Modeling Diverse Communities of Marine Microbes

Michael J. Follows and Stephanie Dutkiewicz

Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; email: mick@ocean.mit.edu, stephd@mit.edu

Keywords
microbes, phytoplankton, community structure, traits, trade-offs, self-selection, adaptive

Abstract
Biogeochemical cycles in the ocean are mediated by complex and diverse microbial communities. Over the past decade, marine ecosystem and biogeochemistry models have begun to address some of this diversity by resolving several groups of (mostly autotrophic) plankton, differentiated by biogeochemical function. Here, we review recent model approaches that are rooted in the notion that an even richer diversity is fundamental to the organization of marine microbial communities. These models begin to resolve, and address the significance of, diversity within functional groups. Seeded with diverse populations spanning prescribed regions of trait space, these simulations self-select community structure according to relative fitness in the virtual environment. Such models are suited to considering ecological questions, such as the regulation of patterns of biodiversity, and to simulating the response to changing environments. A key issue for all such models is the constraint of viable trait space and trade-offs. Size-structuring and mechanistic descriptions of energy and resource allocation at the individual level can rationalize these constraints.
1. INTRODUCTION

The ecology and biogeochemistry of the world’s oceans are tightly interconnected. Microbial community structure is shaped by the highly variable physical, chemical, and predatory environment (Margalef 1968). Microbial community structure, in turn, regulates the environment and biogeochemical pathways including the export of organic matter to the deep ocean, critical for the global ocean sequestration of carbon and the modulation of atmospheric carbon dioxide. The diversity of marine microbial populations is both ecologically and biogeochemically important. For example, within the phytoplankton there is a sharp functional contrast between blooms of aggregating and sinking diatoms, armored against predators, and populations of picoplankton locked in a tightly coupled microbial loop with protist grazers (e.g., Pomeroy 1974, Laws et al. 2000). This contrast regulates the efficiency of nutrient utilization and the quality of exported organic matter on a regional and seasonal basis. Likewise, the depth at which organic matter is respired is, in part, regulated by mesopelagic and deep-sea communities of bacteria, archaea, and larger heterotrophs. The composition of these communities crucially determines the depth at which organic material is remineralized and thus where and when inorganic nutrients and carbon are returned to the surface ocean to fuel primary production again (Burd et al. 2002).

The feedback between biogeochemical cycles and microbial communities has stimulated the explicit representation of the latter in models used for climate and carbon cycle studies. Mathematical and numerical models provide tools with which to untangle the feedbacks and complex interconnections between marine ecology, biogeochemistry, and physiological diversity. They provide a platform to organize and quantify conceptual understanding, and by which to synthesize disparate observations, from environmental metagenomics to satellite observations of the ocean’s optical properties.

The marine microbial environment is extraordinarily diverse (Figure 1), with many thousands of algal species identified by microscopy, and much finer scales of richness have been identified with metagenomic methods (Venter et al. 2004, Delong et al. 2006). Whereas a few species typically dominate the population numerically in any local set of observations, they are complemented by a large number of minor players (Pedrós-Alió 2006). A useful starting point for interpreting populations of marine microbes is thus the Baas-Becking (1934) paradigm that “everything is everywhere but environment selects.” Consider the marine environment as a soup of microbial diversity with an enormously diverse background population of genotypes and associated phenotypes. Relative fitness in relation to the recent history of the local environment leads to selection for a particular subset of physiologies (Figure 2). Other physiologies are excluded in that environment but may be fitter in other regions or seasons. There is an emergent ecosystem structure and function that, in turn, modifies the environment. The background of diversity is maintained by physiological acclimation, genetic adaptation, and dispersal between environmental regimes. Long-standing questions include: Why are marine microbial populations so diverse (“The paradox of the plankton”; Hutchinson 1961)? What are the mechanisms that regulate biodiversity? Are diverse ecosystems more stable with regard to perturbations (e.g., McCann 2000)? These fundamental questions have serious implications for the application of numerical ocean ecosystem and biogeochemistry models to climate change questions. What is the appropriate representation of diverse ecosystems necessary to reflect and understand the response of the marine ecosystem to climate change, from both ecological and biogeochemical perspectives? Mathematical models provide a means to address these ecological questions, and there is a long history of theoretical work in this area (e.g., Armstrong & McGehee 1980, Huisman et al. 2001, Hubbell 2001). However, there have been few studies to date using numerical ocean circulation and ecosystem models to address these issues.
Light microscope images of representatives of several functional groups of phytoplankton at the same magnification scale. The phytoplankton population is extremely diverse in terms of function and genetic makeup. As illustrated, phytoplankton span many orders of magnitude in cell volume or mass. Cell size places some important restrictions on their ability to acquire resources and their interaction with predators. Diatoms form silica-based cell structures (here, a beautiful honeycomb pattern in *Coscinodiscus radiatus*) that may provide an energy savings relative to a cellulose structure. *Prochlorococcus* are the smallest photosynthesizing cells on the planet. Their small size makes them highly suited to nutrient-starved conditions. Coccolithophorids, e.g., *Emiliana huxleyi*, form calcium carbonate mineral structures affecting seawater chemistry. Many dinoflagellates can mix both photosynthesis and predation as energy sources. In coastal waters, some dinoflagellates, e.g., *Amphidinium carterae*, produce toxins, possibly as a defense strategy, which lead to harmful algal blooms and can affect shell fisheries.

Here, we review some current issues in, and examples of, models of diverse marine microbial communities. In part reflecting a historical bias, the discussion and examples shown here are skewed toward models of photoautotrophic primary producers, developed for biogeochemical applications. However, many of the issues discussed could, and should, also be pertinent to heterotrophic and mixotrophic organisms in the general sense. After a brief overview of ecosystem modeling and established approaches, our review focuses on recent developments in modeling diverse microbial populations, the significance for marine ecology and biogeochemistry, and the representation in ocean models.

