Report of the 2nd National Ecosystem Modeling Workshop (NEMoW II)

Bridging the Credibility Gap - Dealing with Uncertainty in Ecosystem Models

J. S. Link, T. F. Ihde, H. M. Townsend, K. E. Osgood, M. J. Schirripa, D. R. Kobayashi, S. Gaichas, J. C. Field, P. S. Levin, K. Y. Aydin, and C. J. Harvey (editors)

U.S. Department of Commerce
National Oceanic and Atmospheric Administration
National Marine Fisheries Service

NOAA Technical Memorandum NMFS-F/SPO-102
January 2010

Report of the 2nd National Ecosystem Modeling Workshop (NEMoW II)

Bridging the Credibility Gap - Dealing with Uncertainty in Ecosystem Models

J. S. Link, T. F. Ihde, H. M. Townsend, K. E. Osgood, M. J. Schirripa, D. R. Kobayashi, S. Gaichas, J. C. Field, P. S. Levin, K. Y. Aydin, and C. J. Harvey (editors)

NOAA Fisheries

NOAA Technical Memorandum NMFS-F/SPO-102
January 2010

U.S. Department of Commerce
Gary Locke, Secretary

National Oceanic and Atmospheric Administration
Dr. Jane Lubchenco
Under Secretary for Oceans and Atmosphere

National Marine Fisheries Service
James W. Balsiger, Acting Assistant Administrator for Fisheries

Suggested citation:

J. S. Link, T. F. Ihde, H. M. Townsend, K. E. Osgood, M. J. Schirripa, D. R. Kobayashi, S. Gaichas, J. C. Field, P. S. Levin, K. Y. Aydin, and C. J. Harvey (editors). 2010. Report of the 2nd National Ecosystem Modeling Workshop (NEMoW II): Bridging the Credibility Gap - Dealing with Uncertainty in Ecosystem Models. U.S. Dep. Commerce, NOAA Tech. Memo. NMFS-F/SPO-102, 72 p.

A copy of this report may be obtained from:

Office of Science & Technology
NOAA, NMFS
1315 East-West Highway, F/ST7
Silver Spring, MD 20910

Or online at:

http://spo.nmfs.noaa.gov/tm/

Executive Summary

The NMFS held a National Ecosystem Modeling Workshop (NEMoW) on August 25-27, 2009. The workshop was held at the Chesapeake Bay Foundation, Merrill Center, in Annapolis, MD. This 2nd NEMoW was held as a national workshop analogous to National Stock Assessment Workshops and National Economists Meetings for the purpose of engaging the ecosystem modeling community within NMFS and how that community can best help the NMFS to meet its goals and obligations. There were 33 participants and 6 observers.

NEMoW II had the following overarching theme: "Bridging the credibility gap." Or, more specifically, how can the NMFS ecosystem modeling community deal with uncertainty in ecosystem models (EM, denoting a broad range of ecosystem modeling)? There was particular focus on the appropriate incorporation of uncertainty into EMs for the provision of living marine resource management advice.

There was a common theme that there are data and information gaps hampering the NMFS EM efforts. The most important information gaps were identified as (i) a lack of trophic ecology data, (ii) a lack of spatially explicit data, (iii) a lack of data for non-target species, and (iv) a lack of socioeconomic data. Other specific areas have been identified by each Center, documenting specific areas that could be explored as a way to remove major sources of EM uncertainty.

There was recognition that there are some common types of modeling uncertainty and some common approaches to address that uncertainty. The main types of uncertainty were noted as (i) estimation, (ii) model, (iii) implementation, and (iv) communication uncertainty. Establishing and refining our list of best practices to address EM uncertainty should be continually re-evaluated. This workshop provided a strong basis for identifying those best practices.

A key conclusion from the workshop was that we need to better engage our stakeholders in terms of communicating, interacting and discussing ecosystem model rationales, uses, applications, and benefits. Several suggestions to that end are noted herein.

This report provides nine recommendations for future National EM efforts in NMFS. Four major recommendations are to: 1) establish distinct EM review panels, 2) identify and note sources of EM uncertainty as a must for EM use, 3) bolster the value of strategic advice, and 4) bolster Ecosystem Modeling Capacity.

The benefits of exchanging best practices and ecosystem modeling experiences among NMFS ecosystem modelers was a subjective, but no less valuable goal and outcome of the workshop.

Given several forthcoming initiatives and copious calls for ecosystem-based management, NEMoW II was quite timely and most attendees thought NEMoWs should persist. The NMFS is in a favorable position as the need to apply EM to key living marine resource issues continues. While the development of expertise and technical capacity is still needed, there exists a reasonably established foundation for NMFS to build upon for future EM efforts.

Acknowledgements

On behalf of the NEMoW II Steering Committee and all NEMoW II attendees, we thank the local hosts of the workshop. Peyton Robertson and his staff at the NOAA Chesapeake Bay Office (NCBO) were extremely helpful in supporting this endeavor. Staff at the Chesapeake Research Consortium, Inc. (CRC) and Chesapeake Bay Foundation provided excellent facilities and support.

We especially thank Howard Townsend for his dedicated efforts at setting up the local arrangements, handling a wide range of workshop logistics, doing a lot of behind the scenes work, and making this meeting run so smoothly.

We acknowledge and thank our external observers. There were four experts outside of NOAA who participated and helped the discussion. We thank Randall Peterman, Nick Bond, Andre Punt, and Thomas Miller for their level of preparation and engagement at this meeting. We also thank Hendrik Tolman from the NOAA Weather Service National Centers for Environmental Prediction and Charlie Stock from the NOAA Research Geophysical Fluid Dynamics Laboratory who provided perspectives from other NOAA line offices and ecosystem modeling approaches.

We thank the NMFS Office of Science and Technology for partial funding support to help defray the costs of this workshop.

Finally, we thank the NMFS Science Board for their enthusiastic support of this workshop.

Table of Contents

Introduction, Context and Background

There has been a recognized need for those scientists within the NMFS who are involved with ecosystem modeling (in its various forms) to routinely gather and discuss best practices, tips, and operational tricks, similar to National Stock Assessment Workshops (NSAWs) and National Economists Meetings. Thus NEMoWs (National Ecosystem Modeling Workshops) in general, and this 2nd NEMoW in particular, were established in large part to address this goal of getting said scientists together in a forum conducive for networking, sharing of ideas, and evaluating how the various regional Science Centers are progressing on their efforts in ecosystem modeling. There is no tangible output from this primary goal, but its importance should not be understated.

The topic of "bridging the credibility gap" was chosen as the theme for NEMoW II for several reasons. As the wording for major enabling legislation for many of the NMFS mandates continues to move towards an EAF (ecosystem approach to fisheries), as NOAA's mission, vision, and policy statements continue to espouse an ecosystem approach, and as many of the national initiatives also note the need for ecosystem-based management, it is clear that ecosystem models (EMs) will be required to help provide management advice for living marine resources (LMRs). With the general terminology EMs, we are covering a full range of models from minimal realistic models (MRM; multispecies and extended stock assessment models (ESAMs)), bulk biomass (network and aggregate) and full system (ecosystem and biophysical) models. For these EMs to provide LMR management advice, their credibility will need to be established and the rigor of quality control/assurance and peer review will need to be at a comparable level as what is done for single species and protected species stock assessments. Thus, one of the keys for EMs to be used in providing ecosystem-based LMR management advice is to ensure that all stakeholders, reviewers, managers and scientists using them have full confidence in what the models are doing in general and that the models have been applied appropriately in specific instances. There have been prior works that established the appropriateness of when to use certain models for particular cases (Plaganyi 2007, FAO 2008, Townsend et al. 2008; Appendix F). There have also been nascent attempts to establish best practices for EM (Plaganyi 2007, FAO 2008, Townsend et al. 2008; Appendix E). Yet how that information gets communicated to stakeholders, reviewers and managers merited some focused attention. Many of the terms of reference (TOR) for NEMoW II were established to elucidate these items.

More specifically, one of the major ways to facilitate credibility in EMs is to adequately and transparently characterize model uncertainty. NEMoW II was focused on descriptions of these items, the processes institutions would establish to implement and address them, and how the determinants of EM uncertainty may be specific to EMs or may be generic modeling uncertainty issues. Thus, NEMoW II explored and examined the range of uncertainties in EMs and how the NMFS ecosystem modeling community can best address those elements of uncertainty such that EMs can provide an appropriate level of LMR management advice.

The workshop format followed a series of keynote presentations with plenary discussions in the morning and breakout groups with plenary discussion of the breakout group reports in the afternoon. This approach fostered a range of interaction formats and allowed for the revisiting of any particular topic from multiple perspectives, building upon the strength of having the NMFS ecosystem modeling community gathered from the different regions. The primary objective was to address the TOR such that we could explore the facets of EM uncertainty and make pragmatic suggestions of how the NMFS could proceed in its EM endeavors by dealing with uncertainty using a suite of "best practices" recommended herein.

Expanded Terms of Reference (TOR)

NEMoW II

Theme: "Bridging the credibility gap" -- or dealing with uncertainty in ecosystem models

Objectives: NMFS will organize and hold a national ecosystem modeling workshop with the following objectives:

1. How to handle uncertainties in ecosystem models (3 types, MRM (multispecies and ESAMs), bulk biomass (network and aggregate) and full system (ecosystem and biophysical) models)
   a. What are the sources of uncertainty
   b. How can they be identified
   c. How can they, or can they, be mitigated
   d. What are the one to two top unknowns in each Center's EM activities?
   e. How uncertain are some of the MS/EM results compared to SS results
2. How to utilize strategic (as opposed to tactical) model output and advice
   a. What models can produce precise BRPs and related outputs that are specific, precise and point estimates (tactical advice)
   b. What is the (expected and reasonable) variability about those estimates from different models classes
   c. What models can not produce precise BRPs, but rather accurate, directional (strategic) outputs
   d. What is the (expected and reasonable) variability about those estimates
3. Development of ecosystem model review venues and protocols
   a. What is the place for strategic advice
   b. What are case studies/examples where this would have helped
4. A re-examination of the ecosystem modeling external review criteria developed in NEMoW I
   a. What are the key "checklists" that external reviewers need for reviewing EMs
   b. Are there appropriate standardized questions and model applications for the broad range of model classes
   c. How can we make these review criteria more widely available to reviewers
   d. How do these review criteria compare to other model (e.g. assessment) review contexts
5. Prepare a report on the above, to be delivered to the NMFS Science Board within six months of the workshop.
   a. Tech Memo
   b. White Paper

Abstracts of Keynote Addresses

Review of major types of uncertainty in fisheries modeling and how to deal with them

Randall M. Peterman
School of Resource and Environmental Management,
Simon Fraser University, Burnaby, British Columbia, Canada, V5A 1S6

To meet the objective of providing ecosystem-level advice to fisheries managers, aquatic ecosystem scientists can build upon the extensive experience in single-species stock assessment and the more general fields of risk assessment, risk communication, and risk management. I will discuss these fields in the context of five major sources of uncertainty faced in both single-species and ecosystem-level assessments. These sources are natural variability, observation error, structural complexity, outcome uncertainty (sometimes called implementation uncertainty), and inadequate communication among scientists, decision makers, and stakeholders. These uncertainties create biological, economic, and social risks, each of which has various magnitudes and probabilities of occurrence. I will give examples of how fisheries scientists have dealt with each major source of uncertainty and how those lessons could be applied in an ecosystem context, regardless of whether the purpose of ecosystem modeling is viewed as increasing understanding, providing broad qualitative strategic advice, or providing more specific quantitative tactical advice.

Both natural variability and observation error have been addressed most recently in fisheries research by two main methods. First is by fitting state-space and errors-in-variables models that explicitly estimate parameters of both process variation and observation error. Second is the method of fitting hierarchical models to reflect spatial covariation among nearby populations arising from common environmental drivers. Both types of estimation methods have been evaluated using “operating models”, which stochastically generate simulated data from known “true” models and parameter conditions, against which the parameter estimates can be compared. In both hierarchical and state-space models, there are many situations in which the magnitude of bias in parameter estimates can be large enough to have important and detrimental management implications. If this is true for simple, single-species models, the chance of complex ecosystem models performing any better at correctly fitting an underlying “true” system seems low.

In the last decade, structural uncertainty about components of an underlying true fisheries system has been recognized as a larger source of uncertainty than natural variation and observation error. Two main approaches have been used to deal with structural uncertainty: choosing a single “best” model or retaining multiple models in analyses. The first approach is reflected by applying informal or formal model selection criteria to identify the “best” model. However, asymmetric loss functions, which pervade environmental management, mean that what is statistically the “best” model based on quadratic (and symmetric) loss functions may often lead to inappropriate management advice. This result will also likely apply to ecosystem models. The most widespread approach now is to retain many alternative models and conduct extensive sensitivity analyses. One way to do this is to analyze each model separately and present results for each to decision makers and stakeholders. Another way is to combine results from the alternative models, with or without weightings. In addition, some models can be discarded if rigorous statistical methods are applied by fitting them to data for a given system. As well, the

statistical performance of the remaining models can be evaluated with “operating models” of assumed hypothetical true systems (as described above) to determine how sensitive the bias and precision of parameter estimates are to changes in structure of the estimation model. Of course, the true, real-world structure is unknown, so such analyses are repeated across a large number of sensitivity analyses that use different hypotheses about that true structure in the operating model. Results often show that certain assessment models are more robust to uncertainties than others. Also, “improved” models sometimes fail to show any benefit over simpler ones, and recently cases have emerged in which outcome uncertainty (the fourth source of uncertainty) has as much effect on results as structural uncertainty. Finally, the most popular and preferred method of dealing with structural uncertainty is to use closed-loop simulations (management strategy evaluations, or MSEs). These models of an entire fishery system help identify the management procedure (data collection system, assessment model, and decision-making process that uses output from the assessment) that is most robust to uncertainties. MSEs include five parts, (1) the natural system plus (2) the simulated “observed” data (together constituting the operating model), (3) the simulated stock assessment, (4) decision making that is based on results from that stock assessment, and (5) harvesting or other human actions, including outcome uncertainty. The 2007 meeting in Tivoli (FAO 2008) and the first NEMoW workshop in 2007 recognized the MSE approach as the preferred method for dealing with multiple ecosystem models. I completely agree with this general conclusion, but I will raise some concerns about the challenges of taking this approach. These concerns include (1) unclear reliability of forecasts from some ecosystem models, (2) non-stationary future environmental conditions, (3) few analyzed data on outcome uncertainty, (4) excessive demands on computer time for simulating both a complex underlying operating model such as Atlantis and a complex ecosystem assessment model, (5) high dimensionality and large amount of output to interpret, and (6) lack of clear ecosystem-based operational objectives, which creates lack of clarity about both indicators for ecosystem scientists to produce and the basis of difficult tradeoff choices by decision makers.

