

> SECTION VI. GLOBAL SYNTHESIS, MODELING, AND PREDICTION
> CHAPTER 11. MODELING AND PREDICTION

Modeling and Prediction of Marine Microbial Populations in the Genomic Era

BY RALEIGH R. HOOD, EDWARD A. LAWS, MICHAEL J. FOLLOWS, AND DAVID A. SIEGEL

The articles in this special issue attest to the fact that we are in the early stages of a scientific revolution in marine microbiology that is being fueled by vast quantities of new information derived from advances in microbiological techniques and genomic studies. Many new species, metabolic processes, and pathways in marine systems have recently been discovered. Subtle metabolic variations within and among species have also been revealed, and previously known genes and metabolisms have been detected in new environments. And this is just the tip of the iceberg. Using genome shotgun sequencing techniques, Venter et

al. (2004) report finding 148 previously unknown bacterial phylotypes and over 1.2 million previously unknown genes, including more than 782 new rhodopsinlike photoreceptors (Figure 1)—just from surface-water samples from the

possible approach to answering some of these questions. Traditionally, these are prognostic numerical models composed of a set of state variables representing various ecosystem and biogeochemical constituents with exchanges between

As the twenty-first century dawns, the application of new microbiological and genomic techniques in marine studies is creating an avalanche of new information.

RALEIGH R. HOOD (*rhood@hpl.umces.edu*) is Professor, University of Maryland Center for Environmental Science, Cambridge, MD, USA. EDWARD A. LAWS is Professor and Dean, School of the Coast and Environment, Louisiana State University, Baton Rouge, LA, USA. MICHAEL J. FOLLOWS is Principal Research Scientist, Department of Earth, Atmosphere, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. DAVID A. SIEGEL is Professor and Director, Institute for Computational Earth System Science, University of California, Santa Barbara, CA, USA.

Sargasso Sea (see also more recent articles by Rusch et al., 2007, and Yooseph, et al., 2007). The overwhelming challenge we face is how to make sense of all of this emerging information. What role do all these new genes and proteins play in driving marine-ecosystem dynamics and biogeochemical cycles? Which are important and which are not? What role are they likely to play in the evolution of marine microbial communities, how might they have influenced global biogeochemical cycles over Earth's history, and how might they do so in the future?

Marine-ecosystem and biogeochemical modeling techniques provide one

them specified by a set of partial differential equations (Hood and Coles, in press). These are the primary tools that we currently employ to synthesize emerging information and fold it into a coherent framework that can be used for prediction. Efforts to develop biogeochemical models that can be used to predict how the oceans will respond to global warming are a prominent example (e.g., Moore et al., 2002a, 2002b, 2004). These models attempt to encapsulate, in a mechanistic framework, the state of our knowledge of the marine ecosystem and of microbial dynamics that control marine biogeochemical

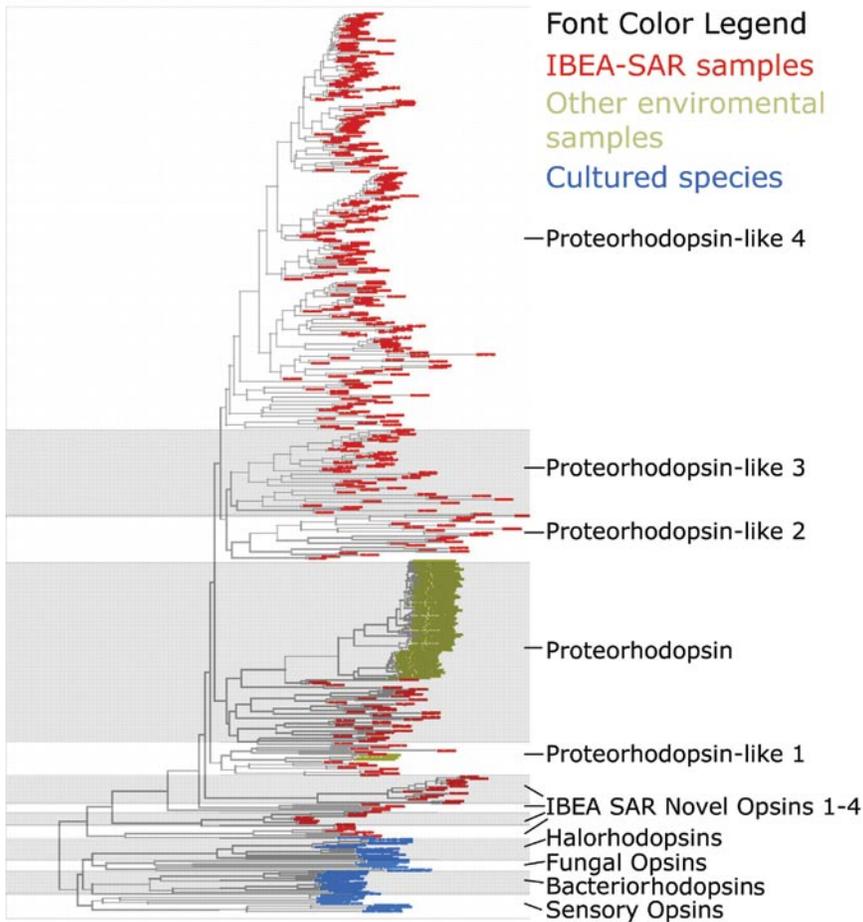


Figure 1. Phylogenetic tree of rhodopsinlike genes in the Sargasso Sea along with all homologs in GenBank. The sequences are colored according to the type of sample in which they were found: blue, cultured species; yellow, sequences from uncultured organisms in other environmental samples; and red, sequences from uncultured species in the Sargasso Sea. The tree is also divided into proposed distinct subfamilies of sequences on the right. *Figure and caption modified from Venter et al. (2004). Reprinted with permission from AAAS*

cycles. Such models have been gainfully applied as exploratory and predictive tools. However, it is not clear that these traditional modeling approaches will be sufficient in the face of all this emerging microbiological and genomic information; such models need to be “told” exactly what organisms and metabolisms exist in the ocean, and the rate coefficients that govern their parameterizations must be specified a priori. As such, they cannot tell you what is important

and what is not. Moreover, these models cannot “evolve” over time and cannot therefore account for changes in species composition and metabolic capacity that may happen in association with changing climate and environment. Given the emerging realization that there are vast and previously unknown quantities of genetic information in the ocean, it would appear that these deficiencies are serious.

