Ocean Modelling 143 (2019) 101462

Contents lists available at ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

A framework for seamless one-way nesting of internal wave-resolving ocean models

Justin S. Rogers [a,*], Matthew D. Rayson [b], Dong S. Ko [c], Kraig B. Winters [d], Oliver B. Fringer [a]

[a] Environmental Fluid Mechanics Laboratory, Stanford University, 473 Via Ortega, Stanford, CA, 94305, USA
[b] Oceans Graduate School, University of Western Australia, Australia
[c] United States Naval Research Laboratory, Stennis Space Center, MS, USA
[d] Scripps Institution of Oceanography, University of California San Diego, CA, USA

ABSTRACT

To resolve nonlinear internal wave motions in regional domains, we propose a one-way ocean model nesting framework in which a nonhydrostatic model is nested within a regional model. The nesting scheme produces internal tides that propagate freely into and out of the nonhydrostatic model domain with no reflections while nudging the low-frequency motions. We demonstrate the method with idealized test cases and a field-scale simulation of the South China Sea. The method relies on four parameters, the time scale of low-pass filtering, the time scale of low frequency forcing, the time scale of sponge damping and the width of the sponge damping layer. We present guidelines for selecting these four parameters based on the background stratification. The internal tides from the large-scale model steepen through nonlinear and nonhydrostatic effects and reasonably match observations at a fraction of the computational cost needed with a single uncoupled model. The model critically relies on the background stratification inherited from the large-scale model, which in this case leads to an underpredicted phase speed and amplitude in the nested model result. This highlights the need for large-scale ocean models to incorporate observational stratification data, especially in the ocean interior.

1. Introduction

Internal gravity waves are typically on the scale of hundreds of kilometers in space, hours in time, and are ubiquitous in the world's oceans. They are thought to contribute an estimated one terawatt of deep-sea turbulent dissipation which is understood to play a crucial role in the ocean's global redistribution of heat and momentum (Alford et al., 2015), and serve important ecological functions to mixing in the nearshore environment (Walter et al., 2012). Internal waves are generated by tidal motions over topography and can steepen into nonlinear solitary-like wave trains (Alford et al., 2010). When these internal waves encounter shallow depths such as the shelf, they steepen and break, often forming dissipative bores (Arthur and Fringer, 2016; Walter et al., 2012).

Recently, global ocean models such as the HYCOM model (Martin, 2000), have been improved to explicitly resolve the internal tidal motions (Arbic et al., 2018), and thus contain long-wavelength internal tides in their solution. At a regional scale, hydrostatic ocean models can reasonably estimate net dissipation from internal tides (Carter et al., 2008; Suanda et al., 2018). However, the coarse spatial resolution O (km) and hydrostatic pressure formulation of these models limit the development of nonlinear solitary wave trains, and computation of breaking and dissipation kinematics (Vitousek and Fringer, 2011).

To model high resolution internal wave processes, a grid nesting approach is often employed where boundary conditions from the large-scale model are used to force a higher resolution nested model (Blayo and Rousseau, 2016; Chen et al., 2013; Kumar et al., 2015; Pickering et al., 2015; Warn-Varnas et al., 2015). For grid nesting problems with waves, the fundamental problem is that wave information propagates differently in the parent grid versus the child grid, and thus the two models contain inconsistent information (amplitude, wavelength, timing, etc.) at the outgoing model boundary (Blayo and Rousseau, 2016; Vitousek and Fringer, 2011). This difference typically leads to reflection of spurious wave information back into the domain. By design, the child grid and parent grid compute different physical processes appropriate for their scale (*i.e.* non-hydrostatic vs. hydrostatic pressure, resolved vs. parameterized turbulence) and thus this is a fundamental challenge for modeling waves in nested domains. A related problem relates to the need to match the outgoing direction of propagation at open boundaries: imposing an inconsistent wave form on a model boundary can send incorrect wave information into the domain. For example, if we wish to model a wave exiting the nested domain, and naively apply these wave velocities on the outgoing model boundary, a spurious wave will be forced into the domain when there is a mismatch between the large-scale and nested domain internal tide solution.

Previous work has attempted to solve this basic nesting problem with several approaches. The first is two-way nesting which attempts to solve for a uniform solution between model types (velocity, pressure, scalars), and has been successfully demonstrated for the nonhydrostatic Navier–Stokes equations (Blayo and Rousseau, 2016). While two-way

* Corresponding author.
E-mail address: jsrogers@stanford.edu (J.S. Rogers).

https://doi.org/10.1016/j.ocemod.2019.101462
Received 28 March 2019; Received in revised form 10 July 2019; Accepted 1 September 2019
Available online 16 September 2019
1463-5003/

nesting is often employed for hydrostatic ocean models with different models including NCOM (Ko et al., 2008), and ROMS (Warner et al., 2010), solving a two-way nested large-scale nonhydrostatic ocean model is under development with the CROCO model, (Auclair et al., 2018) and would require orders of magnitude more computational power.

The second approach is to use adaptive mesh refinement (AMR) methods which adaptively refine the grid in specific areas, and have been demonstrated in stratified flows (Koltakov and Fringer, 2013) and large-scale oceanic flows using the ROMS AGRIF (Debreu et al., 2012) and ICOM (Piggott et al., 2008) models. The disadvantage of a single grid with grid stretching is that grid quality can compromise accuracy when there is significant stretching and the time step is dictated by the finest grid spacing, which can become limiting in some cases.

The third approach is to employ one-way nesting and radiate wave energy out of the domain using a radiative function, which requires solving for the wave speed and properties which are then propagated out of the model domain. This has been applied to both surface gravity waves (Chapman, 1985), atmospheric gravity waves (Klemp and Durran, 1983), and hydrostatic oceanic boundary conditions including internal waves (Carter and Merrifield, 2007; Marsaleix et al., 2006; Mason et al., 2010; Suanda et al., 2018). A variant of this approach uses an adaptive algorithm where inward and outward information fluxes are treated separately (Marchesiello et al., 2001). The limitation of these approaches is that the approximation of wave speeds and properties (vertical and horizontal wavenumbers) inevitably leads to errors in the radiative boundary conditions, which lead to spurious waves reflecting into the domain. It can also be computationally expensive to solve for spectral wave number properties (frequency, wavenumber, amplitude) and estimate the propagation speed.

We present a novel approach which splits the parent grid solution into low-frequency (mean) and high-frequency (wave) components at some dividing frequency. The low-frequency components are applied as Dirichlet boundary conditions on the nested grid and nudged over the model domain. The wave component is forced for wave energy propagating into the domain as Dirichlet boundary conditions on the nested grid boundary and with a sponge damping layer near these boundaries which damps outgoing wave energy. This combination provides an efficient and practical method to conduct one-way multiscale model nesting. Our aim here is to model specific wave events; although the approach is similar to laboratory-scale modeling studies of stratified fluids (Winters, 2016; Winters and de la Fuente, 2012), it has not been applied to a regional scale ocean model. We show that the accuracy of the nesting strategy is dependent on four time scales that are related to physical properties of the system. Most importantly, we show that for realistic applications, the results depend critically on accuracy of the large-scale stratification.

In this manuscript, we present the theoretical framework (Section 2) and illustrate the methods in a series of idealized test cases (Section 3). We then present a realistic model application in the South China Sea, nesting the SUNTANS model (Fringer et al., 2006) into the US Navy NCOM model (Martin, 2000) focusing on modeling internal waves (Section 4). A discussion (Section 5) and summary (Section 6) follow.

2. Theoretical framework

The general setting we wish to model is a three-dimensional nested model domain Ω_N within a large-scale model Ω_{LS} and overlapping boundary Γ (Fig. 1a) with oceanic flows at a range of temporal and spatial scales including internal wave propagation and steepening. The nested model domain has a higher resolution grid and bathymetry than the large-scale model, and may have enhanced physics appropriate for the higher resolution (*i.e.* non-hydrostatic vs. hydrostatic pressure). To accomplish the nesting, we implement a method to nudge the SUNTANS model velocity, temperature and salinity $(\boldsymbol{u}, T, S)$ to a large-scale model velocity temperature and salinity $(\boldsymbol{u}_{LS}, T_{LS}, S_{LS})$ at appropriate time scales using a horizontal source term in the momentum equations.

The SUNTANS model solves the three-dimensional, nonhydrostatic, Reynolds-averaged Navier–Stokes equations under the Boussinesq approximation, along with the scalar transport equations and the continuity equation (Fringer et al., 2006),

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} = -2\boldsymbol{\Omega}\times\boldsymbol{u} - \frac{1}{\rho_0}\nabla p - \frac{g}{\rho_0}\rho\hat{k} + \nabla_H\cdot\left(\nu_H\nabla_H\boldsymbol{u}\right) + \frac{\partial}{\partial z}\left(\nu_V\frac{\partial \boldsymbol{u}}{\partial z}\right) + \boldsymbol{F}_M, \tag{1}$$

$$\frac{\partial T}{\partial t} + u\cdot\nabla T = \nabla_H\cdot\left(\kappa_H\nabla_H T\right) + \frac{\partial}{\partial z}\left(\kappa_V\frac{\partial T}{\partial z}\right) + F_T, \tag{2}$$

$$\frac{\partial S}{\partial t} + u\cdot\nabla S = \nabla_H\cdot\left(\gamma_H\nabla_H S\right) + \frac{\partial}{\partial z}\left(\kappa_V\frac{\partial S}{\partial z}\right) + F_S, \tag{3}$$

$$\nabla\cdot\boldsymbol{u} = 0, \tag{4}$$

where $\boldsymbol{u} = (u, v, w)$ is the velocity vector, $\boldsymbol{\Omega}$ is the earth's angular velocity vector, g is gravitational acceleration; (ν, κ, γ) are the eddy-viscosity, eddy thermal diffusivity, and eddy mass diffusivity respectively, $\hat{k}$ is the vertical unit vector, and $()_H$ and $()_V$ are the horizontal and vertical components of a variable or operator. Notation is summarized in Table 1. To compute density ρ, an equation of state is employed as a function of temperature T and/or salinity S.

A horizontal source term in the momentum equations $\boldsymbol{F_M}(x, y, z, t)$ is introduced, such that

$$\boldsymbol{F_M} = \boldsymbol{F_L} + \boldsymbol{F_D}, \tag{5}$$

where $\boldsymbol{F_L}$ is the low-frequency forcing and $\boldsymbol{F_D}$ is the sponge layer high-frequency damping term (Fig. 1b). Likewise, the temperature and salinity equations are nudged at low-frequency with $\boldsymbol{F_T}$ and $\boldsymbol{F_S}$, respectively.

To perform our nesting approach, it is necessary to separate velocity, density and pressure fields into reference, background and spatially varying components. Following Kang and Fringer (2012), the total density is decomposed with

$$\rho(x, y, z, t) = \rho_0 + \rho_b(x, y, z, t) + \rho'(x, y, z, t), \tag{6}$$

where ρ_0 is the constant reference density, $\rho_b = \overline{\rho} - \rho_0$ is the time-averaged background density with time low-pass $\overline{()}$ filtering, and ρ' is the perturbation density due to wave motions. The pressure is split into its hydrostatic (p_h) and nonhydrostatic (q) parts with $p = p_h + q$, and p_h is further decomposed into constant, background and perturbation components with

$$p_h = p_0 + p_b + p' = \rho_0 g(\eta - z) + g\int_z^\eta \rho_b dz + g\int_z^\eta \rho' dz, \tag{7}$$

where η is the free surface. We separate the flow into barotropic and baroclinic velocity components, $\boldsymbol{u} = \boldsymbol{U} + \boldsymbol{u}'$, where $(')$ denotes the deviation from the barotropic velocity, and the barotropic velocity is given by the depth average

$$\boldsymbol{U} = \frac{1}{H}\int_{-d}^{\eta}\boldsymbol{u_H}dz = \frac{1}{H}\langle\boldsymbol{u_H}\rangle, \tag{8}$$

where $\langle\rangle = \int_{-d}^{\eta}()\,dz$, the depth integration of a quantity, $\boldsymbol{u_H}$ is the horizontal component of the velocity vector, the bottom is defined at $z = -d(x, y)$, and the total water depth is given by $H = -d + \eta$.

Adopting a decomposition of variables into time high $\widetilde{()}$ and time low-pass $\overline{()}$ components with filtering time scale τ_{filt}, the velocity can be further decomposed into both temporally-$(\overline{u} + \widetilde{u})$ and vertically-$(U + u')$ varying components with

$$\boldsymbol{u} = \overline{\boldsymbol{u}} + \widetilde{\boldsymbol{u}} = \left(\overline{\boldsymbol{U}} + \overline{\boldsymbol{u}}'\right) + \left(\widetilde{\boldsymbol{U}} + \widetilde{\boldsymbol{u}}'\right), \tag{9}$$

which represent, from left to right, the low-frequency barotropic and baroclinic and the high-frequency barotropic and baroclinic terms, respectively. The free surface is also be decomposed with

$$\eta = \overline{\eta}_0 + \widetilde{\eta}_0 + \widetilde{\eta}, \tag{10}$$

which represent, from left to right, low-frequency barotropic, high-frequency barotropic and high-frequency baroclinic motions. Here, $\overline{\eta}_0$

Table 1
Notation.

Symbol	Description
$()_H$, $()_V$	Horizontal and vertical components of a variable or operator
$\tilde{()}$	High pass time filter
$\bar{()}$	Low pass time filter
$\langle\rangle$	Depth integration of a quantity
C	Wave phase speed
d	Bottom surface depth
$\hat{D}$	Sponge damping operator
$E_{K0}, E_k^{'}, E_{k0}^{'}$	Barotropic, baroclinic, and cross kinetic energy density
f	Inertial frequency
$\boldsymbol{F_T}$, $\boldsymbol{F_S}$	Low frequency temperature and salinity forcing terms
$\boldsymbol{F_L}$, $\boldsymbol{F_D}$	Low- and high-frequency momentum forcing terms
$\boldsymbol{F_M}$	Horizontal momentum source term
$\hat{\mathcal{F}}$	Energy flux operator
$\langle \boldsymbol{\mathcal{F}}_0 \rangle$, $\langle \boldsymbol{\mathcal{F}}' \rangle$	Barotropic, baroclinic energy flux
f_z	Mode-1 velocity eigenfunction
g	Gravitational acceleration
H	Total water depth
H_e	Effective depth of the interface
$\boldsymbol{k}$	Wavenumber
L,W	Domain length and width
N_b	Background buoyancy frequency
$\boldsymbol{n}$	Unit vector directed out of model domain
p	Pressure
p_h, q	Hydrostatic and nonhydrostatic pressures
r	Radial distance from a point to the nearest boundary
R_F	Wave reflection coefficient
S, S_{LS}	Nested and large-scale model salinity
T, T_{LS}	Nested and large-scale model temperature
t	Time variable
$\hat{t}_r, \theta_r$	Time ramping function and time scale
$\boldsymbol{U}$	Barotropic (depth-averaged) velocity
$\boldsymbol{u}'$	Baroclinic velocity
$\boldsymbol{u}\,(u, v, w)$	Nested model velocity
$\boldsymbol{u_{LS}}\,(u_{Ls}, v_{Ls}, w_{LS})$	Large-scale model velocity
$\boldsymbol{x}\,(x, y, z)$	Horizontal and vertical coordinates
α	Wave attenuation coefficient
Γ	Overlapping boundary of nested and large-scale models
ΔB	Buffer region of wave influence
ΔD	Sponge damping layer width
Δt	Time step
$\Delta x, \Delta y, \Delta z$	Horizontal and vertical grid resolution
ζ	Isopycnal displacement
η	Free surface
$\bar{\eta}_0, \tilde{\eta}_0, \tilde{\eta}$	Low-freq. barotropic, high-freq. barotropic and hi-freq. baroclinic free surface
λ	Wavelength
λ_G	Grid lepticity
ν, κ, γ	Eddy viscosity, eddy thermal diffusivity, and eddy mass diffusivity
ρ	Density
ρ_0, ρ_b, ρ'	Constant reference, background, perturbation density
τ_{filt}	Filtering time scale
τ_L, τ_D	Low- and high-frequency forcing time scale
ϕ	Wave phase
$\boldsymbol{\Omega}$	Earth's angular velocity vector
Ω_{LS}, Ω_N	Large-scale and nested model domains
ω	Radian frequency

is the low-pass time filtered free surface, $\tilde{\eta}_0$ is the spatially low-pass filtered free surface, obtained by spatially filtering η at a scale much larger than the internal wave wavelength λ, and thus $\tilde{\eta}$ is the high-frequency, high-wavenumber component of the free surface.

For oceanic conditions, the time scale of separation is taken at a frequency less than the inertial frequency f and the primary tidal frequencies. Thus, the high-frequency components primarily consist of propagating internal gravity waves with frequency ω within the range $f^2 < \omega^2 < N_b^2$, where the buoyancy frequency N_b is given by $N_b^2 = -g/\rho_0 \frac{\partial \rho_b}{\partial z}$.

The separation of time scales is treated separately in the large-scale grid (LS) and the SUNTANS computations. For the large-scale grid Ω_{LS}, the simulations are run in a one-way nesting routine, and therefore the large-scale results are known a priori to running the coupled model. Therefore, $\bar{\boldsymbol{u}}_{LS}$ is calculated with a simple low-pass filter using a Fourier transform, and $\tilde{\boldsymbol{u}}_{LS}$ is the deviation from this mean. For the SUNTANS domain Ω_N, $\bar{\boldsymbol{u}}$ is computed using a running average of previous results. Following Wolfram and Ringler (2017), the ordinary differential equation for a two-level impulse function is

$$\frac{d\bar{u}}{dt} = \frac{1}{\tau_{filt}}\left(u - \bar{u}\right), \tag{11}$$

where τ_{filt} is the low-pass time filtering scale. A running mean at time step n+1 is thus approximated with the backward in-time discretization

$$\frac{\bar{u}^{n+1} - \bar{u}^n}{\Delta t} = \frac{1}{\tau_{filt}}\left(u^{n+1} - \bar{u}^n\right). \tag{12}$$

This filter is equal to a standard single-pole filter (Smith, 1997), and as long as $\Delta t \ll \tau_{filt}$, (12) typically has significantly lower error than a moving box filter.

