Internal Gravity Waves☆

Maarten C Buijsman, University of Southern Mississippi, Hattiesburg, MS, United States
Brian K Arbic, University of Michigan, Ann Arbor, MI, United States
Samuel M Kelly, University of Minnesota, Duluth, MN, United States
Amy F Waterhouse, Scripps Institution of Oceanography, La Jolla, CA, United States

Introduction

A person, standing on a boat or a beach, will observe an ocean surface perturbed by "surface gravity" waves. It is much harder to see that beneath this restless surface, the ocean interior is also filled with "internal" gravity waves. Like waves at the ocean surface, which occur at the interface between air and water, internal waves occur in the ocean interior, at the interface between layers of different water densities. Due to the solar heating and the precipitation of "fresh" rain water and the discharge of river water, the water density gradually increases from the ocean surface towards the sea floor, where the water is cold and salty. This vertical gradient in density facilitates the propagation of internal waves. Since the density difference between air and water is so large, surface gravity waves have heights of tens of centimeters to a few meters. In contrast, the density differences in the ocean are small, and therefore, internal gravity waves have heights of tens to hundreds of meters. Like surface waves, internal gravity waves have a restoring force of gravity. Internal gravity waves are ubiquitous in the global ocean. They are relevant for ocean mixing and the dispersal of matter, and they impact ocean structures, ocean acoustics, and ship navigation.

Internal waves were "discovered" by mariners, whose ships' progression was slowed down by internal waves. The first record of such an event was made by the Norwegian explorer Nansen in the autumn of 1893, while steaming with the Fram north of Siberia in the sound between the isle of Taimur and Almqvist Islands (Nansen, 1897). Nansen writes: "We approached the ice to make fast to it, but the Fram had got into dead-water, and made hardly any way, in spite of the engine going at full pressure." The Fram was steaming in a layer of fresh water that had melted from the ice, and its progression was not arrested by the sea floor. The dead-water phenomenon was brought to the attention of the Norwegian oceanographer Ekman, who could explain the phenomenon after doing extensive laboratory experiments (Ekman, 1904). In his laboratory, Ekman filled a tank with a surface layer of fresh water and a bottom layer of salt water dyed with ink. Ekman pulled a Fram boat model, attached to a rope, with a constant force through the upper layer. He observed that when dead-water occurred, the force that should have propelled the ship, instead generated internal waves on the interface between the fresh and salty water layers (Fig. 1), slowing the ship.

The first measurements of internal waves were made in 1900 during an oceanographic cruise in the Norwegian Sea led by Norwegian oceanographers Helland-Hansen and Nansen (1909). When they repeatedly measured vertical profiles of salinity and temperature, they noted that "the numerous observations taken at 20 metres show that there must have been much vertical movement of the water at this depth." Helland-Hansen and Nansen also wrote that "the oscillatory movement may possibly have raised and lowered the water with a temperature of 4.8°C., to the extent of about 10 or 15 metres, or even more". Just like waves at the ocean surface, internal waves propagate, and to an observer who is fixed in space, the heaving of the interface may be attributed to the passing of the crests and troughs of the internal waves.

☆*Change History*: June 2018. MC Buijsman, BK Arbic, SM Kelly, and AF Waterhouse updated the text, added Acknowledgements section; added new figures.

This is an update of Maarten C. Buijsman, Brian K. Arbic, Samuel M. Kelly, Amy F. Waterhouse, Internal Gravity Waves, *Reference Module in Earth Systems and Environmental Sciences*, Elsevier, 2018

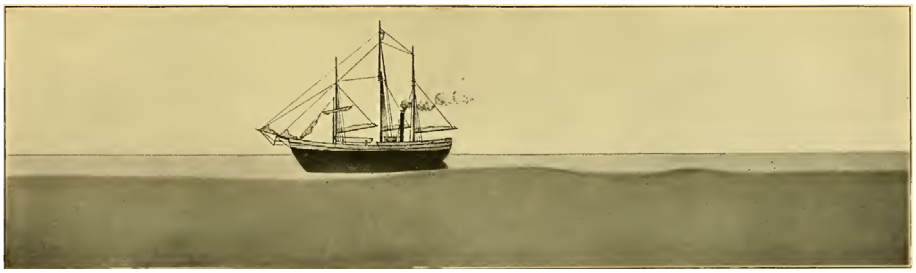

Fig. 1 To understand the "dead-water" phenomena, Ekman performed a series of experiments in a tank filled with fresh surface and ink-dyed salty bottom water layers. Ekman pulled a boat model, attached to a rope, with a constant force through the upper layer. This photograph shows the side of the tank, the Fram boat model, completed by a sketch of the upper part of the Fram, and internal waves on the layer interface behind the ship. Reproduced with permission from: Ekman, *V.* W. (1904). On dead water. In: Nansen, F. (Ed.), The Norwegian north polar expedition 1893–1896. Scientific Results, Vol. V. pp. 83–278. Christiania.

Internal waves impact numerous human activities. Satellite observations of the ocean surface have allowed for the estimation of average sea-surface height, currents and depths of particular density surfaces. However, internal waves complicate the interpretation of the satellite observations as these waves exist over a wide range of horizontal scales, from tens to hundreds of km, and have sea-surface height expressions of a few cm. For example, it is one of the objectives of the Surface Water and Ocean Topography (SWOT) mission (Fu et al., 2012) to predict and remove the tidal internal waves' surface expression from the SWOT altimetry, the measurement of sea surface height. Internal waves have also been the objective of intensive military-funded research because of the possibility that the strong vertical movement of density layers due to internal waves adversely impacts submarine navigation and that the wakes of internal waves generated by submarines might be detectable by remote sensing, thus betraying the submarine's location. More conventional acoustic means of submarine detection are complicated by the deflection of acoustic rays by the rather random variations in sound speed induced by internal waves. In civilian activities, the currents and buoyancy changes associated with internal waves are a matter of concern in offshore oil drilling.

Most importantly, perhaps, is the role that internal waves play in setting the global ocean climate. The difference in the vertical distribution of horizontal currents and their vertical shear that is associated with propagating internal waves, including those of tidal frequency, can eventually lead to instability and turbulence as these internal waves shoal. Therefore, these waves are the main agent for vertical mixing in the ocean interior (Munk, 1966). This mixing plays a major role in determining the strength of ocean circulation, and hence the poleward heat flux and climate. The mixing, along with the associated circulation, also provides nutrient fluxes into the sunlit upper ocean where primary biological production occurs. Moreover, internal waves can stir up sediment and transport sediment and nutrients on the shallow continental shelf and slope. Understanding internal waves is thus of vital importance, particularly since they occur at too small a scale to be treated explicitly in numerical models of the ocean. Their effects must be "parametrized" or represented by formulas that involve only the quantities that are used in the model.

This short article will first describe the theory of internal gravity waves, which can occur in the ocean as wave beams and interfacial waves. This will be followed by a discussion of their generation mechanisms, evolution, and relationship to ocean mixing. We use frequency–wavenumber spectra to summarize many of the concepts presented here. Throughout we will refer to observations and simulations of internal waves. The quality of these observations and numerical simulations, and their interpretation, has improved significantly in recent decades, leading to rapid advances in internal wave research.

