Challenges and Prospects in Ocean Circulation Models


Over the past decade, ocean circulation models evolved. New diagnostics being used for evaluating models parameterizations of the mixing and overturning processes in boundary layers and its consequences are assessed through perturbed physics models. Relatedly, atmospheric models has stimulated new ideas and driven improvements in numerics. Observations have provided insight into turbulence and mixing around the globe at large scales and submesoscale effects are parameterized in most global models. The scales routinely resolved in process models and permitted in a few global models, are simulations of the Southern Ocean, the Meridional Overturning Circulation and its variability, and regional sea level change. Submesoscale variability is now routinely resolved in process models and permitted in a few global models, and submesoscale effects are parameterized in most global models. The scales where nonhydrostatic effects become important are beginning to be resolved in regional and process models. Coupling to sea ice, ice shelves, and high-resolution atmospheric models has stimulated new ideas and driven improvements in numerics. Observations have provided insight into turbulence and mixing around the globe and its consequences are assessed through perturbed physics models. Relatedly, parameterizations of the mixing and overturning processes in boundary layers and the ocean interior have improved. New diagnostics being used for evaluating models parameterizations of the mixing and overturning processes in boundary layers and its consequences are assessed through perturbed physics models. We revisit the challenges and prospects for ocean circulation models following Griffies et al. (2010). Over the past decade, ocean circulation models evolved through improved understanding, numerics, spatial discretization, grid configurations, parameterizations, data assimilation, environmental monitoring, and process-level observations and modeling. Important large scale applications over the last decade are simulations of the Southern Ocean, the Meridional Overturning Circulation and its variability, and regional sea level change. Submesoscale variability is now routinely resolved in process models and permitted in a few global models, and submesoscale effects are parameterized in most global models. The scales where nonhydrostatic effects become important are beginning to be resolved in regional and process models. Coupling to sea ice, ice shelves, and high-resolution atmospheric models has stimulated new ideas and driven improvements in numerics. Observations have provided insight into turbulence and mixing around the globe and its consequences are assessed through perturbed physics models. Relatedly, parameterizations of the mixing and overturning processes in boundary layers and the ocean interior have improved. New diagnostics being used for evaluating models parameterizations of the mixing and overturning processes in boundary layers and its consequences are assessed through perturbed physics models. We revisit the challenges and prospects for ocean circulation models following Griffies et al. (2010). Over the past decade, ocean circulation models evolved through improved understanding, numerics, spatial discretization, grid configurations, parameterizations, data assimilation, environmental monitoring, and process-level observations and modeling. Important large scale applications over the last decade are simulations of the Southern Ocean, the Meridional Overturning Circulation and its variability, and regional sea level change. Submesoscale variability is now routinely resolved in process models and permitted in a few global models, and submesoscale effects are parameterized in most global models. The scales where nonhydrostatic effects become important are beginning to be resolved in regional and process models. Coupling to sea ice, ice shelves, and high-resolution atmospheric models has stimulated new ideas and driven improvements in numerics. Observations have provided insight into turbulence and mixing around the globe and its consequences are assessed through perturbed physics models. Relatedly, parameterizations of the mixing and overturning processes in boundary layers and the ocean interior have improved. New diagnostics being used for evaluating models...
alongside present and novel observations are briefly referenced. The overall goal is summarizing new developments in ocean modeling, including: how new and existing observations can be used, what modeling challenges remain, and how simulations can be used to support observations.

Keywords: ocean circulation, model, parameterization, climate, ocean processes

1. INTRODUCTION AND SCOPE

The ocean circulation is a critical part of modeling the overall earth system (Chassignet et al., 2019). The oceans are the major reservoir of thermal energy important for climate sensitivity and thermosteric sea level rise (Flato et al., 2013; Palmer et al., 2018), and ocean and ice modeling is increasingly important as weather forecasts in polar and marine climates are extended (Saha et al., 2014; Belcher et al., 2015; Hewitt et al., 2017). The ocean is a primary reservoir for anthropogenic carbon (Khatiwala et al., 2009, 2013), and the oceans contain and affect many important ecosystems and resources for society.

The equations for oceanic motions have been long known (Navier, 1822; Laplace et al., 1829; Stokes, 1845; Onsager, 1931), and subtleties of seawater thermodynamics are also understood (IOC et al., 2010). At present only the ACCESS (MOM5 ocean) and the IPSL-CM6 (NEMO ocean model) have implemented and plan operational use in CMIP6 of the new TEOS-10 equation of state with the exception of the preformed salinity. Note that other modeling centers using MOM and NEMO are not using TEOS-10 for CMIP6. Appendix D of Griffies et al. (2016) features a detailed discussion of present equation of state considerations for modeling, but also state, “there remain unanswered research questions raised by IOC et al. (2010), in particular regarding the treatment of salinity. For CMIP6, we cannot impose strict standards defining what it means to be ‘TEOS-10 compliant’ when research remains incomplete.”

However, the fluid and thermodynamics equations are not yet directly useful, as their discretization without further approximations and parameterizations would require computers about 10 billion times faster and bigger in storage than present supercomputers. These remain about two centuries in the future (Fox-Kemper et al., 2014; Fox-Kemper, 2018), see especially the figure illustrating scaling behavior of CMIP models in (Fox-Kemper et al., 2014)1.

Thus, numerical and parameterization improvements will continue to define the state of the art in ocean circulation modeling, in concert with integration and co-analysis of observations. Unlike nature, a computer model rarely indicates if processes are left out; it simply provides an incorrect answer. Experiment and observation are how model biases and unrepresented processes are revealed. Furthermore, forecast and state estimation systems are an increasingly valuable tool in providing context and inferences from observations. These tools expand the reach of observations and can improve observation plans, but they rely on the fidelity of their underlying model.

This article is divided into 9 sections covering: (1) introduction; (2) equations, numerics, and discretization; (3) coupled ocean-cryosphere modeling including particularly sea ice and ice shelves; (4) coupled ocean-atmosphere modeling; (5) coupled ocean-surface wave modeling; (6) ocean modeling parameterizations; (7) ocean model diagnosis and evaluation; (8) novel applications of ocean models; and (9) what to expect by 2030, including an assessment of the Griffies et al. (2010) prediction of ocean modeling improvements since 2010. The emphasis throughout is on improvements over the last decade of modeling.

2. EQUATIONS AND DISCRETIZATION OF OCEAN MODELS

The continuum equations of motion for seawater are known (Müller, 2006). However, present ocean models tend to make the Boussinesq, traditional and hydrostatic approximations in addition to discretizing the continuum equations and parameterizing unresolved processes (Fox-Kemper, 2018).

2.1. Vertical Discretization

As of Griffies et al. (2010), three primary approaches to vertical discretization in ocean models were typical: z-coordinate models, which discretize the geometric distance below the geoid as a vertical coordinate; σ-coordinate models, which scale the distance between the sea surface and the terrain into discrete intervals; and isopycnal-coordinate models where vertical discretizations are based on the density stratification. At that time, hybrid vertical coordinate models existed, but the term “hybrid” had multifarious meanings: to some it denoted a linear combination of two or more conventional coordinates (Song and Haidvogel, 1994; Ezer and Mellor, 2004; Barron et al., 2006) or to others it was a truly generalized coordinate, i.e., aiming to mimic different types of coordinates in different regions (Bleck, 2002; Burchard and Beckers, 2004; Adcroft and Hallberg, 2006; Chassignet et al., 2006; Song and Hou, 2006).

This century significant advances in ocean and atmospheric model numerics have occurred by integrating ideas from Arbitrary Lagrangian-Eulerian (ALE) numerical approaches. Adcroft and Hallberg (2006) classify generalized coordinates ocean models as either Lagrangian Vertical Direction (LVD) or Eulerian Vertical Direction (EVD) models. In LVD models, the continuity (thickness tendency) equation is solved forward in

1Historically, from the Intergovernmental Panel on Climate Change (IPCC) reports FAR (First Assessment Report) to AR6 (Sixth Assessment Report) the median-resolution and highest-resolution IPCC-class atmosphere-ocean coupled models has increased exponentially, doubling every 12.5 and 7.5 years, respectively. However, if the models for AR6 are left out, the doubling rate is revealed to have been about 10% faster in preceding reports: 10.2 and 6.9 years, respectively. This indicates a slow-down in recent ocean model resolution improvements.
time throughout the domain, while an Arbitrary Lagrangian-terrain (ALE) technique is used to re-map the vertical coordinates and maintain different coordinate types within the domain (HYCOM: Bleck, 2002; MOM6: Adcroft and Hallberg, 2006). In contrast, EVD models use the continuity equation to diagnose the velocity crossing the coordinate surfaces, with examples being the z-coordinate models of Bryan (1969), Cox (1984), and Pacanowski et al. (1991) and successors (Smith et al., 2010); σ-coordinate models like Haidvogel et al. (2008), and generalized EVD approaches facilitating a variety of vertical coordinates like MOM5 (Griffies, 2012), MPAS (Ringerler et al., 2013), and NEMO (Gurvan et al., 2017). The added features of these improved modeling platforms provide the opportunity to increase the model time step (via ALE removal of the vertical CFL), reduce spurious numerical mixing (via ALE remapping to continuous isopycnal coordinates), and directly simulate freshwater addition and removal at the ocean surface (enabled in both LVD and EVD through replacing the rigid lid approach of Bryan (1969) with a free surface). Recent examples of successful applications of these techniques are Hofmeister et al. (2010), Leclair and Madec (2011), Petersen et al. (2015), and Zanowski et al. (2015).

2.2. Horizontal Discretization

Traditionally, ocean models have relied on finite difference or finite volume schemes based on staggered rectilinear grids and the hydrostatic approximation which aided in accuracy and conservation of key properties (Arakawa and Lamb, 1977), but newer approaches enable moving away from these simpler schemes.

2.2.1. Structured Meshes

Nesting of rectilinear structured models—particularly ROMS (Shchepetkin and McWilliams, 2005), NEMO-AGRIF (Debreu et al., 2008) and the MITgcm (Adcroft et al., 2008)—have shown the utility of being able to simultaneously represent small-scale processes and also large basins. Many studies of submesoscale dynamics have taken this approach (e.g., Capet et al., 2008; Gula et al., 2016), recently focusing on air-sea coupling on small scales (Seo et al., 2009; Small et al., 2015; Renault et al., 2016). A few global and basin-scale simulations with resolution near 1/50° are approaching submesoscale resolution throughout the domain e.g., HYCOM (Chassignet and Xu, 2017); MITgcm (Qiu et al., 2018); and an ongoing effort using NEMO.

2.2.2. Unstructured Meshes

Recent progress in developing unstructured models such as ADCIRC (Luettich and Westerink, 2000; Dietrich et al., 2012), FESOM (Wang et al., 2014; Sidorenko et al., 2015; Danilov et al., 2017; Wang Q. et al., 2018), FVCOM (Chen et al., 2003), ICON (Korn, 2017), MPAS (Ringerler et al., 2013), and SLIM (Delandmeter et al., 2018; Vallaereys et al., 2018) is making multi-resolution approaches more flexible and natural. One challenge is that new users find it difficult to write analysis software, and modeling centers are addressing this problem through the development and distribution of analysis packages in higher-level languages such as Python and Matlab. A second challenge is the lack of well-tested scale-aware and flow-aware parameterizations capable of transitioning between coarse-resolution and fine-resolution (Haidvogel et al., 2017).