We argue that if marine-ecosystem

models are to be in the vanguard of the ongoing revolution in microbial oceanography, there need to be radical changes in modeling approaches and strategies. The challenge we face is daunting. Our conceptual understanding is evolving much faster than model development. This is starkly illustrated by the fact that even though hundreds of new bacterial phylotypes have been discovered in recent years, most large-scale biogeochemical model formulations still do not include explicit representations of bacteria (Hood et al., 2006). We can anticipate that the accelerating rate of discovery will tend to encourage more rapid model development, but this poses a dilemma for the modeling community because it will also accelerate the current trend toward the development of increasingly complex and potentially intractable model formulations (see Doney et al., 2004, and Rothstein et al., 2006, for additional discussion).

In a perfect world, modeling and theory should help lead the way as we venture into this brave new world of microbial and genomic discovery in the twenty-first century. At the very least, these tools should be employed in combination with field and laboratory studies to help make sense of all this emerging information. To achieve this integration, we will have to augment our traditional ecosystem and biogeochemical modeling approaches with new methods. We argue that these methods should include (1) application of overarching ecological theories that can help guide model development, (2) development of alternative modeling approaches and analysis methods that can overcome some of the limitations

of traditional models, and (3) the adoption of an entirely new perspective on marine microbial interactions that can provide more realistic conceptualizations and models of detailed microbial community interactions.

ECOLOGICAL THEORY AND FUNDAMENTAL LAWS

In contrast to physical oceanographers, marine-ecosystem and biogeochemical modelers have no overarching theory to guide research (i.e., we do not have anything equivalent to the Navier-Stokes equations that provide a framework for the development of models with varying levels of simplification). Ecologists have hypothesized for many years that the evolution of biological systems is driven by fundamental principles or forces. Lotka (1922), for example, argued "...that natural selection tends to make the energy flux through the system a maximum, so far as compatible with the constraints to which the system is subject." Odum (1983) expanded on Lotka's ideas and formulated the "maximum power principle," suggesting that systems prevail that develop designs that maximize the flow of useful energy. Odum argued that theories and corollaries derived from the maximum power principle can explain much about the structure and processes of ecosystems. Although the maximum power principle has drawn some sharp criticism (e.g., Fenchel, 1987), these are important early examples of efforts to formulate laws that can provide an overarching framework for ecological studies and modeling that could be used to help us understand the role of newly discovered microorganisms and metabolic processes

in marine-ecosystem dynamics and biogeochemical cycling.

A number of authors have explored the implications of using various goal functions to help guide the development

of ecosystem models. Cropp and Gabric (2002), for example, used a genetic algorithm to simulate the adaptation of the biota in a simple linear food chain consisting of a limiting nutrient, autotrophs, and heterotrophs. They also concluded that ecological systems exist within the constraints of thermodynamic laws that prescribe the transfer of energy. But their simulations suggested the hypothesis that, within the constraints of the external environment and the genetic potential of their constituent biota, ecosystems will evolve to the state most resilient to perturbation (i.e., toward "maximum resiliency"). Interestingly, in their simple linear food chain, the selection pressures suggested by Lotka (1922) and Odum (1983) led to essentially the same system behavior as did maximum resiliency. Fath et al. (2001) likewise noted an equivalency of system behavior governed by seemingly disparate ecological goal functions. Although there are counter examples, such as Månsson and McGlade (1993), thermodynamic approaches have met with considerable success in

estimating rate parameters in real ecosystems (Jorgensen and Straskraba, 2000). We suggest that thermodynamic goal functions like these can provide a means to estimate unknown rates and

We argue that the traditional modeling tools...will be insufficient to allow an informed synthesis...and an assessment of what is important and what is not.

rate parameters, identify organisms or processes that might be missing in a model, and providing a means to guide community evolution under changing environmental conditions.

For example, Laws et al. (2000) used the simple box model depicted in Figure 2 to explore the implications of applying a resiliency goal function, based upon concepts developed by Steele (1974) and May (1974), to the regulation of sinking carbon export production in open-ocean ecosystems. Prior to the publication of that model, the seminal paper by Eppley and Peterson (1979) provided the conceptual basis for our understanding of the regulation of export production in the open ocean. Based on empirical observations, Eppley and Peterson (1979) reasoned that the ratio of new production (Dugdale and Goering, 1967) to total primary production (the *f*-ratio) could be described by a hyperbolic function of primary production. In subsequent years, accumulating evidence made the relationship postulated by Eppley and Peterson look more

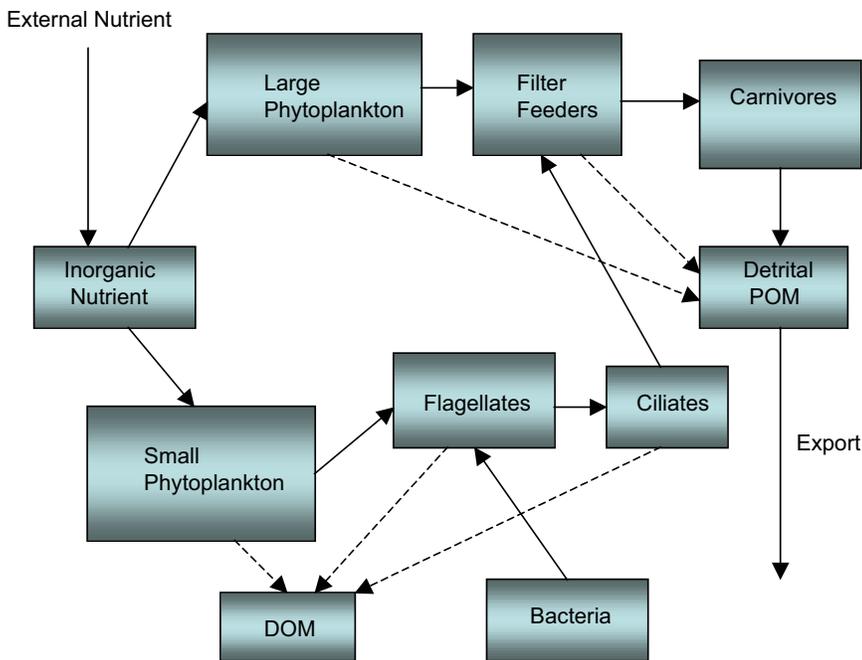


Figure 2. Feeding and excretion relationships in a model pelagic food web in which photosynthetic production is partitioned between small and large phytoplankton cells. DOM and POM are dissolved and particulate organic matter, respectively. Redrawn from Laws et al. (2000)

like a scatter diagram than a hyperbola. It seemed apparent that something was missing in our understanding of the mechanisms that regulate export production (see also new findings on export flux variability reported in Buesseler et al., 2007). The Laws et al. (2000) study was an effort not only to estimate unknown rate parameters but also to identify processes that might be missing in the Eppley and Peterson (1979) model, and develop a means to guide changes in pelagic ecosystem structure and biogeochemical cycling under changing environmental conditions.

