On the Limits to Knowledge of Future Marine Biodiversity

Jesse H. Ausubel

Scripps Institution of Oceanography
Center for Marine Biodiversity and Conservation
Conference on
The future of marine biodiversity:
The known, the unknown, and the unknowable
22-25 April 2005 La Jolla CA

Jesse H. Ausubel
Program Director, Alfred P. Sloan Foundation
630 5th Avenue
New York NY 10111
Ausubel@sloan.org

Nearly everyone in this room has attended conferences about woes of the oceans in the last few years. A prominent example is the 2003 conference titled Defying Oceans End, which usefully lifted awareness of the menu of challenges, for example, for conservation of life on seamounts. Repetition has power, like the refrain of a song. Repetition emphasizes and makes things stick. Congressman John Brademas liked to say when he presided over hearings in Washington DC, "Everything has been said, but not everyone has had a chance to say it." But my premise is that we are not here for the next four days for repetition. Rather, we hope to improve our forecasts, their scope, detail, and likelihood or humility. Prediction is after all the true test of science. One of the main strategies for winning improvement is by exploring the limits to our knowledge, that is, by asking what we know and why, what we could quite readily know, and what may be unknowable or very hard to learn.

We tend to fill conferences, not to mention magazines and airwaves, with what we know. We much less often explore and disclose the limits to our knowledge. Few experts like or bother to write *terra incognita* on their maps. Yet, disclosing the limits to our knowledge is often among the most useful of acts. It helps people choose where to explore, and it helps people to hedge their bets.

Let me offer some generic comments and illustrations about the Known, Unknown, and Unknowable and how they might bear on the many disciplines and forms of expertise present in this room.

Scaling and simplification

I will begin with paramount challenges in ecology, oceanography, and meteorology: scaling and simplification. System dynamics involve interactions among processes acting on diverse scales of space, time, and organizational complexity.

To what extent in various fields do we have ways to scale from small to large and back? Can we represent the dynamics of aggregates, for example, in terms of the statistical dynamics of populations of individual agents or units? Simon Levin of Princeton has done fascinating work about scale in ecology, exploring how much knowing about a tree tells about the forest and vice versa. Or a fish and a school. Or a small eddy and a large gyre. One might ask analogous questions about individuals, households, and larger human societies. Does knowing the attitudes about the ocean in detail of a UCSD student help predict the actions of California?

As analysts and modelers, we must hope that not every detail of interaction is important to "know" important behavior. Otherwise, many behaviors, both macroscopic and microscopic, appear unknowable. At the microscopic scale, the multiplicity and complexity of interactions can make detailed knowledge impossible. What do we understand in various fields about how to define, identify, and suppress irrelevant detail?

How confident are we that we have selected the right variables? We love temperature in these days of global warming but dare we omit acidity or even noise? Microbes, though they weigh 90% of ocean biomass, were unrecognized until recently and still tend to be unmentioned, though not at Scripps.

Also, at all scales, unknowable – or at least stochastic - perturbations can hit systems. But can we characterize the perturbations probabilistically? And what can we know about whether systems are sufficiently adaptive to absorb these influences and thus survive in roughly similar form?

At a more theoretical level, what we see or live with may reflect the capricious influence of historical events that cause bifurcations and thus represent but one realization of stochastic processes that admit many possibilities. A much debated example is whether the genetic code is the only code possible in some broad sense. But are some regimes of marine life also ecosystems that just happened because of some bifurcation? If we believe such a situation to be common, then are future ecosystems largely unknowable, especially as humans introduce more perturbations?

The flip side of simplicity is Heiki Lotze's question of complexity, of trying to model several sets of non-uniform interactions at once. To what extent could we be right about the whole even if unsure about the some of the parts?

I hope several speakers and working groups will consider scaling and simplification. I look forward in particular to Art Miller speaking about what knowledge useful for marine biodiversity a feasible, affordable ocean observing system might produce by 2020, and how its development might be staged. And how issues in marine biodiversity should shape the ocean observing system.

Data Limitations

Models without data are empty balloons, and a second, large set of urgent questions concern data limitations. I need only say the word technology to establish the most fundamental point. Humans need prosthetic devices to see, sniff, and feel the deep, the dark, the tiny, the swift, the shape-changing. Or to capture it. Our forecast of technology will determine much of what we might know as well as much of what might survive. Suppose we could, this year, instantaneously and continuously see from one or two ships the fish in the water over 10,000 square kilometers of continental shelf?

Snapshots of the ocean, even great synoptic gulps, hint at a different limit to knowledge. In some fields, including parts of biology and physics, data come from controlled experiment, allowing close matching between theory and experimental results. Experimental design is, of course, subject to a multitude of biases that may limit knowledge. At least as important, in many fields, including oceanography, macroevolution, and many social sciences, controlled experiments are impossible. Facts obtainable represent samples of what we would like to know in ways whose biases themselves may be hard to know.