Internal Wave Theory

The Basic Physics

A particle displaced vertically over a distance η in a continuously stratified fluid, experiences a restoring buoyancy force $\mathbf{F}_b$. If a whole vertical fluid column is displaced, the vertical uniformity of the motion means that there is no change in the hydrostatic vertical pressure gradient and the restoring force on each particle is just gravity g times the density perturbation, which is minus the vertical displacement times the vertical density gradient $\partial\rho/\partial z$. This leads to a simple harmonic oscillator equation for the motion of the column, with the frequency of oscillation given by the "buoyancy frequency" N, where $N^2 = -(g/\rho)(\partial\rho/\partial z)$. Note that the buoyancy force driving this oscillation is proportional to $N^2\eta$. The buoyancy force pulls the displaced fluid column back to its resting position. However, due to inertia, the column will overshoot its resting position, and in the absence of friction, the oscillation will go on indefinitely. The buoyancy frequency is the upper limit for internal wave frequencies. Internal "wave" motions that occur at this frequency do not portray any wave propagation.

If the fluid columns are now allowed to oscillate obliquely, at an angle θ to the vertical (see Fig. 2), the component of the vertical restoring force is reduced by a factor $\cos\theta$ along the pathway of motion. The relation between the vertical and diagonal displacement ξ is $\eta = \xi\cos\theta$. The restoring force along the diagonal path of motion is then proportional to $N^2\ \xi\cos^2\theta$. In this example, we assume N is invariant with depth. In the absence of rotation, the frequency associated with this oblique oscillation is $\omega^2 = N^2\cos^2\theta$. Note that the oblique oscillations have frequencies that are lower than the buoyancy frequency. At these lower

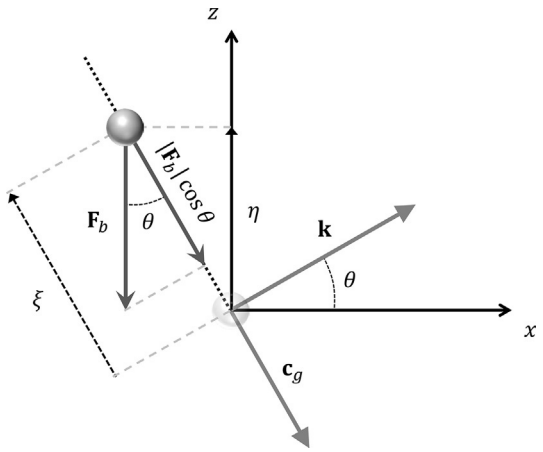

Fig. 2 Schematic of internal wave motion along the diagonal dashed line, which makes an angle θ with the z axis. The vertical displacement of the spheric particle along the z axis is η and the diagonal displacement is ξ. The buoyancy force acting on the particle is $\mathbf{F}_b$. The wavenumber vector $\mathbf{k}$ is perpendicular to the plane of motion, which is parallel to the wave crests. The wavenumber vector indicates the direction of wave-crest (phase) propagation. The group speed vector $\mathbf{c}_g$, which is along the plane of motion, indicates the energy propagation.

frequencies the internal waves are now allowed to propagate. Internal waves are also called "transverse" waves. This means that the particle motions occur parallel to the wave crests and perpendicular to the wave's propagation direction, which is the direction the crests move in. This propagation direction is indicated by the wavenumber vector $\mathbf{k} = (k, l, m)$ in Fig. 2, which has three components along the x, y, and z directions. Note that wave period is related to frequency by $T = 2\pi/\omega$ and the wave length is related to wave number by $L = 2\pi/K$, where K is the absolute value of $\mathbf{k}$.

In a rotating world the oblique motion is acted upon by the Coriolis force, so that fluid oscillations in inclined sheets now develop a crosswise motion, within the sheet but orthogonal to the motion with no rotation. The relationship between the frequency of the oscillations now involves the Earth's rotational frequency. Provided that N is sufficiently greater than the Coriolis frequency f, which is twice the vertical component of rotation, the connection between frequency ω and orientation θ becomes

$$\omega^2 = N^2 \cos^2\theta + f^2 \sin^2\theta \tag{1}$$

which is also referred to as the dispersion relation for internal gravity waves. Eq. (2) is an alternative expression, in terms of the wavenumber components (k, l, m)

$$\omega^2 = \frac{N^2(k^2 + l^2) + f^2 m^2}{k^2 + l^2 + m^2} \tag{2}$$

While the frequency can be as high as N when the particle motion is vertical, it cannot be lower than the Coriolis frequency f. In this limit, the particle motion is horizontal in "inertial" circles, expressing the tendency for steady rectilinear motion with respect to a nonrotating reference frame. Any frequency of motion between these limiting frequencies is possible, depending on θ, or, equivalently, the ratio of vertical to horizontal wavenumber. The group velocity vector (the velocity with which a wave packet, or energy, propagates) is given by $\mathbf{c}_g = (\partial\omega/\partial k, \partial\omega/\partial l, \partial\omega/\partial m)$. For internal waves this is easily shown from Eq. (2) to be at right angles to the wavenumber vector $\mathbf{k}$. In other words, energy propagates parallel to the wave crests (Fig. 2), rather than at right angles as for surface waves! In terms of vertical propagation, waves with upward phase propagation have downward energy flux, and vice versa. In the following two sections, the two identities of internal waves are discussed: internal wave beams and interfacial waves.

Internal Wave Beams

Internal waves can have a "beam"-like appearance. The internal-wave energy propagates along this beam, while the wave crests propagate perpendicular to the beam as described in the previous section. For the same wave frequency, these beams are straight if the buoyancy frequency is independent of the water depth. Internal waves in this identity are difficult to observe in the ocean. One reason is that the velocity shear in these beams is very high, causing them to dissipate rapidly close to their source. Another reason is that these beams quickly scatter in incoherent features because of the spatial inhomogeneity of stratification and currents in the ocean. However, internal wave beams are easily generated and observed in a wave tank, as shown in Fig. 3.

Interfacial Waves

A more common manifestation of internal waves in the ocean is interfacial waves (Fig. 4). In a two-layer fluid, as in Ekman's laboratory experiments (Fig. 1), these waves occur on the interface separating the water layers with different densities and they

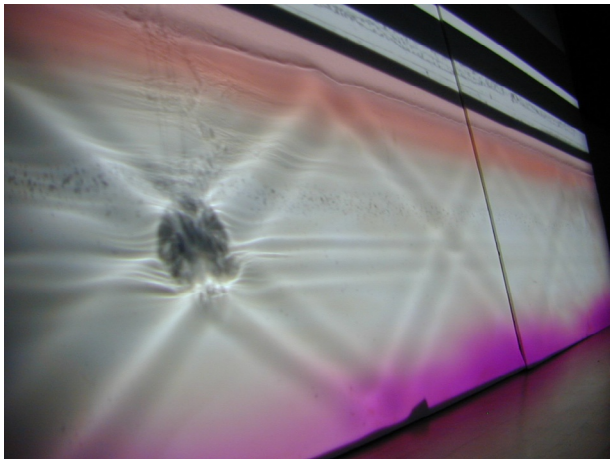

Fig. 3 Internal wave beams visualized with a "shadowgraph", where a point light source illuminates a white screen behind the wave tank. The waves are generated at a single period of 11 s by a cylinder oscillated vertically (the black circle). The buoyancy frequency is constant with a period of about 5 s. The beams comprise parallel crests and troughs of alternating shading. Notice how the wave beams reflect off the bottom and the base of the mixed layer. Image from: https://www.ocean.washington.edu/courses/oc512/lab8.html.