To obtain the energy flux which is needed in the sponge layers, the kinetic energy density is decomposed into the barotropic horizontal,

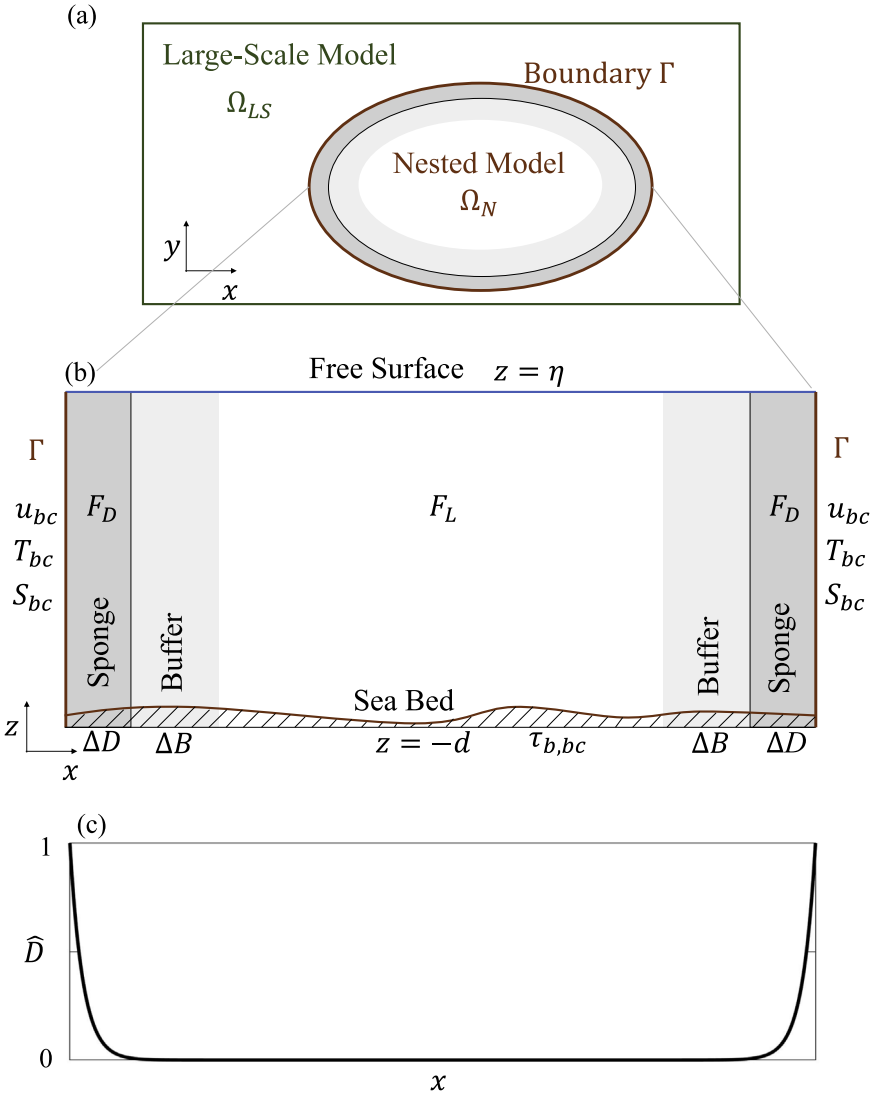

Fig. 1. Conceptual framework model domain showing (a) large-scale model space Ω_{Ls}, nested model space Ω_N, and boundary Γ, (b) $x-z$ plane of nested model showing sponge damping regions, buffer region, sea bed, lateral boundary conditions for horizontal velocity, temperature, and salinity, and (c) spatial variation of damping coefficient $\hat{D}$.

baroclinic, and cross terms (Kang and Fringer, 2012)

$$E_{k0} = \frac{1}{2}\rho_0 \left(U^2 + V^2\right), \tag{13}$$

$$E'_k = \frac{1}{2}\rho_0 \left(u'^2 + v'^2 + w^2\right), \tag{14}$$

$$E'_{k0} = \rho_0 \left(Uu' + Vv'\right). \tag{15}$$

The depth-integrated barotropic and baroclinic energy flux terms ignoring small unclosed and diffusion terms are given by

$$\langle \boldsymbol{\mathcal{F}}_0 \rangle = \boldsymbol{U}_H \langle E_{k0} \rangle + \boldsymbol{U}_H H \rho_0 g \eta + \boldsymbol{U}_H \langle p' \rangle + \boldsymbol{U}_H \langle q \rangle, \tag{16}$$

$$\langle \boldsymbol{\mathcal{F}}' \rangle = \langle \boldsymbol{u}_H E'_k \rangle + \langle \boldsymbol{u}_H E'_{k0} \rangle + \langle \boldsymbol{u}'_H p' \rangle + \langle \boldsymbol{u}'_H q \rangle, \tag{17}$$

where in oceanic conditions the dominant terms in (16) are $\boldsymbol{U}_H H \rho_0 g \eta$ and $\boldsymbol{U}_H \langle p' \rangle$ due to hydrostatic pressure work, and the dominant term in (17) is $\langle \boldsymbol{u}'_H p' \rangle$, or the energy flux due to the wave-induced hydrostatic pressure work (Kang and Fringer, 2012).

2.1. Boundary conditions

We apply Dirichlet boundary conditions along Γ using a decomposed horizontal velocity ($\boldsymbol{u}_{bc}$), from the large-scale grid (Figs. 1b, 2). The boundary velocities were chosen to be consistent with the lateral sponge layer, such that boundary is forced by the barotropic and low-frequency baroclinic components of the large-scale forcing, and the baroclinic wave velocity associated with waves propagating into the domain,

$$\boldsymbol{u}_{bc} = \left(\overline{\boldsymbol{U}}_{LS} + \overline{\boldsymbol{u}}'_{LS} + \widetilde{\boldsymbol{U}}_{LS} + \hat{\mathcal{F}}\widetilde{\boldsymbol{u}}'_{LS}\right)\hat{\imath}_r, \tag{18}$$

where the operator

$$\hat{\mathcal{F}}(x,y,t) = \begin{cases} 1 \text{ if } \overline{\langle \boldsymbol{\mathcal{F}}' \rangle}_{LS,lp} \cdot \boldsymbol{n} < 0 \\ 0 \text{ if } \overline{\langle \boldsymbol{\mathcal{F}}' \rangle}_{LS,lp} \cdot \boldsymbol{n} \geq 0 \end{cases} \tag{19}$$

selects only the wave velocities associated with energy flux propagation into the domain, and $\boldsymbol{n}$ is the unit vector directed out of the nested model domain, normal to the model boundary. Thus, the baroclinic high frequency motions are modified depending on the wave direction, but all other components of the flow are specified by the large-scale model. For the computation of (19), the low-pass filtered flux (in time) $\overline{\langle \boldsymbol{\mathcal{F}}' \rangle}_{LS,lp}$ is spatially smoothed (lp) using a Gaussian filter to a scale much larger than the internal wave wavelength λ to eliminate variations in $\hat{\mathcal{F}}$ smaller than the sponge layer width. We also initialize the boundary velocities smoothly in time with a ramping function $\hat{\imath}_r = 1 - \exp\left(-t/\theta_r\right)$, where t is the model run time and θ_r is a ramping time

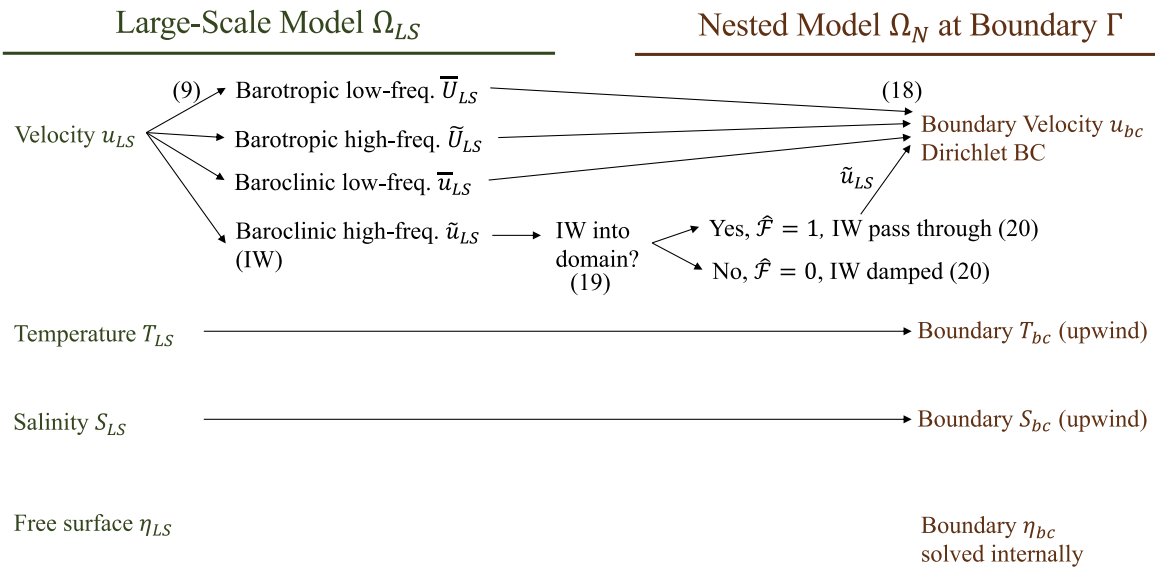

Fig. 2. Development of lateral boundary conditions (BC) for horizontal velocity u_{bc}, temperature T_{bc} and salinity S_{bc} at each grid cell in the nested model domain Ω_N obtained from the large-scale model Ω_{LS}. Selective treatment of internal wave (IW) motions, leads to incoming waves passing through the boundary/lateral sponge layer and damping of outgoing waves. Note for volume conservation in realistic applications, a very small number of boundary grid cells have Neumann boundary condition with an imposed free surface η_{bc} (and no imposed velocity).

scale. The boundary conditions for tracers (T_{bc}, S_{bc}) are upwind, set to the large-scale value for flow entering the domain (T_{Ls}, S_{Ls}), and the free surface is not constrained as we apply Dirichlet boundary conditions on the horizontal velocity (Fig. 2). To ensure long-term volume consistency, very few boundary grid cells in a corner of the domain are specified as Neumann velocity boundary conditions [instead of the Dirichlet condition in (18)] with a specified free surface $\eta_{bc} = \overline{\eta}_0 + \tilde{\eta}_0$ which is the instantaneous spatially low-pass filtered free surface from the large-scale model (*i.e.* removing IW motions).

2.2. Lateral sponge layer

To minimize wave reflections from the nested model boundaries, a sponge damping term is implemented in the horizontal momentum equation (Eq. (1)) of the form

$$\boldsymbol{F}_D = -\frac{\left(\tilde{\boldsymbol{u}}' - \hat{\mathcal{F}}\hat{\imath}_r\tilde{\boldsymbol{u}}'_{LS}\right)}{\tau_D}\hat{D}, \tag{20}$$

$$\hat{D}(\boldsymbol{x}) = \exp\left(-\frac{4r(\boldsymbol{x})}{\Delta D}\right), \tag{21}$$

where τ_D is the damping time scale and $\hat{D}$ modulates the strength of damping spatially which decreases exponentially at a length scale of ΔD from the nearest boundary, and r is the distance from the nearest boundary (Fig. 1c). The damping timescale is thus a local value given by $\hat{D}(\boldsymbol{x})/\tau_D$, which means that the strength of the damping of outgoing high frequency baroclinic components decreases exponentially when moving from the boundary into the domain. By nudging the SUNTANS baroclinic wave velocity $(\tilde{u}')$ to the large-scale baroclinic wave velocities associated with incoming waves only $(\hat{\mathcal{F}}\tilde{u}'_{LS})$, incoming waves and flows are allowed to propagate into the domain, while all outgoing baroclinic wave motions are damped (Fig. 2). This approach is very flexible and allows waves to propagate at different directions and frequencies within the domain while damping energy near the boundaries, thereby avoiding spurious reflections from the boundaries.

2.3. Low-frequency forcing

The low-frequency motions in the SUNTANS model $(\overline{\boldsymbol{u}})$ are forced by the large-scale low-frequency motions $(\overline{\boldsymbol{u}}_{\boldsymbol{LS}})$ with the forcing term $\boldsymbol{F}_L$ in Eq. (5) over the entire model domain using

$$\boldsymbol{F}_L = -\frac{\left(\overline{\boldsymbol{u}} - \overline{\boldsymbol{u}}_{\boldsymbol{LS}}\right)}{\tau_L}\hat{\imath}_r, \tag{22}$$

where τ_L is the low-frequency forcing time scale. Similarly, the low-frequency SUNTANS model temperature and salinity $(\overline{T}, \overline{S})$ are nudged towards the large-scale low-frequency variables $(\overline{T}_{LS}, \overline{S}_{LS})$ using

$$F_T = -\frac{\left(\overline{T} - \overline{T}_{LS}\right)}{\tau_T}\hat{\imath}_r, \tag{23}$$

$$F_S = -\frac{\left(\overline{S} - \overline{S}_{LS}\right)}{\tau_S}\hat{\imath}_r, \tag{24}$$

where (τ_T, τ_S) are the low-frequency forcing time scales for temperature and salinity, respectively. This framework assumes the model is initialized with T_{LS} and S_{LS}, but zero velocity and free surface. Forcing terms for the low-frequency components ((22)–(24)) are initially zero because of $\hat{\imath}_r$. The sponge damping term (20) has $\hat{\imath}_r$ applied to $\tilde{u}'_{LS}$, and initially damps all baroclinic high frequency motions $(\tilde{u}')$ arising from spurious motions from the initial conditions, and slowly transitions to forcing to the large-scale boundary motions $\tilde{u}'_{LS}$. To force wave velocities (and not temperature or salinity), the method proposed here relies on four parameters, the time scale of low-pass filtering τ_{filt}, the time scale of low-frequency forcing τ_{L}, the time scale of sponge damping τ_D and the width of the sponge damping layer ΔD.

3. Idealized test cases & model performance

We apply a series of test cases to demonstrate the model capabilities and the effects of the four boundary parameters under different conditions including a two-dimensional (x-z) idealized internal wave generation case, and a three-dimensional idealized South China Sea.

3.1. Base case: Internal wave propagation in two dimensions

We compute an idealized model simulation of internal wave propagation in one direction (x-z plane) based on an idealization of observations from the South China Sea (Ramp et al., 2010; Zhang et al., 2011). Specifically, we include a typical South China Sea density field given by $\rho_b = -2.1\text{x}10^{-4}\rho_0 T$, and temperature $T = 23.36\exp$

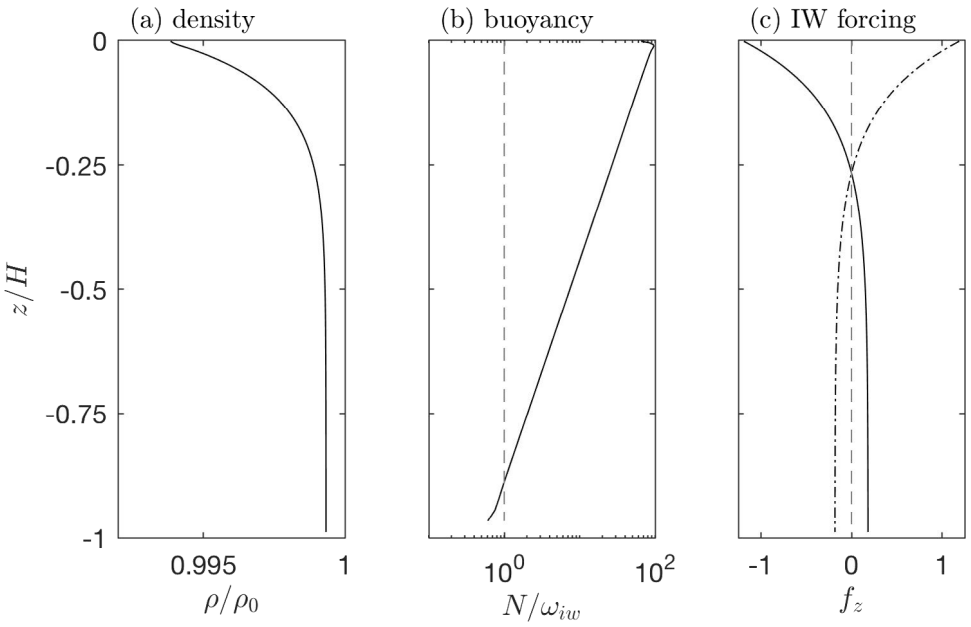

Fig. 3. Vertical profiles of idealized conditions over depth H for (a) density ρ normalized by background density ρ_0, (b) buoyancy frequency N normalized by forcing radian frequency ω_{iw} (at M_2 tidal frequency), with values above dashed line supporting internal wave propagation, and (c) mode-1 velocity forcing function $f(z)$ in the wave-maker region.