For marine questions our data are themselves very patchy. This fact turns into nuts-and-bolts analytical limits. For example, how do we avoid in-sample overfitting? Split-sample analysis is helpful, but it is far from a universal solution. A variety of recent work investigates these issues, incorporating bootstrap and other resampling methods, shrinkage estimation, and explicit consideration of the estimation efficiency loss from split-sample analyses. How do we acquire information from unknown structures? These are questions on which people such as Andy Rosenberg and RAM Myers might shed light.

The Ocean Biogeographical Information System offers a powerful instance of the bias of our observations. Figure 1 shows the vertical distribution of a sample of about 4 million spatially referenced data of ocean life. Figure 1 and Figure 2 remind us that while almost all observations come from near the surface, most of the habitat is down below. The diversity, distribution, and abundance of life below a few tens of meters remain largely unknown or extrapolated from very sparse observations.

One reason to initiate the microbes-to-mammals, bottom-to-top, pole-to-pole, shore-to-mid-ocean Census of Marine Life was the high fraction of data about marine life relating to the 200 or so commercially most important seafood species. I would hypothesize 80%. Frustratingly, experts engage in bloody fights about numbers even for the most observed species such as salmon and tuna. Shockingly, in 2002 the US National Oceanic and Atmospheric Administration (NOAA) concluded that existing data and life history information are too sparse to provide useful assessments for more than 60 percent of the 900 or so regulated fish populations. In 2003 the CoML reported about 217,000 species of animals have been identified in the oceans. If records abound for only 200, the list about which Status Unknown or Status Little Known is an honest description very long. And some experts believe 2 million marine animals remain to be identified.

Other kinds of bias matter. Systems may be simply too large or long-lived to observe. It is hard to study the life history of animals that live two hundred or even fifty years.

Rarity also makes difficulties, whether in ecological or financial systems. If a tsunami happens only once every two hundred years, how can we consider its impact or price insurance for it?

Bias of a more obvious nature must also be confronted. For example, knowledge of classical history depends substantially on one man, Herodotus, and we do not know how additional accounts would change understanding of Greece, Persia, Babylon, and Egypt. Strongly socially constructed observations, accidental experiments, and often a spare historical record form much of the known.

With regard to marine life, we seem to know predominantly about what we eat and trade. Much work lies ahead to provide reliable information on the rest, for which we need clever tricks from statistics and modeling. Otherwise, practical limits of cost of sampling will keep us forever ignorant.

Culture also excludes information. Western science, in large part a product of Catholic monasteries, excluded women until very recently. In turn, scholarship largely excluded study of the history or even the health of women. Academia in the USA and most other places now excludes almost everyone who has not taken identical political orders. I survive in academia because I am a registered Democrat who often votes Green.

Surveys of party registration of faculty across many departments in Stanford and Berkeley show that 90% of faculty are like-minded politically (Figure 3). Surveys of voting behavior in social sciences and humanities show that anthropology and sociology have mastered cloning well ahead of the biologists (Figure 4). We should not be surprised that we find it easy to write consensus reports, nor should we be surprised that much of the rest of America and the world rejects or ignores them. Naomi Oreskes survey of climate papers merely shows the genetic poverty of her sample, and her conclusion expresses the bias of our own tribe to perfection and accordingly got published.

Everyone in this room basically thinks alike. If this were 600 years ago, we would have been a gathering of Benedictine monks from dozens of different monasteries, impressed by our pseudo-diversity because some of the monasteries were in Spain, some in Transylvania, and some in Bavaria. How can we recognize our own profound and pervasive biases and access the knowledge of those whose cultures we reject?

We must take care about rejecting the knowledge of those outside our monasteries and about ignoring uncomfortable facts. Karl Marx wrote about the plight of Silesian weavers but never talked to a weaver of any description nor visited Silesia. As far as we know, Marx never set foot in a mill, factory, or mine, nor talked to a peasant or landowner. Some of what he failed to understand was knowable, if only he made the effort to observe, but his belief system, a total structure of thought, like Leviticus, rejected all abominations, as academic ecologists reject Lomborg. I stumbled into the fertilizer industry a few years ago, and got a smelly shock when I learned that the facts *Nature* and *Science* cite over and over about nitrogen use are simply wrong. And the usual, alarming nitrogen projections are about as exaggerated as were the Pentagon estimates of the size of the Soviet economy.

I hope we will be very open in talking about the data we do or do not have, could or could not have, and the information we reject and exclude.

Behavior

Let me briefly mention a final area that limits knowledge and also responds to the limit. I want to remark on behavior.

One of the hardest limits to knowledge is knowing what is in the mind of another. Probably nothing is more powerful than being inside the mind of the enemy. While nations, companies, and faculty engage in spying of various kinds, we outlaw torture and drugging. They might work too well. And even if we permitted torture and drugging, we probably could not predict financial markets, which are the outcome of behaviors that may be fundamentally unknowable. Financial markets are just the most obvious example of the difficulty of anticipating behavioral responses. Major tax reforms rarely have the expected outcomes. In the US the well-intended Clean Air Act of 1970 and its concept of New Source Review perversely caused scores of filthy coal-burning powerplants to operate to this day.