2.2.3. Hydrostatic Approximation

Traditional ocean modeling has relied on the hydrostatic approximation, which makes accurate estimation of the pressure field, and thus geostrophic flows, easier. This approximation is valid when the aspect ratio of the motions are shallow, which can be guaranteed by using shallow grid cells or found in high stratification. In contrast, oceanic boundary layer Large Eddy Simulations (e.g., McWilliams et al., 1997) cannot use the hydrostatic approximation. As horizontal resolutions increase, nonhydrostatic effects become increasingly strong and presently are detectable at the small-scale end of the submesoscale (Mahadevan, 2006; Hamlington et al., 2014; Suzuki et al., 2016) and in deep convection (Marshall et al., 1997). Few ocean circulation models presently have nonhydrostatic capability. Boundary layer LES models have always been nonhydrostatic (e.g., McWilliams et al., 1993), but some ocean circulation models—beginning with PSOM (Mahadevan et al., 1996a,b) and the MITgcm (Marshall et al., 1997), but more recently SUNTANS (Fringer et al., 2006) and CROCO-ROMS (Cambon et al., 2018)—have this capability.

2.3. Numerical Artifacts

2.3.1. Stability, Conservation, Convergence, and Performance

A good numerical scheme can be shown to converge as resolution increases to the solution of the partial differential equations it approximates discretely. Great numerical schemes converge efficiently, using fewer processor operations, less communication, and/or less memory.

Convergence requires both consistency (i.e., the discrete equations represent the correct differential equations) and stability (i.e., the numerical scheme does not “blow up” upon iteration) (Lax and Richtmyer, 1956). Model crashes can be very costly in a supercomputing setting where model initialization costs are large, so robust stability is a requirement for ocean model numerics. The easiest numerical schemes to implement, and thus most widely used, are explicit methods which are typically conditionally stable: stability occurs only for sufficiently small time steps. Unconditional stability, such as from some implicit schemes, can be desirable but can come at the expense of accuracy. Although unconditional stability allows very large time steps, processes faster than the time step will be neglected or poorly simulated resulting in inconsistency. Griffies and Adcroft (2008) summarize the variety of time step constraints encountered in ocean circulation modeling.

Consistency is easily demonstrated in simple cases, but frequently is difficult to assess in complex oceanic flows. Some methods may be very well-suited to particular phenomena, such as fronts, while another may be better for general purposes but fail at fronts. Most equations of motion are budgets for conserved properties. Mimetic numerics offer the means to exactly conserve discrete analogs of the continuum budgets and continuous equation symmetries—such as conservation
laws, solution symmetries, and the fundamental identities and theorems of vector and tensor calculus—thus mimetic methods frequently assure consistency (Hyman et al., 2002). Spectral schemes, on the other hand, are difficult to make conservative and may spontaneously generate local extrema, but have high order accuracy. Spectral element, discontinuous Galerkin methods, and related schemes seek a happy medium of mimicry and differential accuracy.

Recent improvements in numerics have greatly enhanced the stability of ocean climate models. ALE schemes have shown better stability, parameterizations have led to improved stability (Fox-Kemper and Menemenlis, 2008), and Newton-Krylov and implicit schemes have improved (Nadiga et al., 2006; Lindsay, 2017). Newton-Krylov approaches have also increased the efficiency of spinning up the deep ocean to simulate important tracer distributions (Bernsen et al., 2008; Khatiwala et al., 2009; Bardin et al., 2014).

Scheme efficiency differs with computer architectures. At present, most ocean modeling is carried out on systems featuring hundreds to thousands of computing cores (CPUs), organized with shared memory and extremely rapid communications among nodes with roughly a dozen cores and somewhat slower inter-node communication. Graphical processing units (GPUs) are much faster and power-efficient than normal cores, but are not as general-purpose, do not perform as well with branches, and are typically limited in memory. Cloud computing is an exciting innovation, offering ready access to diverse and affordable heterogeneous hardware, but communication over the internet is much slower than among nodes in a supercomputer so new software engineering and algorithms are needed (Hamman et al., 2018).

2.3.2. Spurious Diapycnal Mixing

A classic ocean numerics challenge is respecting the tiny diapycnal diffusion acting on tracers in the presence of rapid advection and diffusion oriented along neutral direction (Veronis, 1975; Griffies et al., 2000). This spurious mixing problem is exacerbated in the mesoscale eddying regime emerging in ocean climate models. The numerical difficulty arises from the cascade of tracer variance to the grid scale engendered by transient eddy features in an advectively dominant flow. Numerical tracer advection operators must satisfy two competing goals: avoiding false extrema manifest as grid-scale noise and minimizing spurious diapycnal mixing.

There are two sources of spurious mixing: numerical mixing associated with vertical advection across the model vertical coordinate, and numerical mixing in the horizontal direction due to lateral processes where the inclination of isopycnals projects into a diapycnal effect (Gibson et al., 2017). Isopycnal coordinates minimize both sources (Bleck, 1998) for regions where isopycnals approximate neutral directions. However, purely isopycnal models are limited in other ways such as in the representation of weakly-stratified boundary layers, nonhydrostatic flows, and overturning, as well as in capturing the full suite of watermass transformations (IOC et al., 2010; McDougall et al., 2014).

Reducing spurious mixing in Eulerian vertical coordinate models requires high order accuracy of the numerical advection operators in all three dimensions. This approach can be sufficient for idealized model studies (Griffies et al., 2000; Hill et al., 2012); however, the accuracy of advection operators is degraded in limiting fluxes to avoid false tracer extrema in realistic models. Griffies et al. (2000) and Ilicak et al. (2012) point to the importance of tuning lateral friction to ensure that all admitted flow features are properly resolved by the chosen grid spacing, thus allowing for the reduced reliance on flux limiters. Bachman et al. (2017a) show that physically-based, scale-aware parameterizations can achieve the needed flow regularization if suitably matched to the class of flow being simulated.

The ALE algorithm provides a means to use arbitrary vertical coordinates, thus offering a promising middle ground between the isopycnal layer approach and the Eulerian approach (Bleck, 2002, 2005). The general strategy is to make use of a high order vertical remapping scheme (White and Adcroft, 2008; White et al., 2009) within a continuous isopycnal coordinate interior, thus allowing for the full suite of nonlinear equation of state effects while retaining the integrity of transport along neutral directions. Further work is needed to fully assess the ability of ALE to resolve the spurious mixing problem in the presence of an active eddying flow while retaining other capabilities needed for realistic ocean modeling (e.g., high latitude processes).

2.3.3. Implicit Dissipation

In order to maximize stability, many schemes are designed to tend toward dissipation of energy or tracer variance when numerical errors are inevitable. Dissipative artifacts are typical byproducts in this class of numerical schemes, including upwinding, monotonic, flux-limiting, and shock-capturing schemes. This implicit dissipation can add to physical parameterizations of diffusivity leading to overestimation (Bachman et al., 2017a). For this reason, it is sometimes necessary to select the parameterizations and numerics in tandem, thereby selecting numerics reducing the implicit dissipation if a dissipative parameterization is expected to be sufficient (Bachman et al., 2017a), or a dissipative numerical scheme if explicit parameterizations are unknown (Shchepetkin and McWilliams, 2005).

3. OCEAN-CRYOSPHERE MODELING

3.1. Sea Ice

Sea ice is an important component of an state-of-the-art ocean–ice modeling framework. It strongly modifies atmosphere–ocean interactions inciting important climate feedbacks. Griffies et al. (2010) did not address sea ice, but recent progress in those models induced a treatment following Smith et al. (2018), who divides to sea ice studies into four subtopics: (1) sea-ice fractures and cracks, (2) sea-ice interactions with solids, (3) sea-ice interactions with fluids, and (4) multi-scale sea-ice modeling. Large- and mesoscale resolutions and ice-ocean coupling are the topic here, which make the recent findings on topics 3–4 most important (see also Lemieux et al., 2018). Effects on smaller than geophysical scales are neglected here, such as ice micro-structure and off-shore structures in ice-covered oceans.
3.1.1. Sea-Ice Thickness Distribution
Multiple category sea-ice Thickness Distribution (ITD) parameterization has become standard in large-scale sea-ice modeling (Hunke, 2010; Rousset et al., 2015). It better represents the seasonal cycle of sea-ice volume evolution, particularly the spring melt, when compared to single category parameterization (Uotila et al., 2017). The nonlinear dependence of heat conduction through vertical layers of ice and snow, and thus the thermodynamic growth rate at the bottom layer, also makes the implementation of ITD necessary in multi-level models (Castro-Morales et al., 2014). Hunke (2014) conclude that the commonly-used five ice thickness categories are not enough to accurately simulate ice volume.

3.1.2. Sea-Ice Salinity
Following developments such as Vancoppenolle et al. (2009), Turner and Hunke (2015) develop a sea-ice model with a fully prognostic and variable ice salinity vertical profile. After adjustments of gravity drainage parameterization, their model produced good agreement with salinity observations being also sensitive to the melt-pond area and volume. The new thermodynamics produced realistically thicker Arctic ice than the old, constant salinity thermodynamics.

3.1.3. Melt Ponds
A melt pond theory was developed by Flocco et al. (2010) and incorporated in a state-of-the-art sea-ice model with an ITD. They recommend the inclusion of melt ponds in sea-ice models with increased number of thin ice thickness categories. Simulation based results by Kim et al. (2018) indicate that the salinity of melt ponds may control their heat balance–an aspect not yet taken into account in model parameterization.

3.1.4. Ice-Waves Interaction
Bennetts et al. (2015) implement an idealized wave–ice interaction model taking into account both wave-induced ice fracture and wave attenuation due to ice floes. Bennetts et al. (2017) analyze wave-induced breakup of Antarctic ice using a sea-ice model that is a common component of many climate models. Williams et al. (2017) couple a waves-in-ice model with a sea-ice model and studied the effects of ocean surface wave radiation stress compared to wind stress. The wave scattering in the Marginal Ice Zone (MIZ) has similar properties for a large range of ice types and wave periods which allows the computation of temporal evolution of wave packets through the MIZ (Meylan and Bennetts, 2018). As sea-ice models poorly represent ice break-up and formation with ocean waves, satellite observations are particularly beneficial in quantifying wave impacts (Stopa et al., 2018).

3.1.5. Sea-Ice Floe Size Distribution (FSD)
A sea-ice floe size distribution (FSD) improves process precision, mainly for ice–waves interactions in the MIZ and seasonal ice zones. Zhang et al. (2015) develop a theory where FSD and ITD jointly evolve taking into account FSD changes due to ice advection, thermodynamic growth, lateral melting and wave-induced ice fragmentation. Horvat and Tziperman (2015) develop and test idealized cases of a model for the evolution of the joint FSD and ITD, including lateral freezing and melting, mechanical interactions of floes, and floe fracturing by waves. Horvat et al. (2016) show that sea-ice melting-related processes at floe boundaries provoke ocean mixed-layer submesoscale eddies that transport heat horizontally. They emphasize that these processes, FSD thermodynamics, and the associated ocean eddies are neglected in present climate models. Roach et al. (2018) argue that FSD is important for the accurate simulation of polar oceans and climate. They developed a model where floe sizes are prognostic and depend on new ice formation, welding of floes in freezing conditions, lateral growth and melt, and wave fracture of floes.