$\left[-(-z-44.12)/293.12\right]+3.13$ (Fig. 3a) with buoyancy frequency N, Coriolis frequency $f = 5.7\times10^{-5}$ rad/s typical for this latitude, and an imposed forcing frequency at the M_2 tidal period $\omega_{M2} = 2\pi/T_{M2}$ (T_{M2} = 12.42 h). For most of the depth $f^2 < \omega_{M2}^2 < N^2$, thus this environment is capable of sustaining propagating internal waves (Fig. 3b). We apply this test case on a 520 × 0.75 × 3 km [length (L), width (W), depth (H)] domain with grid resolution $\Delta x = \Delta y = 250$ m, which meets the requirements of Vitousek and Fringer (2011) for minimal numerical dispersion, and 100 vertical z-levels (Table 2) with enhanced resolution near the surface and exponentially-increasing vertical spacing with a growth rate of 1.08. The time step Δt is dictated by stability of internal gravity wave propagation at speed C_{IW}, and given by the Courant number constraint $C_{IW}\Delta t/\Delta x < 1$. Boundary conditions are Dirichlet at the x (East–West) boundaries and periodic in y (North–South). An idealized large-scale velocity field $\boldsymbol{u}_{LS}$ is applied to the model as described above using boundary conditions, a lateral sponge layer, and low-frequency forcing and contains the four components in (17),

$$\boldsymbol{u}_{LS} = \left(\overline{\boldsymbol{U}}_{LS} + \overline{\boldsymbol{u}}'_{LS}\right) + \left(\tilde{\boldsymbol{U}}_{LS} + \tilde{\boldsymbol{u}}'_{LS}\right),$$

with

$$\overline{\boldsymbol{U}}_{LS} = \overline{U},$$
$$\overline{\boldsymbol{u}}'_{LS} = \overline{u}',$$
$$\tilde{\boldsymbol{U}}_{LS} = u_{tide}\frac{\boldsymbol{k}_{tide}}{|\boldsymbol{k}_{tide}|}\cos\left(\boldsymbol{k}_{tide}\cdot\boldsymbol{x} - \omega_{M2}t + \phi_{tide}\right),$$
$$\tilde{\boldsymbol{u}}'_{LS} = u_{iw}f_z\frac{\boldsymbol{k}_{iw}}{|\boldsymbol{k}_{iw}|}\cos\left(\boldsymbol{k}_{iw}\cdot\boldsymbol{x} - \omega_{M2}t + \phi_{iw}\right), \quad (25)$$

where f_z is the mode-1 velocity eigenfunction, $\overline{U}$ = 0.1 ms^{-1}, $\overline{u}' = 0$, u_{tide} = 0.06 ms^{-1}, u_{iw} = 0.50 ms^{-1}, and $\phi_{tide} = \phi_{iw} = 0$. Waves in this system consist of both barotropic internal tides with wavenumber $\boldsymbol{k_{tide}}$ and baroclinic internal waves with wavenumber $\boldsymbol{k_{iw}}$. Waves propagate in the positive x direction ($\theta = 0$) and have a horizontal wavelength, $\lambda_x = 2\pi/k_{iw,x}$ = 130 km, consistent with the stratification and forcing frequency. We apply forcing on the baroclinic flow consistent with mode-1 horizontal oscillations for the imposed stratification solved using the eigenmode problem (Fringer and Street, 2003; Kundu and Cohen, 2008) (Fig. 3c). The mode-1 eigenfunction is approximated as $f_z = \alpha_1\exp\left(-\frac{-z+h_2}{\delta}\right)+\alpha_2$, where $\alpha_1 = -1.4248$, $\alpha_2 = 0.1831$, $\delta = 394.4$ m, $h_2 = 9.995$ m in (25) which is a result specific to the density field, and sponge layer width ΔD = 100 km, damping time scale τ_D = 800 s in (19), and low-frequency forcing time scale and filtering time $\tau_L = \tau_{filt} = 2T_{M2}$ in (21). In the y direction V_{LS} = 0.05 ms^{-1}, and to obtain geostrophic balance in the periodic domain, a forcing $F_{Ly} = -f\overline{u}\hat{t}_r$ is applied in (1) and (5).

The results for the free surface and isopycnal displacement show internal waves emanating from the left wavemaker region and propagating in the positive x direction with steepening leading to the formation of a train of rank-ordered solitary-like internal gravity waves (Fig. 4a, b). These waves are subsequently damped within the right sponge layer. The barotropic velocity ($\overline{U}+U'$) is nearly uniform within the domain and is not influenced by the wavemaker or damping layers (Fig. 4c). The Long2DNoDamp model is twice as long as the Base2D model but with no lateral sponge layers ($\Delta D = 0$) (Table 2), which serves as a reference solution for this domain (*i.e.* freely propagating waves with no damping layers or boundary effects). The results are similar to the Base2D model, but the waves are more nonlinear, exhibiting faster steepening. This is because the sponge layer in the Base2D model is forced with a linear wave (25), which delays wave steepening for x < 100 km in this case. The Long2D model is twice as long at the Base2D model, but does contain lateral sponge layers (Table 2), which serves as a reference solution for the case of no sponge layer damping or boundary at the right side of the domain (x > 420 km). The results are nearly identical to the Base2D model, except in the sponge damping region (420 < x < 520 km). This indicates reflections from the sponge damping region are negligible, an important feature of this framework. The density anomaly and wave velocity fields ($\rho', \tilde{u}'$) for the Base2D model show characteristic internal wave generation consistent with u_{iw} = 0.50 ms^{-1} in (20), and subsequent propagation, steepening, and damping at the model edges (Fig. 4d, e). Thus, the expected behavior of the lateral sponge layer (20)–(21) in wave generation and dissipation is confirmed.

The baroclinic mean flow ($\overline{u}'$) within the thermocline is oriented in the direction of wave propagation in the model interior, but is aligned with the wave propagation in the areas of wave dissipation (Fig. 4e). This trend is reversed in the upper and lower parts of the water column. Rotational dispersion due to Coriolis reduces the number of waves in

Table 2
Nested model run details. Notes: (a) at TC1 location, (b) run in both hydrostatic and nonhydrostatic modes, (c) Long2DNoDamp is identical to Long2D but with $\Delta D = 0$. SCS is South China Sea.

Name	Description	Model dimensions L, W, H_{max} (km)	Grid resolution $\Delta x, \Delta y$ (m)	Vertical levels N_z	Forcing	Time step Δt (s)	Sponge τ_D (s), ΔD (km)	Low-frequency τ_{filt}, τ_L	Boundary conditions (x, y)	Grid lepticity ($\Delta x/H_e$)
Base2D	Base Case 2D Idealized	520, 0.75, 3	250, 250	100	Mean BT, Tidal BT, Wave BC	20	800 100	$2T_{M2}$ $3T_{M2}$	x Dirichlet, y Periodic	1.3
Long2D[c]	Base Case 2D Long Domain Idealized	1040, 0.75, 3	250, 250	100	Mean BT, Tidal BT, Wave BC	20	800 100	$2T_{M2}$ $3T_{M2}$	x Dirichlet, y Periodic	1.3
Sponge2D	Sponge Layer Dynamics	293, 0.75, 3	1000, 1000	50	Wave BC	50	varied	none	x Dirichlet, y Periodic	5
LowFreq2D	Low-frequency Dynamics	260, 0.75, 3	1000, 1000	50	Wave BC	50	800 100	varied	x Dirichlet, y Periodic	5
Base3D	Base Case 3D Idealized	520, 520, 3	Stretched to 500, 500	100	Mean BT, Tidal BT, Wave BC	40	800 100	$2T_{M2}$ $3T_{M2}$	Dirichlet	2.5
3DLateral	3D Idealized Lateral Damping	520, 260, 3	800, 800	100	Wave BC	40	800 100	$2T_{M2}$ none	E,W,S Dirichlet N solid wall	4
SUN2300[b]	SCS Realistic 2300	475, 390, 4.38	2300, 2300	100	NCOM May 26–June 5, 2011	100	1600 25	$4T_{M2}$ $16T_{M2}$	Dirichlet, SW corner Neuman	10[a]
SUN1100[b]	SCS Realistic 1100	475, 390, 4.38	Stretched to 1100	100	NCOM May 26–June 5, 2011	50	1600 25	$4T_{M2}$ $16T_{M2}$	Dirichlet, SW corner Neuman	5[a]
SUN525[b]	SCS Realistic 525	475, 390, 4.38	Stretched to 525	100	NCOM May 26–June 5, 2011	25	1600 25	$4T_{M2}$ $16T_{M2}$	Dirichlet,, SW corner Neuman	2.2[a]
SUN250[b]	SCS Realistic 250	475, 390, 4.38	Stretched to 250	100	NCOM May 26–June 5, 2011	8.0	1600 25	$4T_{M2}$ $16T_{M2}$	Dirichlet SW corner Neuman	1.0[a]
SUNlow	SCS Realistic 525 low	475, 390, 4.38	Stretched to 525	100	NCOM May 26–June 5, 2011, low freq	25	1600 25	$4T_{M2}$ $16T_{M2}$	Dirichlet, SW corner Neuman	1.0[a]
SUNhigh	SCS Realistic 250 high	475, 390, 4.38	Stretched to 250	100	NCOM May 26–June 5, 2011, high freq	8.0	1600 25	$4T_{M2}$ none	Dirichlet, SW corner Neuman	1.0[a]
SUNTcor	SCS Realistic 250 Corrected T	475, 390, 4.38	Stretched to 250	100	NCOM May 26–June 5, 2011, T Corrected	8.0	1600 25	$4T_{M2}$ $16T_{M2}$	Dirichlet, SW corner Neuman	1.0[a]

the train. The case without rotation ($f = 0$) shows development of seven to eight solitons (Supporting Material Figure SM1), rather than the four to five as shown in Fig. 4.

The time-averaged barotropic energy flux $\left\langle \overline{F}_{x0} \right\rangle$ increases with x over the model domain due to the applied geostrophic balance in this periodic domain (Fig. 5a), while its variation in time shows oscillations emanating from the boundaries due to the applied barotropic tides (Fig. 5b). The time-averaged baroclinic energy flux $\left\langle \overline{F}'_x \right\rangle$ is relatively constant near the left boundary, decreases slightly over the model domain with small oscillations developing due to poorly resolved high-wavenumber nonlinearities, and subsequently is damped in the sponge damping layers near the right model boundary (Fig. 5a). The time-varying baroclinic energy flux shows wave energy emanating from the left boundary, wave energy propagation across the model domain and damping along in the left model sponge layer (Fig. 5c). We estimate the phase speed C based on the background density profile ρ_b using the first-mode solution to the eigenvalue problem to estimate the base phase speed C_0 (Fringer and Street, 2003; Kundu and Cohen, 2008), and adjusting for Coriolis effects with $C = C_0 \left[1 + (f/\omega)^2\right]^{1/2}$ (Helfrich, 2007). For the M_2 tidal period in this problem, $C = 3.13$ m s^{-1}, which closely follows the propagation speed of individual waves (Fig. 5c).

3.2. Sponge layer dynamics

We apply a test case similar to the South China Sea base case, but with a model domain length of 293 km and horizontal grid size 1000 m, and vary the damping time scale τ_D and sponge layer width ΔD in (20) to evaluate the effects of wave damping and reflection from the sponge layer (Sponge2D, Table 2). The shorter domain length is used to evaluate the sponge layer effects on linear waves, where the generated waves on the boundary do not have sufficient distance to steepen. We also evaluate the sponge layer effects on nonlinear waves by comparing the results of the Base2D model with the Long2D model.

We define the reflection coefficient as a measure of the amount of wave energy reflected by the sponge layer and solid boundary, and is given by $R_F = 1 - \left\langle \mathcal{F}' \right\rangle_{x=L-\Delta D} / \left\langle \mathcal{F}' \right\rangle_{\infty}$, where $\left\langle \mathcal{F}' \right\rangle_{\infty}$ is the wave energy flux in the absence of a sponge layer or reflections at $x = L - \Delta D$. Because $\left\langle \mathcal{F}' \right\rangle$ is a bulk parameter integrating wave energy flux propagating in both the positive and negative directions, $\left\langle \mathcal{F}' \right\rangle$ decreases when wave energy is reflected from the sponge layer. We normalize τ_D by the travel time for a wave to traverse the sponge layer $\Delta D/C$. For $\tau_D < 0.02C/\Delta D$, reflections are significant due to reflected wave energy from the sponge layer itself (Fig. 6). For $\tau_D > 0.08C/\Delta D$, reflections can occur from the model boundary if ΔD is not large enough to dissipate all wave energy. While the parameter space is covered for linear waves, the Base2D model with nonlinear waves also follows a similar trend with near zero R_F (Fig. 6). Thus, while the sponge layer performed well over

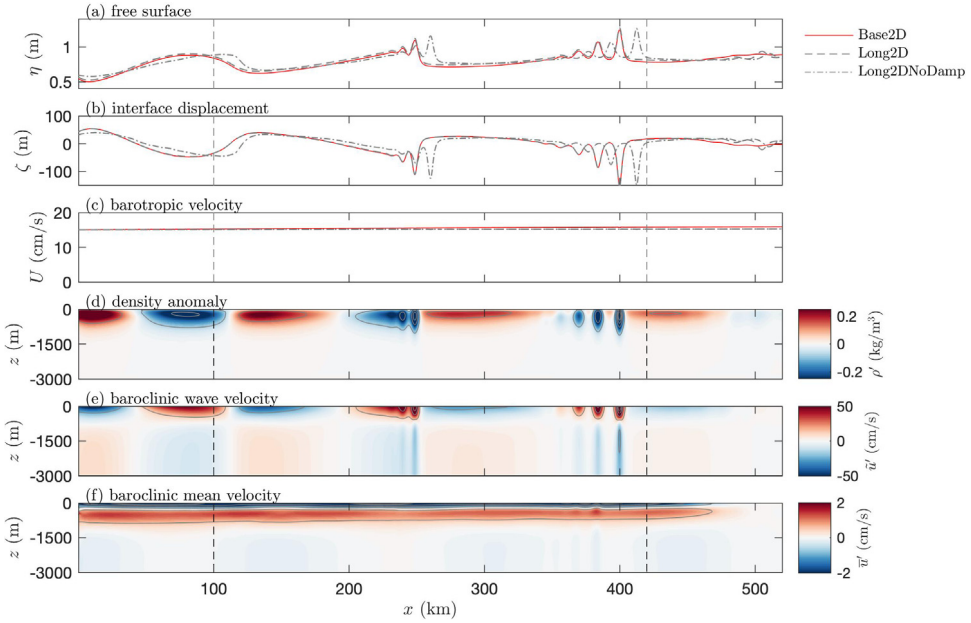

Fig. 4. Example mean flows and mode-1 internal wave propagation and steepening for conditions typical of South China Sea, for different model configurations showing (a) free surface η, (b) displacement ζ of the isopycnal initially located at $z = -381$ m, (c) barotropic velocity U, and Base2D model (d) baroclinic density anomaly ρ', (e) baroclinic wave velocity $\tilde{u}'$, and (f) baroclinic mean velocity $\bar{\tilde{u}}$. Wave forcing is from left to right with sponge layers on the edges of the domain indicated by the vertical dashed lines.

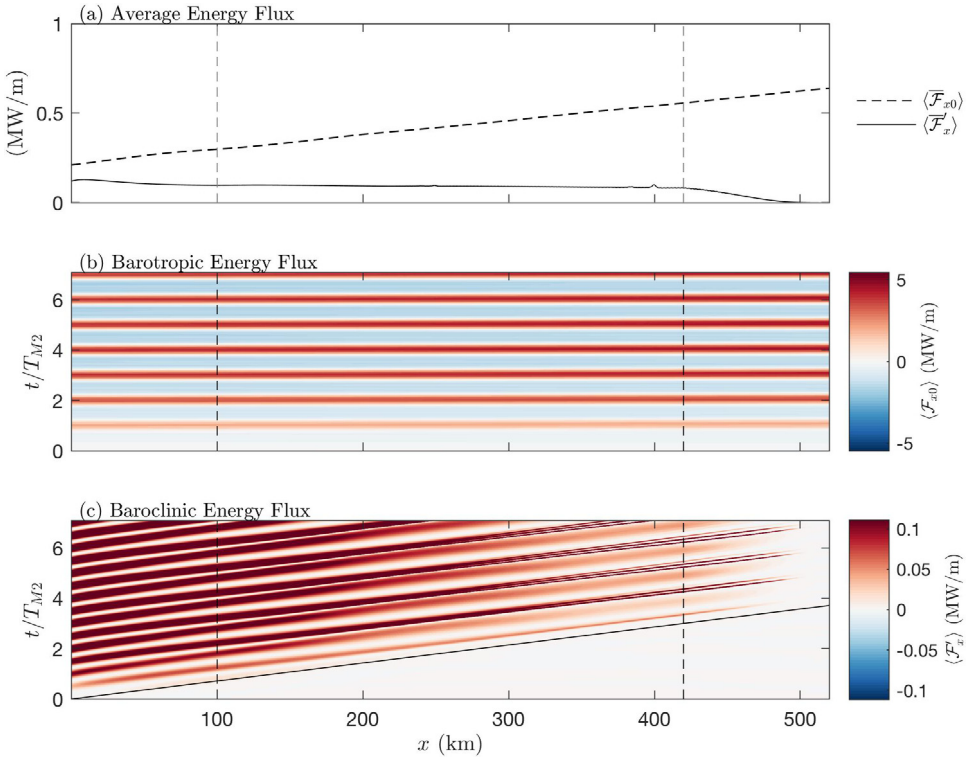

Fig. 5. Energy flux for idealized South China Sea case. (a) time-averaged energy flux for barotropic $\langle \overline{F}_{0x} \rangle$ and baroclinic $\langle \overline{F}'_x \rangle$ flows after spin up, (b) barotropic energy flux variation with time, and (c) baroclinic energy flux variation with time. Solid black line in (c) is the mode-1 phase speed C based on the background density profile.

a range of values, to minimize reflections from both mechanisms, the optimal values appear to be $\tau_D = 0.04C/\Delta D$, and $\Delta D = \lambda/2$. Within this framework, waves with wavelengths smaller than λ (higher modes) will be filtered, while waves with wavelengths larger than λ (lower modes) will not. A larger value of $\Delta D = \lambda$ has a broader region of low R_F, and would be more applicable to a broad spectrum of wave frequencies, but requires more computational space used for the damping region. These results of competing reflection mechanisms are qualitatively similar to wave damping theory for surface gravity waves (Zhang et al., 2014).

A second test case is developed using the same setup as the base case, but with wave energy flux coming from both boundaries with the same period but different magnitudes (wave velocities on the right boundary are half as large as those on the left). The incoming wave energy flux is effectively generated in the model at both boundaries, and the outgoing wave energy flux is efficiently damped (Supp. Mat.