2. MODELING MARINE PHYTOPLANKTON POPULATIONS

The parameterizations of microbes in today’s large-scale marine ecosystem and biogeochemistry models have a direct lineage from the seminal studies of Fleming (1939) and Riley (1946), who used Lotka-Volterra-type, predator–prey models to interpret the blooms and seasonal cycles in phytoplankton abundance in the English Channel and at Georges Bank, Massachusetts. Over the years, and with the advent of efficient computational resources for numerical approaches, models of marine plankton have advanced to include dynamic predator populations (e.g., Steele 1954), to encompass heterotrophic microbes (e.g., Fasham et al. 1990), and have been brought together...
Figure 2
The perspective that in the microbial world “everything is everywhere but the environment selects” (Baas-Becking 1934) provides a useful organizing principle for modeling diverse marine microbial communities. In this paradigm, the ocean is a “soup” that everywhere contains a wide range of potentially viable genotypes and associated phenotypes. At any point, selection occurs according to relative fitness given the recent history of the local environment. Thus, a particular subset of physiologies (genomes) is successful in particular regions and seasons, whereas others, though excluded there, may be fittest elsewhere. The fit microbial community, in turn, shapes the local environment. At the same time, the background variety of organisms with all degrees of fitness is maintained by dispersal between different environments, physiological acclimation, and genetic adaptation. Baas-Becking’s perspective provides a useful starting point for interpreting observed communities and a strategy for modeling diverse populations (Bruggeman & Kooijman 2007, Follows et al. 2007). Here, genetics and physiology are illustrated by a genome of Prochlorococcus ecotype MED4, environment by surface nitrate concentration, and ecosystem structure by species richness (see also Figure 7).

with three-dimensional circulation and biogeochemistry models (e.g., Sarmiento et al. 1993). An excellent historical perspective can be found in the review by Gentleman (2002).

Up to the end of the twentieth century, such models typically resolved only a single, generic photoautotroph: the classic nutrient-phytoplankton-zooplankton-detritus model. Such models were combined into ocean circulation models and provided large-scale estimates of carbon fluxes (e.g., Six & Maier-Reimer 1996). During the era of the Joint Global Ocean Flux Study, however, the biogeochemical modeling community recognized the importance of functional diversity within the phytoplankton and zooplankton communities. Facilitated by the continued expansion of computational resources, three-dimensional ocean biogeochemistry models moved toward the resolution of multiple, biogeochemically significant functional groups of phytoplankton (e.g., Chai et al. 2002, Moore et al. 2002a, Gregg et al. 2003, Le Quéré et al. 2005, Hasioka & Yamanaka 2007).
2.1. Functional Group Models

Such models typically resolve several (between two and six) functional groups or phytoplankton functional types (PFTs) that represent aggregates of many species with common biogeochemical functions. These are differentiated by nutrient requirements and differences in the values of coefficients that govern the basic, underlying physiological parameterizations. These parameters are provided, where possible, by laboratory studies and may be optimized to bring model state variables, such as chlorophyll concentration, into line with field observations (e.g., Friedrichs et al. 2007).

Each functional group provides a distinct control on biogeochemistry and biogeochemical pathways, comprehensively reviewed by Hood et al. (2006). Resolving diatoms, larger cells with silica frustules, as a separate group (e.g., Chai et al. 2002) is significant for capturing the export of organic material and connects to the global silicon cycle (Smetacek 1999, Armstrong et al. 2002). Coccolithophorids, with their dense calcium carbonate platelets, also enhance export of organic matter to the deep and modulate alkalinity, surface ocean carbonate chemistry, and the air-sea equilibrium of CO$_2$ (Zeebe and Wolf-Gladrow 2001). Models have resolved this functional type (e.g., Gregg & Casey 2007, Le Quéré et al. 2005), though the relationship between defining trade-offs and lith production is still rather uncertain. Diazotrophs, or nitrogen fixers, provide a source of new bioavailable nitrogen to the global ocean, fueling new and export production, and are often resolved in functional group models (e.g., Moore et al. 2004, Hood et al. 2004, Coles & Hood 2007), trading off nitrogen fixation against low growth rate and high iron demand. With appropriate parameterizations, large-scale patterns of these modeled phytoplankton functional groups broadly match those observed in situ and suggested from remote ocean color observations (e.g., Gregg & Casey 2007) (Figure 3).

Functional groups help to map out the role of different elements in regulating phytoplankton populations and productivity on a global scale (e.g., Le Quéré et al. 2005, Moore et al. 2004). Models using this approach have provided a means to extrapolate from the laboratory and sparse field observations to global-scale ocean biogeochemistry. For example, the large-scale regulation of productivity by iron, demonstrated by in situ fertilization experiments (synthesized by Boyd et al. 2007, de Baar et al. 2005), has been extrapolated to the global scale using ocean models (e.g., Moore et al. 2002b, Gregg et al. 2003, Aumont et al. 2003, Dutkiewicz et al. 2005a). These models have provided “laboratories” to help elucidate mechanisms controlling interannual variability in ecosystems (e.g., Wang et al. 2006, Dutkiewicz et al. 2001, Henson et al. 2009) and the corresponding impact on the biogeochemical cycles (e.g., Rodgers et al. 2008, Bennington et al. 2009, Ullman et al. 2009, Wang et al. 2005). Such simulations also provide a means to disentangle trends (such as anthropogenically forced changes) from natural variability (e.g., Henson et al. 2010) and explore the impact of changes in both natural and anthropogenic delivery of dust (a source of iron and nitrogen) to the ocean (e.g., Krishnamurthy et al. 2009, Bopp et al. 2003). The philosophies of these approaches, resolving a handful of phytoplankton functional types, are nicely summarized in Hood et al. (2006) and Le Quéré et al. (2005), and such simulations will continue to play a central role in studies of the ocean biogeochemistry.

However, continuing to increase the biogeochemical and ecological resolution of models presents new challenges: What degree of diversity should be represented? How does one represent a highly diverse functional group with a single set of parameters, especially when only a few candidates have been characterized in culture studies (Hood et al. 2006)? A single, optimized set of ecosystem structures (even after intense data parameterization and calibration) can fail to represent the dynamics of ecosystems in environmental conditions for which they were not tuned. Global change will very likely give rise to shifts in marine ecosystem structure that we...
cannot anticipate, perhaps elevating the fitness and abundance of currently rare organisms. Even if we could include every known organism and process in the present-day marine food web in our model structures, they could fail to predict the potential impacts of global change. Emergent features are likely to result from even more complex physiological and ecological interactions such as acclimation and genotypic and phenotypic adaptation. Most such models have many free
parameters that cannot be constrained with the available data and that therefore provide little or no predictive ability (Ward et al. 2010). As additional functional groups are added, the number of these free parameters and error accumulation increases. It is indeed not clear that adding additional functional groups will always lead to additional understanding or predictive skill (e.g., Friedrichs et al. 2007, Thingstad et al. 2010).