3.1.6. New and Updated Rheologies
Ice dynamics modeling has progressed in the last decade. Losch and Danilov (2012) compare solutions for the sea-ice momentum equation based on implicit viscous-plastic (VP) and explicit elasto-viscous-plastic (EVP) methods, and found rather different solutions, yet EVP solutions do converge toward the VP results as time steps approach zero. Bouillon et al. (2013) revisit EVP and confirm unsatisfactory convergence and provide adjustments for the internal ice stress to filter out artificial linear stress features. Kimmritz et al. (2016) continue to test the stability and convergence of the modified EVP method and recommend adaptive variation of modified EVP scheme parameters. Lemieux et al. (2014) develop and implement a new time-stepping method so that sea-ice concentration and thickness evolve implicitly. The new scheme improves tenfold the accuracy of long time steps and is computationally five times more efficient.

Progress has been made in implementing anisotropic rheology in the pack ice field. Hunke (2014) find that an anisotropic rheology slows the ice motion by increasing shear stress between floes, while variable air–ice and ice–ocean drag coefficients lead to thinner ice. An elastic anisotropic plastic rheology produces sharper ice field deformation features and follows a more realistic power law strain rate than the commonly adapted EVP rheology (Heorton et al., 2018).

Arctic sea ice change over the past few decades represents a new dynamical regime with more mobile ice and stronger deformation, challenging traditional sea-ice models. To address this, Rampal et al. (2016) develop a Lagrangian sea-ice model with an elasto-brittle rheology and thermodynamics with healing capability for extreme fracturing. Their new model captures sea-ice drift and deformation, and the seasonal cycles of ice volume and extent improve. Rabatel et al. (2018) test the Lagrangian model of Rampal et al. (2016) for probabilistic forecasting in hindcast mode with uncertain winds driving the ice motion. They show that the Lagrangian sea-ice model (Rampal et al., 2016) is a viable alternative for sea-ice forecasting.

3.1.7. Atmosphere–Ice–Ocean Coupling
Tsamados et al. (2014) implement and test variable atmospheric and oceanic drag parameterizations that take into account ice surface features such as pressure ridges and melt pond edges in ice-only models. As drag coefficients are coupled to the sea-ice properties, they evolve temporally and spatially. Castellani et al.
implement variable air-ice and ice-ocean drag coefficients but in an ocean-ice model. Drag coefficients that vary result in faster moving, thinner and more realistic Arctic sea ice. In the ocean, variable drag coefficients deepen the mixed layer and alter surface salinity and temperature. Barthélémy et al. (2016) present an ITD based sub-grid-scale representation of momentum and buoyancy fluxes between ocean and ice. This effort separates estimates of convection and turbulent mixing by ice thickness category, weighted by fraction per grid cell. For long mixing time scales, the ITD-dependent mixing decreases the under-ice mean mixed layer depth and decreases oceanic heat flux melting ice in summer.

Coupled climate models typically treat sea-ice and the ocean as distinct components that only exchange information via a coupler, and they often lag the application of forces on the ocean relative to the ocean state upon which they are based. This approach has the advantage of permitting the separate development of coupler, ocean, and sea-ice model software. However, key aspects of the rapid sea ice and ocean dynamics are fundamentally coupled, and treating the ice and ocean as disconnected components can lead to coupled numerical instabilities associated with coupled gravity waves and ice-ocean stresses (Hallberg, 2014). To avoid these instabilities, coupled models may be forced to take shorter coupling time steps or make unphysical approximations. A more intimate coupling of the ice and ocean dynamics could avoid these instabilities and improve computational efficiency and physical fidelity.

### 3.1.8. Model Calibration

Massonnet et al. (2014) optimize sea-ice model parameter values with an ensemble Kalman filter method by calibrating the model against sea-ice drift observations in the Arctic. This optimisation resulted in a significant bias reduction of simulated ice drift speed and slight improvement in ITD. Losch et al. (2010) compare different numerical solvers for dynamics and thermodynamics used in sea-ice models and found that the choice of the dynamic solver has a considerable effect on the solution. Hunke (2010) adjusts many sea-ice model parameters to obtain the most realistic representation of global sea ice. In addition to albedo, which is a commonly used tuning parameter, many other parameters related to ice thermodynamics and dynamics can have equally large effect on global sea-ice distribution. More extensive adjustments on an ocean-ice model by Uotila et al. (2012) show that the sea-ice model adjustments matter most in summer, while in winter external factors dominate the evolution of sea ice. Recently, Sumata et al. (2018) used more than two decades of sea ice concentration, thickness, and drift data including their uncertainties in a cost function, to optimize 15 dynamic and thermodynamic process parameters with a genetic algorithm.

### 3.1.9. Other Developments

Exciting progress has also been made in other fields of sea-ice modeling during the last 10 years than listed above. For example, discrete element modeling of sea ice (Herman, 2016), parameterization of snow thermodynamics (Lecomte et al., 2011) and wind-blown snow drift (Lecomte et al., 2015), and landfast ice implementations (Lemieux et al., 2015) can now be treated in more advanced ways in ocean-ice climate models. There are also different subgrid issues at hand when modeling sea ice. Recent work on ice rheology (Feltham, 2008), wave-ice interactions (Williams et al., 2013), and submesoscale-floe size interactions (Horvat and Tziperman, 2017) are making their way to implementation. These improvements to ocean-ice climate model configurations are expected to increase their realism. Future work will assess the benefits of complexity in sea-ice models and enhanced resolution for climate projections. Satellite and in situ observations of sea ice and its atmospheric and oceanic environment will aid in evaluating the realism of these improvements.

### 3.2. Ice Shelves and Icebergs

High-latitude precipitation flows to the ocean often via land ice. Accumulation of snow forms ice sheets which flow slowly to the oceans. The land ice stores significant mass which makes understanding the evolution of the ice sheets, and the interaction between land-ice and oceans, paramount to understanding and predicting future sea level. The land ice-ocean boundary is difficult to access and thus poorly observed. Only recently has this interface been modeled and considered for inclusion in climate models. Where glaciers enter the ocean in fjords, there is often injection of melt water in plumes at depth at the base of the glacier (Jenkins, 2011; Strane and Cenedese, 2015). This configuration is common for Greenland and Arctic glaciers. In the Southern hemisphere, the Antarctic ice sheet feeds floating ice-sheles which form ocean-filled sub-ice shelf cavities. Ice shelves lose mass to the ocean by basal melting, controlled by ocean circulation and properties, or by calving of icebergs (Dinniman et al., 2016; Asay-Davis et al., 2017). The loss of grounded ice is what contributes to sea-level, while the freshwater input to the ocean is determined by the relative flux of basal melt to calving. Basal melting is an in situ freshening of seawater while icebergs can carry the freshwater offshore to the open ocean.

Gladstone et al. (2001) develop a model of icebergs trajectories in the ocean. Jongma et al. (2009) couple an iceberg model to an ocean model to examine iceberg-ocean interactions. Martin and Adcroft (2010) make the icebergs an integral part of the hydrologic cycle in a fully coupled model. Systematic improvements of the representation of icebergs and their interaction with the ocean continue (Marsh et al., 2015; Stern et al., 2016; Wagner et al., 2017; FitzMaurice and Stern, 2018). A challenge for modeling of icebergs is to simulate the observed distribution of size and shape. Using satellite altimetry, Tournadre et al. (2016) find a power-law distribution for small icebergs which has subsequently been used as an input to iceberg models. It is not clear whether the distribution is determined by the poorly understood calving process or the subsequent evolution of the icebergs. However, the majority of the iceberg mass flux into the ocean is due to the large and tabular icebergs. Large icebergs (say larger than 10 km) live longer and thus can transport fresh water further offshore but their size is comparable to ocean grid cells and thus the representation as point particles used in many iceberg models becomes a poor approximation. The problem is worse for the giant tabular icebergs which can
be 100’s of km wide. The detailed distribution and occurrence of giant tabular icebergs is poorly understood and potentially unpredictable and their representation in climate models is an active area of research (Stern et al., 2017).

Sub-ice shelf melting and ice sheet grounding line retreat/advance both amount to a changing ocean geometry, a process difficult to capture—and not represented—by most current-generation GCMs, but one that carries important ocean-ice feedbacks (thinning ice shelves reduce ice sheet buttressing, while grounding line migration impacts the ice sheet’s stress balance). Progress has nevertheless been made, beginning with approaches referred to as discontinuous (Gladish et al., 2012; Goldberg et al., 2012a,b; De Rydt and Gudmundsson, 2016; Mathiot et al., 2017). More recently, approaches have employed either asynchronous or synchronous coupling. In the former case, information between the components are exchanged once every few ice time steps (e.g., monthly), during which the ocean geometry is held fixed. Moving from one fixed ice-shelf topography to another at the coupling step leads to continuity issues with mass, heat, salt, and momentum in the ocean that have to be solved with ad-hoc techniques (Asay-Davis et al., 2016; Seroussi et al., 2017). In the latter case, mass continuity is evolved at the ocean model’s time step, thus ensuring that the freshwater (volume) flux translates conservatively into a corresponding change in geometry (vertical coupling). Furthermore, ice sheet retreat or advance (horizontal coupling) is achieved through maintenance of a thin or massless ocean layer below the (grounded) ice sheet, controlled by a porous flux (Goldberg et al., 2018; Jordan et al., 2018). Faithful representation of glacier-fjord (Greenland) and ice shelf-ocean (Antarctica) interactions remains among the premier challenges in climate modeling in support of sea level science.

3.3. Global Ocean/Sea-ice Simulations Based on CORE-I, CORE-II, and JRA55-do

The Coordinated Ocean-sea ice Reference Experiments (CORE) provide a framework for global ocean/sea-ice model simulations and inter-comparison. The CORE-I simulations documented by Griffies et al. (2009) made use of the atmospheric state from Large and Yeager (2004) to derive their bulk formula forcing. Even though all models made use of the same atmospheric state, the CORE-I simulations revealed previously unknown sensitivities across a suite of global ocean/sea-ice models. In the years since its design, CORE-I has proven to be a useful framework for model assessment and model-model comparison. It has thus become a standard method to develop and to assess global ocean/sea-ice model configurations.

Large and Yeager (2004) developed an artificial “Normal Year Forcing” (NYF) for indefinitely repeating the forcing year. This approach simplifies the task of spinning up a model to quasi-equilibrium and for measuring model sensitivity to changes in numerics and physics. However, the NYF precludes direct comparison of simulations to real-time data such as that increasingly produced from observational campaigns. These limitations led the modeling community to embrace the CORE-II interannual atmospheric state of Large and Yeager (2009). This dataset extends from 1948 to 2007 and forms the basis for nine papers that assess the physical integrity of roughly 20 global ocean/ice models: two papers focused on the North Atlantic circulation (Danabasoglu et al., 2014, 2016), one on sea level (Griffies et al., 2014), three on the Arctic (Ilicak et al., 2016; Wang et al., 2016a,b), two on the Southern Ocean (Downes et al., 2015; Farneti et al., 2015), and one on the Pacific (Tseng et al., 2016).

CORE-II has proven to be a useful means to benchmark global model simulations against observational measures, thus engaging the observational community in a manner unavailable from CORE-I. This collaborative assessment of global simulations is essential for improving both the models and the observations, particularly in a world where datasets available from simulations and observations are beyond the ability of a single research group to master.

The success of the CORE-II studies has left the modeling and observational community wanting more. It also revealed the limitations of the Large and Yeager (2009) dataset, most notably termination in year 2009 and by today’s standards a relatively coarse space and time resolution. To address this community need, Tsujino et al. (2018) have produced the next-generation global ocean/sea-ice forcing product known as JRA55-do, building from the Japanese reanalysis (Kobayashi et al., 2015) and inspired by the bias correction approaches of Large and Yeager (2009). JRA55-do offers refined temporal and spatial resolution, plus the capability of being extended consistently into the future as the atmospheric reanalysis is extended. Consequently, it is being used for the CMIP6 Ocean Model Intercomparison Project (OMIP) (Griffies et al., 2016; Orr et al., 2017). We thus anticipate that JRA55-do will soon become the community standard.