The final source of uncertainty is communication among scientists, decision makers, and stakeholders. Unfortunately, this source is also the one to which the least research attention has been paid, yet it can seriously degrade transmission of high-quality scientific advice. Single-stock as well as ecosystem scientists have developed many creative visualization methods to show multiple variables/locations/times that are output by their models, but few formal “user studies” have been done to help guide future developments in this area. Ecosystem scientists need to collaborate with (1) cognitive psychologists (who have worked for over a decade on how people think and communicate about uncertainties and risks), and (2) computer visualization scientists (where a leading-edge topic is how to better communicate uncertainties in high-dimensional data -- exactly what we face in ecosystem advice to managers). Many lessons can also be learned from how the Intergovernmental Panel on Climate Change communicates its findings from a wide array of complex, highly uncertain models by using a hierarchical information system.

The already-developed best practices for ecosystem modeling should be extended to include a protocol for dealing rigorously with multiple models, numerous sources of uncertainty, and challenges of communicating results.

Generic Levels of EM Uncertainty – Lessons from geophysical modeling and current studies of climate impacts

Nicholas A. Bond
University of Washington / JISAO
(Joint Institute for the Study of the Atmosphere and Ocean)

This paper has two objectives: (1) to review current practices in weather and climate prediction and (2) to summarize some of the author's own findings related to the effects of climate on marine ecosystems. In both cases, the focus is on how models can be used to identify the sources and magnitudes of uncertainty, and how to account for this uncertainty in making projections. An individual model simulation has uncertainty from two sources: initial condition sensitivity and with model formulation (e.g., parameterizations), sometimes termed "structural uncertainty". With regard to the former source, the chaotic nature of non-linear systems must be recognized, which limits fundamentally the interval over which phase changes are predictable. The second source of predictability, the structural uncertainty in models, probably represents a more vexing challenge. This component tends to dominate the uncertainties associated with initial conditions for forecasts with long time horizons. As a means of dealing with structural uncertainty, there is an increasing use of multi-model ensembles. They represent a means for reducing the errors and uncertainties from individual models, especially if they are quasi-random. A complicating factor in the development of model ensembles is that different models have different strengths and weaknesses. In particular, the relative performance of models based on comparisons between hindcast simulations and observations varies substantially with parameter and region. The ambiguity in the evaluation of competing models, and that past performance does not guarantee future skill, may mean there is no clear "best" method for handling model error and uncertainty.

How has Strategy Advice Been Used in a Global LMR Context: A Global Perspective on How to Deal with Ecosystem Model Uncertainty

Andre E. Punt
School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195
CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart, TAS 7001, Australia

A variety of fisheries management jurisdictions have evaluated the performance of management strategies for fisheries and ecosystem management, using what is often referred to as the management strategy evaluation (MSE) approach. A management strategy consists of specifications for monitoring and assessment schemes as well as harvest control rules that take the results of monitoring and assessment and provide recommendations for how tactics should be decided upon. Evaluation of management strategies requires specifications for the goals for management, preferably in the form of quantitative performance criteria, and a set of scenarios, or operating models, which represent the situations (biological, economic, environmental and operationally) to which robust management performance is expected. The operating models aim to explore the four major sources of uncertainty related to achieving fishery management goals: (a) model uncertainty, (b) process uncertainty, (c) parameter uncertainty, and (d) implementation uncertainty.

The bulk of the applications of the MSE approach have been based on single-species operating models and have addressed each of the four sources of uncertainty to varying extents. However, there are an increasing number of applications of this approach which have considered operating models in which productivity is related to climate indices, natural mortality is related to predator numbers and in which an attempt is made to characterize all of the major components of the ecosystem. These "ecosystem" MSEs therefore focus on model uncertainty given that most management strategies are based on single-species considerations only. However, they also lead to increased demands regarding consideration of uncertainty because operating models must be plausible and ideally fit to data for the system under consideration. Furthermore, most "ecosystem" operating models involve many assumptions, thereby making the possible set of alternative operating models needed to fully explore model structure and parameter uncertainty almost computationally impossible.

Five case studies based on pollock in the Gulf of Alaska, the Eastern North Pacific stock of gray whales, anchovy and sardine off southern Africa, and Australia's Southern and Eastern Scalefish and Shark Fishery are examined. These studies are typical of the state-of-the-art application of the MSE approach in an "ecosystem" context. The case studies indicate that, in general, it is possible to apply standard statistical (and MSE) techniques to evaluate model fit and to quantify parameter uncertainty. This is not the case for MSEs based on whole-of-ecosystem models owing primarily to their complexity, although it is possible to identify areas of major uncertainty for such MSEs. Whether MSEs based on whole-of-ecosystem models should be used to evaluate management strategies or only provide directions for broad policy remains unclear.

It is hard to identify cases in which a strategic evaluation has been the only basis for the adoption of a management strategy and only South Africa and a few regional fisheries management organizations (RFMOs) have the ability to formally adopt a management strategy (rather than say a harvest control rule or set of biological reference points). However, there are some common features of MSE which have been used "to guide management decision making": (a) inclusion in the study of the current management strategy (or a reasonable proxy for it), (b) accounting for uncertainty in monitoring and assessment, (c) involvement of stakeholders at all stages of the evaluation process, (d) an attempt to keep management strategies sufficiently simple that stakeholders can understand them (a simple strategy which is almost optimal is always preferable to a very complicated but slightly more optimal strategy), and (e) careful consideration of the scenarios included in the evaluation to ensure that implausible scenarios are not explored.

Tools and Expectations: Bridging the divide from single-species to ecosystem-based approaches to management

Thomas J. Miller
Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, P. O. Box 38, Solomons, MD 20688

An extensive body of knowledge, expertise and experience in the application of single species (SS) assessment models has developed over the last 50 years. This period has seen the development of new statistical approaches to modeling, and an increasing comfort of managers and importantly stakeholders in the application of these approaches. Even with the new powers granted to the Scientific and Statistical Committees at each Fisheries Management Council, the discussion has focused on the specifics of implementation, and not on questioning the concept of reference points more generally. A similar body of experience has yet to develop for an ecosystem approach to management (EAM). Uncertainty remains over how EMs can be used in management, who are the appropriate stakeholders, and how the findings of EMs will be integrated with traditional SS reference points.

There are widely different expectations for EM in management. Some in the management realm have expressed the view that EM results will be of no utility unless they are framed in traditional SS advice. Other stakeholders have viewed EAM as their chance to have a seat at the table. Here I discuss some lessons from a series of case studies involving mid-Atlantic, and north Atlantic ecosystems that sit at the uncomfortable interface between classical single-species management and ecosystem-based approaches.

Management of the Chesapeake Bay has been focused at the ecosystem level since the first compact between State and Federal Governments in 1983. The first ecosystem model was published in 1989 and a fishery-ecosystem plan was published in 2004. Two issues have challenged managers in the region: inclusion of ecosystem services in menhaden management and implementation of the plan recommendations within an existing reference point framework. Efforts to address ecosystem services have included bioenergetics modeling, multispecies VPA, and EwE models. However, managers have struggled with identifying appropriate management goals, incorporating opposing stakeholder views, and deciding how to integrate output from different models. Menhaden is but one example of how EM model advice has presented challenges to the managers in the region. As a result Maryland Sea Grant has been charged with developing a plan to implement EAM for the region. This project involves assessment scientists, oceanographers and managers. Project outcomes have included specific management indicators and goals that bridge between single species and EM results.

Changes in the fish community on Georges Bank has often been cited as an example of a regime change induced by over-exploitation. Recently, Mike Frisk and I have challenged this view, suggested large scale movements from outside the system as a cause. We have used both single species and EM approaches to document patterns and explore the involvement of potential mechanisms. This effort was a retrospective analysis. It is less clear how to apply these approaches in a forward looking analysis.

Science Center and Office Responses to Questions about Ecosystem Modeling Uncertainty

These were collated by each Center Representative into concise, written responses, which are provided below. These were used to develop topical presentations at the meeting.

Pacific Islands Fisheries Science Center (PIFSC)

What are the one to two top unknowns in each Center's EM activities?

Among the top unknowns are a lack of basic data on the abundance and distribution of living marine resources and their forage. Horizontal, vertical, and temporal changes in occupancy and overlap are critical in the Pacific Region. Habitat is either extremely patchily distributed (e.g., insular metapopulations) or consists of large continuous expanses (e.g., wide ranging pelagic stocks), necessitating a wide range of time/space scales of oceanographic study in conjunction with understanding organism movement. Larval and adult connectivity remain largely unknown. The high species diversity in these systems adds difficulty towards identifying important trophic and other interspecific interactions. Another serious unknown at PIFSC is the lack of goals and a plan for integrating PIFSC research to assess ecosystem status. The usual taxonomic structuring and compartmentalizing of center scientists does not facilitate efficient progress towards study of ecosystems or ecosystem modeling.

How is each Center addressing EM Uncertainty: e.g., Sensitivity Analysis, Multi-model inference, Risk Analysis, MSE, etc.?

If uncertainty is addressed, and it is not addressed in all cases, it is through sensitivity analyses or risk analyses.

What are the venues that each Center is using to present and review EM? Are they appropriate/sufficient?

No formal venues exist. Limited partial review may occur through scientific meetings and associated presentations and discussions.

Are SS and EM model outputs being compared/complementary in each Center? If so, how?

No comparisons are currently conducted.

How is each Center engaging Stakeholders, particularly Councils, wrt Strategic Advice and Guidance from EM?

Some presentations are provided to stakeholders but there is no structural process for vetting strategic advice and getting feedback. This relates to the earlier comment regarding lack of a plan. However, PIFSC does enjoy a good working relationship with the Council and its various bodies (e.g., Plan Teams, SSC, Advisory Committees), and works together with them in development of various Fishery Ecosystem Plans (FEPs), a site-based management scheme meant to gradually replace the species-based Fishery Management Plans (FMPs). EM is seen as critically important towards development of these FEPs.

If a Center is not doing EM and evaluating EM uncertainty, why not?

Why is this not being done? There are many reasons, but principally due to limited resources, lack of a goal-oriented plan, lack of collaboration, and inertia.

What are the sources of data availability/accessibility/uncertainty? That is, what is the availability of data that might be used in a Ecosystem model? And how accessible is this data?

Vague question but in general, there is a lack of fishery-independent information except for coral reefs, lobsters, and protected species, and those data are not readily available to others. Fishery dependent data and remotely-sensed ocean characteristics are the most readily available data and these have limited spatial and temporal resolutions. Trophic linkages are understood at a basic level for the main ecosystem components, but clearly there are many data gaps. Pelagic lower trophic level interactions in particular are not well known.

Alaska Fisheries Science Center (AFSC)

What are the one to two top unknowns in each Center's EM activities?

At AFSC, the answers were related to the modeler's specialty, but fell into two general categories where data are lacking: sources of variability in (single species and ecosystem) productivity, and spatial issues. We have good data streams for evaluating fisheries associated mortality (including incidental catch) and to a lesser extent predation mortality, but it is still difficult to evaluate the relative importance of bottom up versus top down effects on ecosystem and single species population dynamics because our information on bottom up effects is weaker. Food web modelers identified low trophic level benthos and pelagics ("forage") as key uncertainties in Alaskan systems and single species modelers identified environmental sources (including transport, larval behavior and mortality) of recruitment variability as top unknowns. Spatial modelers emphasized missing spatial information, including spawning locations, movement and migration, and both benthic and pelagic habitat information (including physical and biotic features).

Two key uncertainties related to modeling included the relationship between fishing mortality and predation/natural mortality and modeling consumption (therefore estimating predation mortality): To what extent is fishing mortality additive onto natural mortality versus being compensatory with natural mortality, i.e., how does fishing affect ecosystem productivity? And, what is the best way to calculate consumption rates (i.e. ration) and do these rates vary seasonally?

Finally, there were two unknowns related to EM activities in general. First, it is not clear at present what the goals of the North Pacific Fishery Management Council are with respect to EM. Second, recent budget decisions have traded the maintenance of broader, fishery independent monitoring programs for grant-specific, focused short term projects. This may hamper future ecosystem-based modeling programs dependent on time series of data from each ecosystem.

How is each Center addressing EM Uncertainty: e.g., Sensitivity Analysis, Multi-model inference, Risk Analysis, MSE, etc.?

AFSC is addressing uncertainty in models intended for management use in a mostly *ad hoc* manner, although there are elements of formal risk analysis built into Allowable Biological Catch recommendations from the most information rich single species stock assessment models.

Single-species assessments of commercial groundfish typically produce estimates of uncertainty for estimated parameters and derived quantities such as stock size. Tier 1 of the groundfish harvest control rules has an explicit risk-based adjustment that increases the buffer between OFL and ABC as a function of statistical uncertainty. This adjustment was derived in the context of decision theory, where an optimal harvest rate is computed by minimizing the expected loss (= maximizing the expected utility) given a specified loss/utility function, where the degree of risk aversion is an attribute of the loss/utility function. Some efforts (presented at the NSAW but as yet unpublished) have been made to extend this sort of approach to community-level models.

In existing AFSC ecosystem models, sensitivity analysis has been carried out through ecoSense routines for static and age structured parameters for EBS and GOA, and for static parameters for AI. The dynamic predator-prey parameter space has been explored for the GOA. Multiple models are in various stages of construction/use/comparison, including single species and ecosim-type where *ad hoc* comparisons have been made, and one model has been developed to switch between single species and multispecies framework (See below). However, we have not yet compared different model structures for the same set of species (e.g., Ecosim-type vs. Atlantis model). An MSE for GOA pollock had difficulty reconciling single species and ecosystem modeling assumptions. That needs a lot more work. In addition, uncertainty due to spatial issues has been addressed in only one ecosystem level model (where smaller scale models were built and compared with the larger scale AI model). In a coupled bio-physical model used for single species recruitment, qualitative sensitivity analyses has demonstrated the importance of correctly characterizing spatiotemporal patterns of spawning as well as larval behavior with respect to water column position. One source of uncertainty that is not being addressed is that sensitivity analyses can be conducted on the biological model, but not on the physical model, due to the way models are run (decoupled) and the cost of running an ensemble of physical models to address physical uncertainty.

What are the venues that each Center is using to present and review EM? Are they appropriate/sufficient?

AFSC EM methods and results are presented in the peer reviewed literature, in NMFS Tech Memos and other technical publication series (e.g., PICES), at scientific conference presentations, and during internal technical workshops. They have also been presented to the Marine Stewardship Council reviewers (as part of pollock certification), and to single species stock assessment reviewers from the Center for Independent Experts (CIE). EM results are presented in the Ecosystem assessment, an annual document for North Pacific Council meetings (including Plan Team, SSC, etc.) The Ecopath/sim type models of the EBS, GOA, and AI were reviewed by the CIE several years ago. Other forms of review include peer reviewed papers, and some review by Plan Team/SSC, but the Plan Team review is not considered very rigorous as many on those bodies are not themselves EM experts. (Relative to the STAR panel review process used on the West Coast, Plan Team review is considerably less rigorous for single species assessments as well. However, we do 20+ assessments every year so substantial reviews are saved for the CIE process).