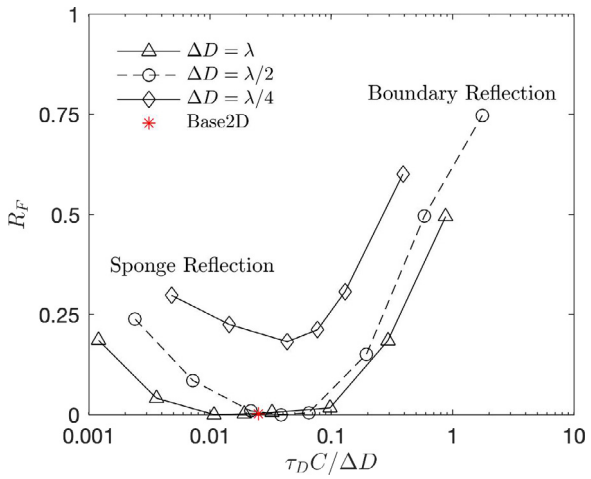

Fig. 6. Sponge region dynamics showing wave energy flux reflection coefficient R_F for linear (black) and nonlinear (red) waves as a function of damping time scale τ_D normalized by the travel time within the damping layer $\Delta D/C$. Low τ_D results in reflections from the sponge layer itself, while high τ_D can result in reflections from the boundary if damping distance ΔD is not long enough relative to the wavelength λ.

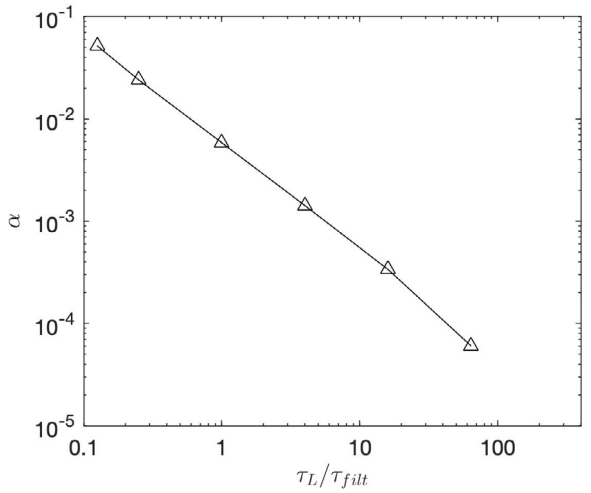

Fig. 7. Internal wave energy flux attenuation coefficient α as a function of low-frequency forcing time scale τ_L normalized by the low-pass filter time scale τ_{filt}. α is the rate of energy loss over one wavelength of propagation λ.

Figure SM2). Additional test cases varying the wave period (T_{iw}) and magnitude (u_{iw}) produced similar results. This demonstrates that the sponge layer formulation is robust to efficiently damp outgoing waves while allowing incoming waves to propagate into the domain over a wide range of wave conditions.

3.3. Low-frequency forcing variables

The primary parameter governing the low-frequency forcing is the time scale τ_L. To test how τ_L affects the high-frequency waves, we modify the base case to have a 260 km domain with 1000 m horizontal grid spacing (Table 2), setting all velocities to zero except the wave velocity $U_{iw} = 0.5$ ms^{-1}, propagating in the positive x direction. We then vary τ_L in (22) and record the wave energy flux one wavelength into the domain. The wave attenuation coefficient is $\alpha = 1 - \langle \mathcal{F}' \rangle_{x=L-\Delta D} / \langle \mathcal{F}' \rangle_{\infty}$, and $\langle \mathcal{F}' \rangle_{\infty}$ is the wave energy flux without low-frequency forcing ($\tau_L = \infty$). The expression for α is similar for the reflection coefficient R_F, but in this case is a measure of wave energy damping since all wave energy flux is in the positive x direction and there are no reflections from the damping layer or boundary. Wave attenuation α decreases as τ_L increases, and for $\tau_L/\tau_{filt} > 1$, $\alpha < 0.01$ (Fig. 7). The total propagation distance must be considered when setting τ_L, since α here is the damping over one wavelength. In practice, τ_L should be as large as possible so as not to influence the wave energy, but small enough to control the low-frequency flow.

3.4. Internal waves & mean flow in three dimensions

We extend the base case model to three dimensions in a 520 × 520 × 3 km (x, y, z) domain with a refined stretched grid in the model exterior sponge damping region ($\Delta D = 100$ km, $\tau_D = 800$ s) and a constant interior horizontal resolution of 500 m (Table 2). Dirichlet boundary conditions are applied on all four boundaries. Forcing parameters are the same as the base case, except the internal waves and barotropic tide are oriented at $\theta = 45°$. The barotropic mean flow $\overline{U}$ is spatially constant and the mean free surface $\overline{\eta}_0$ is tilted consistent with geostrophic flow (Fig. 8a). The barotropic tide $\tilde{U}$ is directed at an angle $\theta = 45°$ with the tidal surface $\tilde{\eta}_0$ consistent with a Poincaré wave with rotation (Fig. 8b). Both $\overline{U}$ and $\tilde{U}$ and associated components of the free surface show little sign of interaction with the sponge layer in the outer damping region. The wave velocity $\tilde{u}'$at the surface and wave component of free surface $\tilde{\eta}$ show internal waves emanating from the south and western boundaries, propagating and steepening towards the upper right and dissipating along the top and right boundaries (Fig. 8c). The wave velocity $\tilde{u}'$ and the interface displacement ζ (defined as deviation of the isopycnal originating at 366 m depth) of O(100 m) show similar trends (Fig. 8d). These results show that the model framework can effectively simulate a range of flow conditions in an idealized, three-dimensional environment with several complex flow elements.

An important consideration in the three-dimensional framework is a lateral damping effect which occurs when waves propagate parallel to a sponge layer boundary (Fig. 9). We illustrate this effect with waves propagating parallel to a sponge layer and a solid wall on the opposite boundary (3D Lateral model, Table 2). The waves are distorted near the edge of the sponge layer (Fig. 9b), which leads to a slowed propagation speed and loss of amplitude when compared to a two-dimensional simulation with the same resolution (Fig. 9a). The two-dimensional simulation serves here as a reference solution for waves propagating without lateral effects. Thus, there is a buffer region of influence ΔB, which equals approximately one wavelength for waves propagating parallel to a sponge layer. When waves are propagating normal to a boundary, this effect does not occur, assuming τ_D and ΔD are selected in a way that eliminates reflections.

4. Realistic model application test case

4.1. Model setup

The previous section showed how the model framework is implemented in two and three dimensions under a range of idealized flow conditions. We now apply the framework to a realistic application in the South China Sea and nest SUNTANS within the US Navy NCOM model (Martin, 2000) (Fig. 10a). The large-scale NCOM model is implemented here at 1/50° horizontal resolution (approximately 2.3 km), with 41 vertical sigma/z-levels with 12 sigma levels on the shelf, and hourly output. The model domain location and 475 x 390 km extent is chosen because of the strong internal tides over the Luzon straight, which propagate westward and steepen and are well-documented in the literature (Alford et al., 2015; Farmer et al., 2009; Lien et al., 2014; Ramp et al., 2010; Zhang et al., 2011). The SUNTANS model is run with 100 stretched vertical z-levels with a stretched grid in the model exterior sponge damping region and a constant horizontal resolution in the interior (Fig. 10b, c). To investigate the effects of grid resolution in the SUNTANS model, interior horizontal resolutions of 250, 525, 1100

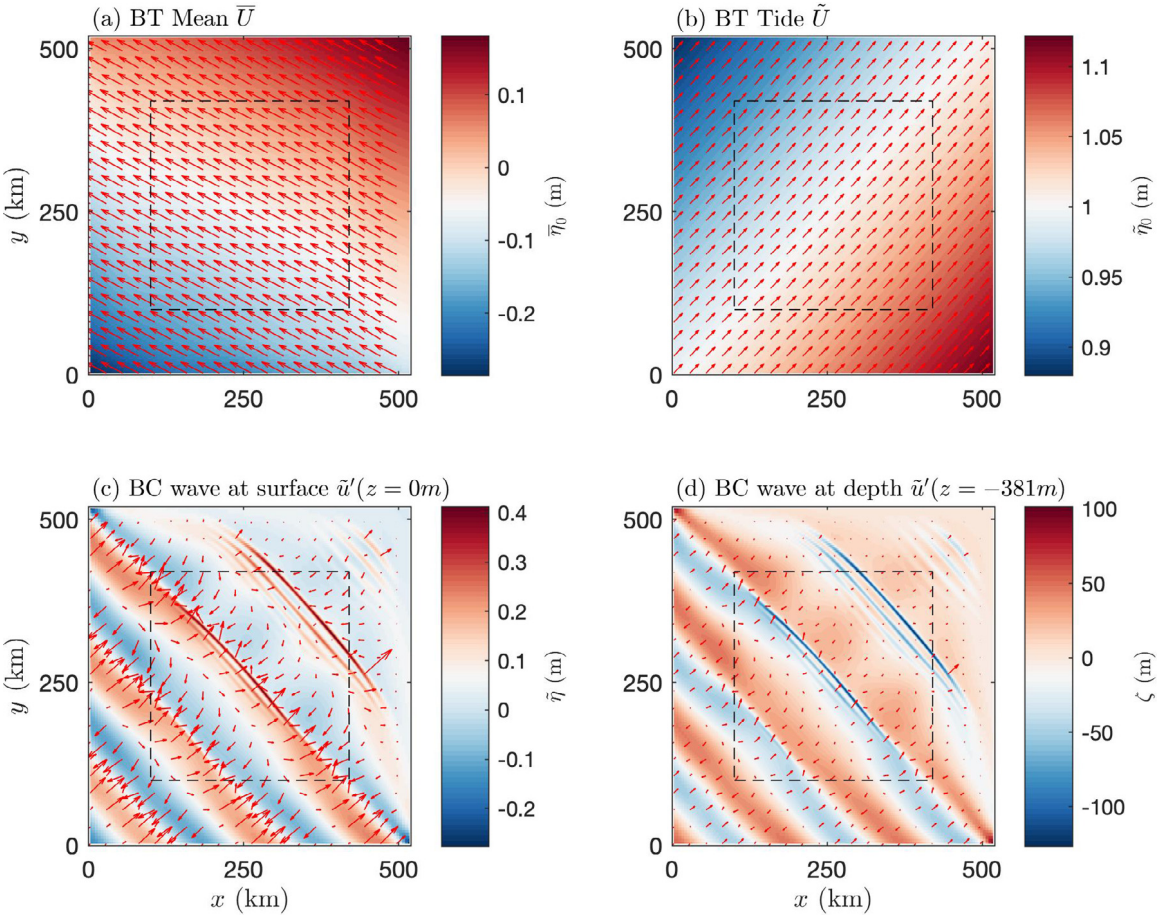

Fig. 8. Idealized three-dimensional simulation showing barotropic flow and internal wave propagation. (a) Barotropic mean flow and free surface, (b) barotropic tidal flow and free surface, (c) baroclinic wave flow at surface and free surface, and (d) baroclinic wave flow at depth and interface displacement. Velocity scale (red vectors) for (a, b) is 10 cm s^{-1}, velocity scale for (c, d) is 50 cm s^{-1}. Dashed black lines are extent of sponge damping layer.

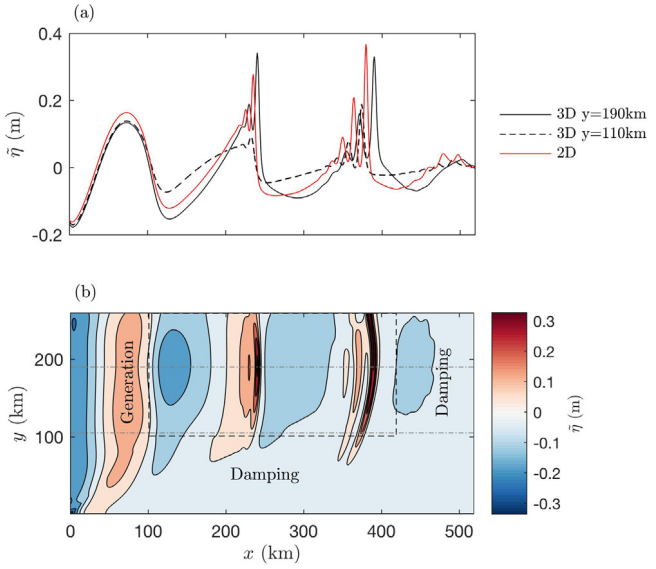

Fig. 9. Lateral effects on wave propagation parallel to sponge layer comparing (a) free surface profiles of three-dimensional simulations along two transects to a two-dimensional simulation, and (b) three-dimensional view of lateral wave effects. In panel (b), black (–) is extent of sponge layer, gray (-.) is location of profiles in panel (a). In panel (b), East, West, and South boundaries are Dirichlet, North boundary is solid wall.

and 2300 m are employed (Table 2). To apply the boundary conditions and sponge layer, the $\hat{F}$ operator (29) is computed from $\overline{\langle \boldsymbol{F}' \rangle}_{LS,lp}$ which uses a low-pass filter in time to obtain the energy flux over multiple waves, and spatially filtered at a scale of the domain size W to smooth out variability at scales smaller than λ (Fig. 10d). The model is run during May 26 to June 5, 2011, a time period when observational data is available from the TC1 mooring in a depth of 450 m (Lien et al., 2014) (Fig. 10).

Due to bathymetric differences and approximations in the spatial and temporal interpolations from NCOM to the SUNTANS open boundaries, volume is not exactly conserved in the model domain. Thus, to ensure long-term volume consistency, several boundary grid cells in a corner of the domain are specified as Neumann velocity boundary conditions with a specified free surface $\eta_{lp} = \overline{\eta}_0 + \tilde{\eta}_0$ which is the instantaneous spatially low-pass filtered free surface from the NCOM model. This eliminates wave effects but retains large-scale dynamics, including the barotropic tide. For this example, these points were chosen as the three grid cells in the southwest corner of the domain as they are within the wave damping region (Fig. 10d). In this way, waves generated by small mismatches in η_{lp} and η are damped.

The SUNTANS model is initialized with the NCOM temperature, salinity, and spatially-smoothed free surface output, and initial velocities are set to zero. Kinematic viscosity is set to $\nu = 10^{-4}$ m^2/s in the vertical direction and $\nu_H = 1$ m^2/s in the horizontal direction, similar to values used in other studies at this scale (Kang and Fringer, 2012; Zhang et al., 2011). Boundary conditions are interpolated at hourly time intervals. Model spinup is approximately seven days for the barotropic tides to equilibrate to the large-scale model from initial

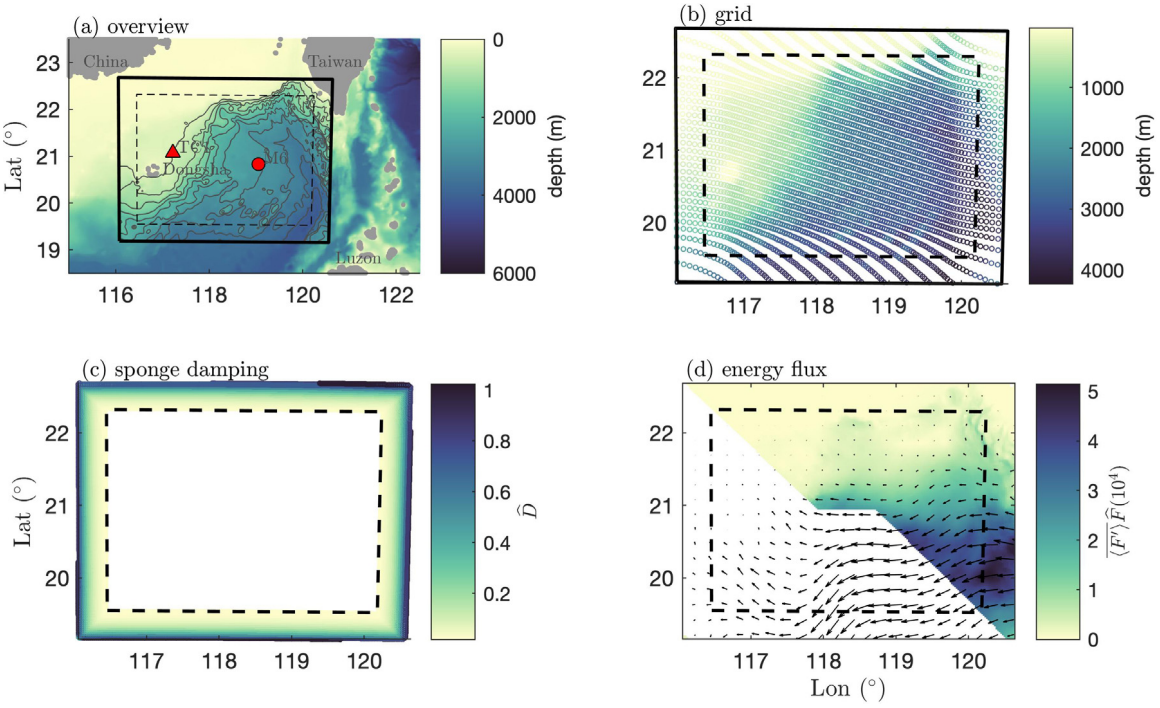

Fig. 10. South China Sea example nested grid. (a) Bathymetry overview showing the SUNTANS model extent (-) and sponge boundary (–) and validation points TC1 (Δ) and M6 (o), (b) bathymetry on the SUNTANS grid showing every 30 points for clarity, (c) sponge damping term $\hat{D}$, (d) example low-pass filtered wave energy flux $\overline{\langle F' \rangle}$ (vectors) and $\hat{F}$ operator activated for NCOM wave energy entering the domain. Within the damping region, colored areas nudge waves to NCOM, and white areas nudge waves to zero (damping).

velocities set to zero (Supp Mat Figure SM3). The results from initializing the model with interpolated velocities show initial high-frequency waves from inconsistencies in the model solution, and a slightly faster equilibration to the barotropic tides of approximately five days (Supp. Mat. Figure SM3). Additional testing with 15 min boundary conditions gave nearly identical results as the 1 h boundary conditions. Based on these results, the hourly boundary condition forcing and zero velocity initial conditions were chosen, and we focus on results of the last three days of the ten-day simulations for the remainder of the runs.

