grids is offered by Beckers and Deleersnijder (1993).

d. Time-marching of the advection equation: Constraints on Δt

In the preceding example, a well-chosen combination of space and time discretizations reproduces the essential character of the exact equation set. In particular, the solutions to the ADE are propagating waves (whose phase speeds may, however, differ from the continuum result) but whose amplitudes are preserved. An alternate way to state this is to observe that the ADR has solutions for ω that are real numbers. This favorable property is a consequence of a clever choice of time marching in which implicit time weighting of the right hand-side terms (i.e., the Euler backward and trapezoidal treatments) are employed.

Unfortunately, in a general setting such forward-in-time treatments may require considerable extra effort in the solution of the resulting coupled ADE. For this reason, explicit-in-time marching schemes are generally preferred. In return for the computational simplicity and efficiency that are thus gained, however, numerical instability (unbounded growth of the approximate solutions) may be encountered.

It is therefore necessary to guarantee that the solution to our ADEs converge to the true solution under space/time grid refinement. Two properties of an approximation are related to its convergence. These are consistency of the ADE with the original PDE and the numerical stability of its solution. The Lax–Richtmyer equivalence theorem prescribes the relationship between convergence, consistency, and stability. The theorem states for linear constant-coefficient PDEs that consistency plus stability together guarantee convergence.

The Von Neumann method for establishing the stability of a difference approximation is the most frequently used and readily applied stability analysis method, though it is not directly applicable to nonlinear equations. In it, we test the stability of a single spatial harmonic of the approximated equation. Stability of all admissible harmonics then becomes the necessary condition for stability of the overall scheme.

The procedure is as follows: assume a separation of space/time variables can be made such that $u^{(n)} = \lambda^n e^{ikx}$, where k is the wavenumber (2π/wavelength) of the trial solution and λ is a (possibly complex) factor specifying the phase and amplitude change in the solution from

one time step to the next. Substitute the trial solution into the difference equation, and solve for λ. Then, requiring $|\lambda| \leq 1$—that is, precluding systematic amplitude growth—ensures stability.

As a next example, set $\beta = \varepsilon = 0$ in the simple time-marching schemes previously shown. The resulting one-step scheme uses time levels n and (n + 1) only. Then, for the advection equation $\frac{\partial T}{\partial t} + U\frac{\partial T}{\partial x} = 0$, and for exact spatial differencing, the Von Neumann method can be applied to show that

$$\lambda = \left\{ \frac{1 - i\kappa\varpi}{1 + i\gamma\varpi} \right\} = \left(\frac{1 - \kappa\gamma\varpi^2 - i\varpi}{1 + \gamma^2\varpi^2} \right)$$

where $\varpi = \mathrm{Uk}\Delta\mathrm{t}$ (note that $\kappa + \gamma = 1$ for consistency with the original equation). Taking the modulus of λ for the various combinations of κ and γ shows the Euler forward scheme to be unstable ($|\lambda| = (1 + \varpi^2)^{1/2}$), the Euler backward scheme to be stable but damping ($|\lambda| = (1 + \varpi^2)^{-1/2}$), and the trapezoidal scheme to be amplitude-preserving ($|\lambda| = 1$). Of these schemes, the time-implicit trapezoidal is best (i.e., unconditionally stable and also of higher order in time) although possibly more costly to implement in a more general setting.

The result of this stability analysis may be anticipated by inspection of the leading-order error terms in the ADE. Note that, from the preceding Taylor series,

$$\frac{T^{(n+1)} - T^{(n)}}{\Delta t} = \left(\frac{\partial T}{\partial t} + \frac{\Delta t}{2}\frac{\partial^2 T}{\partial t^2} + \ldots \right)^{(n)}$$

$$\frac{\partial}{\partial x}T^{(n+1)} = \frac{\partial}{\partial x}\left(T + \Delta t\frac{\partial T}{\partial t} + \frac{\Delta t^2}{2}\frac{\partial^2 T}{\partial t^2} + \ldots \right)^{(n)}.$$

Combining these leading-order terms with the exact advection equation leads us to the following for the Euler-backward method applied at time step "n":

$$\frac{\partial T}{\partial t} + U\frac{\partial T}{\partial x} = \left(\frac{U^2\Delta t}{2} \right)\frac{\partial^2 T}{\partial x^2}$$

where the higher-order terms have been dropped. The ADE with Euler-backward marching thus resembles the original advection equation but with the addition of a spurious diffusion term; after discretization, the modified equivalent PDE is no longer a simple advection equation, but rather has a combined advective–diffusive character. Note that in this case the erroneous diffusive coefficient $\left(\frac{U^2\Delta t}{2}\right)$ is positive; hence, its influence is to damp the amplitude of the solution, anticipating the results of the Von Neumann stability analysis.

If we had selected the Euler-forward method, the diffusive term would again enter but with a negative sign, corresponding to spurious antidiffusion (growth) of the solution. This is again consistent with the conclusion drawn previously that Euler forward is unstable in this setting. Finally, a similar derivation for the leading-order influence of trapezoidal marching confirms that the error is dispersive, not diffusive.

For the two-step schemes (leapfrog and Adams–Bashforth), the identical Von Neumann analysis can be performed. In these cases, the result is a quadratic equation for λ, which may be solved in the usual fashion for the two associated values of λ. For leapfrog ($\beta = 1$, $\kappa = 2$), the two roots are

$$\lambda_2 = i\varpi + (1 - \varpi^2)^{1/2}.$$
$$\lambda_2 = i\varpi - (1 - \varpi^2)^{1/2}.$$

The two-step leapfrog approximation to the wave equation has two possible solutions, and both are conditionally stable, as discussed. But by taking the limit as $\Delta t \to 0$, we note that the former corresponds to the true solution ($\lambda_1 \to 1$) whereas the latter is unphysical ($\lambda_2 \to -1$). These are the physical and computational modes, respectively. The latter can be problematic if it grows too large relative to the physical mode, and hybrid schemes have been devised to keep the computational mode in check (most simply, the occasional use of a trapezoidal correction step).

Finally, if we wish to assure stability of the physical mode ($|\lambda_1| \leq 1$), then ϖ^2 must be less than or equal to 1; this in turn requires $Uk\Delta t \leq 1$. The most severe restriction on Δt occurs for the largest permissible value of the wavenumber k, that is, the finest possible resolved spatial scale. The finest resolvable wavelength is $2\Delta x$; hence, k_{max} is $(2\pi/2\Delta x)$ or $(\pi/\Delta x)$ and $(U\Delta t/\Delta x) \leq (1/\pi)$. More generally, were we to discretize in space as well as time using, say, centered differences, stability would require $(U\Delta t/\Delta x) \leq 1$. This is the well-known "Courant–Friedrichs–Lewy" (CFL) condition for the second-order FD approximation to the advection equation. Finally, note that $Uk\Delta t = (k\Delta x)\left(\frac{U\Delta t}{\Delta x}\right)$, the product of the ratio of grid spacing to wavelength and the Courant number. The latter is the ratio of the time step to the shortest advective time scale ($\Delta x/U$).

The central lesson conveyed in the CFL condition is that, generally speaking, time steps are constrained by the rate at which "information" passes across the spatial grid. For the wave equation, and the second-order methods employed here, the requirement is $\Delta t \leq (\Delta x/U)$. With this lesson in mind for the diffusion equation $\frac{\partial T}{\partial t} = K\frac{\partial^2 T}{\partial x^2}$, simple dimensional analysis suggests that the time step should be limited by a value proportional to $(\Delta x^2/K)$. A detailed stability analysis proves this to be the case.

e. Phase errors and the creation of false extrema

As our simple examples from the advection and bi-directional wave equations demonstrate, solutions obtained by approximating the time and/or space derivatives will generally not preserve the nondispersive character of the true solution for simple gravity waves. Instead, the approximate solution will, in general, display a phase speed error that depends on the wavenumber k (and l, if in two horizontal dimensions). This result holds even for otherwise attractive and robust time-marching schemes such as the trapezoidal method.

The dependence of phase speed error on wavenumber can have significant consequences. Consider uniform advection of a tracer field with an initial distribution that is localized in space. The localized distribution can be thought of as being made up of an overlapping sum of waves of different wavenumbers, adding up in just the right way and moving together in phase. When advanced in the continuum according to the 1d advection equation, the initial distribution is carried along at speed U without change of shape. However, an approximate solution will suffer gradual de-phasing of the component wavenumber contributions with a variety of unpleasant consequences, including the potential production of false extrema (e.g., negative values) in the tracer field. A variety of approaches that attempt to minimize the consequences of dispersive errors have been explored (e.g., Zalesak 1979; Thurburn 1990; Friedrich 1998).

f. Discretization in the "vertical"

A significant choice in the character of numerical ocean models has been the choice of vertical discretization. In the vertical direction, the ocean has two impermeable boundaries: the free surface, which is time- and space-dependent, and the bottom topography, which is usually assumed to be static except when studying near-shore sediment transport processes. The depth of the ocean floor has a significant range from the near-shore shelves at a few meters depth (even vanishing if storm surge and/or tidal wetting/drying are considered) to the deep ocean, reaching over 4,000 meters. Features such as canyons, trenches, sills, and seamounts are known to impose dynamical signatures in the ocean circulation and, therefore, must be properly represented. Hence, the variable ocean depth, the time-evolving free surface, and the space/time dependent internal density structure are all considerations in the treatment of the vertical coordinate.

Historically, ocean models were originally developed tailored to a specific vertical coordinate—e.g., geopotential, terrain-following, or isopycnal—chosen with the intended application in mind. Consider a general vertical coordinate s(i,j,k,t). Then these historically popular formulations are given by $s = z$ (geopotential), $s = z/(H + \eta)$ (terrain-following or "sigma"), and $s = \rho$ (isopycnal).

More recently in ocean models, the discretization of the vertical direction has been posed as an arbitrary Lagrangian–Eulerian (ALE) coordinate, as originally described by Hirt, Amsden, and Cook (1974) generally and by Bleck (2002) for the ocean. ALE can be configured to be purely Lagrangian (with no transport across the coordinate lines and a grid that moves exactly with the flow), Eulerian (wherein the fluid properties are regridded to a fixed coordinate), or anything in between (White, Adcroft, and Hallberg 2009). A general implementation consists of a Lagrangian step followed by a regridding of the time-dependent vertical coordinate and an Eulerian step used to compute the vertical advection relative to the moving vertical coordinate. ALE can be configured to be any of the traditional methods used in ocean models: geopotential, terrain-following, or isopycnal. Recent examples in ocean modelling are given in Chassignet et al. (2006) and Leclair and Madec (2011).

g. Exact conservation: Yes and no

A brief comment on conservation is in order. The exact equations are predicated on the conservation of properties such as mass, momentum, and heat and imply the conservation of higher-order quantities such as energy and vorticity. If the ADEs are derived via consistent differencing from the exact equations expressed in flux form, then exact conservation (to within machine round-off error) can generally be guaranteed for the lowest-order conserved properties.

Note however that the corresponding ADE for higher-order properties (e.g., kinetic energy [KE]) may not conserve such properties exactly. For illustration, consider the derivation of the conservation statement for KE from the 1d advection equation

$$u\left(\frac{\partial u}{\partial t}+U\frac{\partial u}{\partial x}\right)=\frac{\partial}{\partial t}\left(\frac{u^2}{2}\right)+U\frac{\partial}{\partial x}\left(\frac{u^2}{2}\right)=0$$

where the straightforward outcome in the continuum is an equivalent expression for advection of KE ($\frac{u^2}{2}$ in this 1d example). A similar manipulation of the ADE will depart from strict conservation via the appearance of space and time differencing error terms. These can be reduced in size by appropriate choice of space and time steps and differencing schemes. In general, however, departures from strict conservation of higher-order invariants will remain unless more complex, specialized differencing treatments are employed.

h. Initial conditions, boundary conditions, and forcing functions

Solution of the ADE requires the specification of initial conditions for all prognostic variables[7] as well as boundary conditions at surface, bottom, sidewalls, and on any open boundary segments. At solid boundaries, the condition of vanishing normal flow must be imposed; beyond that, however, boundary conditions become less clear-cut. In principle, the rate and manner by which fluxes of momentum, heat, and tracers are carried across all bounding surfaces must be specified. In some cases, such fluxes may be assumed to vanish (e.g., no heat flux through an assumed insulating bottom boundary). More generally, however, specification of boundary mixing and fluxes becomes a question of parameterization (e.g., mixing and exchange at the ocean surface). Finally, on open boundaries the exchange of properties must also be specified (requiring a different form of parameterization). These topics are discussed further below.