There is no formal framework for presentation and review of ecosystem modeling or management within the current AFSC system, but continued efforts have “made space” for annual updates and reports on ecosystem modeling efforts. The Aleutian Islands Fishery Ecosystem Plan (AI FEP) team noted that there is no formal time or forum where ecosystem considerations can be taken into account in our current system so information is integrated in an *ad hoc* manner. More independent reviews not just of EM but also of the management decision making framework for EM in the North Pacific would be useful. A review of coordination and collaboration with other agencies/regions with respect to EM would also be helpful.

Are SS and EM model outputs being compared/complementary in each Center? If so, how?

Food web modeling results are included in the Ecosystem Considerations section of many SS stock assessment documents. However, no SS models incorporate this information formally within the assessment model to date. For bio-physical coupled modeling, time series of estimated recruitment from age-structured single species assessment models are mainly taken as representing "truth" and are compared with corresponding biophysical model estimates as part of efforts to validate the biophysical models. The exception to this is the OSCURS model; categorization of annual recruitment based on OSCURS results are used to provide a qualitative prediction of annual recruitment for several winter-spawning flatfish stocks in the eastern Bering Sea. Ecosim type models can integrate results of SS models and determine how compatible they are with each other and alternative data sources within the thermodynamic constraints of the food web. D. Kinsey developed a multispecies version of the single species model used in the Atka mackerel and pollock assessments in the Aleutian Islands which models mackerel, pollock, and cod together. Estimates of quantities of interest from single-species and multispecies are directly comparable in a practical if not a "formal, statistical" sense in that model. T. A'mar developed a pollock model in the GOA which used information from ecosystem models to assess changes in M over time for pollock. Plan Teams consider information from the Ecosystem Considerations chapter and food web model results alongside stock assessments, and sometimes use this information (in an ad-hoc manner) during the decision process for recommending preliminary ABCs.

How is each Center engaging Stakeholders, particularly Councils, wrt Strategic Advice and Guidance from EM?

AFSC presents EM results that can lead to strategic advice and guidance at several levels within the Council process from single species assessments to FMP-level documents. Single species strategic advice includes qualitative predictions of winter-spawning flatfish recruitment from the biophysical OSCURS model in chapters dedicated to individual species stock assessments for several flatfish stocks. These are also included in the Ecosystem Considerations chapter of the annual SAFE report. The entire Ecosystem Considerations chapter, including EM/Assessment is presented to the Council's Plan Teams and SSC annually as part of the groundfish TAC specification process. Though the advice has no formal role in the process at present, it has played into decision making along with stock assessments, (we can provide a recent example for EBS pollock). AFSC also participates in Council Ecosystem Committee meetings at Council staff's request. AFSC personnel with both single species and EM expertise played a major role in the development of the AI FEP (considered a strategic planning document for future management actions in the Aleutian Islands) and the Arctic FMP (a legally binding management plan once approved by the Secretary) at the Council's request. The AI FEP included a formal process for engaging local communities in the Aleutian Islands during the development of the plan, so some stakeholders beyond those normally participating in the Council process were reached in that case.

If a Center is not doing EM and evaluating EM uncertainty, why not?

N/A

What are the sources of data availability/accessibility/uncertainty? That is, what is the availability of data that might be used in a Ecosystem model? And how accessible is this data?

Considerable data are available on the abundance/biomass, production, and food habits of major groundfish. However, it is not always clear how some of this data should be incorporated into ecosystem models, or whether the data collected is adequate to meet spatially explicit modeling. Data on benthic organisms, habitat types, phytoplankton, zooplankton, and forage fish are not regularly collected by fisheries agencies in Alaska and are therefore less available in general. The EcoFOCI larval fish databases constitute the principal data source used to parameterize the biological components of the coupled biophysical models at the AFSC; this database is not easily accessible. Scientific literature probably constitutes the second most important source for coupled models.

Data maintained by NMFS (marine mammal and groundfish surveys, observer data) is available and readily accessible to NMFS employees. Data maintained by IPHC is available and accessible by request. Data maintained by the State of Alaska is generally available but more difficult to access because each state region handles data differently. Data maintained by individual researchers is often not available except in summarized/published form. Overall, the key weakness for modeling Alaskan ecosystems is the lack of a central catalog of databases/data and institutions, which means type of data available and its availability remain to a large extent a function of the modelers ability to find and track and request that data. An overall data catalog including state, academic, other federal institutions, NGOs would greatly cut research time on finding and tracking data for EM efforts.

Northwest Fisheries Science Center (NWFSC)

What are the one to two top unknowns in each Center's EM activities?

1) the role of abiotic drivers;

2) the importance of linkages between a particular system and adjacent systems (linkages such as physical forcing, biogeochemical cycling, species migrations, human activities; adjacent systems might include nearby sub-basins, nearby LMEs, and/or terrestrial systems).

3) Diets – especially how they change over time and space

4) Biomass of non-target fish and inverts

5) Larval behavior / dispersal

6) Habitat-specific vital rates (i.e., growth, mortality, movement, fecundity)

7) A few years ago we might have said that the biggest unknown was specifying the actual management questions, goals and targets that the models were intended to inform, but it seems like we're making some progress there.

How is each Center addressing EM Uncertainty: e.g., Sensitivity Analysis, Multi-model inference, Risk Analysis, MSE, etc.?

The NWFSC uses or is planning to use a variety of methods: formal and informal sensitivity analysis, in which ranges of parameter values are input to "bracket" possible model outputs; scenario-driven analyses in which multiple alternatives are considered; examining the same question with more than one model type (multi-model inference—perhaps most obviously in the West Coast groundfish fishery with stock assessment outputs being compared to Atlantis outputs, and ultimately the Puget Sound modeling work will include EwE, Atlantis, and other developing models); and the Puget Sound modeling effort is wrapping Risk Analysis, EM-driven indicator identification, and MSE into the IEA process.

What are the venues that each Center is using to present and review EM? Are they appropriate/sufficient?

The NWFSC has no formal means of presenting and reviewing EM at the NWFSC. The question –"when is an ecosystem model mature enough to inform the management process? i.e. what criteria must be met for a model to move from research product to decision making?" has not really been addressed at a Center level. As EMs mature and begin to enter the management arena, the EMers we polled seemed to agree that a formal review process will be necessary.

Are SS and EM model outputs being compared/complementary in each Center? If so, how?

Some work has started comparing SS and EM (Atlantis) model outputs, but these efforts are nascent.

The current terms of references for the stock assessments do not include any specific EM outputs. Review of assessments during STAR panels or Council reviews do not require EM derived outputs, though how a species interacts with other species, habitat requirements, and environmental impacts are definitely considerations when setting catch levels.

From one of our stock assessors—good discussion fodder, "It is not clear to me how the results of EM and SS models would formerly be compared (this topic definitely caught my eye on the agenda). What you see with our single species assessments is detailed investigation of sources of uncertainty. If this level of detail were to be done with EMs, the results would be ruled by uncertainty. This is one of the reasons we are not doing much EM right now."

How is each Center engaging Stakeholders, particularly Councils, wrt Strategic Advice and Guidance from EM?

I think we are engaging some stakeholder groups, such as the Puget Sound Partnership (PSP), very closely. The collaboration we've established with PSP and partner organizations like WDFW (Washington Dept. of Fish and Wildlife) in the development of our models hopefully has produced a level of credibility that will facilitate such advice and guidance when our models are on-line.

As far as the Councils go, salmon ecologists have long used ocean and climate indicators to inform management.

Other issues emerging include: what are the 'standard' products needed by the council?; what council subcommittees will use these products?; in talking to people it seems that some of the larger issues with integrating ecosystem model products into assessments are a mismatch in scale, whether it is in space, time step, or the population dynamics. It seems that the space and type issues are easier to deal with that the mismatch between ecosystem models that use something like life stage (i.e. juvenile, sub adult, adult) and age structured models that have much more detailed dynamics.

If a Center is not doing EM and evaluating EM uncertainty, why not?

We are evaluating (or planning to) uncertainty to some degree, but in some cases we are somewhat limited by the complexity of some of our models (e.g. Atlantis)

What are the sources of data availability/accessibility/uncertainty? That is, what is the availability of data that might be used in a Ecosystem model? And how accessible is this data?

Data are marginally available and accessible, and primarily for middle to upper trophic levels. The quality of those data is variable, though, so even where data are available, their usefulness is compromised by poor seasonal and spatial coverage. Data are held by many different agencies and institutions, and tracking all of the data sources down has been dizzying. In some cases agencies (e.g. The National Marine Fisheries Service) are not willing to share data. Mechanistic studies tend to be conducted in areas of low human impact, making extrapolation of some parameters dubious. Lower trophic level data are poor and are hard to find. Habitat data are of mixed quality, and are essentially useless from a mechanistic perspective since we lack habitat-specific parameters.

Southwest Fisheries Science Center (SWFSC)

What are the one to two top unknowns in each Center's EM activities?

The top unknowns include 1) uncertainties about the shapes of functional relationships between predators and prey, including how the "rules" that govern ecosystem structure and function which are defined by such relationships might change in response to climate shifts; and 2) uncertainties that arise from the spatial and temporal coverage of survey effort. The appearance of "new" predators in ecosystems (e.g., jumbo squid in the California Current) provides an example of the first source of uncertainty and challenges efforts to use historical data for understanding ecosystem structure and function. With respect to the second uncertainty, quarterly surveys may make resolving changes in phenology (also resulting from climate change) difficult, and survey areas may be small relative to regional management strategies and the migratory ranges of some key stocks.

How is each Center addressing EM Uncertainty: e.g., Sensitivity Analysis, Multi-model inference, Risk Analysis, MSE, etc.?

The SWFSC addresses EM uncertainty using a variety of approaches. For models of the krill-based ecosystem in the Antarctic, multi-model inference, risk analyses, and MSE are used. Multi-model inference is approached in two ways. First, a "reference set" of parameterizations was developed for the EM developed by SWFSC scientists (and their international collaborators). This reference set brackets uncertainty in key system properties: the rates at which krill are advected through space (two parameterizations with no advection and two with advection as passive drifters) and the relationships between foraging success by krill predators and the proportion of adult predators that breed in any given year (two parameterizations with a hyperstable relationship and two parameterizations with a linear relationship). All four parameterizations in the reference set were tuned to a "calendar" of events that characterize observed or suspected changes in the Antarctic marine ecosystem (e.g., that fur seal populations increased throughout the 1980s and 1990s). Results from each parameterization have been compared and plausibility weights have been used to average across these alternative models.

Multi-model inference has also been accomplished by comparing results from the EM developed by SWFSC scientists with results from a completely separate model developed by scientists from South Africa. Inference at this level is generally limited to identifying those results that are robust to the different modeling approaches and contrasting those results that are sensitive to the different approaches. Management strategy evaluations have been used to evaluate spatial strategies for allocating a region-wide catch limit for Antarctic krill among small areas to manage risks to krill-dependent predators. The MSE work includes use of all available models (those in the reference set of parameterizations developed for the SWFSC model and those for the South African model), and is used to produce risk assessments (e.g., describe the probabilities that krill-dependent predators will be depleted below some threshold or that the krill densities will fall below thresholds which require fishing vessels to change their behavior).

What are the venues that each Center is using to present and review EM? Are they appropriate/sufficient?

EM efforts for the Antarctic are presented to and reviewed by two of the scientific working groups within the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR). One working group (the Working Group on Statistics, Assessments and Modeling) reviews methodological approaches and developments, and another working group (the Working Group on Ecosystem Monitoring and Management) reviews the ecological context represented by each EM. Scientists from the international Members of CCAMLR attend the working groups and provide the review. Reviews are documented as part of each working group's Report, and these reports are adopted by consensus (dissenting views are also recorded in these reports). A further level of review is sometimes provided by the CCAMLR Scientific Committee (to which the working groups report), here again the international Members of CCAMLR provide the review and document it in a consensus report. Results from models that "pass" review have been incorporated into the reports of the working groups for presentation to the Scientific Committee and may ultimately be presented to the Commission. These venues for presentation and review are appropriate because the CCAMLR makes management decisions on the basis of consensus, thus all Members need to be sufficiently informed of and invested in the modeling process. The sufficiency of these venues is debatable. In some cases, the representatives that Members send to working group meetings are not well qualified to review EMs; nevertheless, their participation in the process is paramount. A mechanism does exist for the working groups to have "invited experts" attend meetings and potentially increase the caliber of review. This has been done occasionally (e.g., an invited expert attended a meeting that provided a scoping session for EMs) but does not occur every year.

For the California Current, most venues to present and review EMs are scientific forums (e.g., CalCOFI, AGU, PICES) with some relatively modest exposure to management forums (e.g., the Pacific Fishery Management Council, PFMC) of physical and biological exchanges (e.g., regime shifts, variable salmon production and potential causes of salmon disaster). The PFMC is currently initiating action to move forward with incorporating ecosystem-based fishery management principles through an Ecosystem-based Fishery Management Plan (EFMP). This plan, initially approved in 2006, would not replace existing Fishery Management Plans (FMPs), but rather would serve as an "umbrella" plan for the four existing FMPs to deal comprehensively with issues related to ecosystem health, productivity, trophic interactions and spatial management measures. As such, there is a growing interest in the management arena for greater exposure to EM and discussion of establishing an EM committee by the PFMC.

Are SS and EM model outputs being compared/complementary in each Center? If so, how?

For the Antarctic, EM model outputs are complementary to those for SS models. This is a result of the current management strategy that has been adopted for krill. That management strategy uses SS models to determine a region-wide catch limit for krill, and, after that, EM models are being used to advise on how that catch limit might be subdivided among smaller areas in a way that best manages risks to krill-dependent predators.

For the California Current, very few SS models and EMs are compared or used in a complementary fashion within the SWFSC. There is a great need to coordinate various efforts among divisions and researchers. There is quite a bit of exchange of ideas and strategies (and comparison of outputs) for single species models among the west coast centers.

How is each Center engaging Stakeholders, particularly Councils, wrt Strategic Advice and Guidance from EM?

For the Antarctic, stakeholders are engaged through the working group and committee process that has been described previously. Stakeholders such as industry representatives and NGOs are generally members of national delegations and the views of these stakeholders are incorporated into the views, statements, and positions of the Members to CCAMLR. The US delegation to CCAMLR currently contains an NGO representative whose views are integrated into US positions during bi-annual delegation meetings and during the annual meetings of the Commission and Scientific Committee. Industry representatives have been on the US delegation in the past, but no US vessels are currently fishing in the Southern Ocean. The SWFSC has a lead role in providing advice based on EMs to the CCAMLR; this is generally done by first submitting working papers to the working groups for discussion and review followed by active participation in subsequent review and debate held within the working groups and Scientific Committee.