4.2. Validation

At station TC1, the different models show small differences in background density, but large differences in buoyancy frequency (Fig. 11a, b). In general, the observations have a higher buoyancy frequency than the NCOM and SUNTANS models (the SUNTANS model is initiated with NCOM stratification) (Fig. 11a). We approximate the linear mode-1 phase speed C for these average profiles using the eigenvalue solution adjusted for frequency as in the Base Case. The results show the phase speed for both K_1 and M_2 periods is lower for the NCOM and SUNTANS models compared to the observations (Fig. 11c).

The results for η and T at 140 m depth show propagation and steepening of the internal tide from east to west and steepening of the internal waves with propagation (Fig. 12). Surface deflections from the internal waves are $O(0.1$ m), and internal temperature deflections are $O(2$ °C) consistent with previous simulations of this region (Zhang et al., 2011). Outgoing internal wave motions were absorbed in the sponge layers and there is no observable internal wave reflection (Supp. Mat. Figure SM4).

The results of the SUNTANS simulations at a coarse (2300 m) and fine (250 m) grid resolution are compared to both aerial imagery from MODIS [NASA Worldview application (https://worldview.earthdata.nasa.gov/) operated by the NASA/Goddard Space Flight Center ESDIS project], and the NCOM model on June 4, 2011 (Fig. 13a, b, c, d). The wave crest from the NCOM result lags behind the MODIS data due to lack of resolution of nonlinear effects which lead to slower wave propagation. The SUNTANS result matches MODIS better due to improved resolution of nonlinear effects. As discussed in Vitousek and Fringer (2011), coarse resolution gives rise to numerical dispersion which produces the unphysical wave train in NCOM (Fig. 13e). The dispersion is weaker in the SUNTANS model with coarse resolution (Fig. 13f), and is eliminated with higher resolution (Fig. 13g).

The results are also compared to observational mooring data at location TC1 (Fig. 10a) (Lien et al., 2014). The barotropic velocities from the observations compare well to the model results for NCOM and SUNTANS especially in the EW direction (Fig. 14a), while in the NS direction capture the long-term average flow, but miss some of the M_2 tidal flow (Fig. 14b). This mismatch in the NS direction is potentially a result of inaccuracy in the NCOM model, whose results are very similar to the SUNTANS model. Temperature observations show internal waves passing by the TC1 mooring with both broad waves and sharply nonlinear wave trains (Fig. 14c). The NCOM model captures some of the variability, but the steep waves lag the observational result by several hours (Fig. 14d). The SUNTANS model results show increasing variability and wave steepness with increasing grid resolution, and more closely approximate the wave timing when compared to the observations although still with some lag (Fig. 14e, f).

Observed and modeled power spectra of wave velocities 100 m below the surface at Station M6 in 3000 m depth are relatively similar between the NCOM, SUN2300, and SUN250 models (Fig. 15). This indicates that the nested SUNTANS model is correctly reproducing the spectrum of wave velocities on the eastern boundary, which are propagating primarily towards the west into the domain. The spectra show peaks at the inertial period f, the K_1 and M_2 tidal periods as expected for this system. At the TC1 mooring in 450 m of water, the modeled and observed wave velocity spectra 100 m below the surface are similar for periods near the K_1 and M_2 tidal periods (Fig. 15). However, for frequencies higher than one hr^{-1}, the observations show significantly more energy, a trend which is also apparent in the temperature profiles which show higher richness than the model results (Fig. 14). The high-resolution SUNTANS model with 250 m resolution produces roughly 100 times the energy at high frequencies than the 2300 m results, but it is still 100 times lower than the observations.

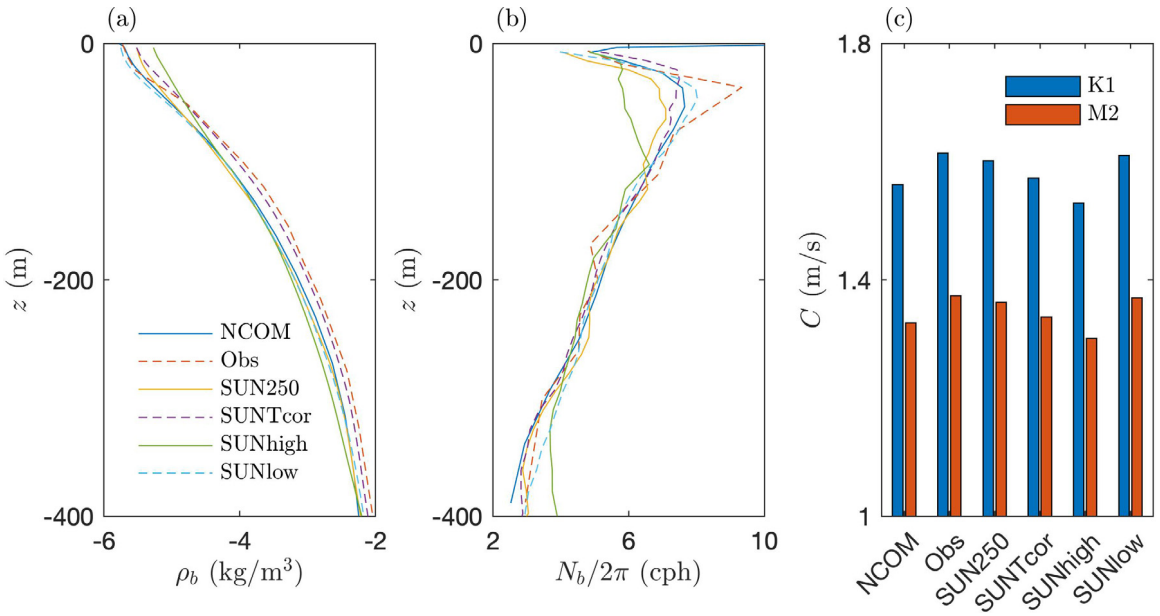

Fig. 11. Background density profiles and its effect on propagation speed at station TC1. (a) background density ρ_b, (b) buoyancy frequency N_b profiles over three-day average for models and observations, and (c) resulting propagation speed C of Mode-1 internal wave for K_1 and M_2 tides.

The energy flux is compared at sites M6 and TC1 (Fig. 16). The energy flux here is computed as $\langle \mathcal{F}' \rangle = \int_{-h}^{0} \left(u'_H p' \right) dz$, where p' is calculated from the density deviation from the time-averaged background state ρ_b. The results at M6 in deeper water show some differences in the time variable energy flux. However, the time-integrated energy flux is similar for the NCOM model and the SUN2300 model indicating the boundary forcing adequately injects the total energy into the nested domain (Fig. 16a, c). At the TC1 site, the wave energy flux is quite variable in time between the models due to differences in timing (Fig. 16b). However, in the time-integrated sense, the energy flux is under-predicted by the NCOM model, over-predicted by the low-resolution SUNTANS model, and well-predicted by the high-resolution SUNTANS model (Fig. 16d).

The amplitude and phase of the waves on June 3, 2011, based on the displacement of the 20 °C isotherm at site TC1, are compared between the observations, NCOM model and SUNTANS model (Fig. 17). The NCOM model underpredicts the amplitude compared to the observations, while the SUNTANS model more closely approximates the observation with increasing grid resolution (Fig. 17a). The phase lag between the NCOM model and the observations is approximately five hours, while the SUNTANS model shows similar phase lag for different resolutions (Fig. 17b).

To evaluate the effect of different boundary and stratification parameters on the model results, a series of runs were conducted including the SUNlow, SUNhigh, SUNTcor, and SUNnof models (Table 2). The SUNlow model has only low-frequency forcing with no internal tide forcing ($\tilde{u} = 0$) for all boundary conditions. The SUNhigh model has only high-frequency forcing such that low-frequency nudging is off ($F_L = 0$) with no low-frequency velocities in the model boundary conditions ($\overline{u} = 0$). The temperature profile of the observations shows denser isotherm spacing and a larger overall vertical temperature difference compared to the NCOM and SUNTANS models (Fig. 14), which could potentially affect the internal wave propagation. To evaluate the effects of the difference in average stratification between the NCOM model and the observational data at TC1, the SUNTcor model adjusts the temperature in the initial condition and boundary conditions by a constant offset $\Delta\overline{T}(z)$ over the model domain, computed as the difference between the average temperature over the observation period and the NCOM model at site TC1, i.e. $\Delta\overline{T} = \overline{T}_{Obs} - \overline{T}_{NCOM}$.

As shown in Fig. 18a, the results for the barotropic velocities for the SUN250 model are in good agreement with the observations, while the SUNlow model contains the mean flows but no tidal oscillations, and the SUNhigh model contains the tidal oscillations but no mean flow (Fig. 18a). The temperature profile of the SUN250 model compares well with the observations (Fig. 18b, c), although, as expected, the SUNlow model temperature profile contains very few wave oscillations (Fig. 18d). In the absence of low-frequency forcing, the SUNhigh model results are similar to the SUN250 model results, but with some lag in the wave arrival times (Fig. 18e). The SUNTcor model shows higher amplitude waves compared to the SUN250 model, and these waves are closer in amplitude to the observations shown in Fig. 18b, indicating that the initial stratification is important at this location.

5. Discussion

5.1. Nonhydrostatic effects

An important consideration in systems with high-frequency internal waves is the importance of nonhydrostatic pressure. Under certain circumstances, internal waves can represent a dynamical balance between nonlinearity and nonhydrostasy (dispersion), and thus may require computationally expensive nonhydrostatic simulations to be well-resolved. Model results at various grid resolutions show little effect of including the nonhydrostatic pressure at low grid resolution, but differences are most pronounced at high resolution in both wave timing and amplitude (Fig. 19). Following Vitousek and Fringer (2011), to lowest order, the ratio of numerical to physical dispersion is $K\lambda_G^2$, where K is a constant dependent on the discretization of the governing equations and $\lambda_G = \Delta x / H_e$ is grid lepticity, where Δx is the horizontal grid spacing and H_e is the effective depth of the internal interface and depends on the normal modes solution for a continuous stratification, although it is given roughly by the upper mixed-layer depth (Fig. 20c). In simulations of a freshwater lake, a grid lepticity of less than roughly two appears to be required to correctly resolve nonlinear internal waves (Dorostkar et al., 2017), a result which is similar to our findings. Over the South China Sea model domain, the grid lepticity is smaller within the interior of the domain where the water is deeper and the grid resolution is finer (Fig. 20a, b). Over the entire model domain, the grid lepticity is less than two for 90% of the grid cells at the highest resolution (SUN250), while the lepticity is larger than two over most of the domain for the coarsest resolution (SUN2300). This explains

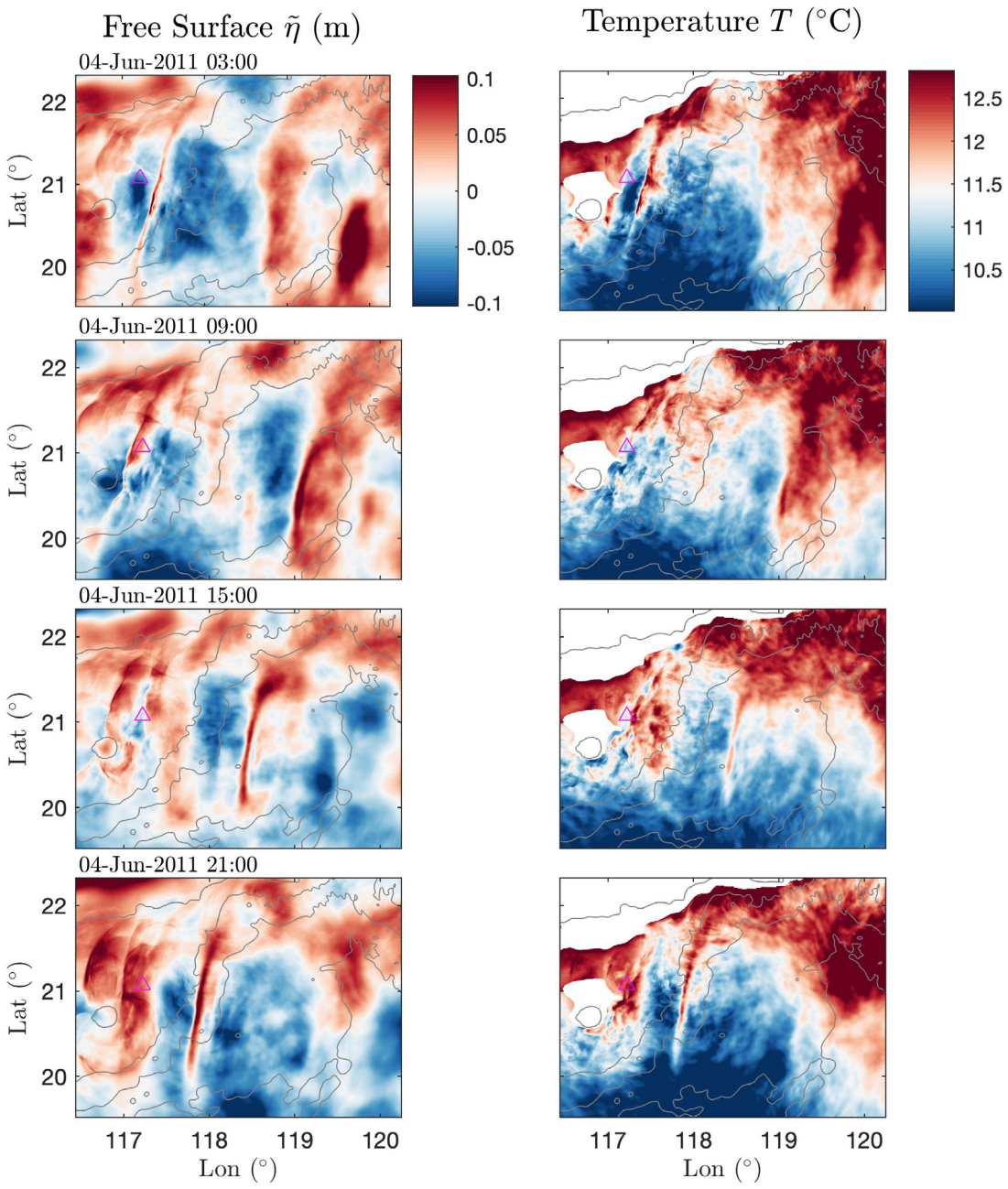

Fig. 12. Evolution of free surface (left) and temperature at 140 m depth (right) at six hour intervals due to internal tide propagation in the SUNTANS model with 250 m horizontal grid resolution (SUN250). Extent is within model interior, exterior sponge layers not shown. Triangle is TC1 mooring, gray (-) show 1000 m depth contours.

why the nonhydrostatic and hydrostatic solutions are nearly identical for the low resolution model SUN2300 (Fig. 19b), and significantly different for the high resolution model SUN250 (Fig. 19e). In addition, the modeled wave energy flux is better approximated at lower grid lepticities (Fig. 16d). Thus, if nonhydrostatic effects are important in the model, creating a grid lepticity less than two is important for correctly resolving the nonlinear wave effects and wave energy flux.

5.2. Frequency effects

The model framework contains several important time scales, the most important being the filtering frequency used to separate high- and low-frequency motions. This frequency should be selected such that it represents an area of relatively low energy. As in this South China Sea example, the value of τ_{filt} was selected to be less than the inertial period (Fig. 15). Thus, in this example, inertial and tidal motions are deemed high-frequency motions, and longer period processes such as mesoscale eddies are deemed low-frequency motions.