Note a common feature of the need to provide initial boundary and forcing information: in any realistic application, all three require access to observational datasets of various types (atmospheric, oceanic, hydrospheric, etc.) and the quality of such datasets will determine to a large extent the success of the simulation. With the increasing use of data assimilation methodologies to improve simulation accuracy, the need for high-quality observational

7. Recall that initial conditions may be required at more than one prior time level, depending on the nature of the time marching algorithm.

products becomes particularly acute. Other contributions to this volume discuss some of these topics (e.g., Brink and Kirincich 2017; Lermusiaux 2017).

i. Beyond simple equations

The systems of equations applied in ocean prediction are of course much more complex than the simple examples used for illustration here. Importantly, they incorporate multiple space dimensions and a simultaneous mix of interacting processes on a wide range of space/time scales. Notwithstanding this additional complexity, and the availability of a wide variety of alternate spatial and temporal discretization methods, the general messages conveyed by these simple examples continue to apply: first, that insufficient spatial resolution can produce significant qualitative and quantitative error; second, that discretization may alter the fundamental character of the equation set; and lastly, that time steps are typically limited by the rates of information flow across the discrete grid.

The details will of course matter in each instance, and conclusions drawn in a simple setting may be overly simplistic or misleading in a more complex situation. For instance, in a multidimensional advection problem, the CFL restriction on Δt becomes more severe, though it remains a function of the advecting velocity components and the grid spacing in each of the coordinate directions (Δx, Δy, and/or Δz). Also, guidance obtained for a single process in isolation may need to be reconsidered when in combination with another, e.g., inertia–gravity waves (Beckers and Deleersnijder 1993) and advection plus diffusion (Beckers 1992).

The simultaneous occurrence of different processes with their own rate of information flow can present significant challenges to efficient numerical implementation. Consider the example of global ocean prediction in the presence of a free sea surface. In the deep ocean, the phase speed of surface gravity waves is particularly rapid ($c_{sgw} = (gH)^{1/2} \approx 200$ m/s). A straightforward time-explicit marching algorithm, like those mentioned previously, would require $\Delta t \leq (\Delta x/c_{sgw})$ for all equations, even for those variables not intimately related to gravity wave propagation. To avoid the associated computational penalty, a variety of time-marching schemes have been devised to treat different processes separately, each with its own unique Δt. Again, a full treatment is beyond our scope. A practical example, used in the Regional Ocean Modelling System (ROMS), is given in Shchepetkin and McWilliams (2005).

The computational requirements of high-resolution ocean modelling can be demanding and in many cases prohibitive. A simultaneous halving of the grid spacing in all 3 dimensions, taking into account the typical requirement to also halve the time step, increases computational cost by a factor of approximately $2^4 = 16$. This rapid cost increase is particularly problematic in global ocean modelling on long time scales. As a consequence, novel approaches that effectively enhance the space/time coverage of ocean models have become essential.

As previously suggested, system improvement may be obtained in several alternate ways, for instance, either by refining spatial resolution or by improving the convergence rate of the numerical approximation. Given their geometrical complexity, ocean models have tended to favor a combination of low-order methods and decreasing spatial resolution.[8] An increasingly valuable supplement in applications with a range of spatial scales of motion is to apply high resolution only in specific subregions of an otherwise less-well-resolved larger domain. Such regional refinement can be either static or adaptive in time; the former is more common, and we focus on that here. The move to a static, multiscale framework may be accomplished in several ways, as described next.

3. Multiscale modelling on structured grids

Representing ocean physics across a wide range of scales remains a significant challenge to climate-scale integrations. Although computer power continues to increase roughly following Moore's Law (Moore 1965), ocean models that simultaneously resolve submeso- and global-scales in a single configuration remain rare—nonexistent for climate timescales—and many decades away (Fig. 4).

One approach being applied to achieve high resolution in a limited spatial domain is the nesting of a high-resolution, limited-area grid within a lower-resolution, larger-scale numerical domain. The fundamental numerical consideration for dynamical nesting is the treatment of the boundaries between coarse (parent) and high-resolution (child) grids. In regional ocean models, the open boundary conditions are typically implemented with variants of the method of characteristics (e.g., Orlanski 1976). Marchesiello, McWilliams, and Shchepetkin (2001) distinguish passive (outflow) from active (inflow) boundaries with the use of Orlanski (1976)-type radiation conditions. For passive boundaries, information is extrapolated from the interior of the domain in such a way as to (approximately) minimize reflections, whereas for active boundaries the interior solution is nudged towards information contained in the exterior solution.

For purely passive boundaries, the radiation condition applied to a field variable, say u(x,y,t), takes the continuous form

$$\frac{\partial u}{\partial t} + c_x \frac{\partial u}{\partial x} + c_y \frac{\partial u}{\partial y} = 0$$

where c_x and c_y are estimates of the rate of phase speed propagation in the directions normal and tangential to the boundary, respectively. The boundary conditions along active boundaries can be obtained from observational data or a larger-scale model. Marchesiello, McWilliams, and Shchepetkin (2001) show that the quality of the solution depends highly

8. As a single example, in the class of finite element models discussed below, refinement of resolution can be of h-type (the size of cells is varied), of p-type (the polynomial order is varied) or of r-type (when the order of reconstruction is varied).

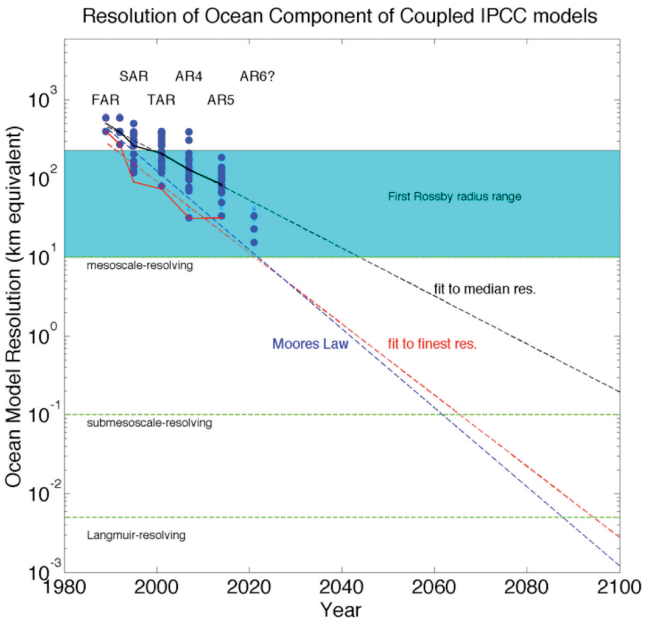

Figure 4. Estimate of the effective nominal horizontal resolution of ocean model components for primary baseline and climate change scenarios as reported in the Intergovernmental Panel on Climate Change (IPCC) reports by year of publication (Fox-Kemper et al. 2014).

on the accuracy of the computation of the normal (to the boundary) phase speed of the most significant wave modes.

For hindcast simulations—that is, those simulating past time periods—information can be downscaled from the coarse- to the fine-resolution region through an overlap in the domains themselves. Downscaling can work well, within the uncertainty of model parameters, when the forcing data are constrained by observations, such as in the reanalysis products (e.g., Curchitser et al. 2005; Hermann et al. 2009). The high-resolution nest can then explicitly resolve features missing from the large-scale model, and it is constrained by the large-scale circulation patterns via the lateral boundary conditions and/or internal nudging. However, if nesting is within a coupled climate model, such as when making a future projection, the forcing functions are not necessarily constrained by data and the coupled model can be expected to respond differently if provided with an alternative (high-resolution) nested ocean.

Multiscale nested coupling can be between two oceans or between a multiscale ocean and the atmosphere. One of the challenges, then, is to not only downscale information to the regional window, but also to understand how regional changes affect the global ocean (Biastoch, Böning, and Lutjeharms 2008) or global climate (Small et al. 2015), i.e., the effects of upscaling. Examples of upscaling include Chanut et al. (2008), who use Adaptive

Grid Refinement in FORTRAN (AGRIF) (Debreu, Vouland, and Blayo 2008) for a model of the Labrador Sea. The AGRIF framework permits the simultaneous integration of the parent and child grids. The child grid must be designed with a constant refinement factor from the parent grid and the coupling happens at every timestep. The AGRIF framework is sufficiently flexible to permit two-way coupling between the parent and child grids. Conservative interpolation enforces a continuity of fluxed variables.

Two-way nesting permits a more freely evolving system. This can be viewed as an advantage if the goal is to allow upscaling of information to the larger-scale circulation. However, with the flexibility also comes the potential to drift from a more realistic solution that might be achievable were observational data to constrain the solution on the parent grid (Döscher, Böning, and Herrmann 1994; Gerdes et al. 2001).

A more flexible framework that does not require the parent and child grids to be collocated, nor the parent and child models to be the same, is described by Curchitser et al. (2011) and Small et al. (2015). This new framework offers the advantage of optimizing grid and model design in specific regions. The framework has been implemented in the National Center for Atmospheric Research-Community Earth System Model (NCAR-CESM) using the Parallel Ocean Program (POP) and ROMS global and regional models, respectively (Small et al. 2015). In ocean-only hindcast mode, ROMS, which was designed as a coastal model, has shown improved skill in modelling boundary currents (e.g., Kang and Curchitser 2013, 2015). The embedding of such a regional model in a global configuration allows for isolation of specific processes affecting both the local climate representation and the potential for upscaling effects. The additional flexibility comes at the expense of boundary condition simplicity. Interpolated radiation boundary conditions are needed to pass information from the parent to the child grid. Additionally, overlap regions and sponge layers may be necessary.

Figure 5 shows an example of an embedded regional ocean model in the northwest Atlantic using the CESM multiscale ocean configuration. The global ocean is POP at 1°; the regional model is ROMS at 7-km resolution. The ocean models are two-way coupled. A merged ocean sea surface temperature (SST) is passed to the global atmosphere at each coupling time step, typically at a daily frequency (R. Dussin, pers. comm.). Though conspicuous features such as the Gulf Stream separation are not as skillful as in the pure ROMS hindcast of Kang and Curchitser (2013, 2015), there is an improvement over the coarse-resolution global model. This configuration is now being used to explore the effects of the Gulf Stream position and correction of some of the global model biases on downstream climate. Similar models in other boundary currents (e.g., Small et al. 2015) can be used to improve the representation of specific features in a climate model.

In spite of increasing computational resources, practical implementation of global high-resolution models will remain challenging for some time. This is especially true with Earth System Models that incorporate biogeochemistry. A careful implementation of a nesting strategy is useful to begin exploring the role of the coastal ocean in the climate system and the potential linkages between disparate scales of motion. The technique relies on

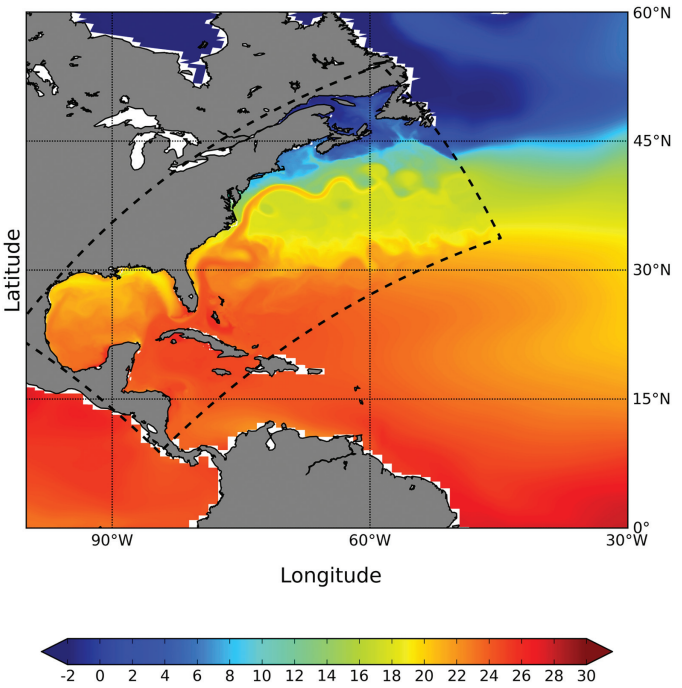

Figure 5. National Center for Atmospheric Research-Community Earth System Model (NCAR-CESM) multiscale configuration in the Northwest Atlantic. The global ocean is Parallel Ocean Program (POP) at 1°; the regional model is Regional Ocean Modelling System (ROMS) at 7-km resolution. The ocean models are two-way coupled. A merged ocean sea surface temperature (SST) is passed to the global atmosphere at each coupling time step (R. Dussin, pers. comm.).

careful and accurate remapping, which permits conservative interpolation of fluxes. Significant challenges remain in designing nesting strategies for model components other than momentum. Radiation conditions may not be appropriate for propagating signals arising from the dominant dynamics in sea ice and biogeochemistry. A final point to be noted is that although high-resolution nests may more accurately represent certain dynamics, such as boundary currents, in and of themselves they cannot correct significant biases that the parent model may contain.