With respect to the Pacific Fishery Management Council, the SWFSC is mostly engaging EM within the Management Team process for coastal pelagic species. In this context, the NGO community and other stakeholders are pressing the issue of “forage set asides” for predators, and the SWFSC is both conducting economics research and collaborating with partners from Mexico and Canada to address this issue. The relationship between Pacific salmon and ecosystem productivity has also been a high priority with stakeholders, particularly the PFMC, as a consequence of sequential fishery disasters for California salmon fisheries in recent years which seem to be a consequence of short-term changes in ocean conditions coupled with long-term degradation of freshwater habitats and reliance on hatchery operations. Consequently, there has been high level engagement with stakeholders with regard to better understanding of the ecosystem causes and consequences of salmon declines, which relate to changes in physical ocean conditions and the composition and productivity of forage species such as krill, juvenile rockfish and coastal pelagic species (which in turn covary with the productivity of seabirds and other higher trophic level indicators). Long term objectives include the development of predictive ability for forecasting fluctuations in salmon productivity as related to changes in the ecosystem using a mix of both physical and community indices, although these methods are largely statistical.

If a Center is not doing EM and evaluating EM uncertainty, why not?

Not applicable for the Antarctic, but for the California Current lack of funds and an explicit mandate makes EM in general (and EM uncertainty even more so) more of a research and/or academic pursuit, for which there is little time and support among the mandated requirements. Nevertheless, the SWFSC is committed to develop an IEA for the California Current, and this may increase EM in the near future.

What are the sources of data availability/accessibility/uncertainty? That is, what is the availability of data that might be used in a Ecosystem model? And how accessible is this data?

Data that can be used to develop EMs for the Antarctic are generally available from published sources, CCAMLR data reports, and various national Antarctic programs. Accessing available data for developing EMs relevant to the Antarctic has generally not been a problem.

In the case of forage species in the California Current, spatially extensive acoustic data sets do exist but need to be processed so that krill can be separated from other mesopelagic forage organisms such as fishes. Furthermore, CalCOFI data have strong potential and have been used for some assessments of EM processes. There are very sparse food habits data; these are not consistently collected or analyzed among regions, time periods, species or assemblages (however there are some bright spots, such as nearly 30 year time series of CA sea lion food habits).

Southeast Fisheries Science Center (SEFSC)

What are the one to two top unknowns in each Center's EM activities?

Most of the ecosystem modeling efforts in the Southeast are being driven by the FMCs and are utilizing Ecopath with Ecosim. SEFSC staff are working with all three councils to develop these models, and some of the major unknowns include trophic interactions (who is eating who) and spatial variability (i.e., how to deal with this variability in a modeling context). The conclusions of the Ecopath model fit to the Gulf of Mexico and presented by Walters et al. (2008) was "based on inference chains that began with untested assumptions about historical impact of trawling on abundance of benthic predators, moved to further untested assumptions about impact of those predators on juvenile (and adult in case of menhaden) survival rates of several species based on very real diet-composition data and prey-preference assumptions about benthic predators, and linked both these uncertain effects with other assumptions about population dynamics responses of ling-lived species." This seems to spell out that the basic energy pathways and rates of exchange, the very drivers of an Ecopath-like model, are major unknowns.

On a similar note, the availability of sufficient fishery independent data over time and space is very limited as well. This would include food habits/preferences, species interactions, predation rates, etc.

In the case of the Gulf of Mexico, the one aspect that probably accounts for the most uncertainty from a global point of view is lack of data from the southern half of the gulf. The lack of data from the southern half of the Gulf has been partially addressed through data sharing and cooperative research (*i.e.* on pelagic longline fisheries) with the Mexican Instituto Nacional de la Pesca (INP) and Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA). The Large Marine Ecosystem project recently initiated by Mexico in cooperation with the SEFSC also will provide data from the southern Gulf of Mexico.

Potential future cooperation with additional species and fisheries may help mitigate some of the data problems.

From a more pragmatic side, since there is no devoted ecosystem division, it remains uncertain who will take the lead to form and maintain a coherent interdisciplinary team; especially given the wide geographical extent of all the various labs. Furthermore, without a devoted division, uncertainty in funding is always an issue.

Another large unknown is what exactly what questions we are attempting to address with these models. Without a clear, well stated question or set of questions, it is impossible to design a model that will provide a clear answer to that question. What are we managing for?

How is each Center addressing EM Uncertainty: e.g., Sensitivity Analysis, Multi-model inference, Risk Analysis, MSE, etc.?

The SEFSC is not explicitly addressing ecosystem modeling uncertainty, but several built-in approaches are used to address uncertainty in EwE. However, single species assessments are conducted in such a manner as to address overall uncertainty in some parameters (bootstrap techniques, Monte-Carlo simulations, etc.). These parameters often have an "ecosystem effect" embedded in them.

The SEFSC is collaborating with academia is developing hierarchical-Bayesian statistical models to deal with the uncertainty surrounding complex ecological systems

What are the venues that each Center is using to present and review EM? Are they appropriate/sufficient?

Workshops on ecosystem modeling have been held through the FMCs in the South Atlantic (contact: Roger Pugliese), Gulf of Mexico (contact: Carl Walters and Behzad Mahmoudi), and the Caribbean (contact: Ron Hill). But relatively little ecosystem modeling expertise is available through the SEFSC to assist in these efforts. The SEFSC also is partnering with SEMARNAT in Mexico in support of Mexico's Large Marine Ecosystem project in the Bay of Campeche and the Terminos Lagoon. The Inception Workshop for this effort was recently held in Merida, Yucatan (June 24-26, 2009).

Are SS and EM model outputs being compared/complementary in each Center? If so, how?

The results of an EcoSim model in terms of red snapper biomass were compared to the red snapper assessment results. However, comparisons between the two types of models are not a regular exercise. This is mostly due to the lack of regular EM outputs.

How is each Center engaging Stakeholders, particularly Councils, wrt Strategic Advice and Guidance from EM?

The following are recommendations from the Gulf of Mexico Fishery Management Council, Ecosystem SSC:

The Ecosystem SSC felt that it would be useful to convene in a standard type meeting before the next modeling workshop (tentatively scheduled to be held this fall), in part to discuss where the focus of the next workshop should be. The next modeling workshop will likely be the last one financed by the Gulf Council under its ecosystem pilot project funds. Since science centers such as the SEFSC and FWRI will be primarily responsible for developing the tools and methods for conducting integrated ecosystem assessments, and for compiling the data needed for such assessments, the Ecosystem SSC recommends that a request be made to the SEFSC to send representatives to all future Ecosystem SSC meetings and workshops.

Frameworks for incorporating ecosystem evaluations of potential Council actions should be developed and integrated into the Council's procedures. As a start, ecosystem evaluations can be integrated into the existing SEDAR process. This would benefit not only the Gulf Council, but the South Atlantic and Caribbean Councils as well. The Ecosystem SSC recommends that it be convened to review the existing SEDAR process and make recommendations to incorporate ecosystem modeling in order to identify potential unintended consequences.

In order to continue to develop and evaluate ecosystem models, and examine their utility to address fishery management issue, the ecosystem modeling workshops should be continue to be held in future 8 years if financially possible. It was noted that ecosystem evaluations are best suited to examine non-traditional management questions. Examples of issues that might be addressed through an ecosystem approach include ecosystem dynamics that drive fisheries such as freshwater flows into the Gulf, nutrient loading, and climate change

The Ecosystem SSC recommends that the Council identify one or two issues each year for the SSC to examine from an ecosystem modeling perspective.

Incorporation of ecosystem considerations into the fisheries management involves an iterative, collaborative, and developmental process that incorporates an adaptive management approach. An ecosystem approach may help to generate adaptability of management to unforeseen circumstances, and to identify management measures that reduce adaptability. The Ecosystem SSC recommends that it begin the process of developing a conceptual framework for advancing an ecosystem approach, including the identification of appropriate policy instruments and data needs.

If a Center is not doing EM and evaluating EM uncertainty, why not?

Statistical uncertainty in modeling is usually a step beyond the initial stages of the effort. The SEFSC is closer to the start of the process of Ecosystem Modeling than some other Centers. Thus, addressing uncertainty has usually taken the form of parameter manipulation as opposed to carrying forward in a more formal manner all the uncertainty associated with each of the variables in the model.

The major reason the SEFSC is not doing more of this is likely due to the SEFSC not having a devoted division for this with its own funding and set of responsibilities. Rather, it is worked in as and when possible in a rather piecemeal fashion.

Unlike with single species assessments and the MSA, we perceive that there are no formal requirements, deadlines, due dates, or benchmarks to address via an ecosystem model.

*What are the sources of data availability/accessibility/uncertainty***?**

At present, most databases that experience regular updating, maintenance, and updated accessibility are those providing data for stock assessments. These data include commercial landings, length and age compositions, commercial logbook data, as well as observer data. The recreational data include MRFSS data, Texas Parks and Wildlife, Headboat and Charterboat logbooks. There is also an annual groundfish survey data (twice a year) as well as video data. While these databases have been designed to serve the needs of traditional stock assessments, there are no formal reasons that they could not be useful and made available for EMs. To the extent that these data are useful from an EM standpoint, they are available, accessible, and of the nature that uncertainty can be estimated from them.

NOAA Chesapeake Bay Office (NCBO)

What are the one to two top unknowns in each Center's EM activities?

For ecosystem modeling in the Chesapeake Bay, the known unknowns are 1) basic biomass estimates and times series for non-fished but ecologically important species (e.g., jellyfish, bay anchovies) and 2) relationship between fish stock proportion in the Chesapeake and larger Atlantic migratory stocks (e.g., are Chesapeake Menhaden a sub-stock of Atlantic Menhaden, what proportion of the bluefish stock is contained in the Chesapeake).

How is each Center addressing EM Uncertainty: e.g., Sensitivity Analysis, Multi-model inference, Risk Analysis, MSE, etc.?

For the most part we use Monte Carlo simulations to depict 'bands of uncertainty' around simulation estimates. We are working towards developing other models for multi-model inference and implement MSE approaches.

What are the venues that each Center is using to present and review EM? Are they appropriate/sufficient?

Currently we are using local experts to help provide data for EM and review of the model overall. More rigorous review would be beneficial, but we have a very limited pool of folks who could review. We have considered using a Center for Independent Experts review.

Are SS and EM model outputs being compared/complementary in each Center? If so, how?

In many instances where we have inadequate time series data on biomass of some species, we use output from SS/ stock assessment models to tune the EM.

How is each Center engaging Stakeholders, particularly Councils, wrt Strategic Advice and Guidance from EM?

We have staff on the Atlantic States Marine Fisheries Commission Multispecies Technical Committee. We do not provide or maintain the models for that but rather work with their MSVPA-X. We would like to suggest using additional models for a multi-model inference approach, but are having trouble gaining traction. We are presenting our modeling work within the Chesapeake Bay Program structure in 1) Modeling Subcommittee (MSC) and 2) Fisheries Steering Committee (FSC). The MSC is primarily focused on the Bay Programs Water Quality Modeling Suite. They are attempting to include fish in some of their models and we are attempting to work with them to use our modeling efforts to do so. The FSC is a group of state fisheries managers and academic biologists that work together to consult on state fisheries management decisions that affect the bay species. Our office provides scientific support and is working to help the states move towards ecosystem-based management. We have provided some strategic advice to this group. This group is evolving to develop a governance structure that supports interstate ecosystem-based management for bay species.

If a Center is not doing EM and evaluating EM uncertainty, why not?

N/A

What are the sources of data availability/accessibility/uncertainty? That is, what is the availability of data that might be used in a Ecosystem model? And how accessible is this data?

The data are from an assortment of survey and monitoring programs run by various state agencies and academic institutions. The data can become unavailable for a variety of reasons (e.g., concerns about academic PIs over data sharing and being scooped, cuts in funding for monitoring programs). In addition, much of the data comes from suboptimal survey methodologies. For example, a beach seine survey was developed for assessing striped bass recruitment; however, multiple other species are caught in the survey. The data from this for other species is now being used to generate relative biomass indices, although the survey methodology may be inappropriate for other species.

Northeast Fisheries Science Center (NEFSC)

What are the one to two top unknowns in each Center's EM activities?

1. Enhanced Diet matrix for the lower and upper trophic levels.
2. Dynamics (i.e., biomass estimates, vital rates, etc.) of under-determined trophic level groups (i.e. bacteria, new production, phytoplankton to benthos, gelatinous zooplankton, other micro/macronekton, some benthos, mesopelagics).

These novel (combined with extant) data would then need to be in the form of consolidated, coordinated, accessible, and distributed data sets from different sources and available for EM activities.

How is each Center addressing EM Uncertainty: e.g., Sensitivity Analysis, Multi-model inference, Risk Analysis, MSE, etc.?

A full range, including sensitivity analysis, multi-model inference, risk analysis, preliminary MSE-like scenario testing.

What are the venues that each Center is using to present and review EM? Are they appropriate/sufficient?

As part of extant stock assessment processes (e.g. SARC/SAW, TRAC, GARM), on an *ad hoc* basis of peer review, publishing in the literature. For some models (e.g. ESAMs, MRMs) the venues are appropriate, but the audiences are not fully receptive. For other models, especially fuller system models, the venues may not be fully appropriate. Discussions with both Councils (MAFMC, NEFMC) are underway to determine how these outputs can be better presented and utilized.

Are SS and EM model outputs being compared/complementary in each Center? If so, how?

There have been some instances of comparing models within the Center, albeit at initial stages. This has been done to a limited extent for ESAMs, MS models, and Ecopath. More could be done. Comparing MS and System models was done formally in our recent GARM III to compare and contrast sums of SS reference points to aggregate and MS reference points.

There have also been (albeit even fewer) instances of comparing similar activities across multiple Centers, but these have not been fully coordinated and have been on an *ad hoc* or project basis.

How is each Center engaging Stakeholders, particularly Councils, wrt Strategic Advice and Guidance from EM?

We have had a series of casual presentations and discussions, formal scoping sessions, and involvement with Marine Resource Education Partnership (MREP). We have presented some of this to various committees of the councils and commissions, and indirectly have provided such information via the stock assessment process. Ongoing discussions with both Councils (MAFMC, NEFMC), particularly their SSCs, are underway to determine how we can better utilize EM outputs. Certainly more could be done.

If a Center is not doing EM and evaluating EM uncertainty, why not?

N/A, as we are doing EM!

What are the sources of data availability/accessibility/uncertainty? That is, what is the availability of data that might be used in a Ecosystem model? And how accessible is this data?

Generally or relatively speaking, good regarding availability and accessibility. Regarding uncertainty, mostly it is reasonable and traceable. However some historic data is not available, but could be with additional resources. Additionally, some of the data on lower trophic levels is not readily available. Further, although the data is accessible, it would benefit from some form of consolidation and integration.