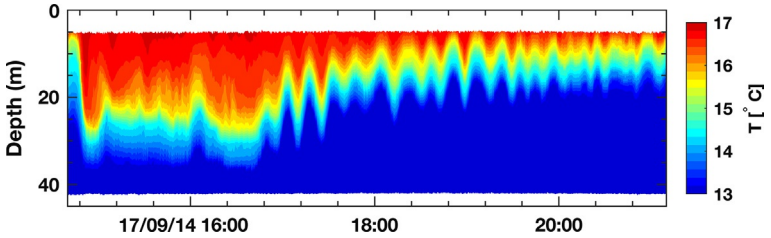

Fig. 4 Internal waves on the interface between warm (top) and cold (bottom) water layers, recorded with water temperature sensors as a function of depth and time in hours. The long wave has a trough that reaches to 35 m depth at 16:30 h and a crest that reaches to 20 m at 21:00 h. Superposed on the long waves are short solitary waves with crest-to-trough heights of more than 10 m. These observations were obtained near Point Sal, California in 45 m of water depth in September 2017 as part of the ONR funded Inner Shelf Direct Research Initiative. The colors represent temperature in °C.

propagate in the horizontal direction. Suppose the upper layer has a density ρ-$\Delta\rho$ and thickness h_1 above the lower layer with density ρ and thickness h_2, then waves that have a wavelength much greater than both h_1 and h_2 travel at a phase speed $[g'h_1h_2/(h_1+h_2)]^{1/2}$ independent of wavelength, where $g' = g\,\Delta\rho/\rho$ is known as the "reduced gravity." If $h_2 \gg h_1$, the phase speed becomes $(g'h_1)^{1/2}$, in clear analogy to the phase speed $(gh)^{1/2}$ for surface waves that are long compared to the water depth h. This formula for the speed of interfacial waves also holds for $h_2 \gg h_1$ even if the wavelength is not long compared with h_2.

The theory behind this requires that the amplitude of the waves is much less than the thickness of the layers. Many observed interfacial waves (Fig. 4) violate this assumption and the requirement that their wavelength is long compared with the layer thicknesses. Finite amplitude is associated with a tendency for waves to steepen, much as in the development of a breaking wave on the beach. On the other hand, a horizontal scale that is not very long, compared with at least the thickness of the thinner layer, leads to dispersion, the break-up of a disturbance into waves of different wavelengths traveling at different speeds. Interestingly, these effects can cancel, leading to the possibility of "internal solitary waves", waves of finite amplitude that can be spatially localized and travel without change of shape. They can occur singly, or in groups as in Fig. 4. Even if these solitary waves occur at a density interface many meters, or tens of meters, below the surface, they are often visible if the upper layer is turbid, so that the crests appear from above as more opaque tubes than the surrounding water. More frequently they are seen because the associated currents cause visible variations in surface roughness (Fig. 5).

While the ocean can be simplified as a two-layer model, the actual ocean is more complicated. In the ocean, the density does not change abruptly from the surface to the bottom layer, but it changes more gradually with depth. In such stratification, vertical mode waves exist. These waves are propagating waves in the horizontal plane, like interfacial waves, and standing waves in the vertical plane, as they reflect off the bottom and surface. A first mode wave has one node in the vertical with opposing horizontal particle velocities above and below the nodes, a second mode wave has two nodes for the horizontal velocity, a third mode has three nodes, etc. If the density change is linear with depth, that is, the buoyancy frequency is constant, the horizontal and vertical velocities, and the vertical water displacement are sinusoidal functions of water depth for modal waves. The wavelength and phase speed of higher mode waves scale with $\sim 1/n$, where n is the mode number. The linearized dispersion relation for vertical mode waves is

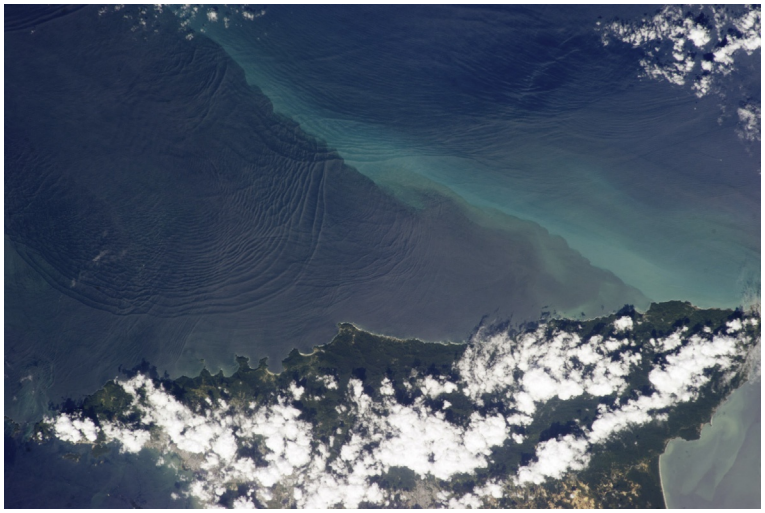

Fig. 5 A wave train with solitary internal waves near Trinidad as seen from space. Image from: http://earthobservatory.nasa.gov/IOTD/view.php?id=80337.

$$\omega^2 = f^2 + c_n^2\left(k^2 + l^2\right) \tag{3}$$

where c_n is the phase speed of internal waves modes not affected by rotation, and k and l are the horizontal wavenumbers along x and y. The eigenpeeds c_n depend on the water depth, buoyancy frequency and mode number, and are obtained as solutions of the Stürm–Liouville problem for a particular N profile and value of f (Gill, 1982). A remarkable aspect of internal waves is that if N is independent of depth and the seafloor is flat, the sum of all vertical mode waves (i.e., their velocities and displacements) yields the wave beams discussed in the previous section.

Internal Waves in the Ocean

The ocean's interior is perturbed by internal waves with different periods (or frequencies) and wave lengths (or wavenumbers), which reflect the generation and dissipation mechanisms of these internal waves. In the following, the internal wave generation, the interaction with topography, background flows, and other waves, and their dissipation in the global ocean are discussed.