4. OCEAN-ATMOSPHERE MODELING

4.1. Fluxes

The forcing products for ocean-ice experiments differ from the behavior observed in coupled models (Bates et al., 2012b) as atmospheric models have biases of their own and because the coupled system can possess positive feedbacks that accentuate ocean and ice model issues and negative feedbacks that limit basin-scale coupled model drift. Even so, a critical application of ocean and sea ice models is their use as components of earth system models for the study of detection and attribution of anthropogenic climate change impacts, as well as tools for designing adaptation, mitigation, and geoengineering strategies. Hence, the fidelity of the coupled model behavior is paramount. This raises the question of whether it is more beneficial to evaluate and tune the ocean and ice components in forced settings, such as the CORE forcing paradigm, vs. in coupled mode. Given underlying biases in atmospheric forcing and the possibility of introducing compensating errors in coupled mode, the optimal strategy is most likely to require both forced and coupled approaches in tandem.

4.2. Boundary Layers

Atmospheric model parameterizations of convection and turbulence, as in the ocean, are simplified representations of
complex interactions. The added complexity of water vapor phase changes, clouds, and cloud and aerosol interactions affecting absorption and reflection of radiation make many aspects of atmospheric parameterizations different from oceanic parameterizations. The basic turbulence closures of the dry atmospheric boundary layer and the upper ocean boundary layer (see section 6.2) have many similarities and often common origination (e.g., Troehn and Mahl, 1986; Large et al., 1994). Large eddy simulation software is commonly shared between the two fluids (e.g., Sullivan and Patton, 2008) and many concepts and discoveries have been transferred between the fields.

5. SURFACE WAVE MODELING

With the exception of nearshore modeling, such as implementations of ADCIRC (Dietrich et al., 2012) and ROMS (Rong et al., 2014) coupled to WaveWatch-III (Tolman, 2009) or SWAN (Booij et al., 1999), surface gravity waves are typically not explicitly included in ocean circulation modeling. However, this attitude is changing with the growing awareness of the importance of wave breaking effects and wave-driven mixing (Melville, 1996; Babanin and Haus, 2009; Huang and Qiao, 2010), such as Langmuir turbulence (Belcher et al., 2012; D’Asaro et al., 2014; Li et al., 2016) and the closely related non-breaking surface-wave-induced turbulence (Qiao et al., 2004, 2016), and variations in sea state and surface roughness effects on air-sea momentum coupling (Kukulka et al., 2007; Chen et al., 2018), sea spray aerosol production (Long et al., 2011) and sea spray effects on typhoon forecasts (Zhao et al., 2017), gas transfer (Miller et al., 2009), wave effects on currents (McWilliams and Fox-Kemper, 2013; Suzuki et al., 2016) and current effects on waves (Gallet and Young, 2014; Ardhuin et al., 2017), wavy Ekman layers (McWilliams et al., 2012), and wave impacts on a variety of other climate processes (Cavaleri et al., 2012). Wave models have long been used in numerical weather prediction systems (Janssen et al., 2002). The recently released CESM2 includes an implementation of the WaveWatch-III model as a new component of the climate modeling system (Li et al., 2016), the newest GFDL boundary layer mixing schemes include Langmuir mixing using a statistical representation of waves (Li et al., 2017) and versions of the ACCESS model (Bi et al., 2013), GFDL model (Fan and Griffies, 2014), and FIO model (Qiao et al., 2013) are coupled to surface wave models.

6. OCEAN PARAMETERIZATIONS

6.1. Mesoscale

The class of ocean parameterizations most characteristic of the challenges particular to the ocean are those describing the oceanic mesoscale and their eddying features. Transient ocean mesoscale eddies are ubiquitous, contain more energy than the time mean ocean currents, and strongly affect critical large-scale behavior such as ocean heat uptake, meridional transport, and carbon storage. Their effects are sufficiently profound that modest changes to mesoscale parameterizations (or mesoscale-resolving models) frequently have a larger effect on ocean climate sensitivity than the total effect of other classes of parameterizations. In some problems, such as the response of the Southern Meridional Overturning Circulation to changing winds, the treatment of mesoscale eddy effects qualitatively affects the answer (Ito and Marshall, 2008; Bishop et al., 2016), although well-crafted parameterizations can capture this effect correctly (Gent, 2016; Mak et al., 2017; Marshall et al., 2017).

There are two main effects of mesoscale eddies that are parameterized: eddy-induced advection and eddy-induced diffusion. The primary example of the former is the Gent and McWilliams (1990) scheme, while the latter is normally treated using an anisotropic diffusion tensor aligned according t0 neutral directions (Redi, 1982). It is frequently convenient in both parameterization (Griffies, 1998) and diagnosis (Bachman et al., 2015) to treat both effects simultaneously as the antisymmetric and symmetric parts of one coefficient tensor in a flux-gradient relation. This basic formulation is almost universally used, but there are major differences in approach in the specification of this coefficient tensor and how it varies in space, with resolution, and as the flow varies. There are reasons to believe it should depend on eddy energy (Eden and Greatbatch, 2008; Marshall and Adcroft, 2010; Marshall et al., 2017), as well as orient itself appropriately along the stratification (Nycander, 2011; McDougall et al., 2014). Likewise, kinematic theory and modeling suggests that the eddy-induced advection and diffusion should be closely related (Dukowicz and Smith, 1997; Bachman and Fox-Kemper, 2013), but not in a trivial manner when there is structure in the vertical stratification (Smith and Marshall, 2009; Abernathey et al., 2013). Additionally, care is required as boundaries are approached and the stratification approaches mixed layer values (Danabasoglu et al., 2008; Ferrari et al., 2008, 2010).

One way to avoid strong dependence on the parameterization of the mesoscale is to resolve the largest mesoscale eddies, which reach near 25 km in the middle latitude open ocean. However, it is rare that mesoscale eddies are well-resolved over all of the world, given the variations in the Coriolis parameter and stratification (Hallberg, 2013). Furthermore, higher vertical modes require finer horizontal resolution, which eventually blends into the submesoscale regime (Boccaletti et al., 2007). Thus, subgrid schemes that expect only part of the mesoscale eddies to be resolved are required as model grid spacing reaches 25 km and finer. One approach is to adopt turbulent cascade theory as Smagorinsky (1963) did for energy, only here apply it to the 2D enstrophy (Leith, 1996; Fox-Kemper and Menemenlis, 2008) or quasieostrophic potential enstrophy (Bachman et al., 2017a). This approach can considerably improve the effective resolution (Soufflet et al., 2016) of eddy-permitting models (Pearson et al., 2017), and is fully flow-aware (i.e., responsive to changing ocean conditions) and scale-aware (consistent across multiple resolutions). Another approach is to introduce new terms in the advection equation that are dispersive or non-Newtonian (Bachman et al., 2018). Under such approaches, the coefficients are highly amenable to flow- and scale-aware treatment (Porta Mana and Zanna, 2014).

Stochastic parameterizations of eddies are also increasing in popularity and sophistication (Grooms and Majda, 2013; Porta Mana and Zanna, 2014; Grooms et al., 2015; Grooms,
2016), although no major climate models use this approach for ocean eddy parameterization yet.

Jansen and Held (2014) introduced a backscatter approach in which a sub-grid scale energy budget is used to inject energy back into a model using a negative, Laplacian eddy viscosity. The approach counters the tendency in some models to dissipate too much energy on the grid scale and increases the effective resolution in both channel configurations (Jansen and Held, 2014; Jansen et al., 2015) and in a shallow water model for a closed basin under double gyre wind forcing (Kloewer et al., 2018). Other approaches arrive at similar results by scaling a positive Laplacian dissipation based on the forward potential enstrophy cascade (Pearson et al., 2017), but this approach requires resolving the whole eddy instability scale (Bachman et al., 2017b). Further work is needed to extend and unify these methods for realistic ocean models at presently affordable resolutions.

6.2. Boundary Layer Mixing

The boundary layers at the surface and bottom of the ocean cannot be ignored, as these layers are where the influence of winds, freshwater, ice, heating and cooling, and relative motion against the solid earth are prescribed. The depth of these layers is directly related to global climate sensitivity (Hasselmann, 1976) and affects weather predictions (Emanuel, 1999). However, because of the proximal forcing, these regions also have distinct stratification—they tend to be largely co-located with the vertically well-mixed layers2—and flow—they tend to be characterized by nearly isotropic three-dimensional turbulence rather than stratified or quasi-2D turbulence of the ocean interior. If the turbulence is isotropic and contained within the 10 to few 100 m mixed layers, then the horizontal scale of this turbulence is also only tens to hundreds of meters, and thus this turbulence will not be resolved in global ocean models until well into the twenty-second century. The simplest parameterizations of the boundary layer are just fixed depth regions of increased vertical mixing. However, even early parameterizations included the ability to deepen the boundary layer through entrainment and restratify through surface fluxes (Kraus and Turner, 1967). Major improvements in boundary layer and mixed layer parameterizations have been driven by improved observational compilations of variability of these layers (e.g., de Boyer Montégut et al., 2004; Sutherland et al., 2014).

Boundary layer schemes take the surface fluxes of wind stress or bottom stress, surface cooling or heating, penetrating solar radiation, and then convert these forcing parameters into a profile of turbulent effects. Many schemes result in profiles of scalar mixing diffusivities and viscosity (e.g., Large et al., 1994) which then can be used in the calling model; other second-moment or higher-moment schemes provide fluxes and covariances directly (Mellor and Yamada, 1982). The new boundary layer scheme for climate models of Reichl and Hallberg (2018) combines the energetic constraints of Kraus and Turner (1967) with the possibility for mixed layer gradients and computational efficiency. A software innovation in ocean modeling is the development of software modules containing many different parameterizations in a portable format so that they can be rapidly substituted within a model, or shared across ocean models that all include the module. The Generalized Ocean Turbulence Model [GOTM: Burchard and Bolding (2001)] led the way to modular community boundary layer turbulence systems, capturing both simple schemes and full second-moment schemes. Now in version 5, GOTM is widely used in coastal and process modeling applications. The Community Ocean Vertical Mixing package (CVMix: Griffies et al., 2015a; Van Roekel et al., 2018) is similar, but is specialized to climate modeling applications and emphasizes parameterizations that produce profiles of diffusivity and viscosity. CVMix is implemented in MOM6, POP, and MPAS.

A significant change in boundary layer schemes over the last decade is the awareness of wave-driven mixing, beyond mixing driven just through wave breaking. The theoretical underpinnings of this idea extend back to Langmuir (1938), but modern theory (Craig and Leibovich, 1976; Teixeira and Belcher, 2002), modeling (Denbo and Skyllingstad, 1996; McWilliams et al., 1997), and observations (D’Asaro and Dairiki, 1997; Belcher et al., 2012; D’Asaro et al., 2014) have led to the development of a variety of recent parameterizations of wave-driven turbulence (Harcourt and D’Asaro, 2008; Van Roekel et al., 2012; Harcourt, 2015; Reichl et al., 2016). Their successful implementation in ocean circulation models is a relatively recent occurrence (Fan and Griffies, 2014; Li et al., 2016) because these mixing schemes depend on the wave state, so an ocean surface wave model is needed as part of the coupled system, although in many cases a statistical approximation may be sufficient (Li et al., 2017). Such an approximation, available through CVMix, is surely warranted to assess whether the development cost of coupling the wave model is justified. With care, mixed layer depth and ventilation biases can be significantly improved (Li et al., 2016; Li and Fox-Kemper, 2017). Related theories following different wave-driven turbulence parameterization assumptions have also shown bias reduction in ocean models (Babanin and Haus, 2009; Huang and Qiao, 2010).