Summary of Major Sources of Uncertainty

Based upon the above information and prior to the workshop, summary tables or bullet points were compiled that attempted to place all of the information provided into a common format and categorization. A large part of the focus was on data needs for constructing and executing EM.

Table 1 shows the results of data or information unmet needs for each Center as appraised prior to the workshop. Table 2 shows a re-evaluation of those needs, done at the workshop, in light of the categorizations and needs from all the Centers and discussions thereof. It should be noted that these tables represent the perceived needs for constructing and executing ecosystem models relative to the data extant in a given region. It does not represent the prioritization of key processes in an ecosystem, nor the relative importance of any one of these factors ranked against anything other than what data are in hand versus what data are needed. This is regardless of what is known or not known about an ecosystem or what major drivers, processes, or state variables are thought to be prominent in an ecosystem. This is also not necessarily germane for what items should be given top priority for funding in a given Center relative to the other priorities at a Center beyond ecosystem modeling. By having a low ranking (high score), that does not mean that an item is not important or not a high priority; rather it means that there is thought to be sufficient information in that region to address, explore or parameterize that item.

From these data or informational needs summaries, it is clear that there needs to be an expanded trophic ecology program across NMFS. Spatially explicit information (and associated ways to address and model spatio-temporal variability in a wide range of state variables and parameters) also came out as a high priority across all the Centers. It is also clear that enhanced sampling of non-target species, particularly those we are now mandated to assess under the reauthorization of the Magnuson Steven Act, was recognized as a major data and information gap relative to what we know. It was also noted that basic abundance data for many groups of species remains an important concern for many Centers. Other items noted that better incorporation of socio-economic information into ecosystem modeling is warranted.

Again, these were prioritized based upon perceived data or information gaps. It was understood that routine and regular sampling programs should not be sacrificed to obtain this additional level of information. Rather, these results from a NMFS-wide perspective can inform novel programs and initiatives that are germane to and/or use ecosystem modeling activities.

Aside from data and information gaps as forming a major basis of uncertainty, it was recognized that many Centers are addressing EM uncertainty in a variety of ways. Five out of seven Centers use risk analysis in one form or another. These directly explore the probability of negative consequences in a particular situation. Similarly, four out of seven Centers use multi-model inference, at various levels of formality, with one Center planning to move towards that end. It was noted that multiple models giving about the same answer lends confidence. Four out of seven Centers use sensitivity analysis in various forms (although it may be six given the differences in terminology used). These encompassed an evaluation of a suitable range of parameters in way or another, including Monte Carlo simulations, reference set of parameters, and similar methods. This approach for handling uncertainty was less common for full system and bulk biomass models than MRMs. Finally, four out of seven Centers use an MSE approach.

The MSE approach provides a simulation of outcomes and scenarios for various management actions, virtually and before they are enacted in an actual ecosystem. The full range of EMs noted have served as operating models in this context. One additional Center is planning to implement an MSE approach.

Additionally, five out of seven Centers directly compare SS and EM model outputs in some way, although most are done on an informal, *ad hoc* basis. Only two out of seven Centers used complementary SS and EM in a LMR assessment context.

The point of these summarized results is that there is a wide range of uncertainty in EMs being used across NMFS. That uncertainty, as one might expect, is being addressed distinctly across the Centers. But contrary to what one might expect, is being addressed similarly using a combination of four or five standard approaches. In some Centers, arguably the most uncertain aspect of EM is the resources to construct and execute EMs, along with the requisite data to build them. The reason we have emphasized data and information gaps as a major source of EM uncertainty here is to demonstrate that the foundational basis for ecosystem modeling is provided by many of the data that NMFS collects. Where those data are lagging behind other regions, a concerted effort might be considered to help augment existing efforts in those regions.

Table 1. Top two data or informational unmet needs for each Center (anonymously labeled A-G) prior to the workshop. Note some Centers chose to highlight the full range of needs aside from the top two. Most categories are self explanatory; sampling gaps means expanded survey coverage or process-oriented studies to fill in a more rigorous (statistically speaking) sampling design to address existing or needed datasets.

Unknown- Data Need	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Trophic dynamics, linkages, pathways	X		X	X	X	X	X
Role of spatial, temporal variability			X	X	X	X	X
Non-target species ecology	X	X			X		X
Questions, goals, plans			X		X	X	X
Environmental drivers				X	X	X	
Population connectivity		X			X		X
Sampling gaps			X	X			
Shapes of functional relationships				X			
Modeling consumption						X	
Is *F* additive or compensatory?						X	

Table 2. Priority of data and informational unmet needs as ascertained at the workshop. It should be noted that this table represents the perceived needs for constructing and executing ecosystem models relative to the data extant in a given region. It does not represent the prioritization of key processes in an ecosystem, nor the relative importance of any one of these factors ranked against anything other than what data are in hand versus what data are needed. This is regardless of what is known or not known about an ecosystem or what major drivers, processes, or state variables are thought to be prominent in an ecosystem. By having a low ranking, that does not mean that an item is not important or high priority; rather it means that there is thought to be sufficient information in that region to address, explore or parameterize that item. Some Centers ranked these in tiers (with ties), others ranked them sequentially. Most categories are self explanatory; sampling gaps means expanded survey coverage or process-oriented studies to fill in a more rigorous (statistically speaking) sampling design to address existing or needed datasets.

	PIFSC	AFSC	NWFSC	SWFSC	SEFSC	NCBO	NEFSC
Data/Information Need							
Trophic dynamics, linkages, pathways	1	3	4	3	2	5	2
Role of spatial, temporal variability	1	1	2	1	1	3	5
Non-target species ecology	1	4	1	3	3	2	1
Environmental drivers	2	2	7	4	3	4	4
Basic Abundance	1	4	2	4	1	2	6
Sampling gaps	2	3	3	3	1	5	3
Exploitation information	3	4	8	4	8	5	6
Economic/social/cultural information	2	2	1	4	3	1	5

Commentary on TOR

Terms of Reference for National Ecosystem Modeling Workshop II

1. How to handle uncertainties in MRM (multispecies and ESAMs), bulk biomass (network and aggregate) and full system (ecosystem and biophysical) models.

It was noted that EM uncertainty can be generally classified into four or five main types of uncertainty. The presentation by Peterman nicely laid these out, with reiteration by Punt. Here we note that there are four types of uncertainty: estimation (combining Peterman's natural or process and observation uncertainties), model, implementation, and communication uncertainty. Estimation uncertainty comprises observation error (imprecision) of the input data used to build an EM and more generally data input bias or error (or lack thereof), parameter estimation error of model components, and errors (bias, precision, and variance) in the estimates of model outputs. Said another way, there is both natural variability in what is being modeled and variability which is carried through in how that data is modeled. There is also uncertainty and variability in what is imposed by a model, which is model uncertainty. Model uncertainty comprises structural (and hence choice of process being modeled) uncertainty of the model, which functional form to choose, and the uncertainty in dealing with optimizing the tradeoff of complexity vs. realism in model structure. Implementation uncertainty is also known as outcome uncertainty, where the process using the EM outputs to provide LMR management advice is not followed for a variety of reasons. Here we characterize those as outcome uncertainty (i.e., deviating from the management targets or goals by any series of performance measures for any set of reasons) and objective uncertainty (i.e., inability to clearly articulate what management targets or objectives are). Finally, we characterize communication uncertainty as misinterpretation uncertainty, where the parties involved have fundamental miscommunications about the entire modeling process, or as objective uncertainty, where scientists, managers and stakeholders have a misunderstanding about how to model and present options relative to the management objectives.

It was noted that to best address these types of uncertainty, simply identifying and characterizing them is a useful—and even requisite—first step. Many of these types of uncertainty are generic for any natural resource management modeling endeavor, so the criticism that these are solely and singularly germane to EMs is inaccurate.

It was also noted that despite the wide range of combinations in which these uncertainties are expressed across the plethora of NMFS EM activities, the way these uncertainties can be addressed generally fell into four or five main categories. Presentations by Peterman, Punt and Bond provided useful examples of these cases. These main approaches to address uncertainty include: risk analysis, multimodel inference, sensitivity analysis, MSE operating models for scenario testing, and visualization. Many of these have been elaborated upon in best practices lists elsewhere (FAO 2008, Townsend et al. 2008; c.f. Appendix E).

It was agreed that, as appropriate for type of model and type of uncertainty, a set of common approaches to best address EM uncertainty should be employed. Table 3 shows that most of the ways to address uncertainty are generally applicable to most model types. The notable exceptions are that many of the sensitivity analysis and possibly risk analysis approaches are not readily feasible for some of the full system class of models, largely due to the copious number of parameters in those types of models, even though there have been attempts to do so (e.g., Pantus 2007, McElhaney et al. in press). Table 4 shows that the type of uncertainty may be best addressed by different methods and that not all ways to address uncertainty in EM are universally appropriate. More so, some methods are more appropriate or were designed to address particular types of EM uncertainty. Again, consulting this table (4, as well as Appendix E) should result in a set of best practices for addressing EM uncertainty.

Table 3. Appropriateness for different methods to address model uncertainty compared to the different types of EMs. x = yes; ? = more difficult, or unlikely; ?? = less relevant.

	Class of Model									
	MRMs			Bulk Biomass				Full System		
	ESAMs	MS	Habitat	Agg	Food Web	Network	Biogeochemical	Biophysical	Bioeconomic	End-to-End
Method to address uncertainty										
Multi-model inference	x	x	x	x	x	x	x	x	x	x
Sensitivity Analysis	x	x	x	x	x	x	?	some	some	??
Risk Analysis	x	x	x	x	?	?	?	?	?	??
MSE	x	x	x	x	x	x	x	x	x	x
Visualization	x	x	x	x	x	x	x	x	x	x

Table 4. Relationship between different methods of addressing model uncertainty and different types of model uncertainty. X = emphatically yes, x = yes, ? = questionable or unlikely, o = no, N/A = not applicable or not relevant.

	Type of EM Uncertainty										
	Estimation				Model			Implementation		Communication	
	Parameter	Output	Obs. Error	Data	Structural	Functional form	Complexity Tradeoff	Outcome	Objectives	Misinterpretation	Objectives
Method to address uncertainty											
Multi-model inference	?	?	?	?	x	x	x	N/A	o	N/A	o
Sensitivity Analysis	x	x	x	o	o	x	x	N/A	N/A	N/A	N/A
Risk Analysis	x	x	x	x	x	x	x	X	x	X	x
MSE	x	x	x	?	x	x	X	X	x	X	x
Visualization	o	x	o	X	o	o	x	o	x	X	x

2. How to utilize strategic (as opposed to tactical) model output and advice.

There was a clear sense that single species models are designed to produce tactical LMR management advice, often in the form of biological reference points (BRPs). These are usually provided for a fishery management council context. It was recognized that MRMs, particularly ESAMs, can replicate those types of BRPs and could be used accordingly. However, the bulk biomass and full system classes of EM were not necessarily designed, nor should they be used, to produce tactical advice. Rather, those types of models are often best suited for providing strategic advice.

Comparing SS models with EM outputs was done at five out of the seven Centers. Yet these comparisons spanned the range of EM classes. Further, only two out of seven Centers noted using EM as complementary to SS approaches in a LMR management context. Those EMs that can produce tactical advice (e.g., MRMs) generally had more conservative advice than SS models without them. Yet this topic has not been fully explored and merits further examination, particularly regarding how EM strategic output complements SS model tactical output.

This NEMoW had copious discussion as to what constitutes strategic advice and how that advice can be used. Definitions of strategic advice ranged from: providing a range of management strategies, as part of the MSE process; to providing a range of robust model outputs that were directional or planar (as opposed to point), also as part of the MSE process; to providing broader context for more specific tactical advice; to simply stating that a particular action, when modeled and as expected to occur outside of the virtual world, will result in something that is bad (or good or no change as the case may be). Presentations by Miller and Punt explored this concept in some detail. Most participants were comfortable with a somewhat ambiguous treatment of the term, as long as it was understood that strategic advice: 1) is not tactical in nature (i.e., not setting this year's quotas), 2) explicitly addresses tradeoffs among biota and management objectives, and 3) has value in terms of contextualizing a particular situation into a broader, ecosystem perspective.

Definitions aside, how strategic advice can be used was a topic this NEMoW wrestled with. Again, the three considerations listed above are probably the most helpful ways to frame strategic advice for the foreseeable future. Providing context and comment on what options are robust are items that need to be more fully utilized.

It was also noted that *de facto* four out of seven Centers are already providing strategic advice in one form or another. Those situations were not solely focused on SS fishery management council (FMC) or protected species issues. Rather, they were instances where multiple ocean-use tradeoffs were being explored. Those situations are explicitly using EM outputs in a strategic manner, as those particular processes are not amenable to specific, tactical BRPs. As calls for EBM continue, it is likely that those fora for multiple ocean-use are going to increase, and strategic advice may be better suited for such venues.

3. Development of ecosystem model review venues and protocols.

There was a sense that EM review, particularly for bulk biomass and full system models, were not most appropriately treated in a stock assessment review process. Conversely, MRMs, especially ESAMs and MS models, have been appropriately reviewed in the stock assessment review process context. Several of the Centers had examples, and the presentation by Miller also described instances where some MRMs were incorporated into the stock assessment review process and how that helped provide more robust and improved LMR advice. Even bulk biomass models have provided strategic context in a couple of instances of a SS assessment review context (largely supporting a FMC process) that were found to be useful. Yet all the Centers do this on an *ad hoc* basis, with no formal EM review process extant.

The summary observations for this review were twofold: 1) that it is a significant (~5 years) process for the outputs of EM to be incorporated and taken up into existing review venues, and 2) we simply need to start with the venues and processes that are established. It was recognized that it would be more preferable to establish a distinct review venue for bulk biomass and full system EMs. Yet in the near-term, lack of a clear review venue should not preclude NMFS ecosystem modelers from incorporating EMs in extant review processes, as long as clear review terms of reference are noted. Similarly, the NMFS should more strongly consider adopting an MSE or similar approach for the implementation and evaluation of LMR management advice.

The Centers ranged from little engagement with council/stakeholders to two that have a significant amount of formal interaction with said stakeholders. Yet almost universally there is not currently a formal mechanism for uptake of EM-derived advice into management. Integrated ecosystem assessments (IEAs) and fisheries ecosystem plans (FEPs) appear to provide a future mechanism for more formal engagement and use of EM outputs. Those would help by establishing processes, protocols and venues for not only the uptake of EM outputs, but the reviews thereof as well.