Generation

Wind plays an important role in the generation of "near-inertial" internal waves. Fast-moving storms set up currents in the upper ocean with the corresponding Coriolis forces unmatched by pressure gradients. Near-inertial motions result, with the current vectors rotating with a frequency close to the local f. The large horizontal scale of these motions means, however, that they experience different f at different latitudes. The current vectors at different latitudes then rotate at different rates, increasing the latitudinal gradients in current vectors and decreasing the horizontal scale. The resulting convergence and vertical motion means that the waves can no longer be purely inertial; they retain their frequency but propagate equatorward to a region where f is smaller. At the same time, they develop an increasing vertical group velocity and propagate downward into the ocean (D'Asaro, 1985). Fig. 6 shows the energy input by the wind and energy fluxes for mode 1 internal waves at inertial frequencies computed from global model simulations with realistic atmospheric forcing (Alford et al., 2016). The wind energy input or "wind work" is the largest under the mid-latitude storm tracks. From these regions low mode energy fluxes radiate equatorward, as near-inertial internal waves can only propagate equatorward of the latitude at which they were generated, i.e., at frequencies $\omega \geq f$.

Higher-frequency waves may be generated by storms that move more slowly, by turbulence in the surface mixed layer, by subtle interactions between surface waves, or as part of the decay process of ocean eddies. They may also arise from interactions between preexisting internal waves, as will be discussed shortly.

The tides are another important source of energy for internal waves observed throughout the ocean (Wunsch, 1975). The barotropic, depth-independent, tidal currents associated with tidal changes in sea level move stratified water over topographic features on the seafloor, setting up internal oscillations much as if the topographic features were oscillating wavemakers in an otherwise still ocean. These "internal tides" are mainly at the tidal frequencies, though there may also be energy at higher harmonics. Fig. 7 shows the stationary in-phase component of the low mode sea-surface height for the M_2 internal tide extracted from a realistically forced global 1/25° HYbrid Coordinate Ocean Model (HYCOM) simulation and from the High Resolution Empirical Tide Model (HRET_v8.1), an altimetry constrained internal tide model (Edward D. Zaron, personal communication). The HYCOM simulation resolves modes 1 to 5, whereas HRET_v8.1 mainly captures the sea-surface height expression of modes 1 and 2. In HYCOM, the energy loss by the surface tide and the resolved low modes to the unresolved high modes needs to be parameterized

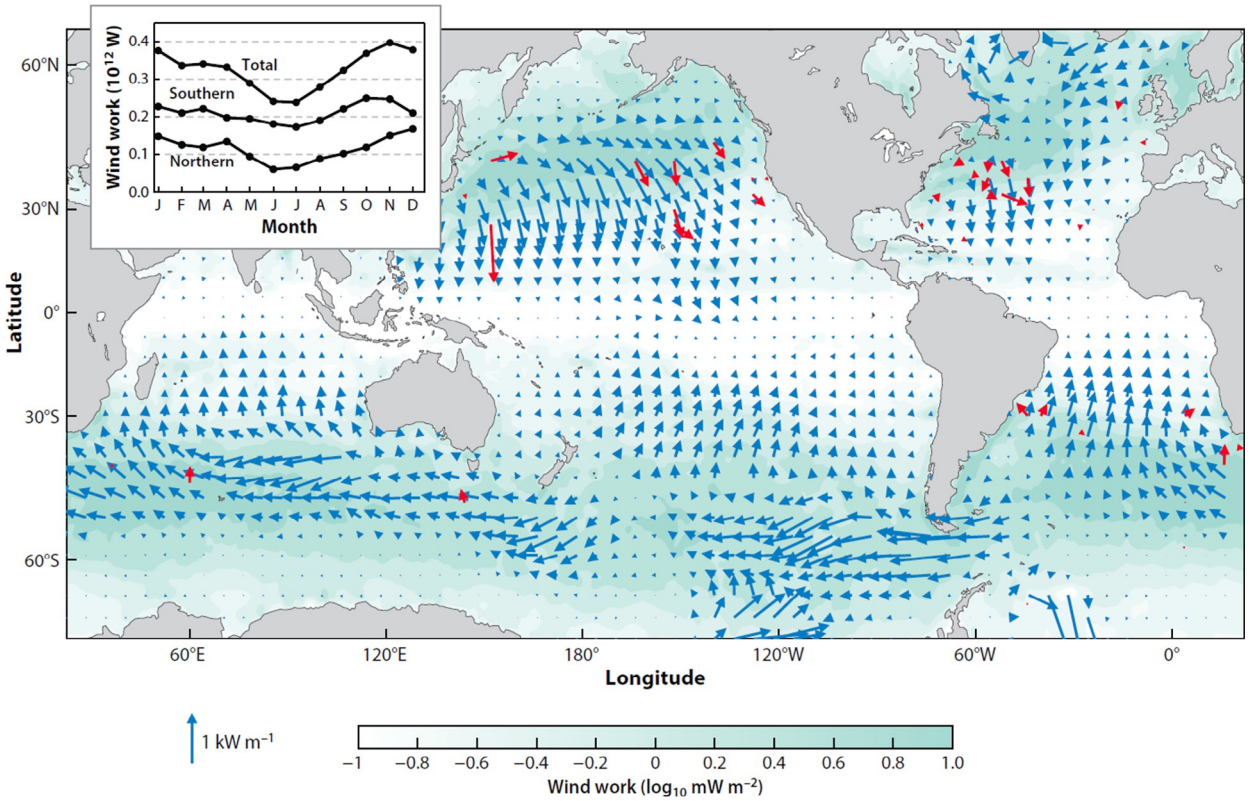

Fig. 6 Modeled (blue arrows) and observed (red arrows) horizontal energy fluxes at the inertial frequency in the first vertical mode. The modeled values are from 2007; the observations span 20 years of measurements. The inset shows the hemispherically and globally integrated wind work computed from a global ocean model. The aqua color shows the annual-mean distribution of the work, illustrating the enhancement under storm tracks. Reproduced with permission from: Alford, M. H., MacKinnon, J. A., Simmons, H. L., Nash, J. D. (2016). Near-inertial internal gravity waves in the ocean. *Annual Review of Marine Science* **8**, 95–123.