6.3. Internal Mixing

Mixing across isopycnal surfaces, although small in magnitude relative to mixing in the surface layer, is an essential component of the global ocean circulation. While earlier models represented diapycnal mixing of tracers as a constant or a depth-dependent function (Bryan and Lewis, 1979), current climate and regional models increasingly attempt to account for the physical processes responsible, allowing mixing to change in response to changing ocean conditions (MacKinnon et al., 2017).

Mixing across isopycnal surfaces in the stratified ocean interior is driven primarily by instabilities of sheared flow. For resolved sheared flows, parameterizations based on the resolved Richardson number (Ri) can be used to estimate diapycnal diffusivity. These parameterizations range from diagnostic equations for Ri-dependent diffusivities (Pacanowski and Philander, 1981; Large et al., 1994) to Ri-dependent structure functions in two-equation turbulence closure schemes used in

2The technical distinction is that boundary layers are in the process of mixing and so are turbulent whereas mixed layers have already been mixed so are weakly stratified in the vertical.
many regional models (Mellor and Yamada, 1982; Umlauf and Burchard, 2003; Jackson et al., 2008).

Much of the shear responsible for mixing is in the internal wave field, on spatial and temporal scales too small for a climate model to resolve, so that mixing parameterizations based on resolved shear are insufficient. Instead, the diapycnal mixing can be formulated in terms of the global internal wave energy budget (Ferrari and Wunsch, 2009; Eden and Olbers, 2014), with sources of energy from the tides (Garrett and Kunze, 2007), surface winds, and flow over topography (Nikurashin and Ferrari, 2011; Scott et al., 2011) including the important feedbacks of wave drag on the flow (Trossman et al., 2013, 2016), and loss of wave energy to dissipation and mixing by wave-breaking at topography or through nonlinear wave-wave interactions in the ocean interior. The energy sources, particularly the tidal energy source, are relatively well understood (Bell, 1975; Garrett and Kunze, 2007; Falahat et al., 2014) compared to the sinks. A model of internal tide energy dissipation (St. Laurent et al., 2002) used in several climate models (Danabasoglu et al., 2012; Danne et al., 2012) represents the local dissipation of internal tides near the wave generation site as a fixed fraction of the energy source, with a vertical distribution which decays exponentially with height above bottom. A similar formulation has been applied to wind-generated near-inertial wave energy, with dissipation decaying with depth below the mixed layer (Iochum et al., 2013), and to lee-waves generated by subinertial flows (Melet et al., 2014). However, the locally dissipated energy fraction is in reality not a constant, but rather depends on the proportion of energy distributed at different wavelengths (St. Laurent and Nash, 2004; Klymak et al., 2010), and hence the small-scale details of the local topography (Lefauve et al., 2015), as well as the effectiveness of latitude-dependent wave-wave interactions such as parametric subharmonic instability (Nikurashin and Legg, 2011; Richet et al., 2018).

The vertical distribution of dissipation due to breaking internal waves also depends on the details of nonlinear interactions, and several new models are attempting a more physical representation of these processes (Polzin, 2009; Melet et al., 2013; Olbers and Eden, 2013; Lefauve et al., 2015). The most difficult component of the internal wave energy budget to represent is the propagation of low-mode waves far from the generation site (Zhao et al., 2016), and the loss of energy from these waves as they encounter different regions of the ocean such as continental slopes (Nash et al., 2007; Martini et al., 2011; Legg, 2014). Several models of low mode wave propagation (Eden and Olbers, 2014; de Lavergne et al., 2016) are being developed and implemented in ocean models, with considerable uncertainty remaining on the relative importance of the processes which determine the final distribution of mixing. The role of interactions between internal waves and geostrophic flow in leading to mixing (Rainville and Pinkel, 2006; Dunphy and Lamb, 2013) remains poorly accounted for, and will require communication between internal-wave driven mixing parameterizations and representations of subgrid-scale mesoscale eddies.

The loss of energy from the internal wave field needs to be translated into a diffusivity, usually accomplished using a mixing efficiency argument (Osborn, 1980). Whereas, current climate models usually use a constant mixing efficiency, with some modifications for low stratification regions (Melet et al., 2013), a growing body of evidence suggests mixing efficiencies vary with process, and may depend on buoyancy Reynolds number (de Lavergne et al., 2016).

The final spatial distribution of interior diapycnal diffusivity ultimately depends on topography, tidal flow, latitude, changing stratification, and variable energy inputs from winds and geostrophic flow. Test simulations with different distributions of diapycnal diffusivity corresponding to the same net energy input indicate sensitivity of overturning circulation, thermocline stratification, and steric sea-level, to the vertical distribution of mixing in particular (Melet et al., 2016).

6.4. Submesoscale

Submesoscale processes are characterized by order one Rossby numbers, and thus sit at the boundary between the largely balanced mesoscale and the unbalanced gravity wave scales (McWilliams, 2016). The first submesoscale process to be parameterized for ocean circulation models was the restratification of upper ocean boundary layers by mixed layer eddies (Fox-Kemper et al., 2008; Fox-Kemper et al., 2011). This parameterization shoals the mixed layer, mostly in wintertime, and is in fairly wide use in ocean models. Observational analyses support the essential formulation of this parameterization (Johnson et al., 2016; du Plessis et al., 2017) and the associated biological impacts (Mahadevan et al., 2012; Omand et al., 2015). Other approaches have been proposed (e.g., Brüggemann and Eden, 2014). However, recent work has shown that when there is competition between boundary layer mixing or convection and submesoscales, the situation is more complex than presently parameterized (Haney, 2015; Smith et al., 2016; Whitt and Taylor, 2017; Callies and Ferrari, 2018). Furthermore, mesoscale fronts can strengthen or weaken the submesoscale restratification effects (Ramachandran et al., 2014; Stamper et al., 2018). Seasonality of these effects has been observed to be significant in some models and observations (Mensa et al., 2013; Callies et al., 2015; Rocha et al., 2016b), but not in the same way in other locations (Luo et al., 2016; MacKinnon et al., 2016; Viglione et al., 2018). It is not clear if these more complex scenarios can be robustly parameterized, but it remains useful to collect observations of a wide variety of submesoscale events to assess the basic and sophisticated versions seen in models.

A variety of other submesoscale processes are under study and awaiting parameterization development or vetting. A parameterization of symmetric mixed layer instabilities has been proposed for submesoscale-permitting models, but not carefully vetted in applications (Bachman et al., 2017b). Submesoscale effects in the oceanic bottom boundary layer are likely to have important consequences for drag and deep stratification (Wennerat et al., 2018). Interactions between submesoscale and boundary layer turbulence are being simulated often in process studies, but they are not yet well understood (Hamlington et al.,
6.5. Vertical Convection
As discussed above, the majority of ocean circulation models have assumed hydrostasy and shallow grid cells so that nonhydrostatic phenomena are not explicitly resolved. Yet many oceanic phenomena have significant vertical velocities and convection, so these effects must be parameterized (Send and Marshall, 1995; Klinger et al., 1996). Surface boundary layer parameterizations incorporate convective mixing due to surface buoyancy loss (Mellor and Yamada, 1982; Large et al., 1994; Reichl and Hallberg, 2018). Prototype “superparameterizations” of these effects, involving 2-dimensional explicit simulations of convection within each horizontal grid-cell, are being studied (Campin et al., 2011). However, in the ocean interior, vertical convection driven by static instability is usually accounted for relatively crudely, through convective adjustment or via the Richardson-number dependent mixing scheme e.g., Jackson et al. (2008). A more physically-based parameterization of mixing due to convective overturns in the ocean interior, e.g., in breaking internal waves, Klymak and Legg (2010), has been implemented in regional and process simulations, but not in large-scale models. Convection due to subsurface input of fresh-water, e.g., from melting glaciers and ice-shelves, can be represented by buoyant plume models (Jenkins, 2011) which have been implemented into larger-scale GCMs (Slater et al., 2015).

6.6. Overflows
Dense water flowing from marginal seas and continental shelves into the deep ocean is an important source of abyssal and intermediate water. These overflows include many processes which occur below the grid scale, e.g., frictional boundary layer processes, entrainment of ambient water, hydraulic control at narrow straits and sills, which therefore must be parameterized (Legg et al., 2009). For coarse-resolution pressure-coordinate models, excessive numerical diffusion as dense water descends the slope is the most challenging issue (Winton et al., 1998); this can be reduced by numerical schemes such as Campin and Goosse (1999), Beckmann and Dööcher (1997), and Bates et al. (2012a), as well as the NCAR overflow parameterization (WuG, 2007; Danabasoglu et al., 2010) which incorporates hydraulic control theory to prescribe the overflow transport and entrainment. Isopycnal coordinates suppress numerical diffusion (White et al., 2009; Ilicak et al., 2012), and when combined with suitable physical parameterizations of the mixing in the bottom boundary layer (Legg et al., 2006) and interfacial shear layer (Jackson et al., 2008), lead to improved Nordic overflows in climate model simulations (Wang et al., 2015). The representation of small-scale topographic control of overflows remains an outstanding issue, which may be solved by the use of partial-barriers and thin walls (Adcroft, 2013) for narrow channels, or enhancements to entrainment (Ilicak et al., 2011) for small-scale roughness.

6.7. Estuaries and Runoff
Freshwater input into ocean models has traditionally been very idealized, involving hand tuning of the region and depth range over which the runoff is introduced to produce a realistic plume. Sun et al. (2017) have implemented a estuary parameterization based on theory (Garvine and Whitney, 2006) that makes these treatments more realistic and less subjective. In the context of biogeochemical modeling, theories, observations and parameterizations for sources and sinks of biogeochemical tracers in estuaries and elsewhere are an important future direction.

7. MODEL DIAGNOSIS AND EVALUATION
7.1. Global Diagnostics
7.1.1. SST and SSS
SST contributes to the local air-sea heat flux and can regulate ocean-atmosphere coupling via SST-wind feedback. SSS does not directly affects the local air-sea heat flux, but it can influence SST and air-sea interactions indirectly where there is an SSS-induced barrier layer. A proper distribution of SST and SSS is needed, not only for air-sea interaction processes, but also to ensure a proper representation of water masses throughout the world ocean. SST is arguably the best known quantity because of high quality AVHRR satellite measurements since the 1980s and accurate in situ measurements for more than a century. Our knowledge of SSS is not as extensive as SST, mostly because satellite measurements of SSS are recent (SMOS in 2010, Aquarius in 2011, and SMAP in 2015) with coarse coverage (Lee and Gentemann, 2018).

In global ocean-ice models uncoupled from the atmosphere, the air-sea turbulent heat fluxes are derived from bulk formulae using a prescribed atmosphere which tend to damp SST differences from the air temperature and acts as a fluid with infinite heat capacity (Griffies et al., 2009). It also does not allow for any ocean feedback to the atmosphere. This feedback takes place via SST (Small et al., 2008) and ocean current/wind shear (Renault et al., 2016a). Interactions of ocean eddies and the atmosphere can regulate western boundary currents (Ma et al., 2016; Renault et al., 2016b). Furthermore, a bulk-forced oceanic uncoupled simulation should prescribe the surface stress using the relative wind using bulk formulae which take into account a parameterization of the partial re-energization of the ocean by the atmospheric response (Renault et al., 2016b).