Related to this process was the topic of engaging stakeholders. It was also noted that it is a fine line in telling our stakeholders what they want us to give them, yet conversely in many respects we have not done a thorough or adequate job of communicating to our stakeholders the value and benefits of EMs. It was noted that statements to the effect of, “Here’s something you need to know that you don’t know yet” would help in the uptake of EM output, particularly strategic advice. As noted above, establishing transparent, dedicated review venues for EM outputs will help. Establishing a similar process (e.g., IEAs, MSE, etc.) for the EM outputs to be utilized will also be important.

As we communicate to stakeholders, and as per the theme of this workshop, we need to accurately and transparently convey the uncertainty associated with these EMs. A way to facilitate this is to adopt common terminology when discussing EM output. Examples from the climate change community are instructive to that end (Table 5, 6). The terminology uses an uncertainty basis for describing model output confidence (Table 5) or model output likelihood (Table 6). If we present our EM outputs, results and management advice in degrees of uncertainty, our stakeholders will be able to make decisions with greater knowledge of the risks associated with their choices.

It was noted that at the least, NMFS ecosystem modelers need to identify and characterize EM uncertainty as part of model review processes. Doing so should also be required for model presentation and description. The particular items to be examined in a review process are discussed in the next section (c.f. Appendices E, F).

Table 5. The standard terms used to define levels of confidence, adapted from the IPCC report, given in the IPCC Uncertainty Guidance Note, for Climate model outputs. Adapted from Le Treut, H., R. Somerville, U. Cubasch, Y. Ding, C. Mauritzen, A. Mokssit, T. Peterson and M. Prather, 2007: Historical Overview of Climate Change. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Confidence Terminology	**Degree of confidence in being correct**
Very high confidence	At least 9 out of 10 chance
High confidence	About 8 out of 10 chance
Medium confidence	About 5 out of 10 chance
Low confidence	About 2 out of 10 chance
Very low confidence	Less than 1 out of 10 chance

Table 6. The standard terms used in the IPCC report to define the likelihood of a modeled outcome or result where this can be estimated probabilistically. Adapted from Le Treut, H., R. Somerville, U. Cubasch, Y. Ding, C. Mauritzen, A. Mokssit, T. Peterson and M. Prather, 2007: Historical Overview of Climate Change. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Likelihood Terminology	**Likelihood of the occurrence/ outcome**
Virtually certain	> 99% probability
Extremely likely	> 95% probability
Very likely	> 90% probability
Likely	> 66% probability
More likely than not	> 50% probability
About as likely as not	33 to 66% probability
Unlikely	< 33% probability
Very unlikely	< 10% probability
Extremely unlikely	< 5% probability
Exceptionally unlikely	< 1% probability

4. A re-examination of the ecosystem modeling external review criteria developed in NEMoW I.

It was recognized that guidance from the FAO Tivoli meeting report (FAO 2008) and the NEMoW I report (Townsend et al. 2008) was useful. Those materials have been reproduced in Appendices E and F, respectively.

Arguably the most important principle that emerged from the discussions at this NEMoW was the need to ensure that the ecosystem model type matches the modeling objectives (and ways to approach a particular question, issue, or problem; see Table F.3).

It was noted that in terms of dealing with EM uncertainty as EMs are undergoing review, there are a set of best practices that need to be followed. Some of these have been noted before (see Appendix E). An additional, albeit recognized as incomplete, list was developed at this NEMoW (Table 7). This list is a useful start and, in combination with the major ways in which uncertainty can be addressed (sensu Table 3, 4), could serve as a useful set of questions for either the initiation of a modeling endeavor or a review of any particular EM application. Further, coupled with prior guidance (e.g. Table F.2), a "checklist" of items was developed at this NEMoW (Table 8). We recommend that both a best practices approach be followed for EMs, and that EM review follow a set of national standards (c.f., Appendix F) and review criteria similar to what has been developed here (Table 8). It may be that additional workshops are warranted to further augment or refine these criteria. Yet as an initial statement of needs, they serve a useful role and in principle likely represent the major themes required for an EM review in a LMR management context.

Table 7. Best practices to generally implement EM applications and to particularly address EM uncertainty would include (but are not limited to):

In general:

1. Don't avoid addressing EM uncertainty just because it seems difficult—remember, this is an emerging field that needs to accumulate wisdom

2. Considering empirical data as a basis for comparison and model improvement (be it by rigorous statistics, or scenario-driven "bounding" of problems, or by tweaking, etc.)

3. Not assuming symmetrical loss functions

4. Being able to identify the cost-benefit tradeoffs of additional model runs

5. Considering/being open to the possibility that your model is either too complex or too simple

6. Complementarity between your modeling effort and ongoing monitoring

7. Stakeholder involvement

8. Careful consideration of the underlying goals and objectives

9. Use multiple models

10. Model nesting: software development that allows you to pull things in and out so that the model's versatility is greater and that more processes or dynamics can be readily tested

11, Practice good quality control, especially wary of propagating questionable parameter values.

In regards to communication:

12. Periodic review (informal: colleagues, stakeholders, managers) throughout model building, to avoid rejection of model at a late stage when problems (model structure, mismatch of objectives, etc.) could be caught earlier.

13. Apply iterative communication with managers; get them used to seeing EM results, start to build credibility

14. Formalize language concerning model uncertainty (follow climate modelers example: "we are highly confident that there is a high chance that..." as per Tables 5, 6 above)

15. Apply most effective communication when presenting results to stakeholders, managers and public; look to cognitive psychologists for best methods of conveying different types of information (e.g., 2 of 10 rather than 0.2)

Table 8. EM performance measures, as probable items to consider as part of an EM review "checklist".

1. Is the model peer reviewed?
2. If not, is it suitably documented for a review committee to evaluate its structure?
3. Is there a statistical fitting criterion?
4. Has the model converged to a solution?
5. Have you tested for multiple initial conditions? What is the number of initial states you've tested?
6. Does it characterize the uncertainty of outputs?
7. Does it hindcast effectively?
8. Is it fitted to data, or to other model output? If the former, okay; if the latter, further explanation is warranted.
9. Which model performance standard did you use? (AIC, "best" management outcome, etc.)
    a. How does it handle extreme conditions or extreme system dynamics?
    b. Asymmetric loss functions?
    c. How does it respond to future perturbations or disturbances?
    d. How does its biomass output compare to biomass output from SS models?
10. Have you selected a model that considers the processes relevant to the management question?
11. Do you have the important functional groups included?
12. Does the model output include indicators that are useful to managers and decision makers? (and/or to test the indicator performance)

Workshop Observations

Randall M. Peterman
Comments on the NEMoW II workshop

Many scientists around the world, including those in NMFS, are developing ecosystem models to help implement ecosystem-based management of marine systems. This workshop of experienced NMFS ecosystem modelers and non-NMFS scientists provided a useful forum in which ideas were actively exchanged. As an outside observer, I was impressed in two ways by this group of scientists and their research. First, there are already numerous initiatives under way to develop and apply ecosystem modeling at most NMFS Centers, and NMFS scientists are using leading-edge, state-of-the-art ecosystem modeling techniques. Second, dedication to high-quality work pervaded the discussions; not only were there presentations of successful case examples, but there were also open discussions of current technical and logistical problems, as well as potential solutions to them.

Several constructive case studies were presented. One, for the Bering Sea pollock fishery, illustrated that ecosystem-based advice could be brought into decision-making criteria informally through taking advantage of opportunities created by unexpected events and linking several sources of evidence about changes occurring in the ecosystem. The case example of king mackerel in Florida demonstrated how successful ecosystem-based management could be developed through multi-year collaborations among scientists, recreational users, and decision makers. Such cooperation is not the norm in all NMFS regions, but based on experience internationally as well as in other NMFS region, it was clear that such collaborations will increase both the rate of implementation and success of ecosystem-based approaches to management. Furthermore, it became clear during the workshop that social and economic processes as well as indicators should be included in ecosystem models. NMFS scientists should also carefully study the successful Australian examples of developing and implementing ecosystem modeling. Finally, the presentation on analyses of climatic models was very useful because those models share similar problems to analyses of ecosystem models -- system complexity, evaluation of multiple models, and high-dimensionality of sensitivity analyses.

The format of the workshop, with interspersion of presentations from non-NMFS people with break-out group sessions, worked reasonably well. The break-out groups differed in their success, though, depending on the particular mix of people and degree of adherence to the suggested questions. Although flexibility in how the groups worked was necessary to avoid stifling creative thinking, in the future such groups should probably be required to address at least some minimum set of questions as well as having time for more open-ended discussions. Similarly, future work on ecosystem models within NFMS might have benefited from a clearer goal of producing specific outputs from the workshop, or at least plans for achieving them. Perhaps the Steering Committee plans to do just this through the synthesis report of the workshop, but that task would have been made easier if workshop participants had focused on, for example, modifying and formalizing specific parts of the "best practices" for ecosystem modeling that have already been drafted (FAO 2008; Townsend et al. 2008). For instance, discussions could have produced a list of specific next steps (with teams of NMFS scientists tasked with pursuing them) to modify best practices for (a) comparison and evaluation of multiple models for a particular ecosystem and across systems, (b) communication among scientists, stakeholders, and decision makers, and (c) general methods for involving stakeholders. Another suggestion for future NEMoW workshops is that participants should read previous relevant reports. At this workshop, it became apparent that several people were not intimately familiar with the contents of the final report from the first NEMoW workshop (Townsend et al. 2008) or the related documents by Plaganyi (2007) and FAO (2008). Because of this, there were some unnecessarily redundant discussions of methods of peer review, best practices, and objectives for ecosystem models. However, in the latter case, a very useful distinction was made between scientific research objectives and management objectives. The former are well understood and well articulated by researchers, but management objectives are often not clear. Several discussions during the workshop therefore highlighted the need to engage decision makers and stakeholders in early and iterative discussions with ecosystem scientists about management objectives. Such discussions will identify appropriate indicators (and uncertainties in them) that need to be produced by ecosystem models. Notably, in Puget Sound, Washington, there is currently a formal structure for engaging stakeholders in development of indicators.

Now is a key time for blending ecosystem concepts into decision making. NMFS scientists should take the initiative to make the case for doing so through, for instance, Scientific and Statistical Committees (SSCs). Such a dialogue is currently under way in New England and perhaps other NMFS regions. Under the Magnuson-Stevens Reauthorization Act, members of SSCs apparently have much more say now than in the past about management targets and limit reference points.

More generally, workshops should be developed throughout NMFS to help educate stakeholders and decision makers about ecosystem concepts and how to understand and effectively use information produced by ecosystem models. As well, although NMFS ecosystem scientists already use a variety of creative methods to show other scientists, decision makers, and stakeholders the complex, multi-indicator outputs from models, those methods could be improved further by working with communications specialists, visualization experts, and cognitive psychologists (Anderson 1998, 2001). Such developments should include formal user studies to determine which communication methods are most effective.

NMFS scientists are well aware of both strengths of ecosystem models and their limitations and challenges, in particular, those arising from the five classes of uncertainty described in my talk. Numerous workshop discussions focused on dealing with these challenges, particularly on statistical and other technical issues. Several ideas emerged that would expand upon the "best practices" for ecosystem modeling. These will help stimulate refinement of a rigorous process for addressing current difficulties with ecosystem modeling and determining the extent to which those problems can be mitigated. An important point at the workshop in this context was that ecosystem modeling is still an evolving technique and the resulting advice will always be imperfect. However, this imperfection should not stop development of ecosystem models; they should be judged in comparison with the next best alternative method for providing ecosystem advice to decision makers, not in comparison with some unachievable ideal of perfect advice. Furthermore, as long as ecosystem scientists are committed to continuous improvement, models will continue to evolve.

Nick Bond
Observations and Impressions of the NEMoW II Workshop from the Perspective of a Weather/Climate/Ocean Professional

As evidenced by the material presented at the workshop, and the associated discussions, modeling marine ecosystems represents an enormous challenge. Perhaps some of the current struggles here are similar to those encountered in weather and climate modeling over the past couple of decades, and hence the experiences and lessons from the latter endeavors are relevant to the topic at hand. For example, it was more-or-less accepted that models should be developed and employed well before their output represented realistic and reliable reproductions of atmospheric and oceanic conditions. The key point here is that important lessons can be learned before all aspects of the model are necessarily on a firm foundation. There is also precedent in geophysical modeling with regards to varying levels of sophistication in the models being used. A very simple one-layer (barotropic) atmospheric model was run operationally into the 1980s long after more complicated models incorporating more of the relevant physics were developed and also run operationally. The benefit of the simpler model was that its behavior was well-understood, and useful insight was often gained into the important physics of the particular situation by comparing the simulations between the simple and more complicated models. Models of varying type and complexity are used in geophysical applications to this day, with the models designed to forecast El Nino/Southern Oscillation representing a good example. To the author's knowledge, the relative skill of these models, in general, still depends very little on their sophistication, amount of data required for initialization, and computational requirements. The climate community as a whole, and the various modeling groups in particular, benefit from the availability of these simulations in quasi-real time. For all kinds of geophysical applications, the benefits of multi-model ensembles are becoming increasingly appreciated, as a means of accounting for errors in initial and boundary conditions, and in model structural uncertainties.

It appears to be an open question regarding what is presently most limiting ecosystem models. Perhaps these limitations can be grouped into three principal categories: understanding of functional relationships between model variables, information on tunable parameters given particular functional forms, and data for initial and boundary conditions. It would seem worthwhile to design tests of various types of ecosystem models to determine where improvements could yield the most "bang for the buck". By way of comparison, recent advances in short-term weather prediction appear to be most efficiently achieved through better assimilation of data for initialization purposes, while for climate models there is more benefit gained in improving parameterization of sub-grid scale processes that cannot be directly modeled. With regards to the data needs for existing ecosystem models, it is unclear (at least to this author) whether current progress is more hampered by the lack of process-study type information for model development and tuning, or by monitoring data for model validation and evaluation. Presumably both needs grow with the complexity of the model.

It seems sensible to continue a multi-pronged approach to ecosystem modeling. Fully-blown, highly complex models arguably are better suited for exploring the sensitivities and emergent properties inherent to a particular ecosystem. In other words, perhaps they are best used now for strategic purposes and are less reliable for specific predictions. Simpler models, such as single-species stock assessment models, may be better suited for tactical purposes and more practical for fully evaluating uncertainties. Again, with numerical weather and climate prediction as a prior example, the ecosystem modeling endeavor should gain from having different groups carry out simulations with a variety of models, and making this output easily available for independent analysis.

André E Punt

Thoughts and comments on NEMoW II

To date, ecosystem models (in their broadest sense) have been primarily academic exercises (although there are some noteworthy exceptions). However, many ecosystem modeling frameworks are now developed to the point where application for tactical or strategic purposes is feasible. The objectives of NEMoW II focused on some of the aspects of how the transition from a primarily scientific enterprise to a key component of the management advisory process, in particular how to handle uncertainties in ecosystem models and review criteria, should occur.