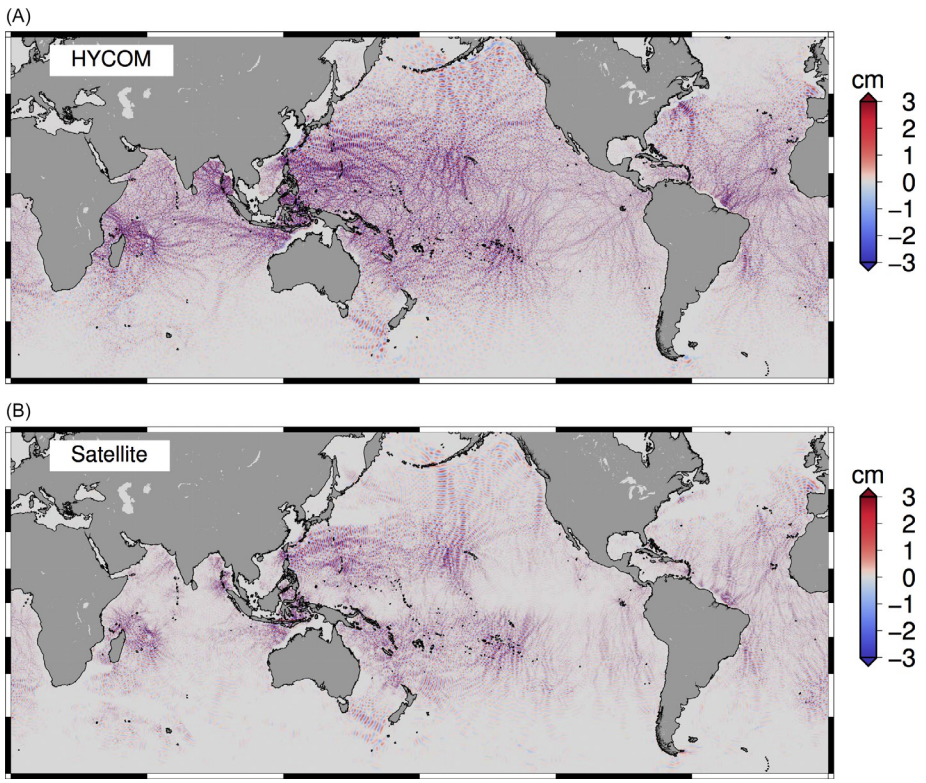

Fig. 7 The stationary in-phase component of the low-mode sea-surface height for the M_2 tide derived from a realistically forced global HYCOM simulation with a 1/25° horizontal resolution (top) and HRET_v8.1, a satellite altimetry constrained internal tide model (bottom). Internal tides radiate from mid-ocean ridges such as Hawaii and French Polynesia and continental shelves such as the Amazon shelf. Unpublished HYCOM results from Gordon Stephenson and altimetry data from Edward D. Zaron.

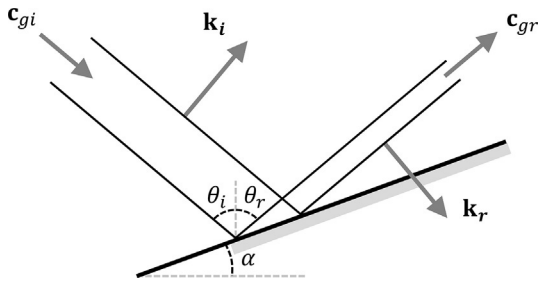

Fig. 8 Internal wave rays are reflected at the seafloor at an equal angle to the normal to a bottom slope. Subscripts *i* and *r* indicate incident and reflected waves.

with a linear wave drag to correctly simulate surface and internal tides. The wave drag is a function of topographic roughness and density stratification. The internal tides are generated at deep mid-ocean ridges, island chains, and shelf regions (Fig. 7). After generation they propagate for 1000s of km along narrow and almost linear pathways. Along the way their energy is believed to be gradually dissipated and scattered due to interactions with topography, the time-variable low-frequency background flow, and amongst the internal waves themselves (see next sections). The HYCOM and altimetry results show good agreement, particularly close to the generation sites, but there are also differences. The internal tides from HYCOM appear to be more energetic and travel farther than in the altimetry. One reason is that the stationary fraction becomes smaller for longer time series: the HYCOM-derived results are based on a 2-week time series while the altimetry-derived results are based on a 25-year time series. The differences are also due to the inability to extract internal tides from altimetry in some low-frequency flows and sub-optimal representation of bathymetry and dynamics in HYCOM.

In areas where low-frequency flows are strong and the bottom topography is rough, such as the Southern Ocean and other regions, the mean flows appear to set up quasi-steady "lee waves", or standing internal waves behind topographic features (as occurs in the atmosphere) (e.g., Nikurashin and Ferrari, 2011). These may propagate upward into the ocean. The relative importance of internal lee wave generation by eddy flows on a global scale has not been fully established, but the current view is that wind-generated near-inertial waves and internal tide generation over rough topography are the most important sources of energy for the global internal gravity wave field.

Reflection and Scattering

Where internal waves reflect off boundaries, they retain their frequency and, therefore, propagation angle with respect to the horizontal (a consequence of their peculiar dispersion relation). On flat boundaries, such as the sea surface, reflected internal waves behave like conventional waves (e.g., sound waves) in that they conserve their propagation angle with respect to the reflective surface. However, incident and reflected internal waves have different angles with respect to the sloping surfaces (Fig. 8).

At subcritical bottom slopes (i.e., slopes shallower than the wave propagation angle), onshore propagating plane waves are always reflected upslope, which focuses their wave rays, increases their wavenumber, decreases their group speed, and increases their amplitude. At supercritical bottom slopes (i.e., slopes steeper than the wave propagation angle), upslope reflection is impossible, and all energy is reflected offshore. In this case, downward reflection focuses waves. For both sub- and supercritical slopes, reflected energy density is most enhanced when the slope is near critical (i.e., almost equal to the wave propagation angle). At critical slopes, the dispersion relation predicts reflected waves with infinite wavenumber and zero group velocity, but analytical solutions in the time domain (Scotti, 2011) and laboratory experiments and numerical simulations are well-behaved, in that they exhibit large but finite wavenumbers. In practice, reflection at critical and near-critical slopes can lead to instability and wave breaking through mechanisms discussed below.

Under the "traditional approximation" of the Coriolis force, internal waves are confined to frequencies between f and N. Therefore, poleward propagating internal waves can reach a "turning latitude" where $|f|$ locally exceeds ω, producing exponential decay rather than wave propagation (i.e., imaginary wavenumbers). At these latitudes, phase speed tends toward infinity and group speed tends toward zero, so that waves refract, like edge waves, and reflect equatorward. Retaining nontraditional Coriolis terms in the momentum equations modifies reflection at the turning latitude.

Waves also reflect where stratification is weak and N falls below ω. These barriers are commonly associated with well-mixed surface and bottom boundary layers (see also Fig. 3). A consequence of internal reflection is that waves are partially insulated from scattering by roughness within the boundary layer. At intermediate water depths, internal waves can "tunnel" through sufficiently thin layers of weak stratification.

The previous discussion considers vertically propagating waves, however, vertically standing waves (i.e., the previously discussed vertical modes) are also common in the ocean, particularly at the tidal and inertial frequencies. These waves propagate horizontally, sometimes for thousands of kilometers, over rough topography before striking large-scale topographic features, such as mid-ocean ridges and continental slopes. As discussed above, subcritical slopes funnel incident waves toward shallower water, where they can steepen and break, while tall supercritical slopes behave akin to vertical walls, reflecting wave energy below the top of the feature and transmitting energy above. Abundant small-scale roughness and occasional tall near-critical slopes can both efficiently scatter waves from low to high wavenumbers, but their relative contributions to global energy balances are presently uncertain.