Most of the parameters in the Laws et al. (2000) model can be assigned values based on information in the literature or are otherwise constrained by the assumption of steady state. However, four parameters—the fractional growth

rates of the large phytoplankton, flagellates, filter feeders, and carnivores—are unconstrained. These growth rates can be determined by selecting values randomly with the criterion for admissibility being a stable steady state such that the characteristic time constant associated with the return to equilibrium must be shorter than a specified value (following May, 1974). This tuning exercise not only determines these four free parameters, but, in the case of Station ALOHA, also leads to the conclusion that the export ratio most likely to be observed lies toward the low end of the spectrum of possible values, a result consistent with field studies (Emerson et al., 1997). Furthermore, when the predictions of the model are examined over a wide range of conditions, the conclusion is that f-ratios are a function of not only

primary production, as postulated by Eppley and Peterson (1979), but also of temperature (Figure 3).

The stability goal function provides a means of simulating changes in pelagic community structure in time and/or space in the Laws et al. (2000) model because it provides an independent criterion for reparameterizing the model under changing physical conditions without resorting to comparisons with observations and retuning. Specifically, the Laws et al. (2000) model was parameterized to maximize the resiliency of the steady-state solution at each point on a global grid in response to large spatial variations in production and temperature. This spatially explicit parameterization changes the size structure of the primary producers and the length of the food chain in the model, and therefore the amount of export versus recycling of organic matter. Thus, the community structure and the f-ratio “evolve” to function properly in the local environment. As a result, the model does a remarkably good job of predicting spatial variations in production and export. In theory, it should also be possible to use the same approach for simulating temporal evolution in community structure, production, export, and perhaps many other ecosystem attributes. This could, for example, provide a means to allow a model to reorganize itself seasonally and/or evolve in response to global-warming-induced changes in oceanic temperatures and primary production.

The success of the Laws et al. (2000) model in explaining the pattern of f-ratios over a wide range of conditions also lends support to May’s (1974) postulates concerning ecosystem resiliency.

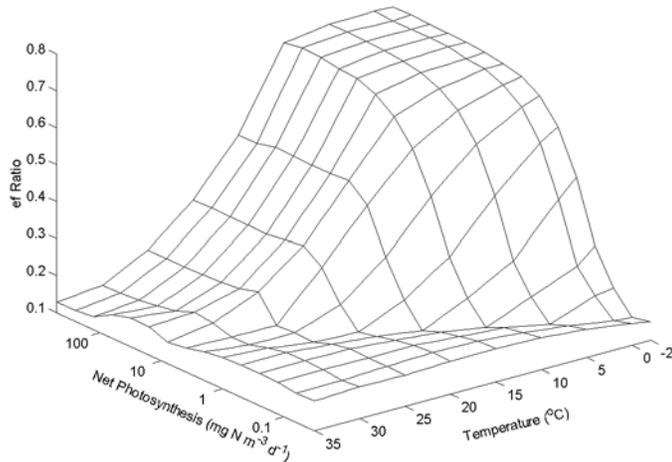


Figure 3. Calculated ratios as a function of temperature and net photosynthetic rate calculated from the Laws et al. (2000) model. From Laws et al. (2000)

It would be misleading, however, to think that resiliency is the sole determinant of ecosystem structure and function. A wide variety of constraints, including genetic plasticity and the physical (e.g., temperature) and resource constraints noted by Patten (1993), will combine to limit the ability of biological communities to adapt to their environments. Identifying the right mix of constraint and adaptability within models of microbial communities remains a challenging task and a significant limitation on our ability to predict the composition and functionality of these communities in a changing world.

LARGE-SCALE, THREE-DIMENSIONAL MODELS

A long-standing challenge in modeling marine ecosystems and biogeochemical cycles is how to represent the diversity of the ecosystem, the spatial and temporal variations of community structure, and their impacts on nutrient and carbon fluxes. The application of new

microbiological and genomic techniques has revealed that there is much greater diversity in marine microbial communities than previously thought, a discovery that has magnified this challenge tremendously. This diversity is currently addressed in ocean biogeochemistry models by resolving, to varying degrees, the size structure of the microbial community at different trophic levels (e.g., by incorporating multiple size classes of phytoplankton and zooplankton) and, more recently, through the inclusion of biogeochemical “functional groups,” where the groups are defined according to specific, biogeochemically relevant functionality (e.g., the formation of a silicious frustule in the case of diatoms, calcification by coccolithophorids like *Emiliania huxleyi*, and nitrogen fixation by cyanobacteria like *Trichodesmium*) (Hood et al., 2006).

These models have provided significant insight into the interactions of the phytoplankton community and biogeochemical cycles, but several major hur-

dles remain. The specification of appropriate functional groups and diversity for a particular problem is difficult and somewhat arbitrary. In addition, most biogeochemically motivated models of marine microbes focus on photo-autotrophs, whereas much less attention has been paid to the role of heterotrophy and even less to mixotrophy. The latter are important determinants of community structure through top-down control effects, and they are also important components of some biogeochemical functional groups (e.g., foram-mediated calcification). Of particular concern in this context is the fact that most biogeochemical models that are being used for regional and global-scale applications do not include explicit representations of bacteria, and those that do are highly simplified, representing bacteria and any related groups (e.g., archaea) as a single, generic state variable. Another problem is that organism types in these models are differentiated by rate constants regulating physiological processes that are poorly constrained (i.e., they are derived from laboratory culture studies, but these are relatively scarce and may not include the appropriate organisms, and the broad functional groups may not be characterized easily by a single parameter value). Finally, there is the overarching issue of dealing with models that are becoming increasingly complex and the question of whether these will ultimately give rise to better predictive skill.