4. Multiscale modelling on unstructured grids

Unstructured meshes are inherently suited to resolve dynamics encompassing a range of scales. By allowing the size of mesh elements to vary, they offer geometric flexibility that goes beyond the functionality allowed by nesting or generalized curvilinear meshes, the two techniques applied commonly to refine on structured meshes. We consider only

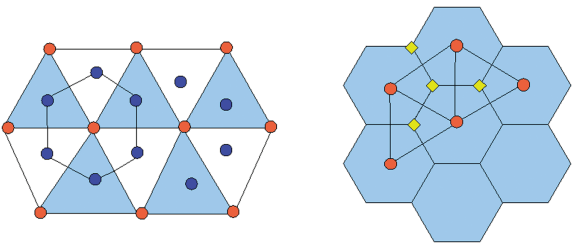

Figure 6. A regular patch of triangular mesh (left) and hexagonal mesh (right). Such meshes are dual to each other. In the triangular case, the degrees of freedom are most frequently located at vertices, centroids, or circumcenters or, in the case of C-grid discretization, at mid-edges. For the hexagonal mesh, it is commonly centers of cells or mid-edges in the case of C-grid discretization.

horizontally unstructured meshes, because the dominance of hydrostatic balance in the ocean demands vertical alignment. Compared with structured meshes, unstructured meshes enable smooth coastlines. Their horizontal density can be adjusted in places where the topographic slope is large, allowing finer topographic details to be better resolved. Many coastal models, designed to use unstructured meshes, benefit from the ability to scale the size of mesh elements with the square root of depth, which allows them to avoid the time step limitation with respect to the speed of surface gravity waves over the deep part of the ocean. Triangular meshes are most flexible and serve as a basis of most models, e.g., ADvanced CIRCulation Model (ADCIRC) (Westerink et al. 1992), Finite-Volume, primitive equation Community Ocean Model (FVCOM) (Chen, Liu, and Beardsley 2003), Semi-implicit Eulerian-Lagrangian Finite Element (SELFE) (Zhang and Baptista 2008), Stanford Unstructured Nonhydrostatic Terrain-following Adaptive Navier–Stokes Simulator (SUNTANS) (Fringer, Gerritsen, and Street 2006) and Finite Element Sea Ice-Ocean Model (FESOM) (Wang et al. 2014).

Figure 6 (left) shows a patch of such a mesh. Most commonly, the variables are located on mesh vertices, centroids of mesh elements (triangles), or mesh edges. However, higher-order representations are also possible when additional degrees of freedom are introduced inside triangles. Meshes dual to triangular, i.e., obtained by connecting circumcenters of triangles, can also be used, as is schematically shown in the left panel, leading to an arrangement shown in the right panel (see Ringler et al. 2013). We will refer to these as quasi-hexagonal, for hexagons will be met most frequently in this case. The circumcenters should lie inside their triangles, which excludes obtuse triangles. Meshes satisfying this property are called orthogonal. The dual mesh in this case presents the Voronoi tessellation. In principle, there is no limitation on the polygon type and generalized meshes combining different polygons, e.g., triangles and quads, are also possible. The question is rather their numerical stability, for mesh heterogeneity may contribute to locally increased errors. In the following, we will use triangular meshes as an example.

a. Examples of unstructured grid techniques

To illustrate the basic approaches for horizontally unstructured meshes we consider the 2D advection–diffusion equation for a tracer T,

$$\partial_t T + \nabla \cdot (\mathbf{u}T - K\nabla T) = 0 \tag{1}$$

with $\nabla = (\partial_x, \partial_y)$ and insulating lateral boundary conditions. Here $\mathbf{u}$ is the horizontal velocity vector and K the diffusivity. For simplicity, we will also consider the 1D version of this equation

$$\partial_t T + \partial_x (uT - K\partial_x T) = 0$$

when describing discretizations.

b. Finite volume methods

There are several possible ways to discretize the equations of motion on unstructured meshes. The first one relies on reformulating equations so that they express balances related to control volumes, hence the name FV method. In the simplest case when variables are located at the centers of triangles, the control volumes (also referred to as cells) are the mesh triangles proper. If they are at vertices, the so-called median–dual control volumes are routinely used. They are formed by connecting centroids to the centers of edges. The other option is to connect the circumcenters, but this is only possible if the circumcenters are inside their respective triangles. On a regular patch shown in Figure 6 these options coincide, but they differ on general meshes.

The FV method relies on the fact that the motion equations have the form of conservation laws, as equation (1). These equations are integrated over control volumes and their flux divergence term is expressed, via the Gauss theorem, in terms of fluxes out of the control volumes. Due to this strategy, local and global balances are ensured on the discretized level. We apply the FV method to equation (1) with cell-centered placement of variables. Indices c, v, and e will be used to designate the cells, vertices, and edges, respectively. Integrating over a triangle c one obtains

$$\partial_t \int T d\Omega_c + \sum_{e(c)} \mathbf{F}_e \cdot \mathbf{n}_e l_e = 0. \tag{2}$$

Here e(c) is the symbolic notation for the indices of the edges of triangle c, $\mathbf{n}_e$ is the outer normal to the edges, and l_e their length. The discrete tracer values are introduced as $T_c = \int T d\Omega_c / A_c$, where A_c is the triangle area. Note that no approximation is involved in deriving equation (2). The essence of the FV approach lies in estimating fluxes in terms of cell-averaged values T_c. The language of fluxes automatically ensures that total tracer is conserved. Assuming the velocity field to be given, fluxes can be estimated if the field T is known at the boundary of the control volume.

This can be done by designing a polynomial reconstruction valid in the cell and its vicinity $T_c(\boldsymbol{x}) = a_0 + a_1 x + a_2 y + \ldots$, where x and y are coordinates measured from the centroid and a_i the expansion coefficients, such that it satisfies the strong constraint $\int T_c d\Omega_c = T_c A_c$ and weak constraints minimizing $L = \Sigma_{n(c)} (\int T_n d\Omega_n - T_n A_n)^2$. Here $n(c)$ is the list of the neighboring triangles. The number of neighbors involved in the reconstruction depends on its order. Only the nearest neighbors (3 triangles sharing edges with c) are needed for linear least-squares reconstruction, and generalizations of this scheme are straightforward. The higher the order of the reconstruction the more accurate is the estimate of fluxes leaving/entering the control volume. However, for each edge the estimates coming from triangles sharing it are generally different. The flux entering equation (2) is therefore some combination of, for example, centered or upwind-biased estimates. On uniform meshes, the first choice will leave a dispersive error, whereas the second one will lead to a diffusive error. This behavior will be preserved if meshes vary smoothly. One can also reconstruct gradients of T on centered and upstream-biased stencils and use them to estimate T at edges, or combine field reconstruction on several stencils to obtain a weighted essentially non-oscillatory (WENO) scheme. The details here depend on the variable placement. For example, for vertex placement of variables the language of gradient reconstruction turns out to be more convenient because gradients are easily estimated on triangles.

Obtaining a monotonic scheme requires limiters or flux-corrected transport algorithms. Because directional splitting is not possible in the horizontal plane, these algorithms prove to be more computationally expensive than their structured-grid counterparts. Note lastly that reconstruction of velocity fields on unstructured grids must be carried out carefully to avoid unintended inaccuracies (e.g., Wang, Zhao, and Fringer 2011).

c. Finite element methods

The FE method relies on expanding the ocean fields in a series of polynomial basis functions defined on mesh elements, and seeking the coefficients of these expansions from the requirement that the governing equations be satisfied in an optimal way.

We introduce a set of basis functions $N_j(x, y)$ defined on mesh elements (in the FE method this name is routinely used instead of volumes or cells in the case of FV) and expand the tracer field as $T = T_j(t) N_j(x, y)$, with summation implied over the repeating indices in this section. The coefficients of expansion are only a function of time, which will be implied below. Depending on the choice of functions, the index j can list mesh elements (triangles) or their vertices, edges, or additional nodes in elements. A simple example is the continuous P_1 representation (P stands for polynomial, and 1 for its degree) in which case $N_j(x, y)$ equals 1 at vertex j and goes linearly to zero at neighboring vertices. In this case $T = T_j(t) N_j(x, y)$ represents a linear interpolation that is continuous across the faces. Continuous quadratic P_2 representation deals with functions defined at vertices and mid-edges. Many other possibilities are described in traditional courses on the FE method such as Zienkiewicz and Taylor (2000). It is important to stress that, despite the nonuniform

placement of the degrees of freedom T_j, the field T becomes defined over the entire mesh. Because this representation is based on a finite set of degrees of freedom, it cannot satisfy the continuous motion equations exactly. Instead, one requires equation (1) be satisfied in a weak sense as

$$\int (M_i \partial_t T - \mathbf{F}_h \nabla M_i) d\Omega = 0 \tag{3}$$

where M_i is an appropriate test function, and integration by parts has been performed to reduce the order of derivatives applied to T. No difficulties occur in this case if the representation for T is continuous, which will be assumed. Clearly, the set of M_i should be sufficient to constrain all the degrees of freedom used to represent T. An obvious possibility is to take $M_i = N_i$ to obtain

$$M_{ij} \partial_t T_j + (A_{ij} + D_{ij}) T_j = S_i \tag{4}$$

where $M_{ij} = \int N_i N_j d\Omega$, $A_{ij} = -\int N_j \boldsymbol{u} \cdot \nabla N_i d\Omega$ and $D_{ij} = \int K_h (\nabla N_i)(\nabla N_j) d\Omega$ are, respectively, mass, advection, and diffusion matrices. Note that derivatives in expressions for matrices A_{ij} and D_{ij} would be singular if N_j were discontinuous. The approach implemented in equation (4) is known as continuous Galerkin (CG) discretization. It is optimal in the sense that the residual of the equation for the field T in the space of basis functions N_j is orthogonal to these functions. A reader should notice that the procedure relies on a scalar product introduced over the space of functions N_j. This proves helpful, for it offers a natural representation for the balance of tracer variance or energy in the case of the primitive equations.

Note that in contrast to FD or FV treatments, the time derivatives are coupled through the mass matrix (M_{ij} previously). It is nondiagonal for the CG discretization and links all degrees of freedom. The presence of mass matrices improves accuracy, as we shall see, by reducing numerical dispersion (for more detail, see Donea and Huerta 2003), but iterative solvers must then be used to disentangle $\partial_t T_j$. Diagonal, or lumped, approximations are sometimes selected for M_{ij} to reduce computational burden, but this has an adverse effect on accuracy.

Discontinuous finite elements can be considered a generalization of both FV and CG FE approaches. In this case, the polynomial representation for T is confined to element interiors, and is discontinuous across the elemental boundaries. The simplest example is P_0, wherein T is elementwise-constant. For the P_1 discontinuous representation, the vertex values are different on each element, so that if six triangles meet at vertex v, there will be six values of T_v there. This leads to clustering of degrees of freedom, so that other variants with internal placement should be preferred.

Because of the discontinuous representation, one writes the weak formulation by integrating over element interiors,

$$\sum_c \left(\int (M_i \partial_t T - \mathbf{F}_h \nabla M_i) d\Omega_c + \int M \mathbf{F} \mathbf{n} d\Gamma_c \right) + P = 0 \tag{5}$$

where the integration in the last term is over the boundary of element c. Because elements are disconnected, equation (5) is incomplete unless certain penalties are added, represented here through P. They include terms that weakly impose the continuity of fluxes and fields. An alternative approach is to consider fluxes $\boldsymbol{F}$ as "numerical" fluxes, combining flux estimates from elements across the face together with relevant continuity constraints needed for accuracy and stability. If $M_i = N_i$, the result is the discontinuous Galerkin (DG) discretization. The reader is advised to consult a regular course (e.g., Li 2006) for details that are numerous here.