2.2. Models Rooted in Diversity

Building a model ecosystem up box by box toward an observed community structure with partially known physiologies thus presents many challenges. Even within the broad functional groups generally resolved in biogeochemical models, marine microbes are incredibly diverse. For example, whereas recent models have represented a single generic diatom, there are in fact thousands of identified species of diatoms spanning many orders of magnitude in cell volume, with a wide variety of morphologies and ecological strategies. The application of genomic characterizations in recent years reveals even finer-scale diversity, for example, the genetically and functionally diverse ecotypes of Prochlorococcus (Rocap et al. 2003) and heterotrophic bacteria (Morris et al. 2005) in the subtropical oceans. Likewise, a diversity of marine nitrogen fixers has been revealed and is being mapped in the oceans by molecular methods (Zehr et al. 1998, Church et al. 2008).

Recent studies have explored alternative approaches in which diversity is a fundamental property of the model ecosystem (Bruggeman & Kooijman 2007, Merico et al. 2009, Follows et al. 2007). In these studies, a diverse set of algal physiologies is initialized, and their interaction with the environment determines their relative fitness. This, in turn, organizes ecosystem structure and the feedback on the resource environment. These models are rooted in the paradigm of Baas-Becking (1934), depicted schematically in Figure 2.

The biodiversity-inspired approach of Bruggeman & Kooijman (2007) explored the organization of microbial communities in a subtropical water column. In their approach, described as a “system of infinite diversity,” a large number of phytoplankton types, distributed evenly over the viable range of two physiological characteristics, was maintained at each location in space. During the integration in time, competition for resources led to the concentration of biomass at the fittest combination of characteristics. With fitness depending on the environment, the modeled spatiotemporal variability induced continual adaptation of the plankton community. This was further ensured by the maintenance of a minimal biomass, even for types with less fit combinations, via an imposed background immigration rate. Elegantly, trade-offs in physiological abilities were imposed by assuming that resources must be invested either in the acquisition of a dissolved nutrient or in light harvesting, reflecting both acclimation and adaptation. The system captured the seasonal dynamics and vertical structure of microbial communities in the subtropical oceans and revealed the seasonal and spatial mapping of fitness conferred by investment in light and nutrient harvesting (Figure 4).

In the “selection-based” approach of Follows et al. (2007), physiological characteristics including sensitivity of growth to light, nutrient, and temperature were assigned stochastically to many tens of phytoplankton types. A randomly assigned size class provided some organizing trade-offs among these and other characteristics. This suite of virtual organisms was initialized in an ocean circulation and biogeochemistry model that resolves nitrogen, phosphorus, iron, and silica cycles. Robustly, across an ensemble of ten integrations, each with its own random assignment of growth parameters, many phytoplankton types with physiologies resembling real world analogs were consistently fit and populated the model ocean with plausible distributions and abundances (Follows et al. 2007). The biogeographical provinces of the model can be described by marking the ecological boundaries, or ecotones, where transitions in the biomass-dominant phytoplankton
(a) Modeled concentration of the biomass dedicated to light-harvesting investment, expressed in nutrient units (analogous to chlorophyll concentration) from the diversity-based marine ecosystem model of Bruggeman & Kooijman (2007). The model environment is based on the seasonal, subtropical water column at the Bermuda Atlantic Time-Series station (32°N, 64°W). (b) The distribution of biomass in the trait space of the model, indicating the fit region of parameter space and the trade-off between light harvesting and nutrient acquisition. The distribution is normalized so its integral equals one (and is therefore unitless). The width of the peak indicates the biodiversity of phenotypes in the solution. Adapted from Bruggeman & Kooijman (2007) with permission of the authors.
Figure 5

Annual mean emergent biogeography from a model in which 78 phytoplankton types were randomly assigned growth parameters. (a) Biogeographical provinces (defined by dominant species) separated by ecotones (transitions from dominance by one phytoplankton type to another): Inside each province, several phytoplankton coexist; only the most dominant is used for the demarcation of the provinces. (b) The large number of phytoplankton types surviving in this model can be collected into four functional groups: (i) diatom analogs (red), (ii) other large phytoplankton (yellow), (iii) other small phytoplankton (blue), (iv) Prochlorococcus analogs (green). This simulation used the ecosystem formulation of Follows et al. (2007) and Dutkiewicz et al. (2009) in a high-resolution (~18-km) version of the MIT general circulation model developed by the Estimating the Circulation and Climate of the Oceans Consortium (ECCO). Simulation performed by Oliver Jahn and Chris Hill, MIT. Figure credit: Oliver Jahn, MIT
Optimum temperature ($T_{opt}$) and light intensity ($I_{opt}$) for growth of all initialized *Prochlorococcus* analogs (circles) from the ensemble of ten model integrations used in Follows et al. (2007). Large circles indicate the analogs that exceeded a total biomass of 10⁶ mol P along the cruise track of AMT13. In the real ocean, several *Prochlorococcus* ecotypes, or strains, have been identified by their genetic diversity that leads to differences in their light and temperature sensitivity and to distinct geographic habitats (e.g., Johnson et al. 2006). Colored circles indicate classification of model ecotypes by geographic distribution: Red denotes significant biomass in the upper 25 m along the transect between 15°S and 15°N, corresponding qualitatively to the habitat of real-world ecotype eMIT9312; blue, analogs that had significant biomass in surface waters poleward of 15° and low biomass within 15° of the equator, broadly reflecting habitat of real-world ecotype eMED4; and green, analogs with a subsurface maximum biomass, similar to real-world habitat of ecotypes eMIT9313 and eNATL2A. Mixed color and solid black circles denote ambiguity in model ecotype classification. Bold diamonds indicate real-world *Prochlorococcus* ecotypes (eMIT9312 in red; eMED4 in blue; eNATL2A in green; and eMIT9313 in yellow). Pleasingly, both the geographic distribution of the wild types and the physiological specialization determined by laboratory studies of cultured types match the model analogs. No *Prochlorococcus* analog initialized with an optimal temperature for growth below 15°C survived. Abbreviation: AMT13, Atlantic Meridional Transect 13. Figure adapted from Follows et al. (2007).

reasons, but the unfit organisms should be excluded for the right reasons too (paraphrasing Rob Armstrong, personal communication).