Errors in SSS arise from inaccurate precipitation fields as well as from the modeled surface advection of the salinity fields. Contrary to the SST, there is no implied salinity restoring in the fresh water flux formulation to minimize the accumulation of flux errors and this will induce a drift in SSS in ocean-ice only simulations. Salinity restoring can minimize this drift, but it does not have any physical meaning (Griffies et al., 2009). The time scale associated with this restoring will also strongly impact the realism of the variability in surface salinity and associated water mass transformations. Furthermore, the SSS field to relax toward is subject to large uncertainties because SSS is still poorly observed when compared with SST, although improving through remote sensing of salinity (Boutin et al., 2018).
7.1.2. Ocean Heat Content

Box 3.1 in the IPCC AR5 report (Rhein et al., 2013) contains an iconic image indicating the predominant role of the ocean in determining the heat balance in the climate system. Largely supported by the advances in ocean temperature measurements made via Argo (Riser et al., 2016), we know that the ocean has absorbed more than 90% of the excess heat trapped in the climate system arising from anthropogenically produced greenhouse gases since 1970. If the ocean did not absorb this excess heat, then uniformly mixing that heat throughout the whole global atmosphere would warm it by roughly 40 K. This dominant role for the ocean in the earth’s heat budget provides a mandate for the oceanography community to accurately represent and parameterize the multitude of processes that affect the transport of heat both laterally and vertically.

At the basin and planetary scales, poleward ocean heat transport is largely constrained by poleward atmospheric transport, which is itself constrained by differential heating between the equator and poles. In contrast, the patterns and rates of vertical heat transport between the surface ocean layer and the ocean interior are largely determined by ocean processes, including small scale mixing induced by breaking gravity waves (MacKinnon et al., 2013, 2017), as well as mesoscale (Fox-Kemper et al., 2013) and submesoscale (McWilliams, 2016) currents, eddies, and filaments. These processes also constrain the meridional transport of salt in the ocean, with mesoscale eddies contributing as much as the time-mean circulation to the transport of salt out of the subtropical gyres (Treguier et al., 2014).

The connection between the planetary heat budget and ocean mixing and stirring has been known for decades, thus maintaining a rationale for conducting ocean research at climate centers. Advances made during the past decade in observations, processes modeling, and global climate modeling have enhanced the constraints on relevant ocean processes and their parameterizations. Even so, there remains significant work required to confidently constrain the rates and pathways for ocean heat uptake. One revealing example concerns the sensitivity of heat uptake to the representation and parameterization of ocean mesoscale eddies as found in a realistic climate model (Griffies et al., 2015b). It will be a decade or more before the O(4–10) km ocean grid spacing used in that study will become routinely available for climate centers, and even longer before the smaller submesoscale processes are routinely resolved.

7.1.3. Tracers

Much attention is paid to active tracers by ocean modelers, such as potential density and potential vorticity, because the distribution of these fields directly affect the circulation. However, many studies have shown that biogeochemical modeling is significantly hindered by the lack of attention paid to passive tracer dynamics (Doney et al., 2004), and anthropogenic ocean heat uptake tends to have marginal impacts on circulation and stratification, but is one of the three most important aspects of sea level rise (Palmer et al., 2018). Recent studies have shown that there is a distinction between Gent and McWilliams (1990) streamfunction and Redi (1982) diffusion coefficients (Smith and Marshall, 2009; Abernathey et al., 2013), but they are linked together in a nontrivial manner rather than totally unrelated (Dukowicz and Smith, 1997; Bachman and Fox-Kemper, 2013).

One method to allow for passive tracers to be mixed at a different rate from active tracers is horizontally anisotropic diffusion (Fox-Kemper et al., 2013) and a matched Gent and McWilliams (1990) operator (Smith and Gent, 2004). It has already been discussed that oceanic diffusion is anisotropic between the along-isopycnal and cross-isopycnal directions, but it is also anisotropic in the horizontal direction largely due to shear dispersion (Oh et al., 2000). If different tracers are assumed to have approximately the same anisotropic diffusivities and gauge, linear algebra methods have been developed to diagnose streamfunctions and diffusivities from multiple tracers at one time (Plumb and Mahlman, 1987; Bratseth, 1998; Bachman and Fox-Kemper, 2013; Bachman et al., 2015). These methods broadly support the Gent and McWilliams (1990) streamfunction and Redi (1982) diffusion operator forms, but they also allow for spatiotemporal examination of these fields and their relationships.

Other methods to allow for passive tracers to obey different behavior than active tracers include separating the Gent and McWilliams (1990) streamfunction and Redi (1982) diffusion or using different diffusivities or rotational gauges for different tracers (Eden, 2010). These methods suffer from underdetermination, however, as the anisotropic flux-gradient relationship for one tracer has more degrees of freedom than constraints, a situation that is made even more severe by freedom of gauge choice.

Observational diagnoses of diffusivities are becoming more common (Marshall et al., 2006; Abernathey and Marshall, 2013; Poje et al., 2014; Cole et al., 2015; Groeskamp et al., 2017), but different diagnostic techniques may result in different diffusivities (Klocker et al., 2012; Pearson et al., 2018). Synthetic “observations” created in a model simulation are recommended for evaluation on an equal basis to the observations. Furthermore, anisotropic diffusion is typically neglected in these diagnostics (Fox-Kemper et al., 2013).

7.1.4. Sea Level

Sea level rise is among the most prominent of societal impacts from anthropogenic climate warming, affecting millions of people who live near coasts and many more affected by migrations away from coasts. Ocean warming has contributed to roughly one-third to one-half of the observed global mean sea level rise during recent decades, with added mass from melting land ice contributing the remainder. Ocean and climate models are the primary tool used to quantify how ocean heating contributes to sea level changes at both global (e.g., Gregory et al., 2013) and regional (e.g., Griffies et al., 2014) scales.

The ocean thermal expansion coefficient is a strong function of temperature, reducing an order of magnitude between the warm tropics and the cold high latitudes (e.g., see Figure 1 of Griffies et al., 2014). Consequently, a unit of heat added to the low latitudes impacts on sea level rise more than the equivalent heat added to the high latitude oceans. Furthermore, ocean circulation moves heat throughout the ocean, thus contributing to regional...
patterns of thermosteric sea level change. Consequently, the accurate projection of global and regional sea level changes involves more than just how much heat crosses the ocean surface. Salinity changes also contribute to regional patterns of halosteric sea level. Indeed, in the Atlantic a negative halosteric sea level change, due to regional increases in salinity, acts to compensate regional thermosteric rise due to warming (e.g., see Yin et al., 2010). Even so, global halosteric sea level changes are negligible since the mass impact on sea level from changes to freshwater content dominate the corresponding changes to halosteric sea level (Lowe and Gregory, 2006).

Ocean and climate models provide a tool to quantify how physical processes impact both regional and global sea level patterns. Griffies and Greatbatch (2012) formulated methods to decompose the global mean sea level budget into physical processes, whereas Gregory et al. (2016) detail an experimental protocol to quantify how changes in heat fluxes, water fluxes, and wind stresses impact on regional sea level patterns (the CMIP6 Flux Anomaly Forced MIP, FAFMIP). These approaches provide frameworks for comparing and contrasting the variety of model projections of sea level, and thus for use in identifying, at a process level, sources for model uncertainties that commonly plague simulations of sea level change particularly at the regional scale.

7.1.5. Ocean Mass and Angular Momentum

The three primary components of global sea level rise are ocean thermosteric expansion, glacier melt and changes in land storage of water, and melting ice sheets. Only the first does not imply a change in the mass of the oceans. Satellites such as the Gravity Recovery and Climate Experiment (GRACE) can directly measure changes in the mass of the oceans through their effects on the earth's gravitational field. Traditional Boussinesq ocean models with a rigid lid struggled to incorporate such changes in mass through freshwater input, because they were models with a fixed volume of fluid so any changes in mass had to come through virtual salt fluxes or temperature-related density changes.

However, modern models using an explicit free surface with a natural water boundary condition overcome the limitations of the rigid lid (e.g., Griffies et al., 2001; Campin et al., 2004), thus accepting changes in ocean volume even while preserving Boussinesq dynamics. This numerical improvement, or post-simulation analysis with similar intent (Griffies and Greatbatch, 2012), has greatly improved the directness of sea level change estimation within the model framework. When coupled with active ice sheet models, ocean models are beginning to directly simulate the needed components of sea level change.

In situ observations of ocean bottom pressure, i.e., the weight per unit area of the ocean above (Hughes et al., 2018), GRACE measurements (Watkins et al., 2015; Save et al., 2016), and data assimilating models spanning annual to decadal time scales (Ponte et al., 2007; Stammer et al., 2010; Köhl et al., 2012; Johnson and Chambers, 2013) are important aspects of continuing to improve ocean models’ capability to simulate all important aspects of climate. Tides (Arbic et al., 2018) and forcing of other high-frequency waves (Callies and Ferrari, 2013) are an important new aspect of high-resolution (mesoscale and submesoscale-permitting) simulations, especially in data assimilating and Lagrangian transport contexts where it has proven very difficult to find effective ways to completely filter these motions from observations (e.g., Beron-Vera and LaCasce, 2016; Buijsman et al., 2017). Gravitational self attraction and loading (i.e., the elastic response of the lithosphere to changes in ocean mass) is now well-known to be important in tide simulations, but their impact on ocean circulation and ocean circulation sensitivity is only recently being considered (Kopp et al., 2010; Slanger et al., 2014; Vinogradova et al., 2015; Arbic et al., 2018). It is through comparison of ocean circulation models, bottom pressure records, and gravity recovery missions that these physical relationships will become understood and effectively simulated.

Finally, ocean angular momentum diagnostics (along with those of the atmosphere) provide valuable information on Earth rotation variability and its relation to oceanic mass transport (Ponte et al., 2001). Calculating long-term trends in ocean angular momentum and implied changes in length-of-day estimates benefits substantially from models that have been rigorously constrained by observations through state and parameter estimation (Quinn et al., 2018).

7.1.6. Meridional Transport

Meridional transport of freshwater and heat is a traditional diagnostic of models, both in mean overturning and eddying components. The advent of continuous monitoring of the RAPID/MOCHA, OSNAP, and SAMOC mooring arrays provides a new tool for model evaluation. The first, RAPID/MOCHA array resulted in significant improvement of models to come into agreement (Msadek et al., 2013). Griffies et al. (2016) present in their table J2 the most important mass transport sections that should be evaluated in models, with recent observation estimates (e.g., Drake Passage transport recently revised to a 30% higher value by Donohue et al., 2016).

7.2. Process Diagnostics

7.2.1. Eulerian

While Eulerian observations and diagnostics are as old as oceanography itself, recent work combining mooring time series for a frequency spectrum evaluation of high resolution ocean and tide models is notable (Arbic et al., 2018).