Compared with the FV method, the DG FE method spares the need for high-order reconstructions if high-order representation is used. This representation is internal to the element, which is beneficial from the standpoint of parallelization. Likewise, mass matrices now connect only local degrees of freedom inside elements, which makes their direct inversion feasible, in contrast to CG FE. This makes the DG FE method appealing for ocean modelling. However, the computational burden is high, and practical applications are still rare (e.g., Dawson et al. 2006; Kärnä, Legat, and Deleersnijder 2012). Besides, the complexity of coastlines and bottom topography is the reason why geometrical refinement and low-order discretizations are preferred.

d. Elementary examples

In order to explain how FV and CG FE methods work (we skip the DG FE case as it requires more lengthy detail) we consider a 1D example assuming the velocity u to be uniform. For the FV method, let T_c be the cell-mean value in cell c. The cell length will be h_c, and its boundaries will be at $x_{c-1/2}$ and $x_{c+1/2} = x_{c-1/2} + h_c$, with the index c increasing in the positive x-direction. We get

$$\partial T_c + (F_{c+1/2} - F_{c-1/2})/h_c = 0.$$

For advection, a local linear reconstruction $T_{c+1/2} = (T_{c+1}h_c + T_c h_{c+1})/(h_c + h_{c+1})$ will lead to a scheme that is equivalent to standard centered differences on uniform meshes. Note that for linear reconstruction the difference between the cell-mean and cell-centered values can be ignored but is essential in higher-order reconstructions. If $u > 0$, the estimate for the advective part of the flux $F_{c+1/2} = uT_c$ will yield the first-order upwind method, and the upwind quadratic reconstruction based on T_{c-1}, T_c, and T_{c+1} will lead to a second-order method on a uniform mesh. In order to increase the accuracy of discretization, instead of accurate representation of fluxes, one has to concentrate on representing the flux divergence and follow the road discussed, for example, by Webb, de Cuevas, and Richmond (1998). No further detail will be provided here because the reconstructions are specific to the mesh geometry. The important statement, however, is that on general triangular or hexagonal meshes similar procedures are possible as on regular meshes and that will yield discretizations familiar from FDs in many cases. The computational effort will be higher, however, for one cannot rely on a regular stencil. As concerns the diffusive part of the fluxes, one

estimates $(\partial_x T)_{c+1/2} = 2(T_{c+1} - -T_c)/(h_{c+1} + h_c)$ in the 1D case. More generally, the centered combination of diffusive fluxes computed on neighboring control volumes is used.

To illustrate the case of CG FE, we place the degrees of freedom T_v at "vertices" located at x_v, with the index increasing in the x-positive direction. The P_1 basis function associated to vertex v is $N_v = 1 - (x - x_v)/h_{v+1/2}$ for $x_v \le x \le x_{v+1}$ and $N_v = 1 + (x - x_v)/h_{v-1/2}$ for $x_{v-1} \le x \le x_v$ and zero otherwise. Here $h_{v+1/2} = x_{v+1} - x_v$ and $h_{v-1/2} = x_v - x_{v-1}$. Performing computations of the entries of the matrices, we get

$$M_{vj}\partial_t T_j = \frac{h_{v+1/2}}{6}\partial_t(2T_v + T_{v+1}) + \frac{h_{v-1/2}}{6}\partial_t(2T_v + T_{v-1})$$

$$A_{vj}T_j = (u/2)(T_{v+1} - T_{v-1}),$$

and

$$D_{vj}T_j = \frac{k}{h_{v+1/2}}(T_v - T_{v+1}) + \frac{k}{h_{v-1/2}}(T_v - T_{v-1}).$$

Taking for simplicity the case of uniform mesh, $h_{v+1/2} = h_{v-1/2} = h$,

$$\partial_t(T_{v-1} + 4T_v + T_{v+1})/6 + (u/(2h))(T_{v+1} - T_{v-1}) - K(T_{v-1} - 2T_v + T_{v+1})/h^2 = 0.$$

We now easily recognize that the expressions for advection and diffusion are just the traditional centered differences. The novel feature is the presence of the mass matrix with the time derivative. Because of the mass matrix, the time derivative is weighted over the same stencil as the space derivative. To understand why this is important, we turn to the von Neuman analysis discussed previously, taking $T = T_0(t)e^{ikx}$. In this case, we get $(2 + cos\,kh)\partial_t T_0/3u(sinkh)T_0 + K(2 - cos(kh))T_0 = 0$ meaning that the phase velocity becomes $c_p = 3(u/kh)sin(kh)/(2+cos(kh))$, which ensures much more accurate behavior for small kh compared with the estimate $c_p = (u/kh)sin(kh)$ that will follow in the absence of the mass matrix and will also be the result for the FV case shown previously.

To conclude, on uniform meshes both FV and linear FE lead to expressions recognizable from FDs. This situation persists in 2D, and one would derive the same statements on regular quadrilateral meshes. The methods are thus generalizations of common technology to rather arbitrary polygonal meshes. They are more expensive, because the information on neighbors has to be retrieved from look-up tables, as well as information on coefficients of differential operators, which have to be computed in advance to minimize run-time effort.

e. Numerical considerations in the primitive and shallow water equations

Although the previous considerations focus on explaining the essence of unstructured numerical methods, it leaves aside an important question of the consistency between the representation of velocity and scalars (pressure and tracers).

Similar to the difference in numerical properties of solutions on the Arakawa A, B, and C grids, the properties of solutions on unstructured meshes also depend on the placement of variables. The collocated placement of velocities and scalars is equivalent to the

A grid and shares a similar difficulty, namely the presence of a pressure mode on uniform meshes. A pressure mode is the possibility of nontrivial pressure distribution that corresponds to zero pressure gradient. It occurs because of gradient averaging implied by collocated meshes. Additionally, on such meshes the discrete curl of pressure gradient in the discretized momentum equation is not necessarily zero, which introduces errors in the discrete vorticity balance. Both call for stabilization in geostrophically dominated regimes, which introduces errors in the energy balance. Although collocated placement of variables is popular in computational fluid mechanics because the same infrastructure is shared by all variables, there is a tendency toward the use of codes with staggered placement of variables, which are free of pressure modes if there is no gradient averaging, in ocean modelling.

Staggered triangular grids, however, encounter a geometrical difficulty. The ratio of the number of vertices to cells to edges is 1:2:3, so that, as a rule, a staggered discretization will not be balanced: the number of degrees of freedom in velocity and pressure taken at different locations will be inconsistent, i.e., will deviate from the ratio 2:1. Thus, for example, a cell-vertex (velocity-pressure) FV discretization, which is an analog to the B grid, is characterized by a too-large velocity space. A triangular C-grid discretization with normal velocities at edges has too many pressure degrees of freedom. The consequence of the lack of balance is the presence of numerical modes. A review by Danilov (2013) presents a more detailed analysis and contains references to numerous works that have explored the properties of particular discretizations.

In summary, there is no perfect staggered discretization, but in some cases the numerical modes can be handled relatively easily. Numerical modes involving extra velocities can be controlled by viscosity; however, there is no obvious means to control too large a pressure space. In this respect, the triangular C grid is a suboptimal choice, and preference should be given to its dual implementation, the hexagonal C grid that has too many velocities. A general issue for C-grid codes is the accuracy of horizontal velocity reconstruction. Codes based on triangular C grids are prone to noise in the vertical velocity field in regimes characteristic of the large-scale ocean. They are, nevertheless, popular in coastal or estuarine-scale applications for which the noise presents a lesser problem. A quasi-hexagonal C grid forms the basis of the MPAS approach (Ringler et al. 2013), whereas the FV cell-vertex discretization is used in FVCOM and in the FV approach of Danilov (2012). Triangular C grids are the choice of Unstructured, Residual, Intertidal Mudflat (UnTRIM) and SUNTANS. FESOM (Wang et al. 2014) is an A-grid model relying on CG FE, as does ADCIRC.

Although initial development of unstructured-mesh ocean models was distributed between the CG FE and FV methods, current understanding is in favor of the FV method. The reason is that the hydrostatic approximation used by models and the need to have a clear definition of fluxes both encounter difficulties in the CG FE implementation. The horizontal coupling of CG FE makes the hydrostatic balance horizontally nonlocal, and breaking this coupling destroys energetic consistency. Although CG FE codes can be made perfectly volume conserving, they treat conservation in the weighted sense, without resorting to fluxes.

The interest in DG FE and in spectral element methods persists; yet, at present they are associated with an overly large computational burden.

f. Meshes, resolution, and practical examples

The practical recipes on how the mesh resolution should be varied depend on the application and differ on regional and large scales. On regional scales, if the dynamics are tidally driven, scaling the mesh element size as $(gH)^{1/2}$, with g the acceleration due to gravity and H the fluid thickness, could be advantageous because such meshes are "uniform" in terms of the propagation speed of surface gravity waves. Making cells on the deep water larger helps to circumvent time-step limitations that will occur otherwise. Extra resolution may, however, be needed to resolve steep or sharply varying topography or to resolve estuaries or details of coastlines. Many examples of successful applications demonstrating the utility of this approach can be found, for example, on the web sites of such models as FVCOM or ADCIRC, and the review by Greenberg et al. (2007) discusses many associated aspects.

Such highly variable meshes are not necessarily optimal in other situations involving baroclinic dynamics, instabilities, and eddies. Here the point is that numerical and physical dissipation depends on resolution and smaller dissipative coefficients are used on finer meshes. The implication is that a mesh optimal for simulating tides can be suboptimal for simulating baroclinic dynamics because of enhanced dissipation across its coarse mesh. Such applications would benefit from certain mesh uniformity, and the unstructured mesh character has only to provide a mechanism for effective nesting. The mesh design is then very similar to that sought by nested structured-grid models, with the difference that technical nesting can be avoided.

This situation is also common for large-scale applications. In this case, there are two main factors motivating the use of unstructured meshes. The first is once again the effective nesting motivated by the need to resolve eddies, as demonstrated by Ringler et al. (2013). The other factor is the geometry of important straits, as in the study by Wekerle et al. (2013). In all cases it should be borne in mind that dissipation, especially spurious numerical dissipation, is linked to resolution, and, by refining, one takes into account smaller flow details and simultaneously activates a part of dynamics that was previously damped on resolved scales. From the consideration of numerical stability and accuracy, smooth mesh transitions should be preferred, because this minimizes residual errors in representation of discrete operators. There are physical consequences, too, for baroclinic instabilities and eddies do not saturate immediately, and the presence of a coarse or eddy-permitting mesh upstream may affect the dynamics downstream in the refined domain. Study of all numerical aspects associated with variable resolution is only beginning.

Questions on how to deploy resolution are closely related to time integration. In the nested mode, the time step of an unstructured-mesh code will be defined by the size of the smallest element. Such codes will then be numerically efficient only if the smallest elements use an overwhelming number of degrees of freedom. This is easy to achieve in some situations,

and this may be the reason for reduced numerical efficiency in some others. Generally, this implies the need for careful consideration of many related issues and much experimentation with alternative approaches.

5. Parameterizing unresolved and partially-resolved phenomena

In ocean modelling, an approximation of the effects of unresolved processes in terms of resolved and known quantities is called a parameterization. In other fields, the terms subgrid model, closure, or regularization are used, which reflect that the unresolved processes 1) lie below the resolution of the grid, 2) are required for the equations of motion to be closed, and 3) tend to regularize or eliminate singularities that may arise in their absence. The most important consideration in determining the parameterizations required in an ocean model is determining which processes are resolved, which are not, and which are partially resolved. In unstructured grid models, this determination is even more important. In hierarchies of models of varying resolution or nested grid modelling, these determinations can help to ensure that equivalent scenarios are being simulated among the models.

Parameterizations may be crude or sophisticated, and they are rarely insignificant to the outcome of the simulation. If they were they would be neglected, for reasons explained as follows. Standard fluid viscosity and diffusivity themselves are parameterizations: they approximate the average over many stochastic trajectories of molecules under the assumptions of local equilibrium and typically also isotropy or transverse isotropy (e.g., when vertical and horizontal directions are distinguished). On scales near the mean free path of molecules, these approximations break down. In situations featuring strongly nonequilibrium thermodynamics, thermal diffusivity fails. Non-Newtonian fluids are still treatable as viscous fluids, but with generally nonlinear constitutive relations between stress and rate of strain, rather than a simple proportionality with a viscous coefficient between stress and rate of strain.

In fluid modelling, turbulence is the most common feature requiring parameterization. The idea of an eddy viscosity or eddy diffusivity—a value much larger than the molecular one to treat turbulent eddies similarly to molecular motions—predates computational modelling (e.g., Boussinesq 1877; Ekman 1905). However, turbulence is not necessarily confined to a largest scale that plays the role of the mean free path, and therefore a scale separation between resolved features and unresolved turbulence may not be realized. Furthermore, geophysical turbulence is frequently anisotropic and heterogeneous, and parameterizations may reflect this broken symmetry to greater or lesser degree.

A lack of scale separation complicates the development and generality of parameterizations significantly. When there is a scale separation between the largest turbulent features and the grid scale, the equations can be cast in terms of predicting only the Reynolds average of the motion, which may be steady, slowly varying, and/or easily resolved. Simulations without scale separation can be termed large eddy simulations (LESs), as only the largest of the turbulent features are resolved and smaller ones are parameterized. In Reynolds-averaged

parameterizations, the modest changes in resolution of the model may not affect the parameterization. In LES, the eddy viscosity typically depends on the resolution of the model and also the flow.