In the vast majority of marine-ecosystem and biogeochemical models, the representation of key physiological processes—for example, the dependence of phytoplankton growth on light, temperature, and nutrient availability—are

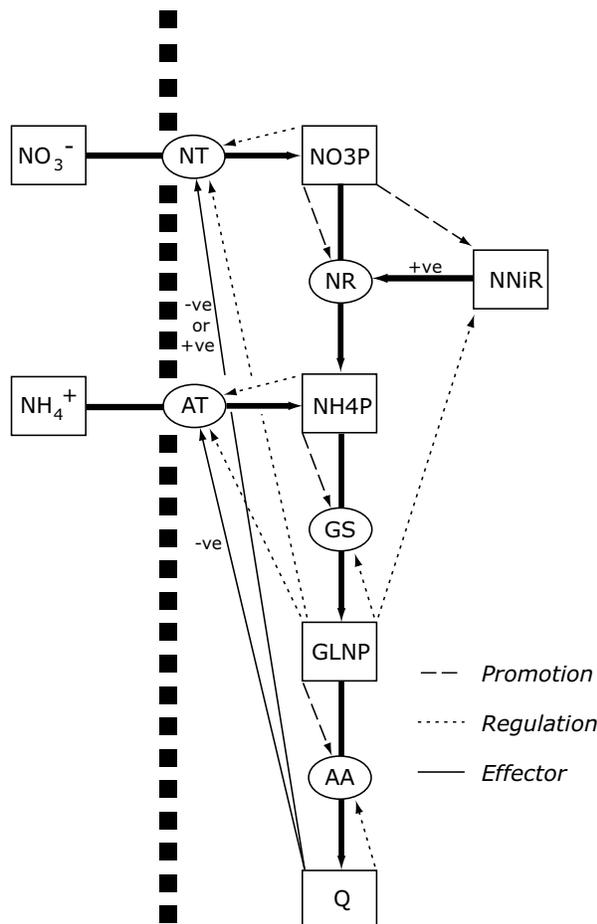
empirical, using a few simple functional relationships (e.g., Holling's type I, II, and III grazing response curves and Q_{10} -type temperature dependence of metabolic function), constrained where possible by laboratory culture and chemostat studies (e.g., Eppley and Coatsworth, 1968; Eppley, 1972). These functional relationships are also typically fixed in the sense that adaptation to changing local conditions is not allowed (i.e., photo and nutrient acclimation). Thus, compared to real-world ecosystems, most marine-ecosystem and biogeochemical models are quite rigid in terms of their physiological responses. This rigidity may be particularly prob-

lematic in efforts to apply these models to assess how plankton community composition and biogeochemical cycles might be impacted by past or future climate change (i.e., we can anticipate that they will tend to predict changes in species compositions and chemical cycling that are too abrupt because the plankton in these models cannot adapt to changing conditions). There has, however, been significant progress in the last decade toward the development of parameterizations of microbial physiology that are informed and underpinned by cellular- and molecular-scale models (e.g., Geider et al., 1996; Flynn et al., 1997) (Figure 4). In these studies, the

environmental sensitivities of particular cell types are represented dynamically, with predictive power based in an understanding of intracellular biochemistry and metabolic pathways. Further development and incorporation of these kinds of physiological "submodels" will not only give rise to more flexible ecosystem and biogeochemical responses, but they should also facilitate incorporation of new metabolic pathways and processes as they are discovered.

A major caveat, though, is that all of this is leading us further down the pathway of developing increasingly complex model formulations that are becoming more and more difficult to parameterize, validate, and diagnose. Will these models ultimately lead to improved understanding and increased predictive skill? Significant progress has also been made in recent years toward addressing this question through the application of data assimilation techniques, which can be used not only to refine model parameter values, but also to provide insights into the impacts of model complexity on predictive skill. For example, Friedrichs et al. (2007) have used the adjoint method to carry out some of the first quantitative intercomparisons of different biogeochemical model formulations that incorporate widely varying levels of ecosystem and biogeochemical complexity. Among other things, this work has shown that complex model formulations can provide better predictive skill than more simplified models and also better portability (Figure 5), where the latter refers to the ability of a model to adapt and function properly in different oceanic environments. But an important caveat is that more complex

Figure 4. Schematic diagram of the Flynn et al. (1997) mechanistic nitrate and ammonium uptake model. NO_3P , NH_4P , GLNP , and Q are internal pools of nitrate, ammonium, glutamine, and other organic cellular N, respectively. NNiR is nitrate-nitrite reductase, and GS glutamine synthetase activities. NT and AT are nitrate and ammonium transporters, respectively. NR describes the process of nitrate reduction through to ammonium, and AA the synthesis of amino acids and all other nitrogenous compounds from GLN . "Promotion," "regulation," and "effector" are used in general terms, with no specific biochemical meaning, indicating positive, negative, or complex feedbacks, respectively. Reprinted from Flynn et al. (1997, Figure 2)



models must be properly constrained with data (i.e., if they have too many degrees of freedom, then they can be tuned to fit noise in the data, which will result in reduced predictive skill). This is a significant caveat because it means that the amount of complexity that we can usefully incorporate into models will be dictated by the availability of the validation data that is needed to constrain them. The implication is that continuing to add new organisms and metabolic processes to models that are discovered through microbiological and genomic studies may not be useful unless validation data relevant to these processes (e.g., time-series or spatial data) can also be obtained.