7.2.2. Lagrangian

Lagrangian particle tracking methods can be used to map ocean circulation pathways determined by the ocean velocity field (e.g., Tamsitt et al., 2017; Lique and Thomas, 2018; van Sebille et al., 2018). Additionally, such methods can be used to objectively quantify how much transport arises from coherent ocean vortices (e.g., see Haller, 2015; Abernathey and Haller, 2018; Tarshish et al., 2018 and references therein). So long as sufficient velocity data is available to compute the particle trajectories, and sampling biases are managed (Pearson et al., 2017), Lagrangian analysis methods offer a powerful tool to diagnostically probe properties of the circulation, including its turbulent properties. These methods, however, are limited in their ability to deduce pathways for tracer transport given that
Lagrangian particle pathways are generally computed without the direct impact from mixing, whereas mixing can be important for determining tracer pathways and timescales.

### 7.2.3. Turbulence

As more sophisticated models of turbulent mixing are developed, the veracity of these parameterizations can be evaluated by comparing the 3D spatial patterns of diffusivity or dissipation generated by these models and their temporal variations with observations (Whalen et al., 2012, 2018; Sutherland et al., 2013; Waterhouse et al., 2014; Kunze, 2017). Parameterizations of turbulent mixing can also be evaluated by examining their impact on the simulated ocean circulation and climate. A series of investigations using the GFDL-ESM2G coupled climate model explore the impact of different components of internal wave driven mixing, including the vertical structure of the local dissipation of the internal tide (Melet et al., 2013), the inclusion of breaking lee-waves due to geostrophic flow over topography (Melet et al., 2014), and different choices for the farfield component of internal tide breaking (Melet et al., 2016). These studies indicate that both the total energy used for mixing and its spatial distribution matters to the circulation. The vertical distribution of mixing is particularly important: more mixing in the deep ocean results in stronger, deeper overturning, while more mixing in the upper ocean leads to stronger shallow subtropical overturning cells and a more diffuse thermocline (Melet et al., 2013, 2016). The horizontal distribution of mixing is particularly important when the sources of deep water masses are differentially affected, for example by lee-wave driven mixing in the deep Southern Ocean (Melet et al., 2014), or by mixing on high latitude shelves and slopes (Melet et al., 2016). Ocean-only MITgcm simulations by Eden et al. (2014) have explored the impact of the energetically-consistent internal wave and mesoscale eddy-driven mixing, and similarly find that enhanced mixing at depth drives a stronger overturning circulation.

### 7.2.4. Submesoscale

The submesoscale is challenging to evaluate diagnostically in comparison to observations, as the features are small (km) and evolve rapidly (hours). However, recent projects have provided observations that are capable of evaluating and constraining submesoscale model behavior through models initialized alongside observations (Mahadevan et al., 2012; Omand et al., 2015), scaling relations (Johnson et al., 2016; du Plessis et al., 2017), and intensive observations (Shcherbina et al., 2013; Poje et al., 2014).

### 7.2.5. Water Masses

As recently reviewed by Groeskamp et al. (2019), the water mass transformation (WMT) framework weaves together circulation, thermodynamics, and biogeochemistry into a description of the ocean that complements Eulerian and Lagrangian methods. In so doing, a WMT analysis offers novel insights and predictive capabilities for studies of ocean physics and biogeochemistry. A WMT analysis is layer-based, thus aiming to remain faithful to the layered structure of the ocean thermodynamic and tracer properties. Since the pioneering work of Walin (1982), there have been numerous studies using both observational and model datasets, aiming to understand and quantify water mass properties and their transformations. Many studies provide inferential estimates of transformations based on surface properties and boundary fluxes (e.g., Large and Nurser, 2001). The incorporation of biogeochemical tracers into the WMT framework remains a cutting-edge use of the analysis method (Iudicone et al., 2011).

Models offer the ability to move beyond the traditional inferential methods commonly employed with WMT analysis. Doing so requires the development of online binned budget diagnostics so that tendencies from all processes within the ocean can be explicitly determined. As a result, the analysis will allow for an accurate deductive rendition of the transformation processes ongoing within the model, and how these processes affect simulated water mass properties.

### 7.2.6. Energetics

A powerful diagnostic tool used since Lorenz (1955) is the evaluation of energy reservoirs and transfers. This approach can be used to evaluate mesoscale eddy parameterizations, as they frequently rely on extraction rate of potential energy (Gent and McWilliams, 1990) or the reservoirs of kinetic energy (Marshall et al., 2017). Global estimates of the energy available for diapycnal mixing from the tide and wind (Ferrari and Wunsch, 2009) are a valuable diagnostic for evaluation of mixing parameterizations (MacKinnon et al., 2017). In some circumstances, the class of turbulent instabilities can be recognized using energetics when other aspects of the structures are not recognizable (Haney et al., 2015).

### 7.2.7. Vorticity

Vorticity budgets are another powerful tool in diagnosing the mechanisms underlying the control of the oceanic general circulation (Fox-Kemper and Pedlosky, 2004; Yeager, 2015). Unlike the energy budget, which can be complicated by non-local radiation effects (Plumb, 1983), the vorticity budget is closed locally (although it can be difficult to assess the cause of vertical stretching terms). Enstrophy or potential enstrophy are useful diagnostics of 2D and quasigeostrophic turbulence (Fox-Kemper and Menemenlis, 2008; Bachman et al., 2017a; Pearson et al., 2017; Pearson and Fox-Kemper, 2018).

Vorticity diagnostics applied to global models of full complexity have revealed the importance of bottom vortex stretching in the balance of western boundary currents (Hughes and de Cuenas, 2001), the Atlantic northern recirculation gyre (Zhong and Vallis, 2007), North Atlantic Deep Water pathways (Spence et al., 2012), and the Atlantic subpolar gyre and meridional overturning circulations (Yeager, 2015). The Argo array of float measurements to 2,000 db has permitted global tests of linear vorticity theory (Gray and Riser, 2014), but deeper measurements are needed to validate model results showing ubiquitous and large deep transports balanced by topographic stretching (Thomas et al., 2014). In eddy-resolving simulations, vorticity analysis highlights the role of mean nonlinear advection and eddy momentum flux in a variety of contexts (Delman et al., 2015; Wang et al., 2017; Wang H. et al., 2018).
8. NOVEL APPLICATIONS

8.1. Southern Ocean

The Southern Ocean is a major sink for anthropogenic energy and carbon, due to the competition of the Ekman and eddy overturning circulations of the SOMOC and the formation of Antarctic Bottom Water. This complex set of phenomena has significantly challenged model capabilities, leading to a variety of model resolution hierarchy studies (Hallberg and Gnanadesikan, 2006; Ito and Marshall, 2008; Munday et al., 2013; Bryan et al., 2014; Marshall et al., 2017). One analysis of a fully-coupled atmosphere-ice-ocean simulation at mesoscale-resolving resolution yielded eddy saturation effects in a natural manner (Bishop et al., 2016).

The DIMES (Sheen et al., 2013), SMILES (Adams et al., 2017), and SOCCOM (Russell et al., 2014) experiments all have matching modeling exercises to explore the impacts of mixing, submesoscales, and processes affecting biogeochemistry of the Southern Ocean. Matched or idealized simulations have led to better understanding of the context of the observations and dynamical insights (Bates et al., 2014; Bachman et al., 2017c; Bronselaer et al., 2017; Stamper et al., 2018).

8.2. Arctic

Progress in modeling the Arctic ocean has occurred through improved sea ice models as well as higher horizontal and vertical resolution, which allow a better representation of currents through the Canadian Archipelago (Hughes et al., 2017) and a better representation of the complex stratification of Arctic watermasses (Wang Q. et al., 2018). New in situ and satellite observations are used to evaluate operational models (Dupont et al., 2015) and reveal processes that are being investigated with models, such as the seasonal cycle in the Atlantic waters (Lique and Steele, 2012). With the decline of summer sea ice and the increased Greenland runoff through fjords, the Arctic freshwater balance is an evolving challenge for models (Lique et al., 2015).

8.3. Atlantic Meridional Overturning Circulation

Through its associated heat and salt transports, the Atlantic meridional overturning circulation (AMOC) significantly influences not only the climate of the North Atlantic and surrounding areas, but also the climate of the entire globe via oceanic and atmospheric teleconnections. In particular, changes in sea surface temperatures (SSTs) linked to AMOC variability can impact climate on interannual to (multi)decadal to even longer time scales, thus making AMOC a central piece of prediction efforts of the earth’s future climate on these time scales. Essentially, AMOC is thought to contain the dynamical memory of the climate system.

The representation of AMOC mean and variability continues to exhibit significant differences among fully-coupled, pre-industrial control simulations of the earth system models, e.g., those participating in the Coupled Model Intercomparison Project phase 5 (CMIP5). In addition to important differences in peak periods and amplitudes of intrinsic variability among them, several aspects of proposed natural variability mechanisms also differ. The roles and impacts of external forcings vs. internal variability in driving variability in AMOC and related, climatically important fields, such as SSTs, during the historical period also remain largely unresolved.

The representation of AMOC mean similarly differs among ocean sea-ice coupled hindcast simulations that were performed as part of the CORE-II ocean model intercomparison effort, despite using the same atmospheric forcing datasets (Danabasoglu et al., 2014). Such differences do not necessarily suggest an obvious grouping of the ocean models based on their lineage, their vertical coordinate representations, or surface salinity restoring strengths. The differences can be partially attributed to use of different subgrid scale parameterizations and parameter choices, to differences in vertical and horizontal grid resolutions in their ocean models, as well as to use of diverse snow and sea-ice albedo treatments. These hindcast simulations, however, tend to show general agreement in their temporal representations of AMOC and SST variabilities, e.g., the observed variability of the North Atlantic SSTs is captured well by the models (Danabasoglu et al., 2016). This suggests that the simulated variability and trends are primarily dictated by the forcings which already include the influence of ocean dynamics from nature superimposed onto external and anthropogenic effects. Nevertheless, there are many important differences among the model solutions, including the spatial structures of variability patterns and where the maximum AMOC variability occurs.

The representation of AMOC mean and variability is also as diverse among various reanalysis products (Karspeck et al., 2016). Arguably, the reanalysis products appear to be less consistent with each other in their AMOC means and its interannual to decadal variability than those of forced hindcast simulations participating in the CORE-II effort.

To resolve these differences in models representations of AMOC mean and variability, including the driving mechanisms, requires basin-wide continuous and comprehensive observations. In particular, identification of robust variability mechanisms has important implications for climate (decadal) prediction efforts both to shed light on sources of skill and subsequently use that knowledge to improve skill. The importance of such trans-basin observational systems has been recognized and several efforts have been in place, starting in 2004. A summary of these observational programs along with recommendations for future observations is provided in Frajka-Williams et al. (2018). A recent review of the AMOC that includes its mean spatial structure, temporal variability, and proposed driving mechanisms is provided by Buckley and Marshall (2016).

8.4. Sea Level Change

As societal impacts respond to relative sea level change, not global mean sea level change, the oceanic and geophysical aspects of the regional sea level problem have been studied heavily in the last decade. One ocean modeling tool that has proven useful is the use of an ocean state estimate (Forget and Ponte, 2015; Stammer et al., 2016) or reanalysis to provide insight into the patterns of regional sea level rise (Vinogradov et al., 2008; Piecuch and Ponte,
Future work in this direction will aid in prediction of ocean dynamics effects on sea level.