Likely the most common eddy viscosity parameterization is that of Smagorinsky (1963) based on the energy cascade idea of Kolmogorov (1941), which has precisely these characteristics:

$$\nu_S = \left(\frac{\Upsilon \Delta x}{\pi}\right)^2 \sqrt{\frac{1}{4}\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\right)\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\right)},$$

where Einstein summation is implied on indices, and the dimensionless constant Υ is near 1. The Smagorinsky viscosity depends on both the velocity u and gridscale Δx. Because of these scalings, the rate at which energy is dissipated can be matched to the rate at which it is flowing from large to small scales in the resolved flow, which mimics the effects of a longer turbulence cascade (i.e., higher resolution).

Although the Smagorinsky viscosity is in wide use in nonhydrostatic simulations, it is not appropriate for use in the large-scale enstrophy cascades of rotating fluids later discovered by Kraichnan (1967) as an integral critical in the derivation of the viscosity diverges in an enstrophy cascade (Fox-Kemper and Menemenlis 2008). The preceding three-dimensional form is the correct one for use in nonhydrostatic models where all velocity components are treated similarly. Smagorinsky (1993) also derives a hydrostatic version that only involves the horizontal velocities, and forms for anisotropic grids require more care (e.g., Ramachandran, Tandon, and Mahadevan 2013). For a derivation of the Smagorinsky closure for hydrostatic and nonhydrostatic models, see Smagorinsky (1993) or Fox-Kemper and Menemenlis (2008).

Even though Smagorinsky developed this parameterization for use in atmospheric models, it is not suitable for most simulations of large-scale atmospheric and oceanic flows that commonly use hydrostatic approximations and anisotropic grids. At large scales, the effects of rotation (quantified by Rossby number Ro = U/fL), stratification (quantified by Richardson number $\mathrm{Ri} = \mathrm{N}^2/(\partial U/\partial z)^2$), and the limited aspect ratio (α = H/L) of motions wider than the depth of the ocean all prevent the occurrence of isotropic three-dimensional (3D) turbulence. In stratified fluids, only scales smaller than the Ozmidov scale (the square root of energy flux divided by N^3) or the Thorpe scale are observed to have turbulence resembling a 3D cascade. Dillon (1982) shows that these scales are frequently related and O(1m) in the upper ocean at a variety of sites—obviously, these scales are much smaller than the grid scale used in large-scale ocean and atmosphere modelling. Perhaps these scales might be resolved by the vertical grid, but not the horizontal: Is an anisotropic form of Smagorinsky appropriate in that case? Even anisotropic forms of the Smagorinsky scheme continue to be derived from Kolmogorov theory for 3D turbulence with a forward energy cascade (e.g., Ramachandran, Tandon, and Mahadevan 2013). In large-scale oceanic turbulence, the forward energy cascade is not usually in effect until very small scales, and a cascade of enstrophy or potential enstrophy, may be more relevant on the scale of the model grid.

Then, the Leith (1996) scheme or adaptations of it (Fox-Kemper and Menemenlis 2008; Bachman, Fox-Kemper, and B. Pearson 2017; Pearson et al. 2017) are preferable. One day, large-scale ocean simulations will reach the point where the grid scale is firmly within the range of scales where Kolmogorov turbulence is expected, as presently occurs in boundary layer LES (Sullivan and Patton 2011), but at present rates of increasing computational power, this day is centuries away for global simulations and decades away for most coastal applications.

In the meantime, it is important to distinguish between partially-resolved processes—requiring parameterizations to be scale-aware, or scale with resolution—and processes that are firmly below the grid scale—for which parameterizations may be less dependent on grid scale. In typical ocean models, mesoscale and submesoscale eddies are partially resolved, justifying a scale-aware LES approach to handling these phenomena (e.g., Pearson et al. 2017). However, boundary layer and breaking internal wave turbulence, which are expected to cascade into the 3D turbulence but only with O(1m) grids and nonhydrostatic models, are far from being resolved. Thus, these processes are likely to be approximated with Reynolds averages and parameterizations that do not depend on modest changes to the scale of the grid, such as a background eddy diffusivity. Sometimes, even Reynolds-average schemes are flow-aware (i.e., they depend on resolved flows in a nontrivial way) and therefore must also become scale-aware to account for the dependence of the resolved flow resolution. For example, the submesoscale restratification parameterization of Fox-Kemper et al. (2011) incorporates the strength of resolved density gradients to infer the strength of unresolved fronts. At coarse resolution, the resolved density gradients are sensitive to changes in resolution, so the parameterization depends on the grid scale to approximately account for this lack of convergence. In this Reynolds-averaged but flow-aware case, the goal is to keep the effect of the *unresolved* phenomena *unaffected* by changes in the grid scale. By contrast, LES schemes are scale-aware to make parameterizations of *partially-resolved* phenomena be *affected* by grid scale, so that the parameterization effect is reduced as the grid is refined.

Therefore, parameterizations must not only account for the effects of unresolved motions, but be sensitive, or scale-aware, as to what kinds of motion are or are not resolved. For example, the majority of present atmosphere and ocean models are hydrostatic, which is a good approximation when the aspect ratio of the resolved motion is small. However, hydrostatic motions cannot exhibit a Kolmogorov 3D cascade, which requires an equality among velocities in all directions. Thus, when the grid aspect ratio is small and the model is hydrostatic, it is not consistent to use a Kolmogorov-based LES parameterization, such as the Smagorinsky scheme, because then the physical parameterization assumes no scale separation from 3D turbulence whereas the numerics demands a scale separation. However, the Smagorinsky scheme is frequently used outside of its realm of theoretical applicability because it is numerically robust, scale-aware, and flow-aware despite being inaccurate in such applications (Griffies and Hallberg 2000). In such models, the effects of 3D turbulence should be included through Reynolds-averaged boundary layer schemes, and other scale-aware and

flow-aware schemes should be used that are suited for the partially-resolved phenomena. Dimensional analysis based on the grid scales, coupled with dominant balances expected by the flow regime, can reveal important aspects of the motion near the grid scale useful in developing or choosing among parameterizations (e.g., Bachman, Fox-Kemper, and Pearson 2017).

On much larger scales, when the Rossby number is small and the Richardson number is large, then the horizontal scale of motion may be larger than the deformation radius (quantified by the Burger number $\mathrm{Bu} = (\mathrm{L_d/L})^2 \approx \mathrm{Ro}^2\mathrm{Ri}$). When a model grid has a small or O(1) Burger number, then the turbulence near the gridscale is quasi-geostrophic (QG), and a different set of turbulence cascades may be used to construct LES parameterizations. At the same time, the effects of 3D turbulence, for example in boundary layer schemes, may also be included with an assumed scale separation. Likewise, on very large scales, the barotropic motions may be approximated as 2D turbulence. Fox-Kemper and Menemenlis (2008) give examples of closures suitable in these regimes, whereas other closures rely on spectral self-similarity but do not specify what kind of self-similarity to expect (e.g., Bardina, Ferziger, and Reynolds 1980; San et al. 2011, San, Staples, and Iliescu 2013).

Examples of processes for which parameterizations presently exist are: mesoscale (Gent and McWilliams 1990; Redi 1982; Treguier, Held, and Larichev 1997; Griffies 1998; Eden and Greatbatch 2008; Jansen and Held 2014; Bachman et al. 2015; Grooms 2016; Bachman, Fox-Kemper, and Pearson 2017; Pearson et al. 2017; Zanna et al. 2017), submesoscale (Fox-Kemper, Ferrari, and Hallberg 2008; Fox-Kemper et al. 2011), boundary layer turbulence (Kraus and Turner 1967; Mellor and Yamada 1982; Price, Weller, and Pinkel 1986; Large, McWilliams, and Doney 1994; Sullivan, McWilliams, and Moeng 1994; Canuto et al. 2001; Harcourt 2013; Li et al. 2016), tidal mixing (Jayne and St Laurent 2001), fine-scale mixing by internal gravity waves (Polzin et al. 2014), symmetric instabilities (Bachman et al. 2017), surface gravity waves (Longuet-Higgins and Stewart 1962; Craik and Leibovich 1976; Lane, Restrepo, and McWilliams 2007), air–sea fluxes (Large and Pond 1981, 1982; Edson et al. 2013), bubbles (Liang et al. 2011), sediment (Warner et al. 2008), estuaries (Garvine and Whitney 2006; Sun et al. 2017), albedo (Payne 1972; Brandt et al. 2005), ice–ocean fluxes (McPhee, Morison, and Nilsen 2008), convection (Killworth 1989; Ilicak, Adcroft, and Legg 2014), abyssal overflows (Yeager and Danabasoglu 2012), skin layers and diurnal layers (Large and Caron 2015), biophysics (Zhang et al. 2009), and many more processes. Process studies—using high-resolution models or observations—of each phenomenon, spanning a hopefully representative set of scenarios, are generally the key to parameterization development and evaluation.

The reader may note that the majority of phenomena requiring closures just mentioned are turbulent or at least nonlinear. This results from the need to parameterize only those motions that can communicate across scales from unresolved to resolved. Scale connectivity is a defining aspect of turbulence, and all nonlinear phenomena have the potential for scale connectivity. Dominantly linear phenomena, such as waves in the absence of mean flow, stratification, or topography variations, do not link strongly across scales,

so unresolved waves of this type do not interact with resolved flows. One important exception occurs at surface and bottom boundary layers, where many phenomena require parameterization to appropriately accept stress and energy forcing (generally using a Reynolds average approach). Mathematically, neglecting these effects produces singular perturbations of the equations of motion thereby eliminating whole categories of possible solutions.

An advantage of scale-aware parameterizations is that they tend to regularize singularities regardless of grid scale and therefore result in robust numerical stability. The Boussinesq, hydrostatic primitive equations common to ocean and atmospheric simulations are well-behaved mathematically, but some mechanisms of regularization are required for this to be true (Cao and Titi 2007; Cao et al. 2013). Even in structured grid modelling, variation in latitude or stratification can change the effective scale of phenomena, and make constant viscosities and diffusivities fail to regularize the flow sufficiently for a given grid scale. In unstructured and variable grid modelling, such considerations are paramount for accurate and stable simulations. However, physical parameterizations are often more complicated, e.g., strongly nonlinear or flow-dependent, which may mean that, although they tend to regularize, it may be impossible to prove that fact. Examples of stability and improved simulation fidelity through physical parameterizations include Griffies and Hallberg (2000), Fox-Kemper and Menemenlis (2008), Chen, Gunzburger, and Ringler (2011) and Ilicak et al. (2012).

There are some disadvantages of physical parameterizations to keep in mind as well. The most obvious is that they may make the simulation more costly (parameterizations cost typically 10% to 25% of the total) and, therefore, reduce the achievable model resolution that is almost always desired. So, parameterizations should be judged based on impact versus cost and designed with attention to balancing increased accuracy versus increased complexity. Additionally, the generally-idealized process studies used to develop a parameterization are typically not representative of all scenarios that may be encountered in realistic modelling. It is common for a parameterization developed in extra-tropical process studies to fail in applications near the equator. In addition, parameterizations introduce imperfectly-known constants or parameters and, thus, additional uncertainty into models. Clever approaches to exploit these uncertainties rely on tuning the parameter values to reduce model to observation mismatch, or better yet to produce ensembles of simulations of varying parameter values to map and quantify the consequences of uncertainties.

Tuning is a dangerous business, because often the right answer can be arrived at for the wrong reasons, which may significantly affect model sensitivity or behavior in unprecedented forcing regimes (such as climate change). A good example of this effect is that often when a new parameterization, with solidly grounded physics based on process modelling or observations, is introduced into a model previously lacking a representation of said physics, it often increases the biases and mismatches in the model solutions versus observations. This is not an indication that the model has been degraded in realism, but that it was previously tuned to mask the neglect of a phenomenon with a canceling error in the values of other

parameters chosen. Once the new phenomenon receives treatment, the erroneous values of the other parameters resulting from tuning are exposed and require readjustment.

Finally, the nonlinear character of most parameterizations may preclude or limit their use in certain circumstances, as do some types of data assimilation that do not require adjoints (e.g., ensemble Kalman methods). Finally, it is easy to oversimplify parameterizations, e.g., substituting a horizontal diffusivity for an along-isopycnal diffusivity or an isotropic scalar diffusivity where an anisotropic tensor diffusivity is more appropriate, and it may be difficult to identify the consequences of these oversimplifications without careful thought and comparison with observations (Veronis 1977; Fox-Kemper, Lumpkin, and Bryan 2013).