We need to recognize that humans are perverse and criminal, that cheating students, plagiarizing professors, crooked corporations, and corrupt UN bureaucrats skimming money from oil-for-food programs are natural. Human nature includes a snake brain, and individually and collectively we are rarely rational.

I look forward to Richard Sandor and Scott Barrett speaking about the limits to knowledge of economic and political institutions and the danger of the fallacies of rational expectations or rational behavior more broadly.

To offer a somewhat different behavioral example, no demographer has produced a model that successfully predicts fertility. I hope Joe Chamie will show some of the wishful thinking that goes into some of the population projections.

Human health and disease of course also depend entirely on behavior. Trailer parks cause tornadoes, said engineer Norman Augustine. Petting zoos kill children by transmitting infections. Sushi kills, too, but only if we eat raw seafood. In a classic line about the suffering from the Sahel droughts of the early 1970s that Mickey Glantz may remember, “Nature pleads not guilty.” Certainly much fascinating marine microbiology is unknown and knowable, but I look forward to Mercedes Pascual also shedding light on what we know about exposures and how behavior might change them.

Consider also the news of recent weeks dominated by the sex life of Michael Jackson, a Papal funeral at which the Roman Emperor Constantine would have felt entirely comfortable, and the wedding of Charles and Camilla. The Millennium Ecosystem Assessment barely registered. The Communications Working Group needs to discuss how to thrive within real limits of human interests.

Indeed, many industries thrive in the face of irrational behavior and an unknowable future. Maybe fields like environmental management could learn from fields like entertainment. If film producers understood behavior, every film would be a hit, while in practice most flop. But the industry understands it needs to make a large portfolio of films, precisely because 90% fail. Turning to the broadest and most fundamental example, the insurance industry copes with limits to knowledge through probability.

The insurance industry even deals with so-called incomplete contracts. In some cases, a fixed sum is payable if an unknowable event happens, for example, for loss of valuables such as paintings where it would be difficult to verify value if the item are destroyed. A more recent variation on the theme is to tie insurance payout to some objective index correlated with the dollar value of the loss. This is often seen in so-called catastrophe bonds. The index maybe a parametric description of the event, such as the Richter scale reading of an earthquake, or some economic index, such as insurance industry losses triggered by the event.

It is an interesting question how well fixed sum or indexed contracts can cover unknowable events. A key point is whether parties can find contract language to trigger a fixed sum payment when the triggering event was unknowable ex ante. Maybe communities should begin seeking coverage for collapse of marine ecosystems. The point is insurance can cover both known and unknowable events.   Obviously we are here at Scripps because we want to lower the chance of collapse, but if we accept, like Jared Diamond, that collapses happen, buying insurance makes sense, too.

I hope these prefatory comments about limits to knowledge with respect to scaling and simplification (or modeling), data, and behavior, and possible insurance against the limits, have awakened you. Because we are homogeneous, the temptation simply to please each other with popular tunes is high. Finally we may benefit from social solidarity, but our solidarity has protected few tuna so far. I hope the speakers and all attendees will make this a meeting where, by defining candidly the limits to knowledge about future marine biodiversity, diverse actors find more powerful ways to better the chances for marine biodiversity to flourish.

Some clusters of Limits to Knowledge

Scaling and Simplification

SCALING (observations at one scale may not help understand or describe another)

SIMPLIFICATION (parameterization)

HOW VARIABLES ARE SELECTED (system identification)

WHAT FEATURES ARE CONSTANT - not knowing them or having them change

WHAT MATHEMATICAL FORMS ARE USED (dynamical)

COMPUTATIONAL LIMITS

ERRORS IN DESCRIBING INITIAL CONDITIONS (magnified over time)

ERRORS IN FORECASTS OF EXOGENOUS VARIABLES (including perturbations)

EXTREME SYSTEM SENSITIVITY (regime shifts)

Data

QUALITY OF DATA, HETEROGENEITY

EXTRAPOLATIONS BEYOND MEASURED RANGES

INTRINSIC VARIABILITY

PROBLEM OF WORKING BACK FROM SURVIVORS (can't figure out what was there based on what is now there, can't see "the disappeared")

LIMITED REALIZATIONS, SMALL SAMPLES

ROLE OF UNMEASURED VARIABLES (no signals reach us…)

TOO MUCH DATA (information pollution)

CULTURE, TABOOS, INHIBITIONS (prevent knowing or researching or measuring)

CONTRACT KNOWING (only know what customers pay for)

COST or DURATION (system is too big to measure, or time required to catch cycles is too long,

or it simply costs too much even if it could be done in principle)

Behavior

ACCESS TO MENTAL STATES (need to know what is in the minds of others)

ASSUMPTION OF RATIONAL BEHAVIOR (when snake brain controls key human behavior)

The hope:

STATISTICAL DESCRIPTIONS, PROBABILITIES (in lieu of "DETAILED KNOWLEDGE")