Wave-Background Flow Interaction

Internal waves interact with background currents and stratification via the nonlinear $\mathbf{u}\cdot\nabla$ terms in the Navier–Stokes equations (Kelly et al., 2016). The interactions are relevant for both predicting wave propagation and diagnosing sources and sinks of wave energy. Simplified equations arise when the waves have small amplitudes and different temporal and spatial scales from the background flow, for example, small-amplitude waves in a steady background flow with large spatial scales, can be modeled via ray-tracing using the "geometric approximation" a term borrowed from optics. In this regime, waves refract and reflect due to mean flow advection and changes in propagation speed (Fig. 9). When the background flow has spatial scales similar to a wavelength the geometric approximation is invalid and additional effects are possible, such as wave trapping at "critical levels" in background shear, instability due to background strain or divergence, and modification of the lower frequency cut-off by background vorticity. When the background flow is energetic, small-scale, and rapidly evolving, wave-background flow interactions are extremely complicated, so that even separating waves from the background flow is difficult.

Wave–Wave Interactions

Several processes cause the energy of the tide- and wind-generated waves to be scattered to different frequencies and wave numbers. One seems to be resonant wave–wave interactions: the nonlinear terms, involving $\mathbf{u}\cdot\nabla$, in the Navier–Stokes equations vanish identically for a single wave, but produce interaction terms if two waves are present (see Müller et al., 2015 for references). These terms, in the momentum and density equations, may be regarded as forcing terms with frequencies and wavenumbers given by the sum and difference frequencies and wavenumbers. For some pairs the sum (or difference) frequency is exactly what would be expected for a free wave with the sum (or difference) wavenumber, so this wave is now resonantly excited, acquiring energy from the original two waves. Detailed calculations for this theory, and using a different approach when the assumptions of weak interaction break down, suggest that there is a cascade of energy to waves with shorter length scales.

The direction in which energy flows in frequency is less clear, though one interaction mechanism, parametric subharmonic instability (PSI), akin to the excitation of a simple pendulum by oscillation of its point of support with twice the natural frequency of the pendulum, can produce small-scale waves with half the frequency of a large-scale parent wave (provided that this half-frequency is still greater than f).

Energetics and Mixing

Throughout the global ocean basins, internal waves either propagate away from their generation sources or break directly where they are generated (Fig. 10). For internal waves that propagate away from generation regions, their evolution has been documented as a

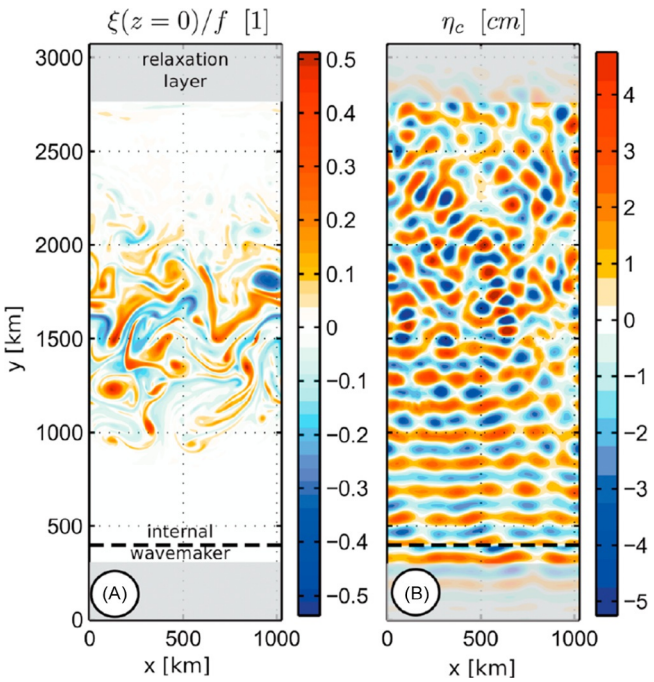

Fig. 9 Numerical simulation of a low-mode internal tide propagating through mesoscale turbulence. (A) Detided surface relative vorticity $\xi(z=0)$, normalized by the Coriolis frequency. (B) Internal-tide sea surface height η_c. The internal-tide wavemaker and relaxation layer are shown. The internal tides are scattered and refracted by the mesoscale turbulence. Reproduced with permission from: Ponte, A. L. and P. Klein (2015). Incoherent signature of internal tides on sea level in idealized numerical simulations. *Geophysical Research Letters* **42**, 1520–1526.

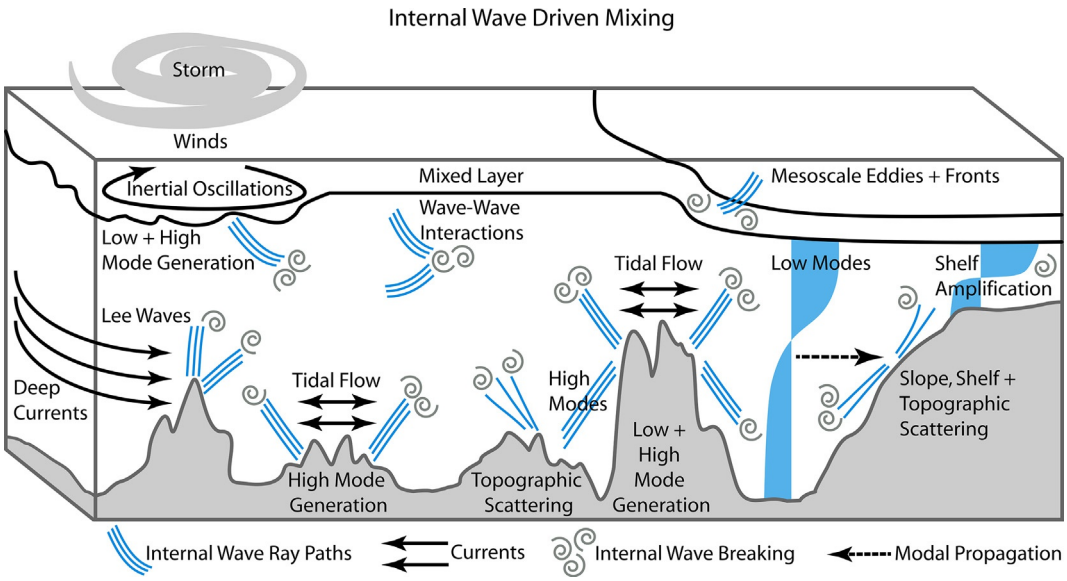

Fig. 10 Schematic showing the various processes generating internal waves such as tides, winds and mean flows (lee waves), and processes leading to the turbulent dissipation and loss of internal wave energy such as local dissipation, interaction with eddies and mean flows, wave–wave interactions, topographic scattering and interaction with the continental slope including shoaling and soliton generation. Turbulent dissipation is represented by curly greylines in the cartoon above. Reproduced with permission from: MacKinnon, J. A., et al. (2018). Climate process team on internal wave–driven ocean mixing. *Bulletin of the American Meteorological Society* **98**, 2429–2454.

cascade of energy from large scales to smaller scales, whereby they eventually break, completely losing their energy (MacKinnon et al., 2018). This energy cascade occurs because of wave–wave interactions as described above. Waves that do not lose their energy while propagating through the ocean basins eventually hit continental margins where they either break, losing their energy via turbulent dissipation or reflect back into the ocean to continue their journey across the basins.