Some numerical schemes seek to regularize, or otherwise constrain solution characteristics, without attempting to reproduce a particular physical phenomenon. Common examples are upwinding, monotonic, and shock-capturing schemes. Generally, these methods arrive at the cost of low-order accuracy, but low-order accuracy is typical in unstructured grid modelling in any case. A closely related idea is that of mimetic schemes, which seek exact mimicry of conservation laws in the continuous equations (for energy, enstrophy, potential density, mass, etc.) in the discrete equations. These schemes inherently offer regularization benefits. For example, if energy is conserved, it is impossible for an instability to cause unbounded growth in energy. However, these schemes are often slower to converge and less formally accurate than nonconservative schemes, e.g., spectral methods. It is easy to prove the effectiveness of such schemes, because conservation is guaranteed. However, it is difficult to prove that choice to be optimal, because other aspects of the solution are degraded in exchange. Furthermore, mimetic and constrained schemes are not a replacement for accurate physical parameterizations. For example, there are many pathways where energy is transferred among scales, and so exact conservation of resolved energy would be incorrect. Consider the example of boundary layers in oceanic and atmospheric flows. These layers have significantly more turbulent energy from surface roughness, waves, and convection—all of which is injected on scales smaller than a typical large-scale model grid, yet these energetic regions tend to perform important mixing of the larger scale property gradients. Conserving energy on the large-scale grid is no substitute for these unresolved phenomena in terms of ocean model accuracy, unlike in engineering LES or DNS applications where small-scale energy is dwarfed by larger-scale flows. Thus, scale-aware physical parameterizations may be used together with mimetic schemes to arrive at realistic flows of energy, enstrophy, or other desired quantities. Examples of accuracy and improved simulation fidelity through such schemes are Lax and Wendroff (2005), Leonard (1979), Zalesak (1979), Friedrich (1998), Nadiga and Bouchet (2011), San et al. (2011), and San, Staples, and Iliescu (2013).

Most modelling approaches blend together the physical aspects of parameterizations of unresolved processes with numerical accuracy, convergence, and regularization concerns such as grid scale and altered vertical coordinate discretization, including Lagrangian or semi-Lagrangian approaches. As such, the desired overall effect may be realized, but care is needed, e.g., in adapting a parameterization that is successful in a pressure coordinate

model into a model using isopycnal coordinates. A general approach to these issues is not yet known, but these steps become an active topic of research each time a new modelling system is developed.

6. Final remarks: Numerical modeling in a multicomponent ocean

Ocean prediction systems require several integrated components. A central element in these systems, discussed here, is a time-evolving numerical circulation model based upon the oceanic equations of motion (Jacobs and Fox-Kemper 2017). Equally important however are, first, in situ and external data collection systems necessary to provide the information required for initial and boundary conditions and, second, data assimilation methodologies to impose internal data-driven constraints. These issues are reviewed in this volume by, among others, Brink and Kirincich (2017) and Lermusiaux (2017).

One overriding constraint on our ability to deliver improved high-resolution numerical forecasts is the competition between the computational cost of our prediction systems and the rate of increase in computer power (i.e., Moore's Law). In addition to the continuing delivery of additional raw computer power, consequential improvement has also been consistently achieved through enhancement in the computational algorithms used for numerical solution, data assimilation, etc. The implementation of multiscale discretization procedures, such as the nested and unstructured grid techniques reviewed here, are examples of the latter means of achieving gains in efficiency. Alternative numerical approaches are continually under development; additional examples include moving grid methods (Koltakov and Fringer 2012), and higher-order FE and mortar element methods (Iskandarani et al. 2002; Maday, Mavriplis, and Patera 1989).

Lastly, we have focused here primarily on numerical treatments for the *multiscale ocean circulation*, i.e., ocean physics. For the reasons noted here, this has been, and will remain, an unavoidable concern in Earth System Model design. Of perhaps equal concern, however, is the increasing cost and complexity of including specialized modules for other geomorphological, chemical, ecological, and cryospheric processes, i.e., *the multicomponent ocean*. Especially over the long time scales of concern to climate modelling and prediction, the linkages between the ocean circulation and these other processes are all of significant concern.

As a consequence, in parallel with the development of multiscale treatments such as those described previously, whole new categories of coupled marine models have been, and are being, developed to address particular concerns. These include integrated models that incorporate the effects of waves and tides, currents and sediment transport (Egbert and Ray 2017; Kirby 2017), geochemical cycling (Hofmann et al. 2008), physical/biological interaction (Curchitser et al. 2013; Hofmann et al. 2009), ice–ocean coupling (Bartino and Holland 2017), and many other processes.

Integrations of multicomponent systems of significant duration are now routinely feasible on many space scales, e.g., estuarine (de Brye et al. 2010; Powell et al. 2012), continental shelf and slope (Previdi et al. 2009), basin-scale (Hermann et al. 2009), and global (Miller

at al. 2017). Despite continued enhancement in computer capacity, the next step forward—end-to-end coupled Earth System Models that allow affordable integration over extended time scales on global, multi-scale domains—will ensure the continuing need for novel new numerical approaches.

Acknowledgments. The authors gratefully acknowledge the constructive suggestions of the anonymous referees. BFK was supported by NSF 1350795.

REFERENCES

Arakawa, A. 1966. Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I. J. Comput. Phys., *1*(1), 119–143. doi: 10.1016/0021-9991(66)90015-5

Bachman, S., B. Fox-Kemper, and F. O. Bryan. 2015. A tracer-based inversion method for diagnosing eddy-induced diffusivity and advection. Ocean Modelling, *86*, 1–14.

Bachman, S. D., B. Fox-Kemper, and B. Pearson. 2017. A scale-aware subgrid model for quasi-geostrophic turbulence. J. Geophys. Res.: Oceans, *122*(2), 1529–1554. doi: 10.1002/2016JC012265

Bachman, S. D., B. Fox-Kemper, J. R. Taylor, and L. M. Thomas. 2017. Parameterization of frontal symmetric instabilities. I: Theory for resolved fronts. Ocean Model., *109*, 72–95. doi: 10.1016/j.ocemod.2016.12.003

Bardina, J., J. H. Ferziger, and W. C. Reynolds. 1980. Improved subgrid-scale models for large-eddy simulation, *in* 13th American Institute of Aeronautics and Astronautics, Fluid and Plasma Dynamics Conference, Snowmass, Colorado, July 14–16, 1980. Stanford: Stanford University, 10 pp. doi: 10.2514/6.1980-1357

Bartino, L. and M. Holland. 2017. Coupled ocean-ice modeling and predictions, *in* The Sea: The Science of Ocean Prediction, special issue, J. Mar. Res., in press.

Beckers, J.-M. 1992. Analytical linear numerical stability conditions for an anisotropic three-dimensional advection-diffusion equation. SIAM J. Numer. Anal., *29*(3), 701–713. doi: 10.1137/0729044

Beckers, J-M. and E. Deleersnijder. 1993. Stability of a FBTCS scheme applied to the propagation of shallow-water inertia-gravity waves on various space grids. J. Comput. Phys., *108*(1), 95–104. doi: 10.1006/jcph.1993.1166

Biastoch, A., C. W. Böning, and J. R. E. Lutjeharms. 2008. Agulhas leakage dynamics affects decadal variability in Atlantic overturning circulation. Nature, *456*, 489–492. doi: 10.1038/nature07426

Bleck, R. 2002. An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. Ocean Model., *4*(1), 55–88. doi: 10.1016/S1463-5003(01)00012-9

Boussinesq, J. 1877. Essai sur la théorie des eaux courantes, Mémoires Présentés par Divers Savants à l'Académie des Sciences, *23*, 1–680.

Brandt, R. E., S. G. Warren, A. P. Worby, and T. C. Grenfell. 2005. Surface albedo of the Antarctic sea ice zone. J. Clim., *18*, 3606–3622. doi: 10.1175/JCLI3489.1

Brink, K. H. and A. R. Kirincich. 2017. Some considerations about coastal ocean observing systems, *in* The Sea: The Science of Ocean Prediction, Part 1, Special Issue. J. Mar. Res. *75*, 161–188.

Canuto, V., A. Howard, Y. Cheng, and M. Dubovikov. 2001. Ocean turbulence. Part I: One-point closure model-momentum and heat vertical diffusivities. J. Phys. Oceanogr., *31*, 1413–1426. doi: 10.1175/1520-0485(2001)031<1413:OTPIOP>2.0.CO;2

Cao, C., S. Ibrahim, K. Nakanishi, and E. S. Titi. 2013. Finite-time blowup for the inviscid primitive equations of oceanic and atmospheric dynamics. Comm. Math. Phys., *337*(2), 473–482. doi: 10.1007/s00220-015-2365-1

Cao, C. and E. S. Titi. 2007. Global well-posedness of the three-dimensional viscous primitive equations of large scale ocean and atmosphere dynamics. Ann. Math., *166*(1), 245–267. doi: 10.1002/cpa.21576

Chanut, J., B. Barnier, L. Debreu, W. Large, T. Penduff, J.-M. Molines, and P. Mathiot. 2008. Mesoscale eddies in the Labrador Sea and their contribution to convection and re-stratification. J. Phys. Oceanogr., *38*, 1617–1643. doi: 10.1175/2008JPO3485.1

Chassignet, E., H. E. Hurlburt, O. M. Smedstad, G. Halliwell, A. H. Wallcraft, E. J. Metzger, et al. 2006. Generalized vertical coordinates for eddy-resolving global and coastal ocean forecasts. Oceanography, *19*(1), 118–129. doi: 10.5670/oceanog.2006.95

Chen, C., H. Liu, and R. C. Beardsley. 2003. An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: Applications to coastal ocean and estuaries. J. Atmos. Ocean. Technol., *20*, 159–186. doi: 10.1175/1520-0426(2003)020<0159:AUGFVT>2.0.CO;2

Chen, Q., M. Gunzburger, and T. Ringler. 2011. A scale-invariant formulation of the anticipated potential vorticity method. Mon. Weather Rev., *139*, 2614–2629. doi: 10.1175/MWR-D-10-05004.1

Choi, B.-J., M. Iskandarani, J. Levin, and D. B. Haidvogel. 2004. A spectral finite volume method for the shallow water equations. Mon. Weather Rev., *132*, 1777–1791. doi: 10.1175/1520-0493(2004)132<1777:ASFMFT>2.0.CO;2

Craik, A. D. D. and S. Leibovich. 1976. Rational model for Langmuir circulations. J. Fluid Mech., *73*(3), 401–426. doi: 10.1017/S0022112076001420

Curchitser, E., H. P. Batchelder, D. B. Haidvogel, J. Fiechter, and J. Runge. 2013. Advances in physical, biological, and coupled ocean models during the US GLOBEC program. Oceanography, *26*(4), 54–69. doi: 10.5670/oceanog.2013.75

Curchitser, E. N., D. B. Haidvogel, A. J. Hermann, E. Dobbins, T. M. Powell, and A. Kaplan. 2005. Multi-scale modelling of the North Pacific Ocean: Assessment of simulated basin-scale variability (1996–2003). J. Geophys. Res.: Oceans, *110*(C11), C11021. doi: 10.1029/2005JC002902

Curchitser, E. N., J. Small, K. Hedstrom, and W. Large. 2011. Up- and down-scaling effects of upwelling in the California Current System, *in* Report of Working Group 20 on Evaluations of Climate Change Projections, PICES Scientific Report No. 40, M. G. Foreman and Y. Yamanaka, eds. Patricia Bay, British Columbia: North Pacific Marine Science Organization, pp. 98–102.