The trend toward the development of increasingly complex models has raised concerns, especially among ecologists who have known for many years that there may be fundamental limits to the level of complexity that can be usefully incorporated into ecological and biogeochemical models (May, 1974). Typically, marine-ecosystem models have progressively increased the complexity and resolution of functional groups, incrementally adding new “species” or functional types with a priori imposed physiological characteristics. In contrast, an alternative approach has recently been explored by Follows et al. (2007): the model is initialized with a very diverse phytoplankton community, explicitly representing many tens of potentially viable functional types whose physiological characteristics are provided stochastically, from plausible ranges. When embedded in a simulated, global, four-dimensional (x , y , z , and $time$) physical and chemical environment, several of the (relatively) fittest

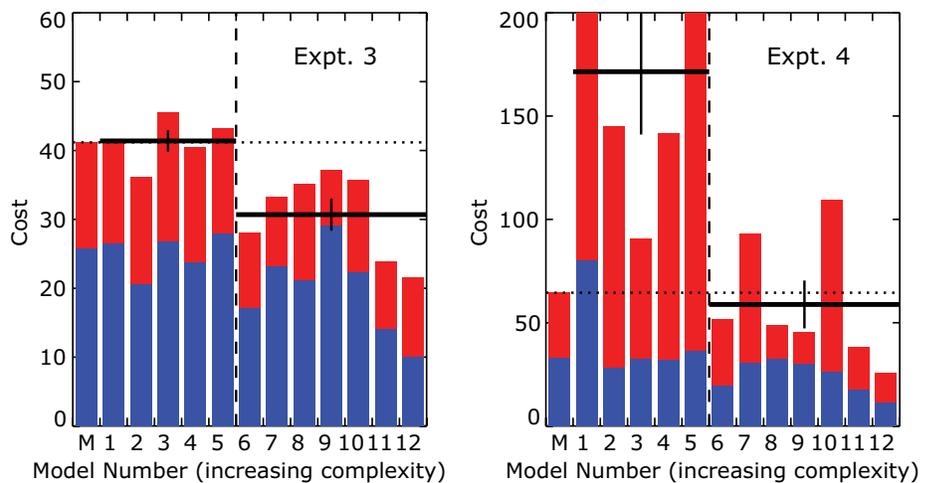


Figure 5. Results from quantitative model intercomparisons from Friedrichs et al. (2007). Cost function (J) is plotted as a function of model number. The taller bars represent higher cost and poorer fit to observations from the Arabian Sea (AS) and the Equatorial Pacific (EP). Red bars represent the EP component of J ; blue bars represent the AS component of J . The vertical dashed line separates the single-phytoplankton models (models 1–5) from the multi-phytoplankton models (models 6–12). Bars lower than the dotted horizontal line indicate that the model-data misfit is lower than that computed from the mean of the observations. Two solid horizontal lines represent mean cost for the single-phytoplankton and multi-phytoplankton models, respectively; error bars indicate one standard error. Expt. 3 results are derived from simultaneous assimilation of AS and EP data. Expt. 4 results are from cross-validation experiments where the models were fit to data from AS(EP) and then cost J calculated at EP(AS). Figures and caption modified from Friedrichs et al. (2007)

organism types grow to dominate each biogeographical “province” while many, less-fit types decline to very low abundance or extinction. The system thus “self-selects” its own community structure in a manner that is conceptually related to natural selection. The resulting community structure and physiological characteristics of the dominant organisms are emergent properties of the model. Follows et al. (2007) demonstrate that the organism types that come to the fore seem to reflect real-world counterparts in the ocean; that is, the model generates phytoplankton biomass distributions and community structure that bear some resemblance to observed, large-scale distributions (Figure 6). When enabled (but not demanded), the model

even supports multiple analogs of strains of the cyanobacterium *Prochlorococcus* that exhibit similar variety of habitat and correlated physiological characteristics to those observed in situ.

The success of this study suggests that the stochastic, self-organizing approach to modeling marine ecosystems has significant potential for wider application. It circumvents many of the obstacles and concerns that are currently encountered with typical “forward” modeling approaches, such as the a priori imposition of increasingly large numbers of organisms and parameterized physiological responses. One could envision this approach ultimately leading to models where the relative fitness of a virtual cell type is determined by

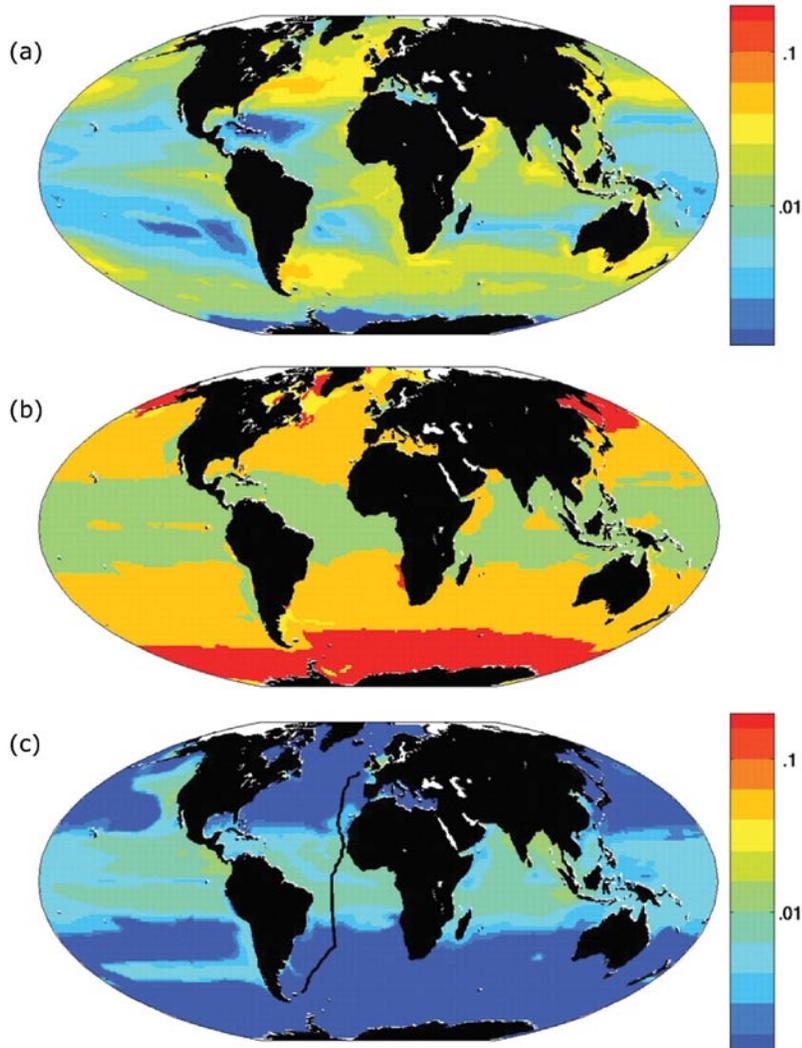


Figure 6. Phytoplankton biomass and community structure from a single integration of the Follows et al. (2007) global model. (a) Modeled annual mean total phytoplankton biomass in terms of phosphorus ($\mu\text{M P}$, average 0–50 m). (b) Emergent biogeography: Modeled photo-autotrophs are categorized into four phytoplankton functional groups, each of which is a composite of several model phytoplankton types. Regions are color coded according to the functional group that dominates local biomass in the annual mean. Green: analogs of *Prochlorococcus*; orange: other small photo-autotrophs; red: analogs of diatoms; yellow: other large phytoplankton. (c) Modeled annual mean total biomass of *Prochlorococcus* analogs ($\mu\text{M P}$, average 0–50 m). From Follows et al. (2007). Reprinted with permission from AAAS