However, it is also important to ask to what extent diversity at finer scales (i.e., within broad functional groups) is important to the understanding and simulation of large-scale marine biogeochemical cycles. Salihoglu & Hofmann (2007) found that incorporating several classes of autotrophic prokaryotes in their model was unnecessary for modeling primary production in a one-dimensional model of the equatorial Pacific ocean. However, models with more functional complexity appear to perform better when ported between different regions than do simpler models (Friedrichs et al. 2007). The degree of diversity to incorporate must be driven by the question to be addressed. The computational costs of diversity-based approaches can be significantly higher than the more phytoplankton functional-type models discussed in Section 2.1. Ultimately, multiple modeling philosophies and platforms will be needed in the quest to understand the links between ecosystems, biogeochemistry, and the environment.

### 2.3. Biogeochemical and Ecological Applications

Functional group models have generally resolved a single diazotroph, based on the cyanobacteria *Trichodesmium*. Recently, however, molecular techniques have revealed a diversity of nitrogen
fixers (Zehr et al. 1998, Church et al. 2008). By modifying the model of Follows et al. (2007) to allow any phytoplankton type to be a diazotroph, with appropriate trade-offs, Monteiro et al. (2010) found analogs of not only *Trichodesmium* but also unicellular diazotrophs and diatom–diazotroph associations to be successful and abundant in the global ocean model. Again, even these broad categories of diazotroph types were composed of several types. For example, several analogs of *Trichodesmium* persisted in the simulations, distinguished by the temperature sensitivity of their growth and consistent with observations of different genotypes with distinct temperature class distributions in the real world (Langlois et al. 2008). The model also suggested that unicellular diazotroph analogs could contribute as much to nitrogen fixation rates as the *Trichodesmium* analogs, in support of previous hypotheses (Carpenter et al. 1999, Zehr et al. 2001).

Diversity-based models also provide an opportunity to explore ecological questions that, to date, have not been addressed in the context of large-scale ocean models. These include the question of what maintains and shapes patterns of biodiversity in the oceans. Observations suggest an equator-to-pole decrease in diversity in marine microbial populations (Pommier et al. 2007, Fuhrman et al. 2008), consistent with records from higher trophic levels both in the ocean and on land. Pleasingly, a similar gradient is a feature of the solutions of a global ocean model with a diverse phytoplankton population (Barton et al. 2010). This gradient is overlain by hot spots of high diversity in some regions of energetic circulation (Figure 7). The model provides a tool with which to inquire why those patterns appear. In the illustrated model, the hot spots are largely a result of large-scale advective transport, bringing together and intermingling populations from distinct biogeographical provinces. The meridional gradient in the numerical model (in which there is no

![Number of phytoplankton species](https://www.annualreviews.org/doi/abs/10.1146/annurev-marine-121510-103821)

**Figure 7**

Phytoplankton species richness in an ocean model. Here, the measure of diversity is the number of phytoplankton types with biomass of more than 0.001% of the type with maximum biomass at any location. Superimposed on a poleward reduction in diversity are hot spots in the regions of the western boundary currents and other regions of energetic circulation. Further description of these large-scale patterns of diversity can be found in Barton et al. (2010). This simulation is the same as that shown in Figure 5. Adapted from figure provided by Oliver Jahn, MIT, with permission.
explicit speciation) appears to be related to neutral coexistence of similarly fit organisms in the stable subtropics (as suggested by Falkowski & Oliver 2007), whereas strong seasonality favors a prime opportunist and beats down diversity in the subpolar oceans (consistent with the simple model of Stewart & Levin 1973). Marine ecosystem models such as these may prove to be useful tools to address a broader range of ecological, as well as biogeochemical, questions.

Such models can also be useful for exploring the controls of microbial biogeography. It is often assumed that temperature constraints prevent Prochlorococcus and Trichodesmium from thriving outside the tropics and subtropics. Indeed, laboratory studies do reveal growth rates reflecting adaptation to warm temperatures, and thus temperature is sometimes used to shape the biogeography of modeled ocean plankton. However, temperature may not be the absolute control on the range of habitat of these organisms: When both warm- and cold-adapted Prochlorococcus analogs were seeded in an ocean model (Follows et al. 2007), the latter never became abundant (Figure 6). It is the nutrient environment that is the fundamental regulator for Prochlorococcus in the model (Bragg et al. 2010). If the organisms are fundamentally adapted to oligotrophic environments, which occur in the subtropics, then temperature dependency of growth should coevolve to be optimal in those waters, too. A similar organization of diazotroph analogs also occurs (Monteiro et al. 2010), and those that were cold adapted did not survive. Phytoplankton have presumably coevolved their nutrient, temperature, and pigment traits for their optimal niche globally and with depth in the water column (Hickman et al. 2010).

Finally, modeling frameworks that embrace diversity may be particularly suitable for studies of global change. It is possible that in a system undergoing significant perturbations, physiologies with previously low fitness and low abundance may come to the fore. In a model with very low intrinsic diversity, there is little opportunity for this to occur, though biogeography will certainly be rearranged. A more diverse model population (e.g., Follows et al. 2007, Bruggeman & Kooijman 2007) carries a “rare biosphere,” a tail of organism types present at low abundance (e.g., Figures 4b, 8) as is observed in nature. As the environment changes, phytoplankton of previously low fitness may become more abundant (Figure 9).

3. THE IMPORTANCE OF TRAITS AND TRADE-OFFS

The requirement of trade-offs, the costs and benefits of particular physiological characteristics, is of prime significance for functional group models, selection-based models, and those with adaptive dynamics (e.g., Bruggeman 2009). Arbitrary combinations of physiological characteristics (encoded as parameters such as growth rates and nutrient affinities) will ultimately generate a “Darwinian demon” or “superbug” that is optimal under all conditions and will exclude all others in all environments. In nature, this is not possible, because all organisms are constrained by conservation of resources and morphological considerations at the individual level.