Danilov, S. 2012. Two finite-volume unstructured mesh models for large-scale ocean modeling. Ocean Model., *47*, 14–25. doi: 10.1016/j.ocemod.2012.01.004

Danilov, S. 2013. Ocean modeling on unstructured meshes. Ocean Model., *69*,195–210. doi: 10.1016/j.ocemod.2013.05.005

Dawson, C. N., J. J. Westerink, J. C. Feyen, and D. Pothina. 2006. Continuous, discontinuous and coupled discontinuous-continuous Galerkin finite element methods for the shallow water equations. Int. J. Numer. Meth. Fluids, *52*(1), 63–88. doi: 10.1002/fld.1156

Debreu, L., C. Vouland, and E. Blayo. 2008. AGRIF: Adaptive grid refinement in Fortran. Comput. Geosci., *34*(1), 8–13. doi: 10.1016/j.cageo.2007.01.009

de Brye, B., A. de Brauwere, O. Gourgue, T. Kärnä, J. Lambrechts, R. Comblen, and E. Deleersnijder. 2010. A finite-element, multi-scale model of the Scheldt tributaries, river, estuary and ROFI. Coast. Eng., *57*(9), 850–863. doi: 10.1016/j.coastaleng.2010.04.001

Dillon, T. M. 1982. Vertical overturns: A comparison of Thorpe and Ozmidov length scales. J. Geophys. Res.: Oceans, *87*(C12), 9601–9613. doi: 10.1029/JC087iC12p09601

Donea, J. and A. Huerta. 2003. Finite Element Methods for Flow Problems. Chichester, UK: John Wiley and Sons, 350 pp. doi: 10.1002/0470013826

Döscher, R., C. W. Böning, and P. Herrmann. 1994. Response of circulation and heat transport in the North Atlantic to changes in thermohaline forcing in northern latitudes: A model study. J. Phys. Oceanogr., *24*, 2306–2320. doi: 10.1175/1520-0485(1994)024<2306:ROCAHT>2.0.CO;2

Dukowicz, J. K. 1995. Mesh effects for Rossby waves. J. Comput. Phys., *119*(1), 188–194. doi: 10.1006/jcph.1995.1126

Durran, D. R. 1999. Numerical Methods for Wave Equations in Geophysical Fluid Dynamics. New York: Springer-Verlag, 465 pp. doi: 10.1007/978-1-4757-3081-4

Eden, C. and R. J. Greatbatch. 2008. Towards a mesoscale eddy closure. Ocean Model., *20*(3), 223–239. doi: 10.1016/j.ocemod.2007.09.002

Edson, J. B., V. Jampana, R. A. Weller, S. P. Bigorre, A. J. Plueddemann, C. W. Fairall, et al. 2013. On the exchange of momentum over the open ocean. J. Phys. Oceanogr., *43*, 1589–1610. doi: 10.1175/JPO-D-12-0173.1

Egbert, G. D. and R. D. Ray. 2017. Tidal prediction, *in* The Sea: The Science of Ocean Prediction, Part 1, Special Issue, J. Mar. Res. 75, 189–237.

Ekman, V. W. 1905. On the influence of the earth's rotation on ocean currents. Arkiv. Mat. Astron. Fysik., *2*, 1–53.

Fox-Kemper, B., S. Bachman, B. Pearson, and S. Reckinger. 2014. Principles and advances in subgrid modelling for eddy-rich simulations. CLIVAR Exchanges, *19*(2), 42–45.

Fox-Kemper, B., G. Danabasoglu, R. Ferrari, S. M. Griffies, R. W. Hallberg, M. M. Holland, et al. 2011. Parameterization of mixed layer eddies. III: Implementation and impact in global ocean climate simulations. Ocean Model., *39*(1–2), 61–78. doi: 10.1016/j.ocemod.2010.09.002

Fox-Kemper, B., R. Ferrari, and R. Hallberg. 2008. Parameterization of mixed layer eddies. Part I: Theory and diagnosis. J. Phys. Oceanogr., *38*, 1145–1165. doi: 10.1175/2007JPO3792.1

Fox-Kemper, B., R. Lumpkin, and F. O. Bryan. 2013. Lateral transport in the ocean interior, *in* Ocean Circulation and Climate: A 21st century perspective. Vol. 103 of the International Geophysics Series, G. Siedler, S. M. Griffies, J. Gould, and J. A. Church, eds. Cambridge: Academic Press, pp. 185–209

Fox-Kemper, B. and D. Menemenlis. 2008. Can large eddy simulation techniques improve mesoscale-rich ocean models? *in* Ocean Modelling in an Eddying Regime. Vol. 177. AGU Geophysical Monograph Series, M. Hecht, and H. Hasumi, eds. Washington, DC: American Geophysical Union, pp. 319–338.

Friedrich, O. 1998. Weighted essentially non-oscillatory schemes for the interpolation of mean values on unstructured grids. J. Comput. Phys., *144*(1), 194–212. doi: 10.1006/jcph.1998.5988

Fringer, O. B., M. Gerritsen, and R. L. Street. 2006. An unstructured-grid, finite-volume, nonhydrostatic, parallel coastal ocean simulator. Ocean Model., *14*(3–4), 139–173. doi: 10.1016/j.ocemod.2006.03.006

Garvine, R. W. and M. M. Whitney. 2006. An estuarine box model of freshwater delivery to the coastal ocean for use in climate models. J. Mar. Res., *64*(2), 173–194. doi: 10.1357/002224006777606506

Gent, P. R. and J. C. McWilliams. 1990. Isopycnal mixing in ocean circulation models. J. Phys. Oceanogr., *20*, 150–155. doi: 10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2

Gerdes, R., A. Biastoch, and R. Redler. 2001. Fresh water balance of the Gulf Stream System in a regional model. Clim. Dyn., *18*(1), 17–27. doi: 10.1007/s003820100164

Greenberg, D. A., F. Dupont, F. H. Lyard, D. R. Lynch, and F. E. Werner. 2007. Resolution issues in numerical models of oceanic and coastal circulation. Cont. Shelf Res., *27*(9), 1317–1343. doi: 10.1016/j.csr.2007.01.023

Griffies, S. and R. Hallberg. 2000. Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. Mon. Weather Rev., *128*, 2935–2946. doi: 10.1175/1520-0493(2000)128<2935:BFWASL>2.0.CO;2

Griffies, S. M. 1998. The Gent-McWilliams skew flux. J. Phys. Oceanogr., *28*, 831–841. doi: 10.1175/1520-0485(1998)028<0831:TGMSF>2.0.CO;2

Grooms, I. 2016. A Gaussian-product stochastic Gent–McWilliams parameterization. Ocean Model., *106*, 27–43. doi: 10.1016/j.ocemod.2016.09.005

Harcourt, R. R. 2013. A second-moment closure model of Langmuir turbulence. J. Phys. Oceanogr., *43*, 673–697. doi: 10.1175/JPO-D-12-0105.1

Hermann, A. J., E. N. Curchitser, E. L. Dobbins, and D. B. Haidvogel. 2009. A comparison of remote versus local influence of El Niño on the coastal circulation of the Northeast Pacific. Deep Sea Res. Part II Top. Stud. Oceanogr., *56*(24), 2427–2443. doi: 10.1016/j.dsr2.2009.02.005

Hirt, C. W., A. A. Amsden, and J. L. Cook. 1974. An arbitrary Lagrangian–Eulerian computing method for all flow speeds. J. Comput. Phys., *14*(3), 227–253. doi: 10.1016/0021-9991(74)90051-5

Hofmann, E., J.-N. Druon, K. Fennel, M. Friedrichs, D. B. Haidvogel, C. Lee, et al. 2008. Eastern U.S. continental shelf carbon budget: Integrating models, data assimilation, and analysis. Oceanography, *21*(1), 86–104. doi: 10.5670/oceanog.2008.70

Hofmann, E., D. Bushek, S. Ford, X. Guo, D. B. Haidvogel, D. Hedgecock, et al. 2009. Understanding how disease and environment combine to structure resistance in estuarine bivalve populations. Oceanography, *22*(4), 212–231. doi: 10.5670/oceanog.2009.110

Ilicak, M., A. J. Adcroft, S. M. Griffies, and R. W. Hallberg. 2012. Spurious dianeutral mixing and the role of momentum closure. Ocean Model., *45–46*, 37–58. doi: 10.1016/j.ocemod.2011.10.003

Ilicak, M., A. J. Adcroft, and S. Legg. 2014. A framework for parameterization of heterogeneous ocean convection. Ocean Model., *82*, 1–14. doi: 10.1016/j.ocemod.2014.07.002

Iskandarani, M., D. B. Haidvogel, J. Levin, E. N. Curchitser, and C. A. Edwards. 2002. Multiscale geophysical modelling using the spectral element method. Comput. Sci. Eng., *4*(5), 42–48. doi: 10.1109/MCISE.2002.1032428

Jacobs, G. and B. Fox-Kemper. 2017. Open ocean dynamics, *in* The Sea: The Science of Ocean Prediction, special issue, J. Mar. Res., in press.

Jansen, M. F. and I. M. Held. 2014. Parameterizing subgrid-scale eddy effects using energetically consistent backscatter. Ocean Model., *80*, 36–48. doi: 10.1016/j.ocemod.2014.06.002

Jayne, S. and L. St Laurent. 2001. Parameterizing tidal dissipation over rough topography. Geophys. Res. Lett., *28*(5), 811–814. doi: 10.1029/2000GL012044

Kang, D. and E. N. Curchitser. 2013. Gulf Stream eddy characteristics in a high-resolution ocean model. J. Geophys. Res. Oceanogr., *118*(9), 4474–4487. doi: 10.1002/jgrc.20318

Kang, D. and E. N. Curchitser. 2015. Energetics of eddy-mean flow interactions in the Gulf Stream region. J. Phys. Oceanogr., *45*, 1103–1120. doi: 10.1175/JPO-D-14-0200.1

Kärnä, T., V. Legat, and E. Deleersnijder. 2012. A baroclinic discontinuous Galerkin finite element model for coastal flows. Ocean Model., *61*, 1–20. doi: 10.1016/j.ocemod.2012.09.009

Killworth, P. D. 1989. On the parameterization of deep convection in ocean models, *in* Parameterizations of Small-Scale Process: Proceedings of the Aha Huliko'a Hawaiian Winter Workshop. Honolulu, HI. DTIC Document, Fort Belvoir, VA: Defense Technical Information Center, pp. 59–74.

Kirby, J. T. 2017. Recent advances in nearshore wave, circulation and sediment transport modelling, *in* The Sea: The Science of Ocean Prediction, Part 1, Special Issue, J. Mar. Res. 75, 263–300.

Koltakov, S. and O. B. Fringer. 2012. Moving grid method for numerical simulation of stratified flows. Int. J. Numer. Meth. Fluids, *71*(12), 1524–1545. doi: 10.1002/fld.3724

Kolmogorov, A. N. 1941. The local structure of turbulence in incompressible viscous fluid for very large Reynolds number. Dokl. Akad. Nauk. SSSR, *30*, 9–13.

Kraichnan, R. H. 1967. Inertial ranges in two-dimensional turbulence. Phys. Fluids, *10*, 1417. doi: 10.1063/1.1762301

Kraus, E. and J. Turner. 1967. A one-dimensional model of the seasonal thermocline. II: The general theory and its consequences. Tellus, *19*(1), 98–106. doi: 10.1111/j.2153-3490.1967.tb01462.x

Lane, E. M., J. M. Restrepo, and J. C. McWilliams. 2007. Wave-current interaction: A comparison of radiation-stress and vortex-force representations. J. Phys. Oceanogr., *37*, 1122–1141. doi: 10.1175/JPO3043.1

Large, W. and J. Caron. 2015. Diurnal cycling of sea surface temperature, salinity, and current in the CESM coupled climate model. J. Geophys. Res.: Oceans, *120*(5), 3711–3729. doi: 10.1002/2014JC010691

Large, W. G., J. C. McWilliams, and S. C. Doney, 1994. Oceanic vertical mixing: A review and a model with a nonlocal boundary-layer parameterization. Rev. Geophys., *32*(4), 363–403. doi: 10.1029/94RG01872

Large, W. G. and S. Pond. 1981. Open ocean momentum flux measurements in moderate to strong winds. J. Phys. Oceanogr., *11*, 324–336. doi: 10.1175/1520-0485(1981)011<0324:OOMFMI>2.0.CO;2

Large, W. G. and S. Pond. 1982. Sensible and latent heat flux measurement over the ocean. J. Phys. Oceanogr., *12*, 464–482. doi: 10.1175/1520-0485(1982)012<0464:SALHFM>2.0.CO;2

Lax, P. and B. Wendroff. 2005. Systems of conservation laws. Selected Papers Volume I, pp. 263–283.

Leclair, M. and G. Madec. 2011. Z-tilde coordinate, an arbitrary Lagrangian–Eulerian coordinate separating high and low frequency motions. Ocean Model., *37*, 139–152. doi: 0.1016/j.ocemod.2011.02.001

Leith, C. E. 1996. Stochastic models of chaotic systems. Physica D, *98*(2–4), 481–491. doi: 10.1016/0167-2789(96)00107-8

Leonard, B. P. 1979. A stable and accurate convective modelling procedure based on quadratic upstream interpolation. Comput. Methods Appl. Mech. Eng., *19*(1), 59–98. doi: 10.1016/0045-7825(79)90034-3

Lermusiaux, P. 2017. Intelligent observing and modeling systems, *in* The Sea: The Science of Ocean Prediction, Special Issue. J. Mar. Res., in press.

Li, B. Q. 2006. Discontinuous finite elements in fluid dynamics and heat transfer. New York: Springer, 578 pp.