The outcomes from the meeting should provide impetus for moving ecosystem models into fora where their results could directly impact decision making. However, it is clear that substantial efforts will be needed to develop (and the road test) review criteria. Such criteria have been developed over decades for single-species assessment modeling frameworks, and it should not be expected that it will be possible to get such criteria right [*sic* correct] immediately. Development of review criteria should be thought of an iterative exercise and these criteria should start with those used for reviewing single-species assessments. In that respect, one particular challenge will be to identify the level of detail for reporting of data sources (and the preprocessing conducted to format the data for use in the ecosystem model) and model fit diagnostics needed when developing “ecosystem assessment reports”. I mention this because the average assessment document for a single-species assessment at the PFMC is now about 200-300 pages, excluding appendices; I would expect that an ecosystem assessment report to be no shorter and possibly much longer. Getting the level of detail in assessment reports right is, however, critical to effective review. There may therefore be value in conducting workshops which develop guidelines for, for example, analyzing diet data for inclusion in ecosystem models.

Development of national review criteria for ecosystem assessment models will not be a trivial task and appropriate resources will needed to be found for this. However, the cost of not doing so will be much less rigor when reviewing ecosystem assessments. Review of ecosystem assessment models, particularly new models, should be expected to take considerable time and this needs to be acknowledged and budgeted for. Finally, there is a now considerable expertise in reviewing single-species stock assessments, particularly where the stock assessment model is “standard” (e.g. ADAPT, XSA, Stock Synthesis). This expertise will need to be developed for the review of ecosystem models.

The amount of work needed to develop the objectives which are to be addressed by ecosystem models should not be underestimated as this will potentially involve discussions with multiple stakeholders (and not just, for example, the Councils). However, *ad hoc* development of BRPs without appropriate stakeholder input risks scientists implicitly defining the management objectives. I strongly support the idea of a future NEMoW focusing on how to determine (and justify) BRPs given stakeholder input on objectives.

While ecosystem models are typically more complicated (and hence arguably realistic) than single-species models, they also make more assumptions and rely on more data than single-species models. If ecosystem models are to be used to provide tactical management advice, simulation studies need be conducted to show that improved performance is possible. Such studies could start with the case in which the ecosystem model being used to provide management advice is actually correct (but naturally the data used to parameterize it are subject to reasonable levels of uncertainty) because if the ecosystem model does not perform better than the single-species equivalent when it is correct, there is little chance that it will perform better when its assumptions are violated.

In closing, I found NEMoW II to be a very organized (and provisioned) meeting. I learnt a considerable amount on where ecosystem modeling is in the US. One aspect which might have led to more progress on the main topics would have been background documents. Most participants flew for several hours to attend the meeting and a small number of documents (no more than say 200-300 pages) could have been distributed before the meeting. Such documents could have included examples of terms of reference for single-species stock assessments because I got the impression that many of the participants at the meeting were not familiar with how stock assessments are reviewed (e.g. in the Council forum) and, the reviewer process differs among Councils.

Thomas Miller
Reflections on NEMoW II

Fisheries ecosystem research and modeling within NOAA has clearly made substantial strides forward since the first NEMoW in 2007. Strategic investments in sample collection and in personnel have paid dividends. Ecosystem modeling approaches have been refined and the understanding of the capabilities of these models has matured. Ecosystem models are now being applied in several regions within NMFS to inform the agency and the relevant Councils of the broader context within which they make management decisions. These advances are to be lauded. However, as always, workshops are designed to identify lacunae in our knowledge more than they are to catalog achievements. What then were the challenges limiting the broader application of ecosystem models in the management context? Several became apparent during this workshop.

There appears to be an increasing expectation among some that ecosystem models will be able to forecast the future state of marine ecosystems. Perhaps they will. However experience suggests that even with sophisticated models, we should still expect circumstances where our models fail and sometimes spectacularly so (Pine et al. 2009). Indeed if civil engineers still have problems building a footbridge with all the advantages they have of deterministic systems, known dynamic properties, known physical properties and advanced mathematical and physical modeling (http://en.wikipedia.org/wiki/Millennium_Bridge_London), how can we expect fisheries ecosystem models to be any more successful? Practitioners and managers alike should develop a clear realization that the most appropriate use of these models is not in forecasting, but rather in understanding trade-offs inherent among alternative choices before management. Ecosystem models can likely inform which policies are more or less likely to avoid an undesirable state of nature. However, ecosystem models are not a tool to determine the appropriate level of precaution necessary to prevent the undesirable state of nature.

It was also evident from the workshop that stakeholder groups have yet to be fully engaged in the ecosystem modeling arena. I believe this is a mistake. It is too late if stakeholders are only invited to the table to view results and select among scenarios (Miller et al. submitted). In fact in much of current fisheries management stakeholders feel like they are “invited, informed and ignored” (Karl et al. 2007). I am defining stakeholders quite broadly here to include the managers, assessment scientists, fishers, and representatives of the NGO community. Engaging stakeholders early may seem an unnecessary burden. Indeed, engaging stakeholders may mean that the goals and objectives of the modeling may be modified from those initially envisaged by the modelers. However, stakeholders often provide insights into the biology and ecology of components of the ecosystem of which the modelers may not be aware. Early engagement of stakeholders offers the benefit of creating a broader acceptance of the objectives, data inputs and inferences drawn from ecosystem modeling. Ultimately, early engagement will likely ensure greater utility and broader acceptance of the results of the model when used in the management arena.

One of the key stakeholder groups for ecosystem modeling is the managers. Over the last 25 years, fisheries scientists have done a very good job in "training" managers to expect very specific, quantified tactical advice relating to the status and trends of exploited stocks. With respect to ecosystem modeling, we may be a victim of our own success. Managers are expecting a similar standard for strategic advice from ecosystem models. However, it is likely that strategic ecosystem advice will be more broadly based (Hall and Mainprize 2004). One example of such broadly based advice is the "colormap" approach involving a matrix in which the rows are species, the columns years and the individual cells are color-coded reflections of the population health of each species in each year (Choi et al. 2005). Other examples of possible advice include the trajectories of multivariate ordinations summarizing changes in the overall ecosystem (Link et al. 2002), and RAPFISH which is an ordination of categorically ranked performance measures (Pitcher and Preikshot 2001). Regardless of the approach, we need to begin now to educate managers in the kind of information that they may be receiving from ecosystem models. Managers need to be comfortable with the concepts and metrics before they are used in critical management decisions. A corollary of this observation is that managers also need to provide clear statements of objectives or expectations for what is or might be expected of ecosystem models. Clearly one model cannot meet all demands. A dialogue must be started between managers and ecosystem modelers in which managers identify the principal challenges they anticipate facing, and ecosystem models respond by educating managers on the capabilities of current model frameworks.

The workshop also highlighted differences between modeling groups focusing on estuarine and coastal ecosystems and those working on oceanic ecosystems. The ecosystem models developed by both groups focus on evaluating trade-offs. However, it was more common for the trade-offs explored in estuarine and coastal models to involve both fishery concerns and ecosystem services, whereas those explored in oceanic models were more typically focused on fishery issues. It is likely that trade-offs among fishery goals will be easier to handle than those involving ecosystem services, because the currency of the fishery goals, that of yield, is consistent. Dealing with trade-offs involving ecosystem services remains a core challenge facing modelers and the agency alike. Concerns over ecosystem services will require interactions among fisheries scientists, ecologists, environmental scientists, economists and social scientists. It is not clear how the agency will respond to this demand.

Differences between the extents to which ecosystem models have been developed among the different regions within NMFS were also evident at the workshop. Perhaps this is to be expected given the stark differences in the ecosystems managed, fishing practices and histories among the different regions. It may not be possible to develop the detailed food web-based models that exist in the structurally relatively simple high latitude systems in the low latitude tropical ecosystems in the Caribbean and Pacific. However, the dissimilarity among the regions is perhaps the strongest reason why it is important that efforts at coordination and information exchange at the national level, such as through NEMoW, remains critical.

Charlie Stock

Summary of my thoughts after attending the NEMoW 2 workshop in Annapolis.

The discussions during the NEMoW 2 meeting suggested that the first step toward bridging the credibility gap in applying ecosystem models to marine management is clearly defining management objectives for each ecosystem model application. While the scientific objectives and rationale for ecosystem models are well defined, this does not appear to be the case for the management objectives. It was recommended that all ecosystem modeling activities have clearly stated objectives during NEMoW 1. A refined recommendation for NEMoW 2 could be the need for clear management objectives for each ecosystem model application. It seems plausible that the role of ecosystem models in marine ecosystem management may be in part defined by those management objectives that cannot be met by traditional single species approaches.

A plausible range of outcomes must also be identified for ecosystem models to be used in management. This requires testing the sensitivity of model results within and between models. Parameter sensitivity becomes more difficult and labor intensive as model complexity increases but quantifying and articulating sensitivities clearly is critical to establishing model credibility. Clearly defined management objectives should help focus sensitivity around those parameters with the strongest impacts on management outcomes. Past observed patterns may help limit parameter ranges beyond constraints applied by direct observations of a process. Grouping parameters with similar effects may help to communicate sensitivities. While general approaches to model sensitivity analysis should be drawn from during these exercises, each application of a model to management will likely require a unique set of sensitivity tests to establish credibility.

An ensemble of different types of ecosystem models should also be considered for management decisions. Building such ensembles is complicated by the fact that not all ecosystem model types are designed to answer the same management questions. Clearly defined management objectives are likely the most effective means of forming a sensible ensemble from diverse ecosystem model types.

I agree with statements made during the workshop that the common perception that there is less uncertainty in single-species stock assessment models is likely erroneous. The structure of these simple models is derived from a number of strong *a priori* assumptions about how the ecosystem works and this "structural rigidity" could lead to a false sense of security. In addition, the strongly empirical basis of the functional relationships in many of these models may prove highly problematic in cases with a changing environmental baseline such as can be expected under climate change. However, while these statements may generally be true, their manifestation for any particular application must be carefully studied through sensitivity and mechanistic diagnosis of the ecosystem model results for the added value of ecosystem models to be realized. If the complexity of an ecosystem model prevents systematic treatment of sensitivity and diagnosis of underlying mechanisms, it may be unwise to use such models for management (i.e., if one can't understand what the model is doing...).

There seems to be no universally accepted approach to weighting alternative models. Highly quantitative methods (such as those discussed by Nick Bond) strive for objectivity through statistical means but are still subject to debate over detailed statistical choices. Caution should also be exercised to ensure that the criteria against which models are weighted have a direct bearing on what you want the model to predict. For example, a climate model's ability to capture specific regional sst (sea surface temperature) or wind patterns may not reflect the climate model's ability to predict global climate change. The delphi method, in contrast, seemed to me to be dependent on a highly subjective steps that required expert panels to decide which models they trust. While weighting models is likely both sensible and necessary to refine predictions, the difficulty of developing weights suggests to me that it 1) should be done sparingly, 2) that the rationale should be clearly articulated and have a firm theoretical basis, and 3) that some sensitivity to the choice of model weights should be conducted.

One clear recommendation was that there was a need for review panels with diverse expertise to review ecosystem model results for potential inclusion in management. It would be a challenge to form such panels, but they (and the "discomfort" they would cause) would be critical for vetting ecosystem model findings. The BEST/BSIERP program is a potential case study for such a process. Perhaps the most challenging aspect would be to communicate to these panels what the standards for inclusion should be, given the lively debate over which ecosystem modeling approaches are useful and appropriate. A critical trait for panel members would be a willingness to consider alternative approaches as long as they are well defended. Clear statements of NOAA's modeling philosophy would be useful in this regard.

Steering Committee

It was clear in many discussions throughout the meeting that there may need to be a clearer statement of objectives for doing EM, or at least reminders thereof. It was recognized that part of this stems from not yet having a clear, direct, legal mandate. Yet many (competing) legal mandates are in fact extant (e.g., Magnuson-Stevens Act, National Environmental Protection Act

[NEPA], Marine Mammal Protection Act [MMPA], Endangered Species Act [ESA], among others), *ad hoc* and indirect. We as an agency simply forget them in the context we are working in. Keeping such a broader perspective will help the NMFS EM community to provide useful information for a variety of LMR management issues. The result of this observation means that we need to continue the discussion about relevance of these approaches, both internally and externally to all of our constituents.

We note that there have been numerous statements of rationales for doing EM (e.g., NEMoW I (Townsend et al. 2008), FAO (FAO 2008, Plaganyi 2007), etc.), including: addressing tradeoffs; mitigation of competing mandates; tracking of emergent properties and system level properties; and addressing climate change and related broad-scale effects. That there are actually these common objectives / rationales is worth noting, but that they are not widely known or understood reiterates the need for better interaction and communication with stakeholders.

The need exists to state the (generic) standards of rejection / acceptance for EMs and use thereof. It was also noted that the NMFS EM community should confront the difficulty in adequately addressing uncertainty in contemporary EM efforts, and initiate the research efforts and pathways necessary to adequately account for such uncertainty in the ecosystem models of the future. In particular, it was recognized that the NMFS EM community should not overly self-flagellate, as many of the areas of uncertainty or need for review are common to any natural resource modeling endeavor. In discussions during the meeting, the point was made several times that single species models were equally challenged in adequately addressing uncertainty in the early stages of their development, and the manner in which such uncertainty is estimated, expressed and conveyed continues to evolve. Thus, as SS models have had a longer history, we can learn from them but also not expect EM modeling and its review processes to be as mature. It was repeatedly noted that the need also exists to clearly establish a review process specifically for EM being used in a LMR context. What that specifically looks like is to be determined.

There was recognition that the 1st NEMoW recommendation to hold regular and routine NEMoWs was well received. There was some sentiment for a range of what future NEMoWs could address, including: a future workshop to evaluate model behavior and performance on "conjured" virtual data, to explore how various diagnostics for particular EM should be established, to explore the utility and mechanisms to calculate system level BRPs, to explore the ways to include socio-economic factors more so than now, to explore how ecosystem models can best address particular objectives and questions (see Table A.3 in Townsend et al.), to explore novel system-level emergent properties as BRPs, etc. There was no shortage of potential NEMoW topics nor enthusiasm for them.

There was a recognized disparity across the Centers in the agency in terms of EM capacity. The question arose, how should we address this? There were several suggestions, including: continued NEMoWs, rotational assignments, model building or issue specific workshops as noted above, etc. It was also noted that as part of the NEMoW objectives, sharing best practices, tips and tricks from Center to Center is a vital part of NEMoWs. To that end, it was recognized that lessons learned could be passed along from those Centers doing more EM now. Conversely, for those Centers with currently less EM capacity, it was noted that they in fact may be able to avoid any "inertia" from pre-established protocols and thereby leap-frog some Centers that have more established sets of data collection and synthesis programs. It was clear that the need to establish novel data collection and data management programs in some Centers is very much needed.