An example of such a trade-off is that between K and r strategists (McArthur and Wilson 1967, Kilham & Hecky 1988). K-strategists, or gleaners, are adapted to compete effectively for limited resources, whereas r-strategists, or opportunists, are adapted to take advantage when resources are abundant. In terms of today’s phytoplankton parameterizations, K-strategists are characterized by high nutrient affinities (low half-saturation), and r-strategists by high maximum nutrient uptake (or maximum growth) rates (Grover 1991). Litchman et al. (2007) compiled results from laboratory studies to show clear empirical evidence for a trade-off between half-saturation and maximum growth (or uptake) in marine phytoplankton that is tied, in part, to cell size but also to trade-offs between encounter and handling of resources (Smith et al. 2009). The simplest functional-type models of phytoplankton reflect these two strategies. Broad classes of gleaners and opportunists,
Figure 8
Rank abundance of global biomass of the model of Dutkiewicz et al. (2009). Curves for 10 ensemble members (independent integrations) are shown, each initialized with 78 phytoplankton types, with random assignment of temperature, light, and nutrient-growth parameters. In all ensemble members, a few cell types account for most of the biomass, but a long tail indicates that many alternative physiologies persist at low abundance. Those types with low relative fitness in all environments are at the far end of the tail and continue to decline in abundance. Only the 50 most abundant types are depicted.

respectively, dominate the biomass of the oligotrophic subtropics and seasonally nutrient-replete subpolar oceans in a more complex selection-type model (Dutkiewicz et al. 2009).

Key physiological characteristics, or traits, provide a means by which to organize the all-important trade-offs (e.g., Margalef 1978, Reynolds 1984). Litchman & Klausmeier (2008) provide an excellent review of key phytoplankton traits and trade-offs, with implications for modeling. The trait space defining a phytoplankton may have axes based on morphology, resource-acquisition abilities, predator susceptibility, motility, and reproduction. However, these characteristics are not generally independent but have correlated variations often constrained by cell size and shape as well as energy and nutrient allocation.

Consider a two-dimensional trait space (Figure 10) in which each type of phytoplankton in a modeled population must be characterized. Although the illustrated traits are nonspecific, they might, for example, represent maximum nutrient uptake rate ($V_{max}$) and associated half-saturation coefficient ($K_s$). A completely random assignment of the two traits may generate an organism anywhere in the two-dimensional trait space. However, energy or resource trade-offs may limit the actual area that can be plausibly occupied and result in a correlation of traits. Of these physically reasonable trait combinations, only a subset is likely to be fit in a particular environment, and thus the observed set of trait combinations does not necessarily fill all possible areas of parameter space in a given environment or set of environments (see Bruggeman & Kooijman 2007) (Figures 4, 10).

To date, global and regional three-dimensional ocean models have typically used crude, though empirically motivated, parameterizations of trait trade-offs. Mechanistic models and laboratory characterizations are being used to provide fundamental constraints for the trade-offs that

www.annualreviews.org • Modeling Diverse Marine Microbes 439
The change in distribution of 4 (of the 100) phytoplankton types in a model simulation of a future-change scenario with no emissions policy. The marine ecosystem is based on Follows et al. (2007) and Dutkiewicz et al. (2009) and is embedded in the MIT Integrated Global Systems Model (IGSM2.3; an earth system model of intermediate complexity; Dutkiewicz et al. 2005b). One hundred phytoplankton were initialized in the model ocean and a quasi-steady-state distribution was obtained for present-day conditions. The model was then run forward to year 2100, by which time the ocean surface temperature has risen on average by approximately 3°C. Numbers were assigned to each phytoplankton type by order of initialization of parameters, and 4 representative types are shown here. (a) Annual mean biomass (mmol C m⁻³) for the 4 phytoplankton types for year 2000 conditions. (b) The same phytoplankton at year 2100. Shifts in biogeographical distribution occur over the 100-year simulation: Whereas some types have a global reduction in habitat as their temperature, light, and nutrient niche shrinks, others (e.g., type 76) with a very low abundance in year 2000 (out on the tail of the rank abundance; see Figure 8) become substantially more abundant in year 2100 (S. Dutkiewicz, J.R. Scott, M.J. Follows, unpublished manuscript).
Traits and trade-offs: Schematic of a generic two-dimensional trait space that characterizes each phytoplankton’s physiology. Energetic trade-offs limit the actual area that can be plausibly occupied, resulting in a correlation of traits. Such correlations have been observed in laboratory studies for, e.g., maximum nutrient-uptake rates and nutrient half-saturation constants and minimum cell quota (e.g., Litchman et al. 2007). Of these physically reasonable trait combinations, only a subset is likely to be fit in a particular environment, and thus the observed set of trait combinations does not necessarily fill all possible areas of parameter space. In a different environment, a different subset will be fit.

regulate microbial responses (Litchman et al. 2007, Finkel 2001, Armstrong 1999). In the following sections, we discuss constraints for modeling microbial physiology related to cell size and resource allocation.

3.1. Allometry

It has long been recognized that microbial growth, respiration, nutrient acquisition, light harvesting, and sinking speed scale with cell volume or mass (e.g., Laws 1975, Banse 1982, Tang 1995, Litchman et al. 2007, Finkel 2001, Kooijman 2001). Using cell size as a key organizing trait can significantly reduce the region of possible trait and parameter space in ecosystem models (Litchman & Klausmeier 2008).

The connection between metabolism, growth rate, and organism size has been noted for centuries (see, e.g., Kooijman 2000). Clear power law relationships hold, spanning many orders of magnitude of body size, including marine microbes (e.g., Banse 1982, Taylor & Shuter 1981, Belgrano et al. 2002, Tang 1995) (Figure 11). These relationships have been interpreted in terms of the optimization of pathways for the internal distribution of substrates (West et al. 1997) and, alternatively (but perhaps not exclusively), the influence of surface area to volume on nutrient uptake and demand (Litchman et al. 2007, Munk & Riley 1952, Aksnes & Egge 1991, Armstrong 2008). Predation may also be related to size (e.g., Hansen et al. 1997) due to the general expectation that predators are likely to tackle prey smaller than themselves (though there are numerous exceptions). The sinking rate of cells and particles can be linked to size and relative buoyancy through Stokes’s Law and elaborations thereof (e.g., Smayda 1970). Light harvesting and photosynthetic efficiency may also be regulated by the packaging of chlorophyll in cells of different size (Morel & Bricaud 1981, Finkel 2001). Thus, cell size provides a useful and mechanistic trait for the organization of some key trade-offs in microbial models.
Ecosystem models in which traits are tied to allometry may be pleasingly elegant. Relationships between phytoplankton cell size, abundance, respiration, and links to environmental controls can be captured by such models (e.g., Laws 1975, Irwin et al. 2006). Size-based models are also used to represent complex trophic networks (Figure 12) with relatively few parameters and can be applied to studies of biomass spectra and fluxes of energy (e.g., Baird & Sutthers 2007, Stock et al. 2008). Combining size with metabolic constraints following the Dynamic Energy Budget (DEB) theory (Kooijman 2000) provides a basis for models that abide by basic principles such as mass and energy conservation (e.g., Maury et al. 2007).