### 8.5. Global Coupled Mesoscale-Permitting

Global ocean-only and ocean-ice forced high resolution simulations have been possible for a few decades (Semtner and Chervin, 1992; Maltrud and McLean, 2005; Chassignet et al., 2009). However, recently coupled simulations have brought climate modeling capabilities into the mesoscale-resolving or mesoscale-permitting regime (McClean et al., 2011; Small et al., 2014; Griffies et al., 2015b; Hewitt et al., 2016). While all aspects of resolving mesoscale features are not yet discovered, it is already clear that air-sea fluxes are fundamentally affected (Bishop et al., 2017) and that dampening of oceanic variability through coupling with the atmosphere is one prominent effect (Ma et al., 2016; Renault et al., 2018). At present, most of these efforts use relatively simple subgrid dissipation schemes and parameterizations of unresolved smaller mesoscale and submesoscale eddies, but a few parameterizations have been proposed for this regime (Bachman et al., 2017b; Zanna et al., 2017) and one has been tested with promising results (Pearson et al., 2017). Interestingly, the lower horizontal dissipation in these simulations makes bottom drag more important energetically (Pearson et al., 2017), which makes returning to a careful consideration of the amount of bottom drag used more important (Sen et al., 2008; Arbic et al., 2009).

#### 8.6. Global Submesoscale-Permitting

Taking advantage of improvements in numerics, computing, and scale-aware parameterizations, a heroic calculation of an ocean-ice configuration MITgcm at nominal 2 km global resolution has been run for 2 years of simulated time including all realistic forcing, including tides. These simulations allow for large region and global assessment of large submesoscale phenomena and tide-current coupling (Rocha et al., 2016a,b; Su et al., 2018), building on earlier work on tides in HYCOM (Shriver et al., 2014).

#### 8.7. Ocean State Estimates and Reanalyses

Numerous advances have been made in ocean state estimation methods. For many quantities, ocean reanalyses have reached a degree of accuracy at which they could be considered in IPCC evaluations (Carton et al., 2018). However, significant model disagreement remains in some features such as the AMOC (Karspeck et al., 2016). Ensemble methods have also become more effective as resolution has increased, permitting fronts and eddies (Penny et al., 2015; Penny, 2017). The first modern-era coupled reanalysis of the global atmosphere, ocean, land surface, and cryosphere was created by NCEP with the Climate Forecast System Reanalysis spanning the period 1979 onward (CFSR; Saha et al., 2010). More recently, ECMWF generated a longer coupled reanalysis spanning the twentieth century (Laloyaux et al., 2018). Large biases have been identified in atmospheric-only surface forcing data sets (Tsujino et al., 2018) and it is expected that coupled reanalyses will become the standard to resolve these issues. An important distinction in methods is that between filter vs. smoother approaches. The former is the method of choice in the context of prediction, whereas the latter has been used for reconstruction (e.g., Wunsch and Heimbach, 2013; Stammer et al., 2016).

### 8.8. Decadal Prediction

The possibility of using coupled climate models for forecasting on decadal timescales was first explored in several perfect model and potential predictability studies (Griffies and Bryan, 1997; Boer, 2004; Pohlmann et al., 2004; Collins et al., 2006) and was soon thereafter tested using ensembles initialized from observation-based reconstructions (Smith et al., 2007; Keenlyside et al., 2008; Pohlmann et al., 2009; Yeager et al., 2012; Meehl et al., 2014; Scaife et al., 2014). The North Atlantic is a region consistently found to exhibit prediction skill on decadal timescales associated with ocean thermohaline circulation memory (Yeager and Robson, 2017), with predictable Atlantic sea surface temperature underpinning skill in seasonal climate over land regions such as the African Sahel (Yeager et al., 2018). With an accurate initialization of the coupled system, particularly the ocean state as determined from monitoring systems, then useful prediction of some aspects of regional climate 10 years in advance appears to be a realizable goal. Hypothetically, mesoscale-resolving ocean models will provide even better prediction system skill and predictability, through improved mesoscale air-sea interaction and reduced ocean bias (Siqueira and Kirtman, 2016). However, mesoscale eddy dynamics generate interannual large scale variability (Penduff et al., 2011) with low predictability, which has been addressed with a large ensemble approach (Leroux et al., 2018). To use mesoscale-resolving models in a forecast system, advanced data assimilation techniques, such as the ensemble Kalman filter, are needed to manage the large volumes of data involved.

### 9. WHAT TO EXPECT BY 2030

Ocean modeling continues to improve. Over the next decade, we expect:

- Resolution approximately a factor of 2 finer: 0.5° to 0.1° models will be common and global 1 km models in prototype evaluation. However, the rate of resolution refinement in ocean models has decreased by about 10% after 30 years of steady improvements. This may represent a reduction in resources, a failure in Moore's law, or a shift toward more complex models or large ensembles.
- Continuing improvement and familiarity using unstructured grids and ALE vertical coordinates.
- Nested and regional downscaling simulations connecting coastal impacts–inundation, ice shelves, sea level–and global changes.
- Increased use of ensembles of high-resolution ocean models and coupled models, which help to distinguish internal variability from forced and climate from weather.
- New parameterizations and improvements of existing parameterizations: Stochastic parameterizations are being explored for eddies (Grooms and Majda, 2013; Grooms, 2016), mixing (Juricke et al., 2017), air-sea fluxes (Williams,
2012), and parameterizations generally (Brankart, 2013; Andrejczuk et al., 2016); energetically-consistent schemes are under development (Eden et al., 2014; MacKinnon et al., 2017; Marshall et al., 2017); submesoscale (Fox-Kemper et al., 2011) and Langmuir (Li et al., 2017) turbulence parameterizations are becoming common; sensitivity to surface wave effects on mixing (Qiao et al., 2016) and air-sea fluxes (Zhao et al., 2017) are being explored; generation of gravity waves over bottom topography (Trossman et al., 2013, 2016); vertical convection is improving (Klymak and Legg, 2010; Campin et al., 2011); and coastal estuaries (Sun et al., 2017) have been parameterized over the last decade. Competition and complementary parameterizations will arise approximating other aspects of these unresolved phenomena.

- Improvements to air-sea and air-ice coupling, through numerical improvements, resolution, tighter coupling between ocean and ice dynamics, and a better representation of the impacts of surface waves.
- More direct simulation of sea level change, through oceanic free surface permitting numerics, actively simulated ice sheets and ice shelves, and improved and efficient incorporation of geophysical aspects of oceanic change such as self-attraction and loading.
- More direct simulation of tides, especially in high-resolution models, following recent progress in HYCOM and the MITgcm (Shriver et al., 2014; Rocha et al., 2016a,b; Ansong et al., 2017; Buijsman et al., 2017; Arbic et al., 2018; Su et al., 2018).
- More quantitative paleoclimate constraints on ocean circulation and change. Proxy modeling is one important step in this direction (Dee et al., 2015), as are simulated proxies (Stevenson et al., 2018). Until recently proxies for some ocean variables have been too uncertain or sparse to quantitatively constrain models (Stevenson et al., 2013; Amrhein et al., 2018; Bereiter et al., 2018), although sea level is an important exception (e.g., Miller et al., 2013).
- Increasing use of artificial intelligence, neural networks, and deep learning as part of the development and improvement of parameterizations—particularly in their numerical optimization for efficiency (Schneider et al., 2017; Gentine et al., 2018; Bolton and Zanna, 2019). These efforts may lead to reduced use of FORTRAN in favor of more modern computing languages such as Python and Julia (e.g., Häflner et al., 2018).

Sea ice modeling is expected to improve through (Notz and Bitz, 2016; Smith et al., 2018):

- Ice floe size prediction with ice–wave interactions.
- Anisotropy and lead orientation prediction.
- Realistic treatment of snow, melt ponds, evolution of salinity, and light-absorbing particles.
- Grid refinement in regions where small scale processes are needed to be resolved.
- Consistent models of sea-ice rheology and microstructure evolution based on first physical principles.

Key observational constraints for the next decade that would be useful in the above lists are:

- Analysis of high-resolution observations using scale-selective diagnostics of ocean tracer and momentum covariances, such as spectra, cospectra, structure functions, and relative dispersion. These analyses are key for constraining high-resolution models, which may share many of the same biases as coarse resolution models, but also are expected to have new biases related to improper small-scale phenomena. The development of subgrid schemes for high-resolution models requires observations that can distinguish good high-resolution models from bad.
- The follow-on mission to GRACE, GRACE-FO, will continue the improvements in modeling ocean and ice mass relocation, as well as other important aspects of the satellite constellation for ocean observation and monitoring. The proposed satellites capable of simultaneously constraining wind, wave, and currents such as SKIM, will be highly valuable in evaluating the formulation of high-resolution the coupled wave-ocean-sea ice simulations. SST, ocean color, and altimetry remain crucial in both data assimilation and model evaluation.
- Continued long-term and deep monitoring of the slow changes to the earth system, many of which are oceanographic. Our history of high quality observations is only slightly longer than a century (Roemmich et al., 2012), but it is well-known that ocean model statistics take 2–20 times longer to manifest (Wittenberg, 2009; Stevenson et al., 2010; Lindsay, 2017). Paleoclimate evidence provides reason to worry about the accuracy of low-frequency model variability (Laepple and Huybers, 2014).
- Increased observations in the polar regions, which challenge both ocean and cryospheric models. Under ice shelf and repeat sampling of evolving sea ice conditions are particularly essential to optimizing the coupled models discussed here.
- Continued progress on continuously-uploaded monitoring systems, such as GOOS, that can be used in ocean state estimates, decadal forecasts, and forcing products.
- Increased observations in the deep ocean, particularly near topography, will be vital for constraining slow changes to the ocean reservoirs, key turbulent mixing, and boundary-confined flows. Deep Argo floats may be one component of this measuring system; others include more routine and possibly automated microstructure measurements as well as new capabilities for measuring turbulent quantities in the bottom boundary layer.
- The use of ocean state and parameter estimation to provide a comprehensive framework for formal model calibration.
- The use of observing system simulation experiments and optimal observing network design approaches in the context of ocean state and parameter estimation to predict the result of instrument deployments to result in better targeting of key areas and instrument accuracy requirements.
- The use of ocean state estimates to provide a context for observations.
• The continued successful partnerships between global modelers, process modelers and observationalists to drive process understanding and innovation in observing and parameterizations.

• Simulations of ensembles of observations similar to those that were collected to estimate bias and systematic errors.

9.1. Did the OceanObs09 “What to Expect?” Correctly Predict?

Griffies et al. (2010) state that “The origins of these biases and model differences may be related to shortcomings in grid resolution; improper numerical algorithms; incorrect or missing subgrid scale parameterizations; improper representation of other climate components such as the atmosphere, cryosphere, and biogeochemistry; all of the above, or something else.” It is clear from this review that progress has indeed been made on all of these fronts in the last decade. They also proposed that deep insight would be learned from the study of AMOC variability and stability, patterns of sea level rise, and the Southern Ocean. Indeed, the simulations and analyses featured above focused on these aspects of the earth system have revealed important new insights, but also new challenges and prospects.

REFERENCES


AUTHOR CONTRIBUTIONS

BF-K with significant input from SL, SG, and SJM, created the outline of topics and wrote the proposal. The authors of Griffies et al. (2010) and the CLIVAR Ocean Model Development Panel members were asked to contribute or recommend other authors. All authors then contributed to the writing and editing process and recommended references and examples.

FUNDING

BF-K was supported by NSF 1350795, NSF 1655221, ONR N00014-17-1-2963, and N00014-17-1-2393. PU was supported by the EC Marie Curie Support Action LAWINE (grant no. 707262).

ACKNOWLEDGMENTS

The CLIVAR and USCLIVAR program offices are thanked for their continued support of the ocean modeling community. NCAR is sponsored by the US National Science Foundation. Laure Zanna is thanked for contributing an unofficial review including helpful suggestions.


Fox-Kemper et al.  Ocean Circulation Models