Observations of this energy loss, or turbulent kinetic energy dissipation rate, can be obtained using a turbulent microstructure profiler (Moum and Osborn, 1986), which directly measures the small-scale turbulent shears associated with internal waves in the water column. Additionally, based upon the principles related to the energy cascade from large scales to small and wave–wave interaction theory, dissipation rates can be inferred from the finescale structure of density and/or velocity shear in the water column from platforms such as the global Argo array among other autonomous, moored and shipboard oceanic systems.

Global maps of turbulent energy dissipation rate in the upper 1–2000 m show considerable variability throughout the ocean basins (Fig. 11). These inferred estimates of dissipation from strain-based inferences using the Argo array have been found to agree well with more direct estimates of dissipation rates obtained from microstructure profiler estimates (Whalen et al., 2015). Enhanced dissipation is associated with regions of rough topography such as over the Mid-Atlantic and Hawaiian Ridges though it is not clear whether this results directly from the increased shear of the reflected and scattered waves, or via stronger wave–wave interactions at increased internal wave energies. The mid-latitudes in both northern and southern hemispheres show regions of elevated dissipation (with a seasonal cycle) associated with winter storms (and the wind-generated near-inertial waves they produce).

Mixing may be considerably more intense throughout the water column in some regions, such as the Southern Ocean where internal waves are generated due to the lee waves at rough topography. Conversely, rather weak mixing in the main thermocline of the ocean, together with much weaker stratification in abyssal areas of strong mixing, means that the overall energy loss from the internal wave field is small enough that it would take many tens of days to drain the observed energy levels. This may tie in with the remarkable feature that observed internal wave energy levels in the ocean seem to be rather uniform in space and time, at least much more so than for surface gravity waves; there is no such thing as an "internal calm". The interpretation is that the decay time of internal waves is, unlike the situation for surface waves, considerably longer than the interval between generation events.

There are exceptions to this picture, of course, with, for example, much lower internal wave energy levels and mixing in the Arctic Ocean (except near some topographic features), perhaps as a consequence of less wind generation, because of the protective ice cover, as well as rather weak tidal currents. However, as the climate shifts and polar temperatures rise, ice cover in the Arctic will be reduced and we may see an increase in wind-generated internal waves, which may result in enhanced mixing and redistribution of heat, which will further reduce the ice-cover. This positive feedback loop may shift our understanding of current internal wave properties in the Arctic (Dosser and Rainville, 2016).

The All-Encompassing Frequency-Wavenumber Spectra

The presence of tidal and near-inertial waves in the ocean and their scattering to smaller scales and different frequencies can be neatly captured in frequency and wavenumber spectra. Measurements of the frequency spectra of internal waves can be obtained

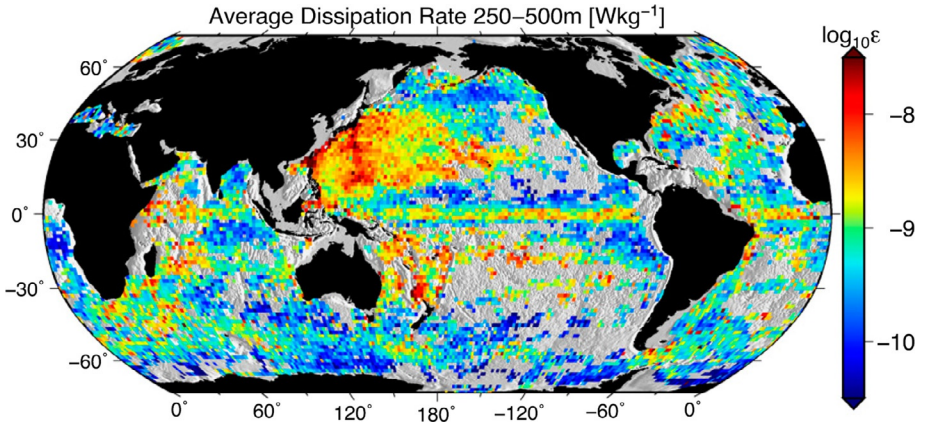

Fig. 11 Average turbulent kinetic energy dissipation rate estimates, epsilon (W/kg), derived by applying a strain-based finestructure method using vertical profiles of density between 250 and 500 m from the global Argo array. Reproduced with permission from: Whalen, C.B., MacKinnon, J. A., Talley, L. D., and Waterhouse, A. F. (2015). Estimating the mean diapycnal mixing using a finescale strain parameterization. *Journal of Physical Oceanography* **45**, 1174–1188.

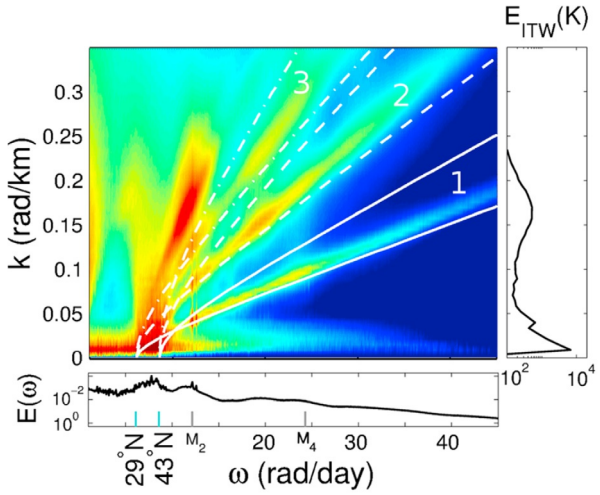

Fig. 12 The frequency-wavenumber (K–ω) spectra of kinetic energy $E(K, \omega)$ with units (m/s)^2 [(day) (km)] on a $\log_{10}$ scale computed using data from the global HYCOM simulations with a horizontal resolution of 1/25°. The magnitude of the horizontal wave number $K = (k^2 + l^2)^{1/2}$. The spectra are computed for a rectangular box in the North Pacific. The white curves represent bounding linear dispersion relations (Eq. (3)) of the first three internal wave vertical modes (solid = mode 1; dashed = mode 2; and dash dotted = mode 3). The mode bands are labeled inside the panel with white lettering. Below and to the right of the color-contoured $E(K, \omega)$, the spectra $E(\omega)$ integrated over all horizontal wavenumbers K, and $E_{ITW}(K)$ integrated over frequencies larger than the inertial frequency ($\omega \geq f_{29N}$), are shown, respectively. In the $E(\omega)$ spectra, cyan lines indicate the inertial frequencies at the bounding latitudes 29°N and 43°N, while M_2 and M_4 tidal frequencies are also indicated. Reproduced with permission from: Müller, M., Arbic, B. K., Richman, J. G., Shriver, J. F., Kunze, E. L., Scott, R. B., Wallcraft, A. J., and Zamudio, L. (2015). Toward an internal gravity wave spectrum in global ocean models. *Geophysical Research Letters* **42**, 3474–3481.

from analysis of time-series of measurements, at a fixed point, by current meters or by temperature sensors that show changes associated with vertical motion of the temperature-stratified water. Such measurements show a block of energy at frequencies between f and N. For currents there is typically a peak (an "inertial cusp") near f (see also bottom subplot of Fig. 12), but this is suppressed in temperature data as near-inertial motions are largely horizontal. Both temperature and velocity variance spectra show tidal peaks (see Fig. 12). At frequencies higher than f, and higher than the tidal frequencies, the spectra of the internal gravity wave continuum fall off with an energy distribution close to ω^{-2}. The continuum spectrum was described by Garrett and Munk (1975) and is thought to arise from wave–wave interactions between internal waves, with wind-generated near-inertial energy and internal tide energy being the primary sources.