explicit models of cellular-scale metabolic processes, modified by its genomic information and the environment in which it is embedded. Trade-offs could be determined by considering the thermodynamics and nutrient budgets of

individual cells. Embedded in a simulated physical, nutrient, and predatory environment, the system would move toward a solution where the relatively fittest cell types dominate. This approach might not only provide tools for orga-

nizing and interpreting genomic data, but also the basis for an entirely new class of genetically informed models that allow community structure to emerge and evolve in response to changing environmental conditions.

PLANKTON POPULATION DYNAMICS AND SUBGRID-SCALE PROCESSES

Marine microbes, including bacteria, archaea, and viruses, are among the most numerous organisms on Earth, with prokaryotic abundances typically greater than a million cells per milliliter of seawater, and viruses often occur with abundances that are ten times that amount (e.g., Fuhrman, 1999; Pomeroy et al., this issue; Breitbart et al., this issue). Yet the distances between adjacent organisms are quite large relative to the sizes of the organisms themselves as well as the scales of the turbulent fluid that surrounds them (e.g., Hulbert, 1970). It has long been suggested that phytoplankton create diffusive spheres of nutrient deficit about them (Pasciak and Gavis, 1974), and their discrete nature has been observed in situ (e.g., Franks and Jaffe, 2001). This notion of individual interaction is built into modern models of zooplankton prey encounter and utilization (e.g., Wiggert et al., 2005). In all, marine microbial interactions are, to first order, among individuals, and this discrete nature will have influence on the population dynamics and community structure changes.

The discrete nature of marine populations suggests that to understand and predict the dynamics of marine microbial communities, the spatial organization and interactions of individuals need

to be assessed (e.g., Blackburn et al., 1997). Even though the microbial abundances are huge, in the fluid mechanical sense they are dilute. Therefore, microbial populations will respond to environmental perturbations as discrete individuals interacting (slowly) through a viscous fluid medium (e.g., Siegel, 1998). This, in turn, should allow the emergence of microscale patterns in the spatial organization of individuals within a community. Figure 7 is a cartoon version of this “spatialization” in an impressionist painting of microbial interactions within a drop of seawater by Farooq Azam (1998). Here, there is a localization of where processes are occurring and the development of “hot spots” of activity. Clearly, the aggregated activity of this entire drop of seawater depends on the combined actions of the various localized hot spots of activity.

This localization of activity suggests that determinations of volume-averaged organism abundance is not sufficient alone for describing the dynamics of microbial communities. These circumstances, where processes acting on an individual scale can regulate the dynamics of the community on population scales, are often called subgrid-scale (SGS) problems. We believe that achieving this level of understanding will lead to increased predictability in microbial systems and ultimately to new ways to represent newly discovered microbes and metabolic processes in marine-ecosystem and biogeochemical models.

The development of individual-based models (IBMs) that represent ecosystem dynamics at these scales is an obvious avenue for future modeling-oriented research. IBMs are models that explicitly

represent individual organisms (i.e., their movement, interactions, and impacts). This type of model is particularly well suited for representing organisms in dilute environments that are influenced by physical and chemical cues in the context of larger-scale circulations, for example, zooplankton dynamics in the California Current (e.g., Batchelder et al., 2002), salmon migration and population dynamics (e.g., Rand et al., 1997), and even jellyfish transport and swimming behavior (Hood et al., 1999; Matanoski and Hood, 2005). Microbial community interactions represent a similar kind

of problem where the consequences of directed and random motility and the slow diffusion of chemical substrates in a viscous fluid medium are particularly important. Microbial community IBMs could be used, for example, to investigate the costs and benefits associated with the production of allelopathic substances or the role of the SGS clustering of organisms on the trajectory of a population. Similarly, these types of models could also be used to try to better understand the ecological and biogeochemical role of newly discovered bacterial phylogenies and metabolic processes.

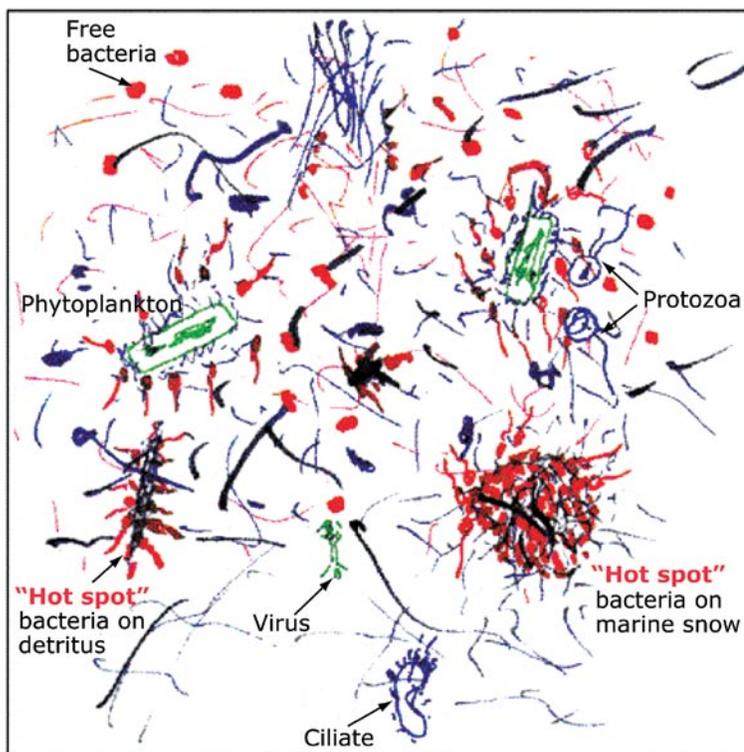


Figure 7: The microbial loop: a bacterium's-eye view of a drop of seawater from the ocean's euphotic layer. Bacteria (red) acting on marine snow or detrital particles (black) or organic matter efflux from phytoplankton (green) creating diverse microniches or “hotspots,” which can support high bacterial diversity and high productivity. Protozoa are also seen aggregating about these “hotspots” of activity. Figure and legend modified from Azam (1998). Reprinted with permission from AAAS