To analyze the effect of the forcing terms on the high- and low-frequency components, we time average the momentum equation (1) and decompose variables into their high $(\tilde{})$ and low-passed $(\bar{})$ components. Ignoring viscous effects, the equation for low-frequency motions given by,

$$\frac{\partial \overline{\boldsymbol{u}}}{\partial t} + \overline{\boldsymbol{u}} \cdot \nabla \overline{\boldsymbol{u}} + \overline{\tilde{\boldsymbol{u}} \cdot \nabla \tilde{\boldsymbol{u}}} = -2\overline{\boldsymbol{\Omega} \times \boldsymbol{u}} - \frac{1}{\rho_0} \nabla \overline{p} - \frac{g}{\rho_0} \overline{\rho} \hat{k} + \boldsymbol{F}_L, \tag{26}$$

where $\overline{F}_M = \overline{F}_L + \overline{F}_D = \overline{F}_L = F_L$ from (5), since the damping term contains the high-frequency forcing so that $\overline{F}_D = 0$. Taking (1) and substituting $\boldsymbol{u} = \overline{\boldsymbol{u}} + \tilde{\boldsymbol{u}}$ and subtracting (26) gives the equation for the high-frequency motions,

$$\frac{\partial \tilde{\boldsymbol{u}}}{\partial t} + \tilde{\boldsymbol{u}} \cdot \nabla \overline{\boldsymbol{u}} + \overline{\boldsymbol{u}} \cdot \nabla \tilde{\boldsymbol{u}} + \left(\tilde{\boldsymbol{u}} \cdot \nabla \tilde{\boldsymbol{u}}\right) = -2\boldsymbol{\Omega} \times \tilde{\boldsymbol{u}} - \frac{1}{\rho_0} \nabla \tilde{p} - \frac{g}{\rho_0} \tilde{\rho} \hat{k} + \boldsymbol{F}_D. \tag{27}$$

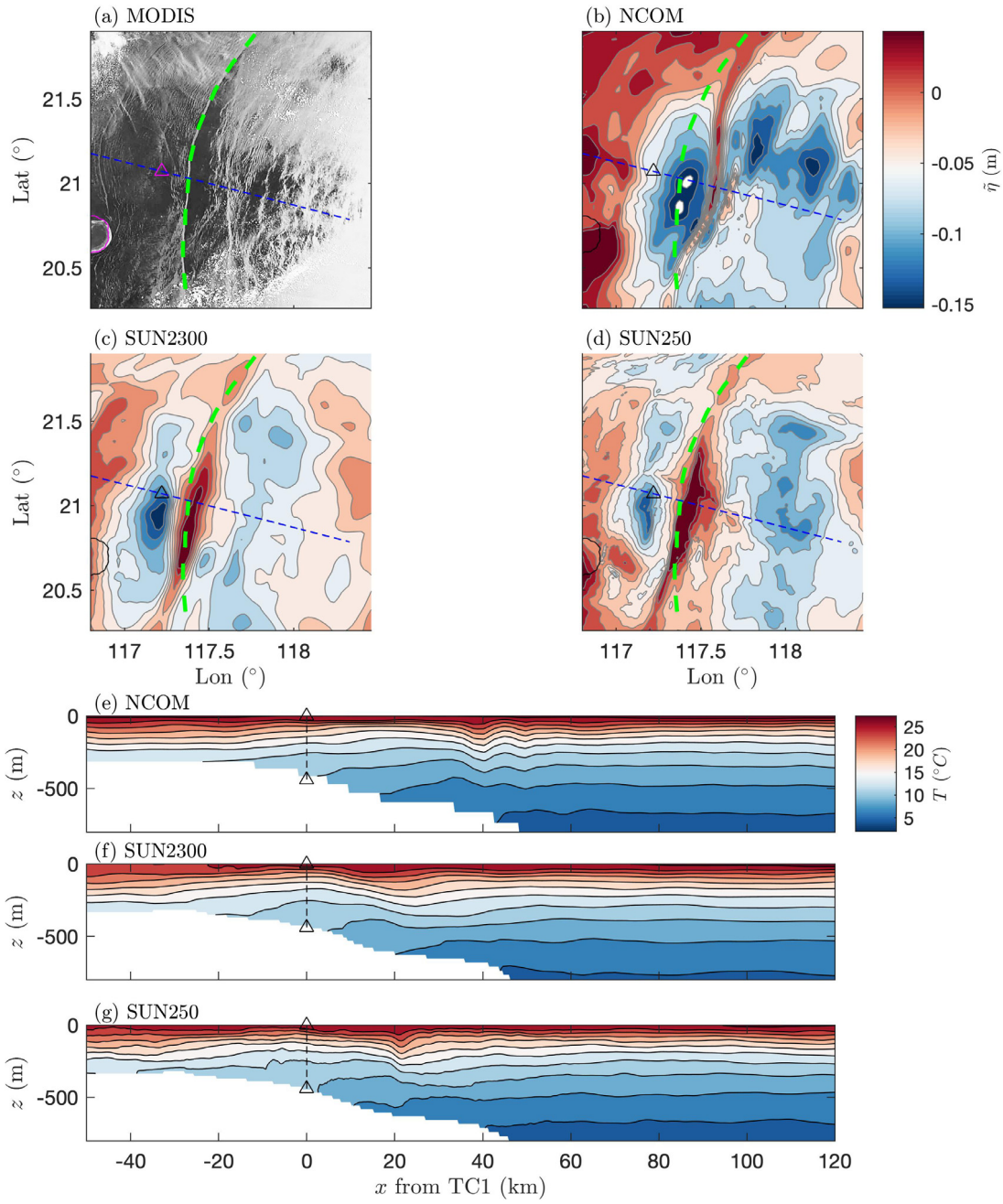

Fig. 13. Comparison of aerial imagery, free surface and temperature profiles at 3:00 on June 4, 2011 GMT comparing (a) MODIS aerial imagery, (b) free surface $\tilde{\eta}$ from NCOM, and (c) free surface from SUNTANS with (c) 2300 m and (d) 525 m grid resolution. Spatial comparison of temperature T from (e) NCOM, SUNTANS with (f) 2300 m and (g) 250 m grid resolution. In (a)–(d), circle is 50 depth contour on Dongsha Atoll, dashed green is wave front. TC1 mooring is shown by triangle in all plots. Blue dash in (a)–(d) is location of profiles in (e)–(g). MODIS imagery from the NASA Worldview application operated by the NASA/Goddard Space Flight Center Earth Science Data and Information System (ESDIS) project.

Eqs. (26) and (27) show that F_L is forcing the mean flow $\bar{u}$, while F_D acts to force and damp the high-frequency flow $\tilde{u}$. However, the nonlinear terms in (26) and (27) are capable of forcing low-frequency flows from the wave motions, and similarly the mean flows can interact with wave flows to force (or damp) wave flows. These effects can be seen in the comparison of runs with full forcing (SUN250), low-frequency only (SUNlow), and high-frequency only (SUNhigh) (Fig. 18c, d, e). The SUNlow model contains primarily low-frequency motions, but some higher frequencies are still present from waves generated within the model domain. The SUNhigh model contains much of the variability related to the full forcing but the timing of the internal waves is delayed compared to the observations.

Internal wave properties such as phase speed, damping, and nonlinear steepening are very sensitive to the background stratification (which we define as the time-averaged density, $\rho_b = \overline{\rho} - \rho_0$), and the buoyancy frequency N_b. In the South China Sea example at station TC1, the different models show smaller differences in background density but large differences in buoyancy frequency (Fig. 11a, b). In general, the observations have a higher buoyancy frequency than the NCOM and SUNTANS models (Fig. 11a). The SUNTANS models show decreased phase speeds for the three models with internal waves present (SUN250, SUNhigh, SUNTcor) a result of the changes to the density and buoyancy frequency profiles due to mixing near the surface (Fig. 17). In addition, nonlinear solitons have an amplitude-dependent

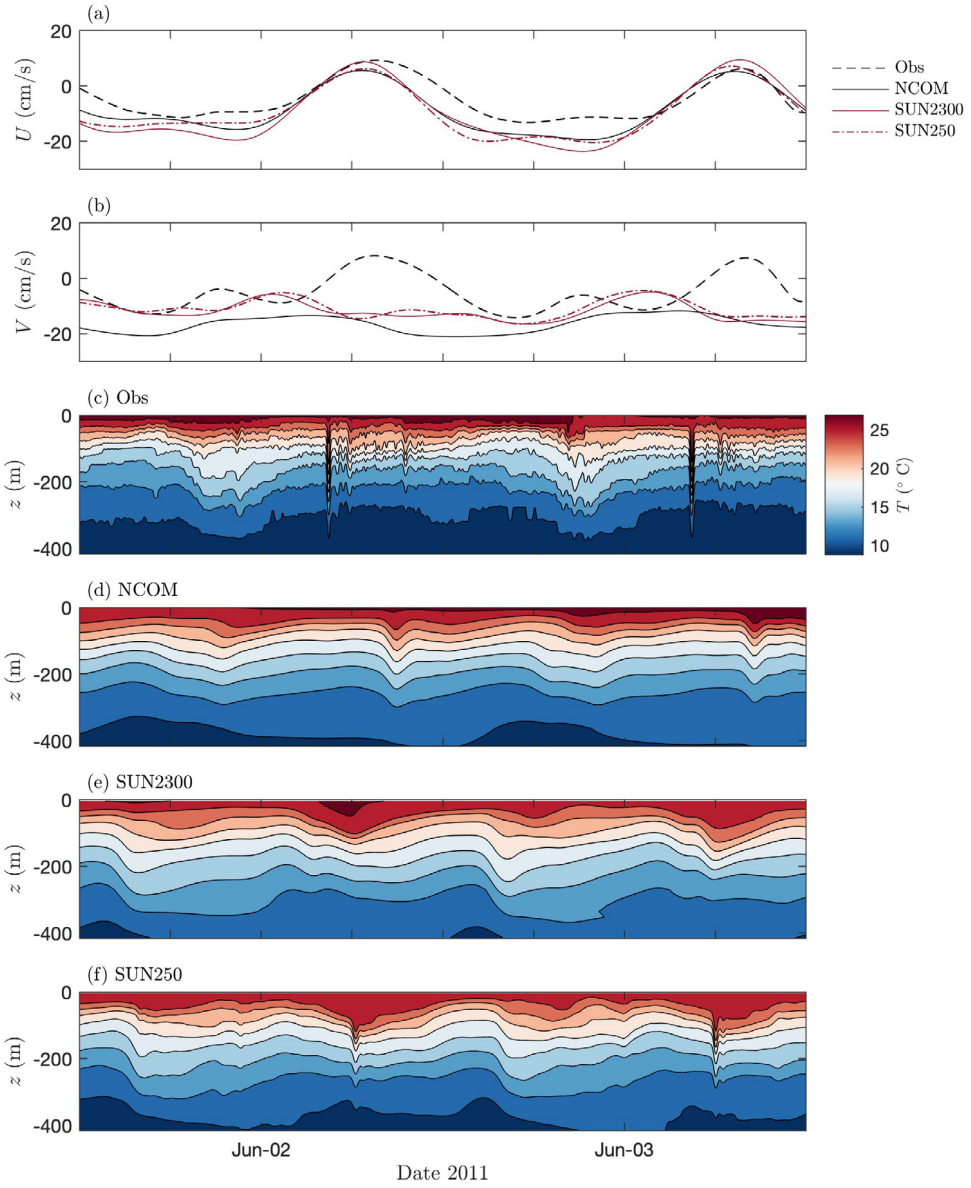

Fig. 14. Model validation of velocity and temperature at site TC1 comparing results from observations (Lien et al., 2014) (Obs), NCOM model results, and SUNTANS results (SUN2300, SUN525). Barotropic filtered velocities in (a) U and (b) V directions. Temperature and baroclinic velocities for (c) observations, (d) NCOM model, and SUNTANS at (e) 2300 m and (f) 250 m resolution. In panels (a) and (b), U and V are 6-hour low-pass filtered.

phase speed, which likely contributes to the observed phase lag for models with lower amplitude waves (Rayson et al., 2018).

These results indicate that correctly approximating the density profiles appears to be critical to modeling internal waves. In this case, the large-scale NCOM model temperature profile was under-predicting the density and buoyancy frequency, which leads to an underprediction of the phase speed. The nested SUNTANS model inherits this under prediction of density and buoyancy frequency through the initial and boundary conditions. Added to this are mixing effects, which are most pronounced at the surface. In order to correct for this difference, a heat model could be added to the SUNTANS model to account for heat fluxes at the surface. Additionally, the low-frequency temperature profile could be nudged using F_T in (2).

The approach proposed here is based on integrated energy flux and allows for only one mean direction of energy flux propagation θ at each grid cell which is applied at the boundaries and damping layer [*i.e.*, $\hat{F}(x,y,t)$ and $\tilde{\boldsymbol{u}}'_{LS}(x,y,z,t)$]. This approach is expected to work well under general oceanic conditions with a well-defined mean direction. However, for cases of strong opposing wave directions of the same frequency, or complex energy flux directed in multiple directions, this approach may not perform well. In this case, an alternate approach would be to decompose the energy flux into frequency ω and direction θ [i.e., $\hat{F}(x,y,t,\omega,\theta)$ and $\tilde{\boldsymbol{u}}'_{LS}(x,y,z,t,\omega,\theta)$], but this approach would likely require a very large computational effort.

By forcing the barotropic velocities to match the large-scale flow at the SUNTANS model boundary (18), there is the potential for barotropic wave reflections off the model boundaries. The velocities perpendicular to the boundary are nearly identical (E/W boundary $\tilde{U}$ and N/S boundary $\tilde{V}$) between the SUNTANS and NCOM models (Supp. Mat. Figure SM5). In the direction parallel to the boundary, there are some small mismatches between SUNTANS and NCOM, most likely arising from free surface waves. The presence of these small spurious barotropic waves on the exterior does not affect the results much in the model interior (Figs. 12, 13, 14), especially since the focus of this present paper is on internal waves. A Flather-type boundary condition (Flather, 1976) could be imposed on the high frequency barotropic velocity to modify (18) and further reduce these small oscillations. However, we found it unnecessary, particularly given the need to

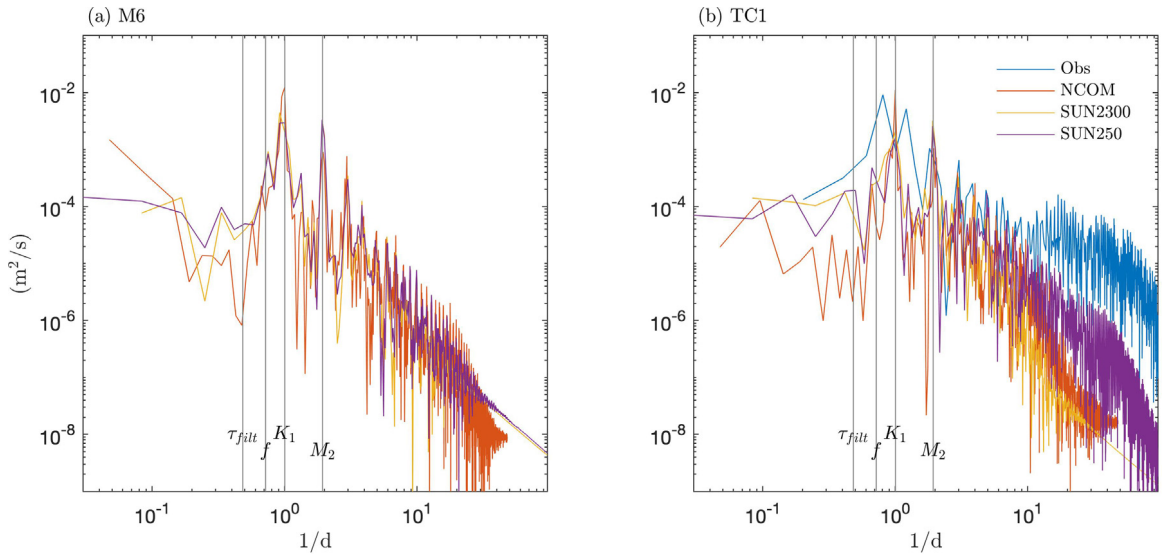

Fig. 15. Power spectra of wave velocities $\vec{u}'$ from the NCOM model, SUNTANS model at coarse resolution (SUN2300), high resolution (SUN525) and observations (Obs) at sites (a) M6 (3000 m depth) and (b) TC1 (450 m depth). Frequencies of M_2 tide, K_1 tide, inertial period f, and selected low-pass filtering timescale τ_{filt} are indicated by vertical lines. Velocities are taken at 100 m below the surface.

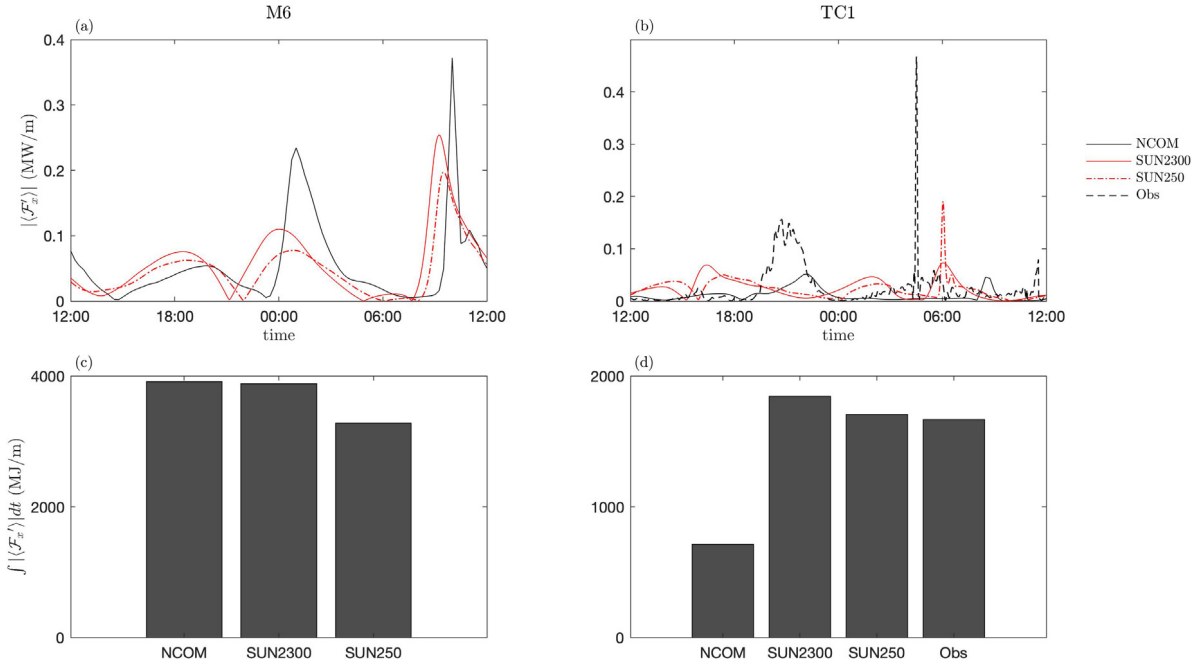

Fig. 16. Energy flux comparison between observations, the NCOM model and the SUNTANS model at location M6 near Luzon in 3000 m depth (left) and shallower location TC1 (right) in 450 m depth over a 24 h period. Comparing wave energy flux $\langle \mathcal{F}_x' \rangle$ as a function of time (top) and time-integrated over 3 days (bottom). The observations (Obs) at TC1 are from Lien et al. (2014). Time in panels (a) and (b) is June 2 12:00 to June 3 12:00, 2011 GMT.

introduce another tunable parameter representing the clamping time scale relative to the Flather condition. Additionally, the application of the barotropic tides could also be applied by computing $\widetilde{U}_{LS}$ in (18) using tidal constituents (Janeković and Powell, 2012). For some applications, this could improve the tidal signal by computing $\widetilde{U}_{LS}$ at each time step, instead of the coarser time step of the boundary condition file. However, for this model, testing with 15 min vs. 1 h boundary file time step yielded nearly identical results, therefore we included $\widetilde{U}_{LS}$ as a time series in (18).

5.3. Recipes for modeling internal waves

We have presented a generalized framework for computation of internal waves in a one-way nested grid. The model features four parameters that must be selected based on the model conditions which include the sponge layer damping time scale τ_D, the sponge layer width ΔD, the low-frequency filtering time scale τ_{filt}, and the low-frequency forcing time scale τ_L. In addition, the nested model grid domain location, model extent (L, W), grid spacing ($\Delta x, \Delta y, \Delta z$), forcing time step, and spin up time scale θ_r are important parameters.