Measurements at a single fixed point do not, however, provide information on the vertical and horizontal wavenumber content of the energy at any frequency. It is possible to compute vertical wavenumber spectra from current meter data sampled at many vertical depths or from electromagnetic velocity profilers (Leaman and Sanford, 1975). An example of a horizontal wavenumber spectrum from a numerical model is shown in the rightmost subplot of Fig. 12. The mode 1 and 2 wavenumbers are the small peaks near 0.04 and 0.09 rad/m.

It is feasible to compute frequency and wavenumber spectra that begin to resolve the internal gravity wave continuum using numerical ocean models, provided they have atmospheric forcing (to generate near-inertial flows), tidal forcing (to generate internal tides), and high grid resolution (to resolve wave–wave interactions), all at once. Such internal gravity wave spectra cannot be computed from satellite altimeter data, which lack sufficient time sampling. Recently, Müller et al. (2015) computed the horizontal wavenumber-frequency spectra from global 1/25° HYCOM simulations with simultaneous tidal and atmospheric forcing. These frequency and wavenumber spectra (black curves), and frequency-wavenumber spectra (colors) are shown in Fig. 12. Because the frequency-wavenumber spectra are computed over a box covering 14° in latitude, the near-inertial frequencies cover a wide band (see $E(\omega)$ curve in bottom subplot), and the Stürm-Liouville problem yields a band of (white) dispersion curves for each vertical mode in the color plot. High variance is seen throughout the near-inertial energy band, and in the narrower semidiurnal tidal band. The most exciting new result of Müller et al. (2015) is the variance along the low-mode dispersion curves at frequencies higher than semidiurnal tidal frequencies, all of which represents strong evidence that the recent high resolution numerical simulations are generating a supertidal internal gravity wave continuum (Arbic, 2018).

Conclusions

Wind and tide generated internal gravity waves are ubiquitous in the global ocean where they occur mostly as interfacial waves. Internal gravity waves are a key ingredient of the behavior of the ocean as they affect the large-scale circulation through mixing. In recent years much progress has been made in understanding the energetics of the internal gravity waves. Although numerical global ocean models obtain realistic solutions for low-mode waves as compared to observations, higher vertical mode wave generation and propagation, wave–wave interactions, interactions with the mean state, reflection and scattering from the rough seafloors, and wave-breaking processes are still not adequately simulated, as the models' resolution is not yet sufficient to resolve all these processes. Hence these processes must be parameterized by subgridscale schemes, such as topographic internal wave drag and wave-breaking parameterizations. We now have a better understanding of many of these processes than 10 years ago, and how to resolve and/or parameterize them, but we are still working towards a complete understanding of internal gravity waves in the ocean, and towards the goal of correctly simulating them in all their complexity.

Acknowledgements

This work is based in part on a previous edition by Chris Garrett. We thank him for laying the groundwork. We acknowledge Edward D. Zaron and Gordon Stephenson for preparing Fig. 7.

References

Alford MH, MacKinnon JA, Simmons HL, and Nash JD (2016) Near-inertial internal gravity waves in the ocean. *Annual Review of Marine Science* 8: 95–123.

Arbic BK, et al. (2018) A primer on global internal tide and internal gravity wave continuum modeling in HYCOM and MITgcm. In: Chassignet EP, Pascual A, Tintore J, and Verron J (eds.) *New Frontiers in Operational Oceanography*. GODAE OceanView.

D'Asaro EA (1985) The energy flux from the wind to near-inertial motions in the surface mixed layer. *Journal of Physical Oceanography* 15: 1043–1059.

Dosser HV and Rainville L (2016) Dynamics of the changing near-inertial internal wave field in the Arctic Ocean. *Journal of Physical Oceanography* 46: 395–415.

Ekman VW (1904) On dead water. In: Nansen F (ed.) *The Norwegian north polar expedition 1893–1896. Scientific Results. Volume V*, pp. 83–278. Christiania.

Fu L-L, Alsdorf D, Morrow R, Rodriguez E, and Mognard N (2012) *SWOT: The surface water and ocean topography mission: Wide-Swatch altimetric measurement of water elevation on earth*. Pasadena, California: Jet Propulsion Laboratory.

Garrett CJR and Munk WH (1975) Space-time scales of internal waves. A progress report. *Journal of Geophysical Research* 80: 291–297.

Gill AE (1982) *Atmosphere-ocean dynamics. International Geophysics Series 30*. London: Academic Press.

Helland-Hansen B and Nansen F (1909) The Norwegian sea—its physical oceanography based upon the Norwegian researches 1900–1904. In: *Report on Norwegian fishery and marine investigations*, Vol. II(2), Kristiania: Det Mallingske Bogtrykkeri.

Kelly SM, Lermusiaux PFJ, Duda TF, and Haley PJ Jr. (2016) A coupled-mode shallow water model for tidal analysis: Internal-tide reflection and refraction by the Gulf Stream. *Journal of Physical Oceanography* 46: 3661–3679.

Leaman KD and Sanford TB (1975) Vertical energy propagation of inertial waves: A vector spectral analysis of velocity profiles. *Journal of Geophysical Research* 80: 1975–1978.

MacKinnon JA, et al. (2018) Climate process team on internal wave-driven ocean mixing. *Bulletin of the American Meteorological Society* 98: 2429–2454.

Moum JN and Osborn TR (1986) Mixing in the main thermocline. *Journal of Physical Oceanography* 16: 1250–1259.

Müller M, Arbic BK, Richman JG, Shriver JF, Kunze EL, Scott RB, Wallcraft AJ, and Zamudio L (2015) Toward an internal gravity wave spectrum in global ocean models. *Geophysical Research Letters* 42: 3474–3481.

Munk WH (1966) Abyssal recipes. *Deep Sea Research and Oceanographic Abstracts* 13: 707–730.

Nansen F (1897) *Farthest North: The epic adventure of a visionary explorer*. Skyhorse Publishing.

Nikurashin M and Ferrari R (2011) Global energy conversion rate from geostrophic flows into internal lee waves in the deep ocean. *Geophysical Research Letters* 38: L08610.

Scotti A (2011) Inviscid critical and near-critical reflection of internal waves in the time domain. *Journal of Fluid Mechanics* 674: 464–488.

Whalen CB, MacKinnon JA, Talley LD, and Waterhouse AF (2015) Estimating the mean diapycnal mixing using a finescale strain parameterization. *Journal of Physical Oceanography* 45: 1174–1188.

Wunsch C (1975) Internal tides in the ocean. *Reviews of Geophysics* 13: 167–182.