SUMMARY AND CONCLUSIONS

As the twenty-first century dawns, the application of new microbiological and genomic techniques in marine studies is creating an avalanche of new information. We argue that the traditional modeling tools that we have used for decades to simulate marine-ecosystem dynamics and biogeochemical cycles will be insufficient to allow an informed synthesis

can help guide modeling efforts. These have already been applied in marine-ecosystem modeling studies to constrain free parameters, identify processes that are missing, and create models that can reorganize themselves and “evolve” in response to spatial and temporal changes in environmental conditions.

New techniques like data assimilation can now be used to help parameterize

informed models that allow community structure to emerge and evolve in both time and space.

It is also important to recognize that we may need to change the way we think about, study, and model microbial interactions. Measures of mean organism abundance may be meaningless in describing the dynamics of microbial communities at very small scales of interaction. If we want to understand these interactions, then investigations at these small scales will be required. We also need to figure out how this information propagates up to and impacts larger-scale ecosystem dynamics and biogeochemical cycles. We already have at our disposal modeling tools, like IBMs, that are well suited for simulating microbial interactions in dilute, viscous environments. These types of models need to be applied more aggressively to try to better understand the ecological and biogeochemical role of newly discovered bacterial phylotypes and metabolic processes.

If the modeling community wants to be in the vanguard of scientific discovery, it will have to develop and/or apply some new tools.

of this information and an assessment of what is important and what is not. The problem is that traditional models are static structures that need to be told what is important, and they cannot “evolve” under changing environmental conditions. These models also tend to be somewhat ad hoc in terms of the processes and organisms that are included, and they are still quite primitive in some respects—especially in terms of their representations of prokaryotic organisms. Dealing with the increasing complexity of these models represents a major challenge in and of itself.

If the modeling community wants to be in the vanguard of scientific discovery, it will have to develop and/or apply some new tools. In this paper, we suggest several alternative modeling approaches, including the identification of overarching ecological theories and “goal functions” based upon thermodynamic analogues and stability considerations that

complex models and determine whether or not they can provide better predictive skill. New adaptive physiological and metabolic submodels are also being developed that can give models more flexibility to adapt to changing environmental conditions. Alternative model-

ing approaches have also been recently developed that allow emergent plankton communities to self-assemble, organized in an analogy of “natural selection” and responding to the simulated physical and chemical environment. We are optimistic that these new approaches can give rise to much more flexible, genetically

...we may need to change the way we think about, study, and model microbial interactions.

ACKNOWLEDGEMENTS

The contributions of R.R. Hood and E.A. Laws to this paper were supported by NSF grant OCE 0622276. Additional support was provided by NSF grants OCE 0628389 and OCE 0308440 to D.A. Siegel, and grant OCE 336839 to M.J. Follows. MJF is also grateful for

support from the PARADIGM consortium of the National Oceanographic Partnership Program and the Gordon and Betty Moore Foundation Marine Microbiology Initiative. 