For the sponge region, τ_D is selected based on the wave speed of the internal waves. Optimal results for dissipation in the sponge layer were obtained for $\Delta D = \lambda/2$ and $\tau_D = 0.04C/\Delta D$ (Fig. 6), where C is the wave speed. An important consideration is the location of the region where most of the internal wave energy is dissipated in the model domain, which can be used to select these parameters. In the South China Sea example, most of the dissipation is located in the shallow western edge

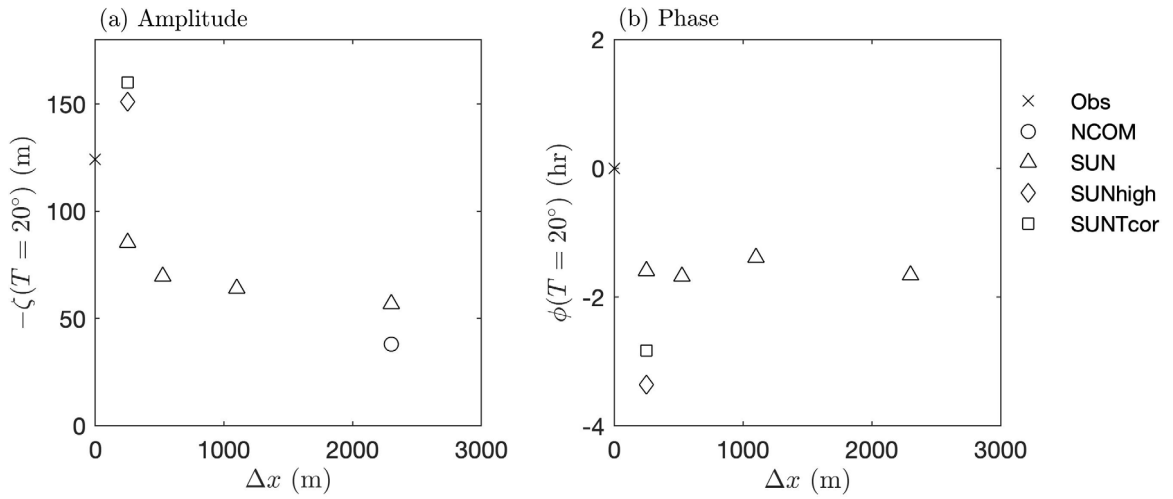

Fig. 17. Analysis of internal wave amplitude and phase of the displacement of the 20 °C isotherm at location TC1 during June 3, 2011, comparing observations, NCOM model results, and SUNTANS model results. In panel (a), the amplitude taken as the maximum deviation during 6 h window, and (b) associated phase lag compared to observations as a function of grid resolution Δx. Amplitude in panel (a) is taken as the difference between the maximum depth of the 20 °C isotherm and its average location for the preceding 3 h.

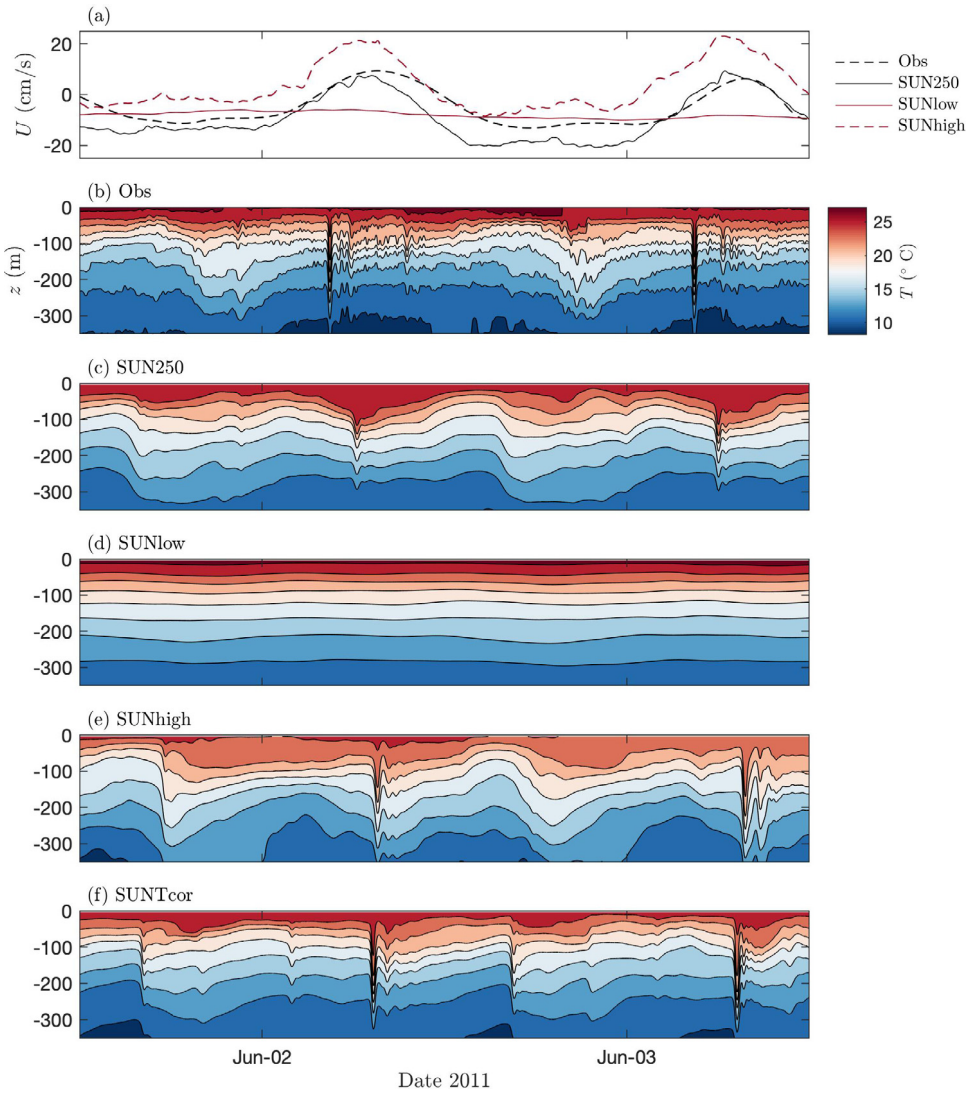

Fig. 18. Effect of different forcing effects on model results at location TC1. (a) comparison of barotropic velocities, and comparison of temperature profiles for (b) observations, (c) full forcing (SUN250), (d) low-frequency forcing only (SUNlow), (e) high-frequency forcing only (SUNhigh), and (f) initial temperature corrected with observations (SUNTcor).

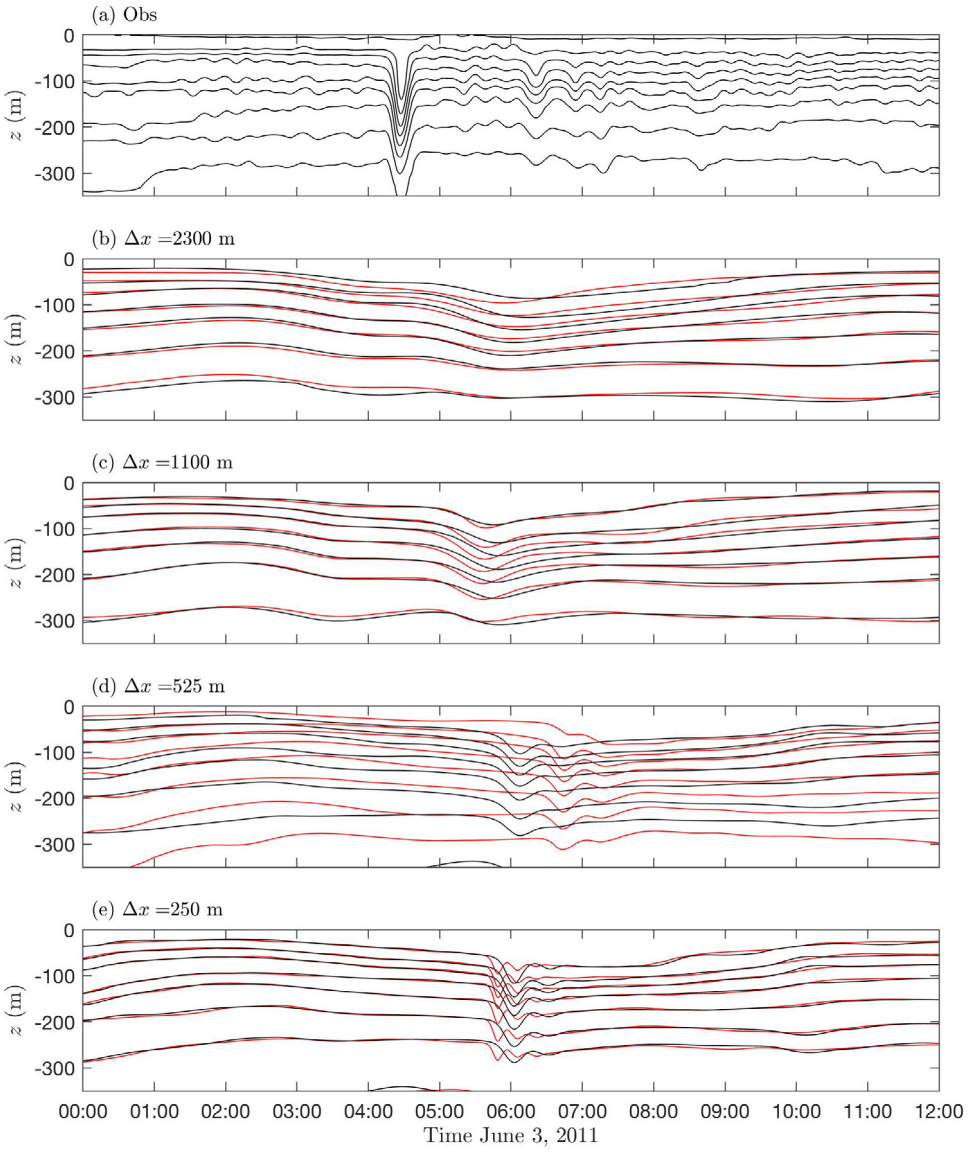

Fig. 19. Comparison of hydrostatic vs. nonhydrostatic effects from temperature results at location TC1. (a) observations. Comparison of model results at horizontal grid lepticity (b) 10 (SUN2300), (c) 5 (SUN1100), (d) 2.2 (SUN525), and (e) 1.0 (SUN250). Contours represent nonhydrostatic (black) and hydrostatic (red) isotherms at internals of 2 °C.

of the domain, where because of the shorter internal tides in shallow water, a much smaller sponge layer width ΔD can be used.

For the low-frequency components, τ_{filt} should be selected so that it separates the low- and high-frequency components of the flow at a location in the frequency spectrum of little energy. For the South China Sea example, τ_{filt} was selected to be less than the Coriolis frequency to render inertial and tidal motions as high-frequency motions that are computed by SUNTANS rather than nudged by NCOM. The low-frequency forcing time scale τ_L should be selected to ensure that it is large enough to have minimal damping effect on the high-frequency motions (Fig. 7), while still being small enough to nudge the flow towards NCOM (Fig. 18a). For the South China Sea example, $\tau_L = 4\tau_{filt}$ was selected to balance these goals.

The model grid has a significant effect on the model results. Horizontal grid resolution affects the grid lepticity ($\Delta x/H_e$), which should be less than about 2 to adequately resolve nonhydrostatic processes. For the South China Sea example at the TC1 mooring, this was important for correctly modeling the wave amplitude, timing and energy flux (Figs. 16, 17). Although not evaluated in this study, the nonlinear steepening length scale may be important in selecting the horizontal grid resolution, as nonlinear internal waves require sufficient distance to steepen (Rayson et al., 2018). Vertical resolution is less important as long as the mode shapes are adequately resolved, and 50 to 100 vertical z levels were used with good results. In addition, the model grid extents should be large enough to minimize boundary effects, which influence the waves within some buffer region ΔB (Fig. 9). When waves are propagating perpendicular to model boundaries (and the damping region is smoothed so as to not create reflections), there is very little effect and $\Delta B = 0$. However, when waves are propagating parallel to model boundaries along a sponge damping region, the waves are slowed and modified near the boundary. Based on model simulations, the distance of this effect is approximately one wavelength $\Delta B = \lambda$.

The boundary conditions are forced by a time series with discrete time step Δt. For an example barotropic tidal signal in deep water (Fig. 21a), the maximum velocity imposed on the boundaries creates the largest internal waves, which steepen in the domain interior. To obtain at least 90% of this velocity magnitude, a record length of at least 12 days and a time step of less than three hours is required for this site (Fig. 21b). In addition, the model must be run for sufficient length to spin up the barotropic tides and dissipate inconsistencies related to the initial conditions. For this South China Sea example, the spin up time was approximately seven days (Supp. Mat Figure SM3).

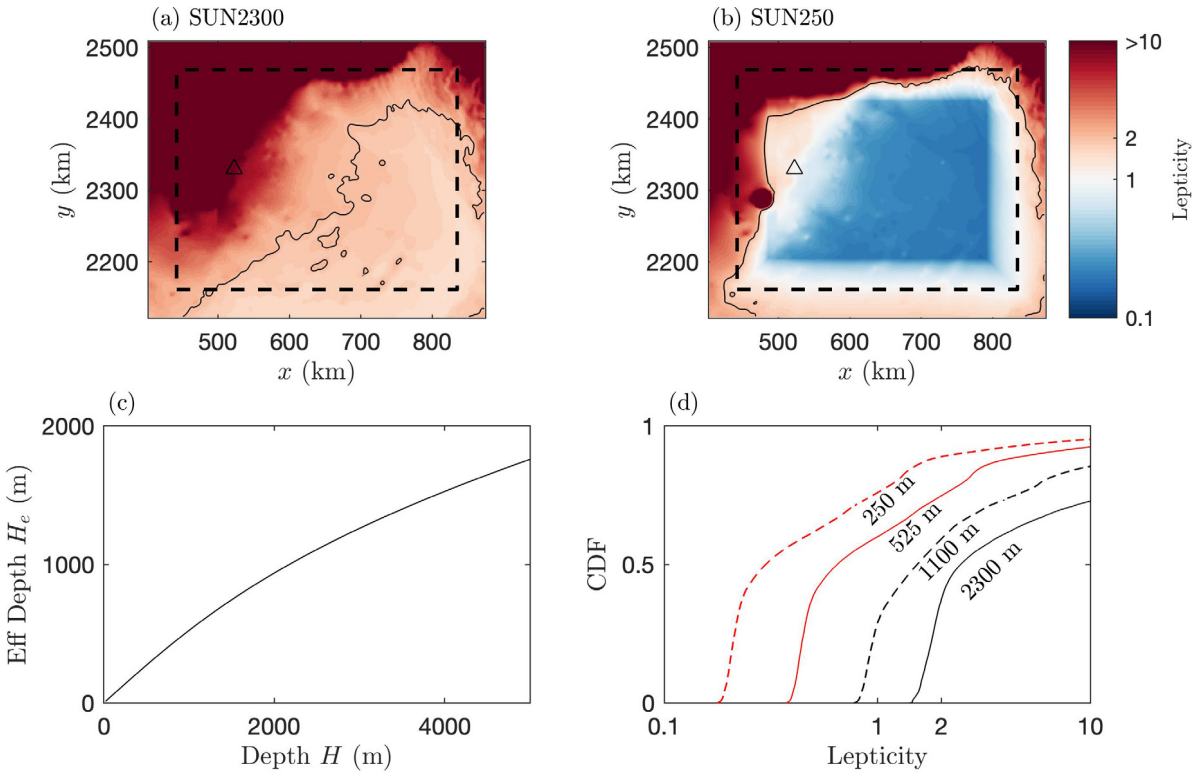

Fig. 20. Grid lepticity for the SUN2300 (a) and SUN250 (b) grids, showing position of TC1 mooring (triangle) and sponge damping region (–) and contour of lepticity equal to two (-). (c) Effective depth H_e for average stratification conditions as a function of depth H, and (d) cumulative distribution of grid lepticity over model domain for interior grid spacing equal to 2300 m (black -), 1100 m (black –), 525 m (red -), and 250 m (red –).

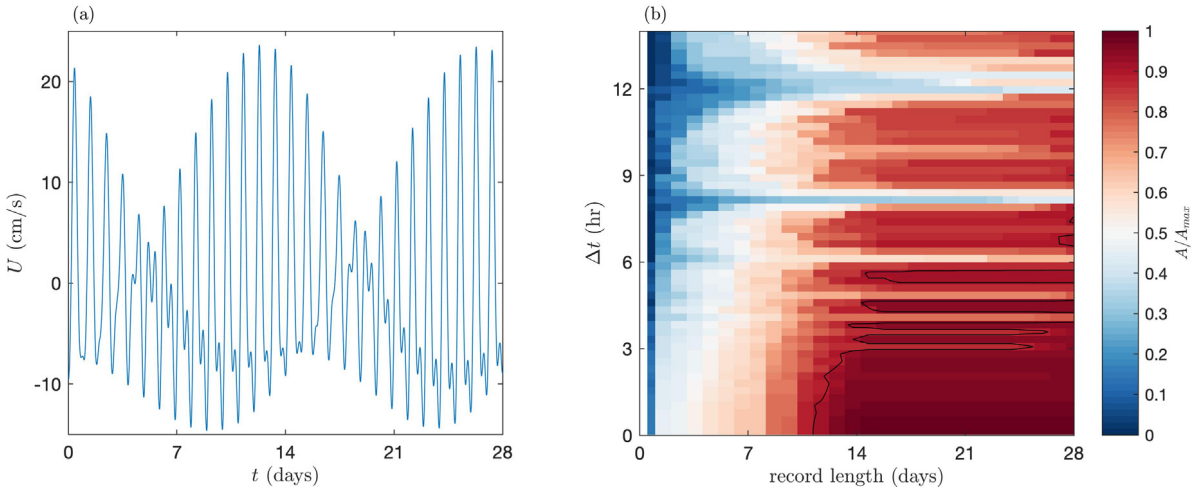

Fig. 21. Tidal velocity and effect of time step resolution and model record length on resolved tidal amplitude. (a) Fully resolved barotropic tidal velocity signal U near Luzon in 3000 m depth over 28 days with 1 min time resolution and maximum amplitude $A_{max} = 24$ cm s^{-1}, (b) maximum resolved amplitude A normalized by A_{max} as a function of the record length and time resolution Δt. Black contour is 90% error.