### 3.2. Resource Allocation

Cell size appears to be a key regulating factor for traits and trade-offs, but allometry alone cannot capture all of the important dynamics of marine microbial populations (Mei et al. 2009, Irwin et al. 2006, Chisholm 1992). Raven et al. (2005) said “biology’ complicates almost all of the arguments made purely on the basis of cell size.” For example, empirically, phytoplankton growth rate does indeed decrease with cell size according to a power law over several orders of magnitude (Figure 11). However, the spread at any particular cell size is as much as an order of magnitude. Within the cells depicted in Figure 11, dinoflagellates have systematically lower maximum growth rates than diatoms and other phytoplankton with the same cell volume (e.g., Banse 1982, Tang 1995). The offset in size–dependent growth between diatoms and dinoflagellates has been interpreted as reflecting a different investment in cellular machinery (Tang 1995). Laboratory studies reveal a higher cellular abundance of chlorophyll as a percentage of total protein in diatoms than dinoflagellates (Chan 1978). This is in accord with the observation that diatoms are more effective (at the same cell size) in pure phototrophic growth, whereas dinoflagellates are investing elsewhere, perhaps in mixotrophic abilities (Smayda 1997, Litchman et al. 2007, Bruggeman & Kooijman 2007).
Protozoans

Phytoplankton

Unresolved higher predation

Figure 12
Schematic of the size-resolved biological model of Baird & Suther (2007). Three trophic levels are shown: phytoplankton, protozoans, and metazoans, each encompassing several size classes. Curved arrows at the perimeter of the figure indicate phytoplankton and protozoans dividing. Dashed gray arrows indicate growth of metazoans into new size classes. Egg spawning is shown by curved dashed-dotted gray arrows. Loss to unresolved higher predation is indicated by purple arrows. All remaining arrows indicate predation between size classes and trophic levels. Parameters such as maximum growth rate, sinking, swimming, carbon content, and egg production, as well as minimum and maximum predator–prey size ratios, are functions of size. Model equations vary between phytoplankton, protozoans, and metazoans. However, within each group, one set of equations and allometric relationships is used for all sizes, simplifying the parameterization of numerous size classes. Figure adapted from Baird & Suther (2007) with permission of the authors.

There are other examples where the general organization of physiological characteristics by cell size and shape is modulated by the demands of specialization and associated resource and energy constraints. For instance, populations of diazotrophs (nitrogen-fixers) appear to grow much more slowly than other cells of comparable size due to the metabolic expense of fixing nitrogen (Geider & La Roche 2002), a trade-off commonly parameterized in functional group models. It is unlikely that allometry alone will sufficiently constrain trait space for marine ecosystem and biogeochemistry models, and resource allocation should be brought to bear explicitly.

4. PHYSIOLOGICAL PARAMETERIZATIONS
Consider the underlying framework of physiological parameterization on which large-scale marine ecosystem and biogeochemistry models are based. Most currently employ the simplest and most economical description of microbial physiology, Monod kinetics (Monod 1949), in which the population growth rate is a Michaelis-Menton-like function of the environmental concentration of the limiting substrate (Figure 13a). In this model, nutrient uptake and synthesis of new cells
Schematic representation of the parameterization of photoautotroph physiology employed in marine ecosystem models: (a) Monod-type, (b) Droop/Caperon-type with individual quota for each element, carbon (Qc), nitrogen (QN), and phosphorus (QP), (c) schematic concept for a model of algal physiology that resolves key biochemical components of an algal cell, including carbohydrates (CH), lipids (LIP), amino acids (AA), nucleic acids (NUC), and proteins (PR). Abbreviations: DIC, dissolved inorganic carbon; DIN, dissolved inorganic nitrogen; DIP, dissolved inorganic phosphorus. Figure inspired by Reynolds (2006), Pahlow & Oschlies (2009), Klausmeier et al. (2004), Shuter (1979), and others.

are directly linked, the limiting nutrient is determined following the law of the minimum, and the elemental ratios of the cells are fixed. Computational economy has been a significant motivation for the use of this simplest of parameterizations, though the idealizations and limitations of this approach have been criticized (e.g., Flynn 2003, Droop 2003) and computational constraints are rapidly becoming less significant.

A more physiologically defensible, yet still highly idealized, alternative is the Droop/Caperon “internal-stores” approach (Droop 1968, Caperone 1968), in which substrate uptake and synthesis are decoupled (Figure 13b). An internal store of each nutrient element is filled and synthesis/growth is linked to the internal concentration of the most limiting element, again found using a law-of-the-minimum approach. Less common in large-scale biogeochemical models (though see, e.g., Moore et al. 2002a), the internal stores approach is more physiologically defensible than Monod kinetics and enables variation in the elemental ratios of organic matter in response to environmental changes (Thingstad & Pengerud 1985, Martinussen & Thingstad 1987). However, although cell quotas and elemental ratios are flexible, they are somewhat arbitrary as the allocation of elements within the cell is unresolved.