Li, Q., A. Webb, B. Fox-Kemper, A. Craig, G. Danabasoglu, W. G. Large, and M. Vertenstein. 2016. Langmuir mixing effects on global climate: WAVEWATCH III in CESM. Ocean Model.. *103*, 145–160. doi: 10.1016/j.ocemod.2015.07.020

Liang, J.-H., J. C. McWilliams, P. P. Sullivan, and B. Baschek. 2011. Modelling bubbles and dissolved gases in the ocean. J. Geophys. Res.: Oceans, *116*(C3), C03015. doi: 10.1029/2010JC006579

Longuet-Higgins, M. S. and R. W. Stewart. 1962. Radiation stress and mass transport in gravity waves, with application to 'surf beats'. J. Fluid Mech., *13*(4), 481–504. doi: 10.1017/S0022112062000877

Maday, Y., C. Mavriplis, and A. T. Patera. 1989. Nonconforming mortar element methods: Application to spectral discretizations, *in* Domain Decomposition Methods. Philadelphia, PA; SIAM, pp. 392–418.

Marchesiello, P., J. C. McWilliams, and A. Shchepetkin. 2001. Open boundary conditions for long-term integration of regional ocean models. Ocean Model., *3*(1–2), 1–20. doi: 10.1016/S1463-5003(00)00013-5

McPhee, M., J. Morison, and F. Nilsen. 2008. Revisiting heat and salt exchange at the ice-ocean interface: Ocean flux and modelling considerations. J. Geophys. Res.: Oceans, *113*(C6), 2156–2202. doi: 10.1029/2007JC004383

Mellor, G. L. and T. Yamada. 1982. Development of a turbulent closure model for geophysical fluid problems. Rev. Geophys. Space Phys., *20*(4), 851–857. doi: 10.1029/RG020i004p00851

Miller, A. J., M. Collins, S. Gualdi, T. G. Jensen, V. Misra, L. P. Pezzi, et al. 2017. Coupled ocean-atmosphere modelling and predictions, *in* The Sea: The Science of Ocean Prediction, part 1, special issue, J. Mar. Res. 75, 361–402.

Moin, P. 2010. Fundamentals of Engineering Numerical Analysis. Cambridge: Cambridge University Press, 256 pp.

Moore, G. E. 1965. Cramming more components onto integrated circuits. Electronics, *38*(8), 114–117.

Nadiga, B. T. and F. Bouchet. 2011. The equivalence of the Lagrangian-averaged Navier-Stokes-alpha model and the rational large eddy simulation model in two dimensions. Phys. Fluids, *23*, 095105. doi: 10.1063/1.3632084

Orlanski, I. 1976. A simple boundary condition for unbounded hyperbolic flows. J. Comput. Phys., *21*, 251–269. doi: 10.1016/0021-9991(76)90023-1

Payne, R. E. 1972. Albedo of the sea surface. J. Atmos. Sci., *29*, 959–970. doi: 10.1175/1520-0469(1972)029<0959:AOTSS>2.0.CO;2

Pearson, B., B. Fox-Kemper, S. D. Bachman, and F. O. Bryan. 2017. Evaluation of scale-aware subgrid mesoscale eddy models in a global eddy-rich model. Ocean Model., *115*, 42–58. doi: 10.1016/j.ocemod.2017.05.007

Pedlosky, J. 2003. Waves in the Ocean and Atmosphere: Introduction to Wave Dynamics. New York: Springer, 264 pp.

Polzin, K. L., A. C. Naveira Garabato, T. N. Huussen, B. M. Sloyan, and S. Waterman. 2014. Finescale parameterizations of turbulent dissipation. J. Geophys. Res.: Oceans, *119*(2), 1383–1419. doi: 10.1002/2013JC008979

Powell, E., D. Kreeger, J. Morson, D. B. Haidvogel, Z. Wang, R. Thomas, and J. Gius. 2012. Oyster food supply in Delaware Bay: Estimation from a hydrodynamic model and interaction with the oyster population. J. Mar. Res., *70*(2–3), 469–503. doi: 10.1357/002224012802851904

Previdi, M., K. Fennel, J. Wilkin, and D. Haidvogel. 2009. Interannual variability in atmospheric CO2 uptake on the Northeast U.S. continental shelf. J. Geophys. Res., *114*(G4), G04003, doi: 10.1029/2008JG000881

Price, J. F., R. A. Weller, and R. Pinkel. 1986. Diurnal cycling: Observations and models of the upper ocean response to diurnal heating, cooling, and wind mixing. J. Geophys. Res.: Oceans, *91*(C7), 8411–8427. doi: 10.1029/JC091iC07p08411

Ramachandran, S., A. Tandon, and A. Mahadevan. 2013. Effect of subgrid-scale mixing on the evolution of forced submesoscale instabilities. Ocean Model., *66*, 45–63. doi: 10.1016/j.ocemod.2013.03.001

Redi, M. H. 1982. Oceanic isopycnal mixing by coordinate rotation. J. Phys. Oceanogr., *12*, 1154–1158. doi: 10.1175/1520-0485(1982)012<1154:OIMBCR>2.0.CO;2

Ringler, T., M. Petersen, R. Higdon, D. Jacobsen, M. Maltrud, and P. W. Jones. 2013. A multi-resolution approach to global ocean modelling. Ocean Model., *69*, 211–232. doi: 10.1016/j.ocemod.2013.04.010

San, O., A. E. Staples, and T. Iliescu. 2013. Approximate deconvolution large eddy simulation of a stratified two-layer quasigeostrophic ocean model. Ocean Model., *63*, 1–20. doi: 10.1016/j.ocemod.2012.12.007

San, O., A. E. Staples, Z. Wang, and T. Iliescu. 2011. Approximate deconvolution large eddy simulation of a barotropic ocean circulation model. Ocean Model., *40*(2), 120–132. doi: 10.1016/j.ocemod.2011.08.003

Shchepetkin, A. F. and J. C. McWilliams. 2005. The Regional Ocean Modelling System (ROMS): A split-explicit, free-surface, topography following coordinates ocean model. Ocean Model., *9*(4), 347–404. doi: 10.1016/j.ocemod.2004.08.002

Smagorinsky, J. 1993. Some historical remarks on the use of nonlinear viscosities, *in* Large Eddy Simulation of Complex Engineering and Geophysical Flows, B. Galperin, and S. A. Orszag, eds. Cambridge: Cambridge University Press, pp. 3–36.

Smagorinsky, J. 1963. General circulation experiments with the primitive equations I: The basic experiment. Mon. Weather Rev., *91*, 99–164. doi: 10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2

Small, R. J., E. N. Curchitser, K. S. Hedstrom, B. Kaufman, and W. Large. 2015. The Benguela upwelling system: Quantifying the sensitivity to resolution and coastal wind representation in a global climate mode. J. Clim., *28*, 9409–9432. doi: 10.1175/JCLI-D-15-0192.1

Stewart, R. H. 2008. Introduction to Physical Oceanography. Gainesville: Orange Grove Texts, 345 pp.

Sullivan, P. P., J. C. McWilliams, and C. H. Moeng. 1994. A subgrid-scale model for large-eddy simulation of planetary boundary-layer flows. Boundary-Layer Meteorology, *71*(3), 247–276. doi: 10.1007/BF00713741

Sullivan, P. P. and E. G. Patton. 2011. The effect of mesh resolution on convective boundary-layer statistics and structures generated by large-eddy simulation. J. Atmos. Sci., *68*, 2395–2415. doi: 10.1175/JAS-D-10-05010.1

Sun, Q., M. M. Whitney, F. O. Bryan, and Y. H. Tseng, 2017. A box model for representing estuarine physical processes in Earth system models. Ocean Model., *112*, 139–153. doi: 10.1016/j.ocemod.2017.03.004

Thompson, J. F., Z. U. A. Warsi, and C. W. Mastin. 1982. Boundary-fitted coordinate systems for numerical solution of partial differential equations—A review. J. Comput. Phys., *47*(1), 1-108. doi: 10.1016/0021-9991(82)90066-3

Thuburn, J. 1990. TVD schemes, positive schemes, and the universal limiter. Mon. Weather Rev., *125*, 1990–1993. doi: 10.1175/1520-0493(1997)125<1990:TSPSAT>2.0.CO;2

Treguier, A. M., I. M. Held, and V. D. Larichev. 1997. Parameterization of quasigeostrophic eddies in primitive equation ocean models. J. Phys. Oceanogr., *27*, 567–580. doi: 10.1175/1520-0485(1997)027<0567:POQEIP>2.0.CO;2

Veronis, G. 1977. Use of tracers in circulation studies, *in* The Sea. Vol. 6: Marine Modelling, E. D. Goldberg, ed. Hoboken: John Wiley and Sons, pp. 169–188.

Vitousek, S. and O. B. Fringer. 2011. Physical vs. numerical dispersion in nonhydrostatic ocean modelling. Ocean Model., *40*(1), 72–86. doi: 10.1016/j.ocemod.2011.07.002

Wajsowicz, R. C. 1986. Free planetary waves in finite difference ocean models. J. Phys. Oceanogr., *16*, 773–789.

Wang, Q., S. Danilov, D. Sidorenko, R. Timmermann, C. Wekerle, X. Wang, et al. 2014. The Finite Element Sea Ice-Ocean Model (FESOM) v.1.4: Formulation of an ocean general circulation model. Geosci. Model Dev., *7*(2), 663–693. doi: 10.5194/gmd-7-663-2014

Wang, B., G. Zhao, and O. B. Fringer. 2011. Reconstruction of vector fields for semi-Lagrangian advection on unstructured, staggered grids. Ocean Model., *40*(1), 52–71. doi: 10.1016/j.ocemod.2011.06.003

Warner, J. C., C. R. Sherwood, R. P. Signell, C. K. Harris, and H. G. Arango. 2008. Development of a three-dimensional, regional, coupled wave, current, and sediment transport model. Comput. Geosci., *34*(10), 1284–1306. doi: 10.1016/j.cageo.2008.02.012

Webb, D. J., B. A. de Cuevas, and C. Richmond. 1998. Improved advection schemes for ocean models. J. Atmos. Ocean. Technol., *15*, 1171–1187. doi: 10.1175/1520-0426(1998)015<1171:IASFOM>2.0.CO;2

Wekerle, C., Q. Wang, S. Danilov, T. Jung, and J. Schroter. 2013. The Canadian Arctic Archipelago throughflow in a multiresolution global model: Model assessment and the driving mechanism of interannual variability. J. Geophys. Res.: Oceans, *118*(9), 4525–4541. doi: 10.1002/jgrc.20330

Westerink, J. J., R. A. Luettich, C. A. Blain, and N. W. Schener. 1992. ADCIRC: An Advanced Three-Dimensional Circulation Model for Shelves, Coasts and Estuaries; Report 2: Users Manual for ADCIRC-2DDI. Contractors Report to the US Army Corps of Engineers. Washington, DC: US Army Corps of Engineers.

White, L., A. Adcroft, and R. Hallberg. 2009. High-order regridding–remapping schemes for continuous isopycnal and generalized coordinates in ocean models. J. Comput. Phys. *228*(23), 8665–8692. doi: 10.1016/j.jcp.2009.08.016

Wunsch, C., D. B. Haidvogel, M. Iskandarani, and R. Hughes. 1997. Dynamics of the long-period tides. Prog. Oceanogr., *10*(1–4), 81–108. doi: 10.1016/S0079-6611(97)00024-4

Yeager, S. and Danabasoglu, G. 2012. Sensitivity of Atlantic meridional overturning circulation variability to parameterized Nordic Sea overflows in CCSM4. J. Clim., *25*, 2077–2103. doi: 10.1175/JCLI-D-11-00149.1

Zalesak, S. T. 1979. Fully multidimensional flux-corrected transport algorithms for fluids. J. Comput. Phys., *31*(3), 335–362. doi: 10.1016/0021-9991(79)90051-2

Zanna, L., P. P. Mana, J. Anstey, T. David, and T. Bolton. 2017. Scale-aware deterministic and stochastic parametrizations of eddy-mean flow interaction. Ocean Model., *111*, 66–80. doi: 10.1016/j.ocemod.2017.01.004

Zhang, R.-H., A. J. Busalacchi, X. Wang, J. Ballabrera-Poy, R. G. Murtugudde, E. C. Hackert, and D. Chen. 2009. Role of ocean biology-induced climate feedback in the modulation of El Nino-Southern Oscillation. Geophys. Res. Lett., *36*(3), L03608. doi: 10.1029/2008GL036568

Zhang, Y. and A. M. Baptista. 2008. SELFE: A semi-implicit Eulerian–Lagrangian finite-element model for cross-scale ocean circulation. Ocean Model., *21*(3–4), 71–96. doi: 10.1016/j.ocemod.2007.11.005

Zienkiewicz, O. C. and R. L. Taylor. 2000. The Finite Element Method. Oxford: Butterworth-Heinemann, 689 pp.

Received: *17 May 2017*; revised: *8 December 2017*.

Editor's note: Contributions to *The Sea: The Science of Ocean Prediction* are being published separately in special issues of *Journal of Marine Research* and will be made available in a forthcoming supplement as Volume 17 of the series.