Finally, the background stratification has a significant effect on the model results. Small differences in the temperature profiles O(1 °C) and hence background density ρ_b can create large differences in the buoyancy frequency N_b^2 and in the propagation speed (Fig. 11). When wave features are propagated over large distances, this results in large errors in wave arrival times. In addition, the effect of internal waves is to induce mixing, which was most pronounced near the surface. To improve the model results, either a surface heat flux model or a low-frequency nudging of the background temperature field could be employed. Thus, while we employ state-of-the-art computational nesting methods, specific wave events are not modeled as accurately as would be desired. This discrepancy is primarily due to the background stratification inherited from the large-scale model, which critically affects the solution of modeled internal waves.

A significant advantage of our method is the reduced computational cost from running a regional model vs. a much larger grid required to simulate the internal tide. As an example, if we wish to model the internal waves at some location in the South China Sea, the first approach would be to model the entire South China Sea basin and Luzon ridge. Zhang et al. (2011) modeled this region with 11 M grid cells on a variable unstructured grid with two km resolution in the South China Sea basin and an 11 s time step due to small grid cells in other parts of the domain. The second approach, as proposed in this study, using nesting, at similar two km resolution (SUN2300) has only

2.4 M grid cells and uses a time step size of 100 s. Assuming a scalable computation, this represents a factor of 40 savings in computational resources for a similar level of resolution.

6. Summary

We present a method for seamless integration of one-way nested ocean model for internal waves. The method relies on a large-scale model which provides initial and boundary conditions to the nested model. The forcing is applied through a decomposition of time scales (low and high) and depth averaging (barotropic and baroclinic), and requires four constants. The barotropic component of flow is applied at the boundaries. A low-frequency nudging is applied to the model interior to nudge the solution towards the low-frequency flow of the large-scale model. Finally, the wave component of flow (high-frequency, baroclinic) is applied at the model boundaries where the wave energy enters the domain. A sponge layer on the model exterior allows incoming wave energy to pass through, but damps wave energy exiting the domain to prevent reflections.

We apply the model framework to idealized two- and three-dimensional model domains in 3000 m depth with stratification and forcing typical of the South China Sea. The results are as expected, showing waves generated along one boundary, propagating through the interior, and damped on the opposing end.

We further apply the model to a realistic model application in the South China Sea. The results show reasonable agreement with field observations of both in situ temperature and energy flux measurements, and also satellite observations of wave timing and shape. The model results are sensitive to horizontal grid spacing, and it is found that a grid lepticity less than two is required to accurately model nonhydrostatic effects which are important in some areas of the domain. The model domain size should be selected such that at a minimum, space is allowed for the sponge damping layers and a buffer region where waves propagating parallel to the sponge layer are distorted.

Differences in wave timing and amplitude are observed between the observations and model results. The most likely reason for this is differences in the background stratification due to either incorrect large-scale initial/boundary conditions or enhanced mixing. The model could likely be improved through inclusion of a surface heat flux model and nudging of the background temperature profile to observational data. Thus, even with state-of-the-art computational methods, the wave results are less accurate than would be desired, and critically depend on the stratification inherited from the large-scale model. This highlights the need for large-scale ocean models to incorporate observational data, especially related to stratification in the ocean interior.

Overall, the results are in good agreement with observations (when the large-scale model provides accurate forcing) and present a flexible framework for application of regional nested models for internal wave simulations. The method relies on four parameters. First, the time scale of low-pass filtering τ_{filt} should be selected so that it separates the low- and high-frequency components of the flow at a location in the frequency spectrum with little energy, here selected as four times the M_2 tidal period. Second, the time scale of low-frequency forcing τ_{L} should be selected to ensure that it is large enough to have minimal damping effect on the high-frequency motions, here selected as $\tau_L = 4\tau_{filt}$. Third, the width of the sponge damping layer ΔD is based on the internal tide wavelength, $\Delta D = \lambda/2$. Fourth, the time scale of sponge damping τ_D is selected based on the wave speed of the wave, $\tau_{\mathrm{D}} = 0.04C/\Delta D$.

A significant advantage of this method is the reduced computational cost from running a regional nested model vs. a much larger grid required to simulate the internal tide over a basin. In the South China Sea example presented here, this represents a factor of 40 savings in computational resources for a similar level of resolution.

Acknowledgments

This research is supported by ONR Grants N00014-15-1-2287 (OF), N00014-16-1-2256 (OF), and N00014-17-WX0-1702 (DSK). This manuscript was improved by helpful discussions with Stephen Monismith, Leif Thomas and Kristen Davis. Validation data was generously provided by Ren-Chieh Lien and Chris Jackson. Eric Mayer, Yun Zhang, Bing Wang, and Kurt Nelson all contributed code for the modeling. We gratefully acknowledge the US Army Research Laboratory DoD Supercomputing Resource Center for computer time on Excalibur, and especially thank the diligent staff at the HPC Help Desk for their support.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ocemod.2019.101462.

References

Alford, M.H., Lien, R.-C., Simmons, H., Klymak, J., Ramp, S., Yang, Y.J., Tang, D., Chang, M.-H., Alford, M.H., Lien, R.-C., Simmons, H., Klymak, J., Ramp, S., Yang, Y.J., Tang, D., Chang, M.-H., Sciences, O., Columbia, B., Landing, M., Academy, N., 2010. Speed and evolution of nonlinear internal waves transiting the South China Sea. J. Phys. Oceanogr. 40, 1338–1355. http://dx.doi.org/10.1175/2010JPO4388.1.

Alford, M.H., Peacock, T., MacKinnon, J.A., Nash, J.D., Buijsman, M.C., Centuroni, L.R., Chao, S.-Y., Chang, M.-H., Farmer, D.M., Fringer, O.B., Fu, K., Gallacher, P.C., Graber, H.C., Helfrich, K.R., Jachec, S.M., Jackson, C.R., Klymak, J.M., Ko, D.S., Jan, S., Johnston, T.M.S., Legg, S., Lee, I.-H., Lien, R.-C., Mercier, M.J., Moum, J.N., Musgrave, R., Park, J.-H., Pickering, A.I., Pinkel, R., Rainville, L., Ramp, S.R., Rudnick, D.L., Sarkar, S., Scotti, A., Simmons, H.L., St Laurent, L.C., Venayagamoorthy, S.K., Wang, Y.-H., Wang, J., Yang, Y.J., Paluszkiewicz, T., David Tang, T.-Y., 2015. The formation and fate of internal waves in the South China Sea. Nature 521, 65–69. http://dx.doi.org/10.1038/nature14399.

Arbic, B.K., Alford, M.H., Ansong, J.K., Buijsman, M.C., Ciotti, R.B., Farrar, J.T., Hallberg, R.W., Henze, C.E., Hill, C.N., Luecke, C.A., Menemenlis, D., Metzger, E.J., Müeller, M., Nelson, A.D., Nelson, B.C., Ngodock, H.E., Ponte, R.M., Richman, J.G., Savage, A.C., Scott, R.B., Shriver, J.F., Simmons, H.L., Souopgui, I., Timko, P.G., Wallcraft, A.J., Zamudio, L., Zhao, Z., 2018. A primer on global internal tide and internal gravity wave continuum modeling in HYCOM and MITgcm. New Front. Oper. Oceanogr. http://dx.doi.org/10.17125/gov2018.ch13.

Arthur, R.S., Fringer, O.B., 2016. Transport by breaking internal gravity waves on slopes. J. Fluid Mech. 789, 93–126. http://dx.doi.org/10.1017/jfm.2015.723.

Auclair, F., Benshila, R., Debreu, L., Ducousso, N., Marchesiello, P., Lemarié, F., Auclair, F., Benshila, R., Debreu, L., Ducousso, N., Dumas, F., 2018. Some recent developments around the CROCO initiative for complex regional to coastal modeling, in: Comod Workshop on Coastal Ocean Modelling.

Blayo, E., Rousseau, A., 2016. About interface conditions for coupling hydrostatic and nonhydrostatic Navier–Stokes flows. Discrete Contin. Dyn. Syst. Ser. S 9, 1565–1574. http://dx.doi.org/10.3934/dcdss.2016063.

Carter, G.S., Merrifield, M.A., 2007. Open boundary conditions for regional tidal simulations. Ocean Model. 18, 194–209. http://dx.doi.org/10.1016/j.ocemod.2007.04.003.

Carter, G.S., Merrifield, M.A., Becker, J.M., Katsumata, K., Gregg, M.C., Luther, D.S., Levine, M.D., Boyd, T.J., Firing, Y.L., 2008. Energetics of M 2 barotropic-to-baroclinic tidal conversion at the Hawaiian Islands. J. Phys. Oceanogr. 38, 2205–2223. http://dx.doi.org/10.1175/2008jpo3860.1.

Chapman, D.C., 1985. Numerical treatment of cross-shelf open boundaries in a barotropic coastal ocean model. J. Phys. Oceanogr. 15, 1060–1075. http://dx.doi.org/10.1175/1520-0485(1985)015<1060:NTOCSO>2.0.CO;2.

Chen, Y.J., Shan Ko, D., Shaw, P.T., 2013. The generation and propagation of internal solitary waves in the South China Sea. J. Geophys. Res. Ocean 118, 6578–6589. http://dx.doi.org/10.1002/2013JC009319.

Debreu, L., Marchesiello, P., Penven, P., Cambon, G., 2012. Two-way nesting in split-explicit ocean models: Algorithms, implementation and validation. Ocean Model. 49–50, 1–21. http://dx.doi.org/10.1016/j.ocemod.2012.03.003.

Dorostkar, A., Boegman, L., Pollard, A., 2017. Three-dimensional simulation of high-frequency nonlinear internal wave dynamics in Cayuga Lake. J. Geophys. Res. Ocean 122, 2183–2204. http://dx.doi.org/10.1002/2016JC011862.

Farmer, D., Li, Q., Park, J., 2009. Internal wave observations in the South China Sea: The role of rotation and non-linearity. Atmos.-Ocean 47, 267–280. http://dx.doi.org/10.3137/OC313.2009.

Flather, R.A., 1976. A tidal model of the North-West European continental shelf. Mem. Soc. R. Sci. Liège X 141–164.

Fringer, O.B., Gerritsen, M., Street, R.L., 2006. An unstructured-grid, finite-volume, nonhydrostatic, parallel coastal ocean simulator. Ocean Model. 14, 139–173. http://dx.doi.org/10.1016/j.ocemod.2006.03.006.

Fringer, O.B., Street, R.L., 2003. The dynamics of breaking progressive interfacial waves. J. Fluid Mech. 494, 319–353. http://dx.doi.org/10.1017/S0022112003006189.

Helfrich, K.R., 2007. Decay and return of internal solitary waves with rotation. Phys. Fluids 19, http://dx.doi.org/10.1063/1.2472509.

Janeković, I., Powell, B., 2012. Analysis of imposing tidal dynamics to nested numerical models. Cont. Shelf Res. 34, 30–40. http://dx.doi.org/10.1016/j.csr.2011.11.017.

Kang, D., Fringer, O., 2012. Energetics of barotropic and baroclinic tides in the Monterey Bay area. J. Phys. Oceanogr. 42, 272–290. http://dx.doi.org/10.1175/JPO-D-11-039.1.

Klemp, J.B., Durran, D.R., 1983. An upper boundary condition permitting internal gravity wave radiation in numerical mesoscale models. Mon. Weather Rev. http://dx.doi.org/10.1175/1520-0493(1983)111<0430:AUBCPI>2.0.CO;2.

Ko, D.S., Martin, P.J., Rowley, C.D., Preller, R.H., 2008. A real-time coastal ocean prediction experiment for MREA04. J. Mar. Syst. 69, 17–28. http://dx.doi.org/10.1016/j.jmarsys.2007.02.022.

Koltakov, S., Fringer, O.B., 2013. Moving grid method for numerical simulation of stratified flows. Internat. J. Numer. Methods Fluids 71, 1524–1545. http://dx.doi.org/10.1002/fld.3724.

Kumar, N., Feddersen, F., Uchiyama, Y., McWilliams, J.C., O'Reilly, W., 2015. Midshelf to surfzone coupled ROMS–SWAN model data comparison of waves, currents, and temperature: Diagnosis of subtidal forcings and response. J. Phys. Oceanogr. 45, 1464–1490. http://dx.doi.org/10.1175/JPO-D-14-0151.1.

Kundu, P.K., Cohen, I.M., 2008. Fluid Mechanics, fourth ed. Academic Press.

Lien, R.-C., Henyey, F., Ma, B., Yang, Y.J., 2014. Large-amplitude internal solitary waves observed in the Northern South China Sea: Properties and energetics. J. Phys. Oceanogr. 44, 1095–1115. http://dx.doi.org/10.1175/JPO-D-13-088.1.

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

Marsaleix, P., Auclair, F., Estournel, C., Marsaleix, P., Auclair, F., Estournel, C., 2006. Considerations on open boundary conditions for regional and coastal ocean models. J. Atmos. Ocean. Technol. 23, 1604–1613. http://dx.doi.org/10.1175/JTECH1930.1.

Martin, P.J., 2000. Description of the navy coastal ocean model. Version 10.

Mason, E., Molemaker, J., Shchepetkin, A.F., Colas, F., McWilliams, J.C., Sangrà, P., 2010. Procedures for offline grid nesting in regional ocean models. Ocean Model. 35, 1–15. http://dx.doi.org/10.1016/j.ocemod.2010.05.007.

Pickering, A., Alford, M.H., Nash, J., Rainville, L., Buijsman, M.C., Ko, D.S., Lim, B., 2015. Structure and variability of internal tides in Luzon strait. J. Phys. Oceanogr. 45, 1574–1594. http://dx.doi.org/10.1175/JPO-D-14-0250.1.

Piggott, M.D., Gorman, G.J., Pain, C.C., Allison, P.A., Candy, A.S., Martin, B.T., Wells, M.R., 2008. A new computational framework for multi-scale ocean model.ling based on adapting unstructured meshes. Internat. J. Numer. Methods Fluids 56, 1003–1015. http://dx.doi.org/10.1002/fld.1663.

Ramp, S.R., Yang, Y.J., Bahr, F.L., 2010. Characterizing the nonlinear internal wave climate in the northeastern South China Sea. Nonlinear Process. Geophys. 17, 481–498. http://dx.doi.org/10.5194/npg-17-481-2010.

Rayson, M.D., Jones, N.L., Ivey, G.N., 2018. Observations of large-amplitude mode-2 nonlinear internal waves on the Australian North West Shelf. J. Phys. Oceanogr. 49, 309–328. http://dx.doi.org/10.1175/jpo-d-18-0097.1.

Smith, S.W., 1997. The Scientist and Engineer's Guide to Digital Signal Processing. http://dx.doi.org/10.1007/BF02834636.

Suanda, S.H., Feddersen, F., Spydell, M.S., Kumar, N., 2018. The effect of barotropic and baroclinic tides on three-dimensional coastal dispersion. Geophys. Res. Lett. 45, 11,235–11,246. http://dx.doi.org/10.1029/2018GL079884.

Vitousek, S., Fringer, O.B., 2011. Physical vs. numerical dispersion in nonhydrostatic ocean model.ing. Ocean Model. 40, 72–86. http://dx.doi.org/10.1016/j.ocemod.2011.07.002.

Walter, R.K., Woodson, C.B., Arthur, R.S., Fringer, O.B., Monismith, S.G., 2012. Nearshore internal bores and turbulent mixing in southern Monterey Bay. J. Geophys. Res. 117 (C07017), http://dx.doi.org/10.1029/2012JC008115.

Warn-Varnas, A., Ko, D.S., Gangopadhyay, A., 2015. Signatures of tidal interference patterns in the South China Sea. J. Oceanogr. 71, 251–262. http://dx.doi.org/10.1007/s10872-015-0282-8.

Warner, J.C., Armstrong, B., He, R., Zambon, J.B., 2010. Development of a Coupled Ocean–Atmosphere–Wave–Sediment Transport (COAWST) modeling system. Ocean Model. 35, 230–244. http://dx.doi.org/10.1016/j.ocemod.2010.07.010.

Winters, K.B., 2016. The turbulent transition of a supercritical downslope flow: Sensitivity to downstream conditions. J. Fluid Mech. 792, 997–1012. http://dx.doi.org/10.1017/jfm.2016.113.

Winters, K.B., de la Fuente, A., 2012. Modelling rotating stratified flows at laboratory-scale using spectrally-based DNS. Ocean Model. 49–50, 47–59. http://dx.doi.org/10.1016/j.ocemod.2012.04.001.

Wolfram, P.J., Ringler, T.D., 2017. Quantifying residual, Eddy, and Mean flow effects on mixing in an idealized circumpolar current. J. Phys. Oceanogr. 47, 1897–1920. http://dx.doi.org/10.1175/JPO-D-16-0101.1.

Zhang, Z., Fringer, O.B., Ramp, S.R., 2011. Three-dimensional, nonhydrostatic numerical simulation of nonlinear internal wave generation and propagation in the South China Sea. J. Geophys. Res. Ocean. 116, 1–26. http://dx.doi.org/10.1029/2010JC006424.

Zhang, Y., Kennedy, A.B., Panda, N., Dawson, C., Westerink, J.J., 2014. Generating-absorbing sponge layers for phase-resolving wave models. Coast. Eng. 84, 1–9. http://dx.doi.org/10.1016/j.coastaleng.2013.10.019.