Several studies point the way toward more biologically meaningful, but still computationally efficient, descriptions of cellular processes (Figure 13c). Shuter (1979) developed a model of algal growth that crudely resolved major biochemical components of the cell and a mechanistic parameterization of growth. Geider et al. (1998) allocated cellular resources toward light harvesting, basic metabolic needs, and storage, providing a dynamic model of photoacclimation. Bruggeman & Kooijman (2007) and Bruggeman (2009), in a similar vein, allocated cellular resources toward light harvesting and resource acquisition (both organic and inorganic), whereas Klausmeier et al. (2004) considered the impact on elemental ratios of the N:P stoichiometry of synthesis and resource acquisition machinery. Pahlow & Oschlies (2009) resolve key cellular components in an idealized yet mechanistic parameterization of algal population growth. Interestingly, in the late 1970s, models of Escherichia coli were developed that had strong analogies with Shuter’s (1979) algal growth model (Shuler et al. 1979). Those models laid the foundation for the increasing biological resolution of E. coli simulations that has continued to this day. This common ground points the way to a unified underlying platform for representing phototrophic, heterotrophic, and mixotrophic microbes and the associated trade-offs.
At the other end of the spectrum from Monod kinetics, whole-genome characterizations of marine microbes are now available (e.g., Kettler et al. 2007), effectively providing a blueprint for organisms at the level of fundamental metabolic pathways and biochemical reactions. In the case of *E. coli*, for which this type of information is most mature, this has led to extremely detailed models of individual cells rooted in conservation of elements and energy at the cellular scale, guided by whole-organism integral constraints (Flux Balance Analysis; e.g., Edwards et al. 2001). There is currently a vast chasm between the ecologically and biogeochemically oriented parameterizations embedded in global ocean models and the very detailed, metabolic-pathway perspective enabled by modern genomic studies.

It is unlikely that we will see the embedding of whole-cell systems biology models in large-scale ocean simulations in the near future, not only because of the obvious computational constraints, but because carrying billions of such detailed individuals would be unlikely to lead to insightful models. By analogy, physical oceanographers may understand and are able to simulate small-scale turbulence but choose to parameterize it in larger-scale simulations. However, we do expect that the emerging wealth of physiological understanding from genomic studies and systems-biological models of whole cells will lead to new and improved parameterizations of microbial physiology and resource allocation that are tractable, useful, and rooted in basic conservation laws at or below the individual level.

We believe that the development, constraint, and application of biochemically informed models, in the spirit of Shuter (1979) and Pahlow & Oschlies (2009), will bridge the gap between these perspectives by moving beyond Monod or Droop/Caperon kinetics toward more physiologically detailed and meaningful (yet still computationally tractable) descriptions of diverse microbial physiology for applications in marine ecology and biogeochemistry.

**SUMMARY POINTS**

1. Ocean models provide a platform to synthesize and quantify conceptual understanding. Ocean models can be used to interpolate and extrapolate sparse observations. In the last decade, ocean circulation and biogeochemistry models have moved toward resolving aspects of the diversity of marine microbial populations. Significant emphasis has been placed on resolving broad functional groups of phytoplankton, each with specific biogeochemical importance. These functional-type approaches have improved, and will continue to improve, our understanding of the ecosystem control on ocean biogeochemical cycles.

2. Recent studies have, however, begun to resolve finer-scale diversity within modeled phytoplankton and bacterial populations. Allowing diversity-based models to “self-organize” according to relative fitness provides a way to circumvent some of the obstacles to box-by-box extension of the functional group approach. An ecological principle that underpins the diversity-based models is Baas-Becking’s (1934) “everything is everywhere but the environment selects.”

3. Such approaches open up new avenues for application of ocean models to marine biogeochemical and ecological questions, including the regulation of global patterns of biodiversity and community structure and how these may alter in a changing environment. Such questions naturally interface with genomic observations of marine microbial populations.
4. The costs and benefits of resource and energy expenditure at the individual level are primary factors in determining relative fitness under different environmental conditions. Marine ecosystem models are just beginning to implement trait-based trade-offs linked to cell size and resource allocation.

FUTURE ISSUES

1. We expect to see molecular and genomic studies, along with individual-scale systems biology models, inform parameterizations of the trade-offs that constrain organisms.

2. Theoretical ecology and modeling of marine biogeochemical cycles have largely been separate activities to date. Diversity-based models are facilitating a bridge across this divide. Specifically, the perception of plankton communities as diverse, competition-shaped collections of types, distributed over trait space, permits the use of “adaptive dynamics” (e.g., Bruggeman 2009, Merico et al. 2009). This theoretical framework provides means for reduction of model complexity and for the integration of acclimation, succession, and evolution in a single modeling framework.

3. Whereas the biogeochemical models resolving marine microbial communities have focused on phototrophs, future models must take a more general perspective in which heterotrophic, mixotrophic, and phototrophic strategies are treated on a common platform. They are equally important for the application of models to ecological questions and for studies of global biogeochemical cycles, end-to-end food webs, and marine resources.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Mark Baird, Jorn Bruggeman, Watson Gregg, Oliver Jahn, and Chris Kempes for providing figures. Additionally, we thank Jorn Bruggeman for helpful comments. We are grateful for support from the Gordon and Betty Moore Foundation, National Oceanic and Atmospheric Administration, and National Aeronautics and Space Administration.

LITERATURE CITED


Gentleman W. 2002. A chronology of plankton dynamics in silico: how computer models have been used to study marine ecosystems. Hydrobiologia 480:69–85


Contents

Geologist at Sea: Aspects of Ocean History
Wolfgang H. Berger ................................................................. 1

Submarine Paleoseismology Based on Turbidite Records
Chris Goldfinger ................................................................. 35

Natural Processes in Delta Restoration: Application to the Mississippi Delta

Modeling the Dynamics of Continental Shelf Carbon
Eileen E. Hofmann, Bronwyn Cahill, Katja Fennel, Marjorie A.M. Friedrichs, Kimberly Hyde, Cindy Lee, Antonio Mannino, Raymond G. Najjar, John E. O’Reilly, John Wilkin, and Jianhong Xue .............. 93

Estuarine and Coastal Ocean Carbon Paradox: CO₂ Sinks or Sites of Terrestrial Carbon Incineration?
Wei-Jun Cai ................................................................. 123

Emerging Topics in Marine Methane Biogeochemistry
David L. Valentine ............................................................ 147

Observations of CFCs and SF₆ as Ocean Tracers
Rana A. Fine ................................................................. 173

Nitrogen Cycle of the Open Ocean: From Genes to Ecosystems
Jonathan P. Zehr and Raphael M. Kudela ........................................ 197

Marine Primary Production in Relation to Climate Variability and Change
Francisco P. Chavez, Monique Messié, and J. Timothy Pennington ............. 227

Beyond the Calvin Cycle: Autotrophic Carbon Fixation in the Ocean
Michael Hugler and Stefan M. Siweert ........................................ 261