REFERENCES

- Azam, F. 1998. Microbial control of oceanic carbon flux: The plot thickens. *Science* 280:694–696.
- Batchelder, H.P., C.A. Edwards, and T.M. Powell. 2002. Individual-based models of zooplankton populations in coastal upwelling regions: Implications of diel vertical migration on demographic success and nearshore retention. *Progress in Oceanography* 53:307–333.
- Blackburn, M., F. Azam, and A. Hagstroem. 1997. Spatially explicit simulations of a microbial food web. *Limnology and Oceanography* 46:613–622.
- Buesseler, K.O., C.H. Lamborg, P.W. Boyd, P.J. Lam, T.W. Trull, R.R. Bidigare, J.K.B. Bishop, K.L. Casciotti, F. Dehairs, M. Elskens, and others. 2007. Revisiting carbon flux through the ocean's twilight zone. *Science* 316:567–570.
- Cropp, R., and A. Gabric. 2002. Ecosystem adaptation: Do ecosystems maximize resilience? *Ecology* 83:2,019–2,026.
- Doney, S., M.R. Abbott, J.J. Cullen, D.M. Karl, and L. Rothstein. 2004. From genes to ecosystems: The ocean's new frontier. *Frontiers in Ecology and the Environment* 2:457–466.
- Dugdale, R.C., and J.J. Goering. 1967. Uptake of new and regenerated forms of nitrogen in primary productivity. *Limnology and Oceanography* 12:196–206.
- Emerson, S., P. Quay, D. Karl, C. Winn, L. Tupas, and M. Landry. 1997. Experimental determination of the organic carbon flux from open-ocean surface waters. *Nature* 389:951–954.
- Eppley, R.W. 1972. Temperature and phytoplankton growth in the sea. *Fishery Bulletin* 70:1,063–1,085.
- Eppley, R.W., and J.L. Coatsworth. 1968. Uptake of nitrate and nitrite by *Ditylum brightwellii*—kinetics and mechanisms. *Journal of Phycology* 4:151–156.
- Eppley, R.W., and B.J. Peterson. 1979. Particulate organic matter flux and planktonic new production in the deep ocean. *Nature* 282:677–680.
- Fath, B.D., B.C. Patten, and J.S. Choi. 2001. Complementarity of ecological goal functions. *Journal of Theoretical Biology* 208:493–506.
- Fenchel, T. 1987. *Ecology—Potentials and Limitations*. Ecology Institute, Inter-Research, Oldendorf/Luhe, Germany.
- Flynn, K.J., M.J.R. Fasham, and C.R. Hipkin. 1997. Modeling the interactions between ammonium and nitrate uptake in marine phytoplankton. *Philosophical Transactions of the Royal Society of London B* 352:1,625–1,645.
- Follows, M.J., S. Dutkiewicz, S. Grant, and S.W. Chisholm. 2007. Emergent biogeography of microbial communities in a model ocean. *Science* 315:1,843–1,846.
- Franks, P.J.S., and J.S. Jaffe. 2001. Microscale distributions of phytoplankton: Initial results from two-dimensional imaging. *Marine Ecology Progress Series* 220:59–72.
- Friedrichs, M.A.M., J.A. Dusenberry, L.A. Anderson, R. Armstrong, F. Chai, J.R. Christian, S.C. Doney, J. Dunne, M. Fujii, R. Hood, and others. 2007. Assessment of skill and portability in regional marine biogeochemical models: The role of multiple planktonic groups. *Journal of Geophysical Research* doi:10.1029/2006JC003852.
- Fuhrman, J.A. 1999. Marine viruses and their biogeochemical and ecological effects. *Nature* 399:541–548.
- Geider, R.J., H.L. McIntyre, and T.M. Kana. 1996. A dynamic model of photoadaptation in phytoplankton. *Limnology and Oceanography* 41:1–15.
- Hood, R.R., and V.J. Coles. In press. Marine Ecosystem Models. In *Encyclopedia of Ecology*, S.E. Jorgensen and A. Voinov, eds, Elsevier.
- Hood, R.R., E.A. Laws, R.A. Armstrong, N.R. Bates, C.W. Brown, C.A. Carlson, F. Chai, S.C. Doney, P.g. Falkowski, R.A. Feeley, and others. 2006. Pelagic functional group modeling: Progress, challenges and prospects. *Deep-Sea Research Part II* 53:459–512.
- Hood, R.R., H.V. Wang, J.E. Purcell, E.D. Houde, and L.W. Harding. 1999. Modeling particles and pelagic organisms in Chesapeake Bay: Convergent features control plankton distributions. *Journal of Geophysical Research* 104:1,223–1,243.
- Hulburt, E.M. 1970. Competition for nutrients by marine phytoplankton in oceanic, coastal and estuarine regions. *Ecology* 51:475–484.
- Jorgensen, S.E., and M. Straskraba. 2000. Ecosystems as cybernetic systems. Pp. 249–264 in *Handbook of Ecosystem Theories and Management*. S.E. Jorgensen and F. Muller, eds, Lewis, Boca Raton, FL.
- Laws, E.A., P.G. Falkowski, W.O. Smith, H. Ducklow, and J.J. McCarthy. 2000. Temperature effects on export production in the open ocean. *Global Biogeochemical Cycles* 14:1,231–1,246.
- Lotka, A.J. 1922. Contribution to the energetics of evolution. *Proceedings of the National Academy of Sciences of the United States of America* 8:147–150.
- Matanoski, J.C., and R.R. Hood. 2005. An individual-based numerical model of medusa swimming behavior and implications for foraging efficiency and spatial distributions. *Marine Biology* 149:595–608.
- May, R.M. 1974. *Stability and Complexity in Model Ecosystems*. Princeton University Press, Princeton, NJ.
- Moore, J.K., S.C. Doney, J.A. Kleypas, D.M. Glover, and I.Y. Fung. 2002a. An intermediate complexity marine ecosystem model for the global domain. *Deep-Sea Research Part II* 49:403–462.
- Moore, J.K., S.C. Doney, D.M. Glover, and I.Y. Fung. 2002b. Iron cycling and nutrient-limitation patterns in surface waters of the World Ocean. *Deep-Sea Research Part II* 49:463–507.
- Moore, J.K., S.C. Doney, and K. Lindsay. 2004. Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. *Global Biogeochemical Cycles* 18:GB4028 doi:10.1029/2004GB002220.
- Odum, H.T. 1983. *Systems Ecology: An Introduction*. John Wiley & Sons, New York, 662 pp.
- Pasciak, W.T., and J. Gavis. 1974. Transport limitation of nutrient uptake in phytoplankton. *Limnology and Oceanography* 19:881–888.
- Patten, B.C. 1993. Toward a more holistic ecology, and science: The contribution of H. T. Odum. *Oecologia* 93:597–602.
- Rand, P.S., J.P. Scandol, and E.E. Walter. 1997. NerkaSim: A research and educational tool of Pacific salmon in a dynamic environment. *Fisheries* 22:6–13.
- Rothstein L.W., J.J. Cullen, M. Abbott, K. Chassignet, E.P. Denman, S.C. Doney, H. Ducklow, K. Fennel, M. Follows, D. Haidvogel, and others. 2006. Modeling ocean ecosystems: The PARADIGM program. *Oceanography* 19:25–51.
- Rusch, D.B., A.L. Halpern, G. Sutton, K.B. Heidelberg, S. Williamson, S. Yooshep, D. Wu, J.A. Eisen, J.M. Hoffman, K. Remington, and others. 2007. The *Sorcerer II* Global Ocean Sampling expedition: Northwest Atlantic through eastern tropical Pacific. *PLoS Biology* 5:doi:10.1371/journal.pbio.0050077.
- Siegel, D.A. 1998. Resource competition in a discrete environment: Why are plankton distributions so paradoxical? *Limnology and Oceanography* 43:1,133–1,146.
- Steele, J. 1974. *The Structure of Marine Ecosystems*. Harvard University Press, Cambridge, MA.
- Venter, J.C., K. Remington, J.F. Heidelberg, A.L. Halpern, D. Rusch, J.A. Eisen, D. Wu, I. Paulsen, K.E. Nelson, W. Nelson, and others. 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304:66–74.
- Wiggert, J.D., A.G.E. Haskell, G.-A. Paffenhofer, E.E. Hofmann, and J.M. Klinck. 2005. The role of feeding behavior in sustaining copepod populations in the tropical ocean. *Journal of Plankton Research* 27:1,013–1,031.
- Yooshep, S., G. Sutton, D.B. Rusch, A.L. Halpern, S.J. Williamson, K. Remington, J.A. Eisen, K.B. Heidelberg, G. Manning, W. Li, and others. 2007. The *Sorcerer II* Global Ocean Sampling expedition: Expanding the universe of protein families. *PLoS Biology* 5:doi:10.1371/journal.pbio.0050016.