Carbon Concentrating Mechanisms in Eukaryotic Marine Phytoplankton
John R. Reinfelder ............................................................ 291
Microbial Nitrogen Cycling Processes in Oxygen Minimum Zones
Phyllis Lam and Marcel M.M. Kuypers .................................................. 317

Microbial Metagenomics: Beyond the Genome
Jack A. Gilbert and Christopher L. Dupont ............................................ 347

Environmental Proteomics: Changes in the Proteome of Marine Organisms in Response to Environmental Stress, Pollutants, Infection, Symbiosis, and Development
Lars Tomanek ......................................................................................... 373

Microbial Extracellular Enzymes and the Marine Carbon Cycle
Carol Arnosti .............................................................. 401

Modeling Diverse Communities of Marine Microbes
Michael J. Follows and Stephanie Dutkiewicz ........................................ 427

Biofilms and Marine Invertebrate Larvae: What Bacteria Produce That Larvae Use to Choose Settlement Sites
Michael G. Hadfield ........................................................................... 453

DNA Barcoding of Marine Metazoa
Ann Bucklin, Dirk Steinke, and Leocadio Blanco-Bercial ........................................ 471

Local Adaptation in Marine Invertebrates
Eric Sanford and Morgan W. Kelly ......................................................... 509

Use of Flow Cytometry to Measure Biogeochemical Rates and Processes in the Ocean
Michael W. Lomas, Deborah A. Bronk, and Ger van den Engh ........................... 537

The Impact of Microbial Metabolism on Marine Dissolved Organic Matter
Elizabeth B. Kujawinski ....................................................................... 567

Errata

An online log of corrections to Annual Review of Marine Science articles may be found at http://marine.annualreviews.org/errata.shtml
New From Annual Reviews:

**Annual Review of Statistics and Its Application**

Volume 1 • Online January 2014 • http://statistics.annualreviews.org

Editor: **Stephen E. Fienberg**, Carnegie Mellon University

Associate Editors: **Nancy Reid**, University of Toronto

**Stephen M. Stigler**, University of Chicago

The *Annual Review of Statistics and Its Application* aims to inform statisticians and quantitative methodologists, as well as all scientists and users of statistics about major methodological advances and the computational tools that allow for their implementation. It will include developments in the field of statistics, including theoretical statistical underpinnings of new methodology, as well as developments in specific application domains such as biostatistics and bioinformatics, economics, machine learning, psychology, sociology, and aspects of the physical sciences.

Complimentary online access to the first volume will be available until January 2015.

**TABLE OF CONTENTS:**

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>What Is Statistics?</td>
<td>Stephen E. Fienberg</td>
</tr>
<tr>
<td>A Systematic Statistical Approach to Evaluating Evidence</td>
<td>David Madigan, Paul E. Stang, Jesse A. Berlin, Martijn Schuemie, J. Marc Overhage, Marc A. Suchard, Bill Dumouchel, Abraham G. Hartzema, Patrick B. Ryan</td>
</tr>
<tr>
<td>The Role of Statistics in the Discovery of a Higgs Boson</td>
<td>David A. van Dyk</td>
</tr>
<tr>
<td>Brain Imaging Analysis</td>
<td>F. DuBois Bowman</td>
</tr>
<tr>
<td>Statistics and Climate</td>
<td>Peter Guttorp</td>
</tr>
<tr>
<td>Climate Simulators and Climate Projections</td>
<td>Jonathan Rougier, Michael Goldstein</td>
</tr>
<tr>
<td>Probabilistic Forecasting</td>
<td>Tilmann Gneiting, Matthias Katzfuss</td>
</tr>
<tr>
<td>Bayesian Computational Tools</td>
<td>Christian P. Robert</td>
</tr>
<tr>
<td>Bayesian Computation Via Markov Chain Monte Carlo</td>
<td>Radu V. Craiu, Jeffrey S. Rosenthal</td>
</tr>
<tr>
<td>Build, Compute, Critique, Repeat: Data Analysis with Latent Variable Models</td>
<td>David M. Blei</td>
</tr>
<tr>
<td>Structured Regularizers for High-Dimensional Problems:</td>
<td>Martin J. Wainwright</td>
</tr>
<tr>
<td>Statistical and Computational Issues</td>
<td>High-Dimensional Statistics with a View Toward Applications in Biology, Peter Bühlmann, Markus Kalisch, Lukas Meier</td>
</tr>
<tr>
<td>Next-Generation Statistical Genetics: Modeling, Penalization, and</td>
<td>Kenneth Lange, Jeanette C. Papp, Janet S. Sinsheimer, Eric M. Sobel</td>
</tr>
<tr>
<td>Optimization in High-Dimensional Data</td>
<td></td>
</tr>
<tr>
<td>Breaking Bad: Two Decades of Life-Course Data Analysis in</td>
<td>Elena A. Erosheva, Ross L. Matsueda, Donatello Telesca</td>
</tr>
<tr>
<td>Criminology, Developmental Psychology, and Beyond</td>
<td></td>
</tr>
<tr>
<td>Event History Analysis</td>
<td>Niels Keiding</td>
</tr>
<tr>
<td>Statistical Evaluation of Forensic DNA Profile Evidence</td>
<td>Christopher D. Steele, David J. Balding</td>
</tr>
<tr>
<td>Using League Table Rankings in Public Policy Formation:</td>
<td>Harvey Goldstein</td>
</tr>
<tr>
<td>Statistical Issues</td>
<td></td>
</tr>
<tr>
<td>Statistical Ecology</td>
<td>Ruth King</td>
</tr>
<tr>
<td>Estimating the Number of Species in Microbial Diversity Studies</td>
<td>John Bunge, Amy Willis, Fiona Walsh</td>
</tr>
<tr>
<td>Dynamic Treatment Regimes</td>
<td>Bibhas Chakraborty, Susan A. Murphy</td>
</tr>
<tr>
<td>Statistics and Related Topics in Single-Molecule Biophysics</td>
<td>Hong Qian, S.C. Kou</td>
</tr>
<tr>
<td>Statistics and Quantitative Risk Management for Banking and Insurance</td>
<td>Paul Embrechts, Marius Hofert</td>
</tr>
</tbody>
</table>

Access this and all other Annual Reviews journals via your institution at [www.annualreviews.org](http://www.annualreviews.org).