As novel initiatives and national programs emerge, it will be important to have established groups at each Center to handle the EM (and more broadly, EAF) issues that we foresee will continue to arise. Compared to NEMoW I, where two out of seven Centers had dedicated EM efforts, there are now five out of seven such groups. Thus progress is being made to build up the EM capacity in the agency. The NMFS EM community needs to develop further EM capacity in a LMR context-- not only for executing and implementing the models, but also for reviewing EM for LMRs in other regions-- to avoid overly burdening the same individuals within Centers and the NMFS. Building up the NMFS EM capacity will continue to be important and we reiterate the NEMoW I recommendation of establishing partnerships with Sea Grant and related organizations for the training of graduate students and post-docs in that regard.

Finally, there was the sentiment that we need to just start (or continue) our ecosystem modeling endeavors. This includes the development, execution and implementation of EMs, as well as the review and presentation of EM outputs. To do so means we need to begin plugging into extant review and presentation venues that engage our stakeholders. There is the need to support and endorse further LMR data servicing (collating, collecting, compiling, organizing). This is not an EM activity in itself, but an activity without which believable EMs can not constructed. There was also the observation that there will be uncertainty in any model, so we need to simply express and admit the sources of uncertainty in our EMs as best as we can, addressing that uncertainty via logical, best practice methods. There was the sense that the use and uptake of EM outputs is an extended process, as noted by a few examples around the country. As we move towards more complete EM implementation across the range of mandates and issues we are faced with, it was recognized that we probably already provide more strategic use of EM output than we probably think by way of a wide range of contextual information that is used routinely by our stakeholders.

References

Anderson, J.L. 1998. Embracing uncertainty: The interface of Bayesian statistics and cognitive psychology. Conservation Ecology 2(1):2 (http://www.consecol.org/vol2/iss1/art2).

Anderson, J.L. 2001. Stone-age minds at work on 21st century science: How cognitive psychology can inform conservation biology. Conservation Biology In Practice 2(3):18-25.

Choi, J. S., K. T. Frank, B. D. Petrie, and W. C. Leggett. 2005. Integrated assessment of a large marine ecosystem: A case study of the devolution of the Eastern Scotian Shelf, Canada. Oceanography and Marine Biology - an Annual Review, Vol. 43 43:47-+.

FAO (Food and Agriculture Organization of the United Nations). 2008. Fisheries Management. 2. The ecosystem approach to fisheries. 2.1 Best practices in ecosystem modelling for informing an ecosystem approach to fisheries. FAO Technical Guidelines for Responsible Fisheries, 4 (Suppl. 2, Add. 1), 79 pp. (report from the meeting in Tivoli, Italy, 2007).

Hall, S. J., and B. Mainprize. 2004. Towards ecosystem-based fisheries management. Fish and Fisheries 5(1):1-20.

Karl, H. A., L. E. Susskind, and K. H. Wallace. 2007. A dialogue not a diatribe - Effective integration of science and policy through joint fact finding. Environment 49(1):20-+.

Link, J. S., and coauthors. 2002. Marine ecosystem assessment in a fisheries management context. Canadian Journal of Fisheries and Aquatic Sciences 59(9):1429-1440.

McElhany, P., E.A. Steel, K. Avery, N. Yoder, C. Busack, and B. Thompson, in press. Dealing with uncertainty in ecosystem models: lessons from a complex salmon model. Ecological Applications.

Miller, T. J., and coauthors. submitted FishSmart: An innovative role for science in stakeholder-centered approaches to fisheries management. Fisheries.

Pantus, F.J. (2007) Sensitivity analysis for complex ecosystem models. PhD Thesis, School of Physical Sciences, The University of Queensland (available online at http://espace.library.uq.edu.au/view/UQ:137017)

Pine, I., W. E., S. J. D. Martell, C. J. Walters, and J. F. Kitchell. 2009. Counterintuitive responses of fish populations to management actions: some common causes and implications for predictions based on ecosystem modeling. Fisheries 34(4):165-180.

Plaganyi, E.E. 2007. Models for an ecosystem approach to fisheries. FAO Fisheries Technical Paper 477, 108 pp.

Pitcher, T. J., and D. Preikshot. 2001. RAPFISH: a rapid appraisal technique to evaluate the sustainability status of fisheries. Fisheries Research 49(3):255-270.

Townsend, H.M., J.S. Link, K.E. Osgood, T. Gedamke, G.M. Watters, J.J. Polovina, P.S. Levin, N. Cyr, and K.Y. Aydin (editors). 2008. Report of the National Ecosystem Modeling Workshop (NEMoW). National Marine Fisheries Service, NOAA Technical Memorandum NMFS-F/SPO-87, 93 pp.

Recommendations

1. Establish distinct EM review panels

We should establish panels with MRMs (ESAM, MS models) in existing review venues (e.g. stock assessment, PR, etc. contexts). However, it was agreed that bulk biomass and full system models need something with distinct, but standard review criteria. These latter two classes of EM need a venue that emphasizes the value of strategic information. It was agreed that it was unclear exactly what these review venues would look like, but in actual practice need clearly defined TORs.

2. Reiterate need for an MSE (or similar) process and facilitate their establishment

Having EMs embedded as operating models in a management strategy evaluation context would free them from the onus of precision and instead, as intended for the IEA process, allow them to do one of five things. First, explore a range of scenarios that bracket the likely reality of what are feasible LMR management options. Second, establish which strategies are most robust to a wide range of uncertainty. Third, help LMR managers and stakeholders make the most robust decisions, learn about the systems under consideration, and avoid potentially very negative choices. Fourth, explore the range of tradeoffs across and between fishery and non-fishery sectors. Finally, establish a time-frame for a process that is more flexible, more timely, but less frequent and intense.

3. Identify and note sources of EM uncertainty as a must for EM use

All applications of EM need to, as best as they can, state the known sources of uncertainty. It was noted that there will be uncertainty in any model, so the need to simply express and admit the sources of uncertainty in our EMs lends to credibility and transparency in the modeling process. If a source of uncertainty is so unique from that found in a typical "best-practices" list, it should be especially noted.

4. Use commonly accepted ways to address EM uncertainty

As appropriate for type of model and type of uncertainty, a set of common approaches to best address EM uncertainty should be employed. These include various forms of: risk analysis, multi-model inference, sensitivity analysis, visualization tools, and MSE operating models for scenario testing.

5. Establish mechanisms to enhance stakeholder interactions and communication

Additional communication, education and interaction among stakeholders and NMFS scientists regarding the use and application of EM is warranted. Employing a cognitive psychological approach to thinking about uncertainty and obtaining professional facilitators will help this process. It was noted that establishing a forum for doing so, especially with council personnel and related constituents, in an informal, non-threatening venue would help to this end.

Developing communication strategies to target different audiences and providing multilevel information systems were also identified as important needs in this context. It will require an investment to do this level of engagement, but may ultimately cost more if that investment is not made.

6. Creatively display, visualize and present multiple indicators

There are a range of approaches to visualize EM outputs and by improving presentation of EM data and outputs, we can better mitigate implementation, process and conceptual uncertainty. The need to visualize the tradeoffs among a range of ecosystem properties is only going to continue to increase. Doing so will require engaging other experts in the field of data visualization and perhaps evolve into situations that provide options for "Gaming" across the multivariate, multi-factorial processes that influence LMRs.

7. Bolster the value of strategic advice

Strategic advice in a LMR management context is highly valuable. The suite of EMs can provide a range of LMR management advice. Core among that advice, and often overlooked, is the role of strategic advice. Exploring and better explaining the utility and value of strategic advice is highly merited.

8. Investigate the utility of system level BRPs and associated control rules

Further exploration of this approach is warranted and has value, for several reasons. As systemic biomasses are more stable, this is a desirable property. As such these EM outputs are more stable, predictable, resilient, and simple forms of LMR advice. Concurrently, they are generally more conservative, different, novel and potentially conceptually difficult LMR advice. Further exploration of this approach as a major use of EM outputs, in the context of wide range of legislative mandates, is needed.

9. Bolster ecosystem modeling capacity

The NMFS EM capacity needs to be buttressed. This would include addressing those regions with obvious data or information gaps. This would also include developing the EM expertise over the next decade via training of graduate students and post-doctoral researchers and by providing institutional support of EM efforts with adequate resources and staffing to address the NOAA-wide commitment to ecosystem-based management.

Appendix A.

NEMoW II Agenda

25-Aug-09		Topic	
830	Welcome & Orientation		Host (Chris Kinkade)
840	Overview, Objectives & Goals	Workshop Purpose & Rationale	J. Link
845	The role of EM in NOAA NMFS	NMFS Context and Forthcoming EM Issues	K. Osgood
915	Keynote Speaker	Review of Major Types of Uncertainty in Modeling & How to Deal with Them	Randall Peterman
1015		Discussion on Types of EM Uncertainty	Plenary
1030	Coffee Break		
1100	Keynote Speaker	Generic Levels of EM Uncertainty- Data Inputs, Parameters, Outputs & Considerations from other disciplines	Nick Bond
1145		Discussion on Types of EM Uncertainty	Plenary
1200	Lunch		
1300	Break-out Groups	Discuss Model Uncertainty	Break-out Groups
1400		Report back to plenary	Plenary
		Discussion on Types of EM Uncertainty	Plenary
1530	Coffee Break		
1600	Break-out Groups	Overriding Workshop Theme: "Bridging the Credibility Gap" -- Nat'l EM Standards of Use & Review	Break-out Groups
1700		Report back to plenary & Wrap up	Plenary
~1730	Adjourn		

26-Aug-09		A ***<u>BRIEF</u>*** summary of how each Center is addressing EM Uncertainty: maybe multiauthored, but 1 presenter for each theme, presented by theme	
830	Keynote Speaker	How Has Strategic Advice Been Used in a Global LMR Context? A Global Perspective on How To Deal with EM Uncertainty	Andre Punt
930		What are the one to two top unknowns in each Center's EM activities?	D. Kobayashi/C. Harvey
945		How is each Center addressing EM Uncertainty: e.g., Sensitivity Analysis, Multi-model inference, Risk Analysis, MSE, etc.?	J. Link/G. Watters
1000		What are the venues that each Center is using to present and review EM? Are they appropriate/sufficient?	K. Aydin/P. Levin
1015		Are SS and EM model outputs being compared/complementary in each Center? If so, how?	J. Field/S. Gaichas
1030	Coffee Break		
1100		How is each Center engaging Stakeholders, particularly Councils, wrt Strategic Advice and Guidance from	P. Levin/G. Watters

		EM?	
1115		If a Center is not doing EM and evaluating EM uncertainty, why not?	M. Schirripa/D. Kobayashi
1130		What is the availability of data that might be used in a Ecosystem model? And how accessible is this data?	K. Osgood/M. Schirripa
		Discussion on Common Approaches to Handle EM uncertainty	Plenary
1200	Lunch		
1300	Break-out Groups	Discuss & demonstrate approaches and experiences in handling EM uncertainty	Break-out Groups
1400		Report back to plenary	Plenary
		Discussion on handling EM uncertainty	Plenary
1530	Coffee Break		
1600	Break-out Groups	Discuss EM output relative to SS (Fish and PR) model output: wrt use, precision, variability, uncertainty, etc.	Break-out Groups
1700		Report back to plenary & Wrap up	Plenary
~1730	Adjourn		
27-Aug-09	Model Classes	How to Utilize Strategic Model Output	
830	Keynote Speaker	What Models Can produce Tactical BRPs, What Models Can Produce Strategic Outputs: Case Studies where Strategic Advice would have helped	Tom Miller
930		Discussion on the Best Venue/s for Providing Strategic Advice	Plenary
1030	Coffee Break		
1100		Discussion on how to provide Strategic Advice to LMR Management Bodies	Plenary
1200	Lunch		
1300	Break-out Groups	Discuss EM Review Criteria wrt Strategic Advice and Uncertainty	Break-out Groups
1400		Report back to plenary	Plenary
1430		Discussion of EM Review Criteria wrt Strategic Advice and Uncertainty	Plenary
1530		Coffee Break	
1600		Discussion of Next Steps	Plenary, Steering Cmte
~1730	Adjourn		

Appendix B.

Participants List

NMFS Personnel

Steering Committee

Kenric Osgood	Kenric.Osgood@noaa.gov	ST
Howard Townsend	Howard.Townsend@noaa.gov	NCBO
Jason Link	Jason.Link@noaa.gov	NEFSC
Kerim Aydin	Kerim.Aydin@noaa.gov	AFSC
Phil Levin	Phil.Levin@noaa.gov	NWFSC
Michael Schirripa	Michael.Schirripa@noaa.gov	SEFSC
Don Kobayashi	Donald.Kobayashi@noaa.gov	PIFSC

Other NMFS Personnel

Abigail Franklin	Abigail.Franklin@nooa.gov	SF
Andrew Leising	Andrew.Leising@noaa.gov	SWFSC
Bill Overholtz	William.Overholtz@noaa.gov	NEFSC
Bern Megrey	Bern.Megrey@noaa.gov	AFSC
Buck Stockhausen	William.Stockhausen@noaa.gov	AFSC
Cameron Ainsworth	Cameron.Aisnworth@noaa.gov	NWFSC
Chris Harvey	Chris.Harvey@noaa.gov	NWFSC
Derek Orner	Derek.Orner@noaa.gov	NCBO
Evan Howell	Evan.Howell@noaa.gov	PIFSC
Frank Parrish	Frank.Parrish@noaa.gov	PIFSC
Jason Cope	Jason.Cope@noaa.gov	NWFSC
John Field	John.Field@noaa.gov	SWFSC
John Manderson	John.Manderson@noaa.gov	NEFSC
John Ward	John.Ward@noaa.gov	SF
Jon Brodziak	Jon.Brodziak@noaa.gov	PIFSC
Jonathan Phinney	Jonathan.Phinney@noaa.gov	SWFSC
Liz Brooks	Liz.Brooks@noaa.gov	NEFSC
Melissa Haltuch	Melissa.Haltuch@noaa.gov	NWFSC
Michelle McClure	Michelle.McClure@noaa.gov	ST
Paul McElhany	Paul.Mcelhany@noaa.gov	NWFSC
Tom Ihde	Tom.Ihde@noaa.gov	NCBO
Tom Minello	Tom.Minello@noaa.gov	SEFSC
Rob Gamble	Robert.Gamble@noaa.gov	NEFSC
Sam McClatchie	Sam.Mcclatchie@noaa.gov	SWFSC
Sarah Gaichas	Sarah.Gaichas@noaa.gov	AFSC
Stephanie Oakes	Stephanie.Oakes@noaa.gov	ST

External Observers

Andre Punt	aepunt@u.washington.edu	UW
Nick Bond	nick@atmos.washington.edu	JISAO-UW
Tom Miller	miller@cbl.umces.edu	CBL
Randall Peterman	peterman@sfu.ca	SFU

Other NOAA Personnel

Charlie Stock	Charles.Stock@noaa.gov	OAR GFDL
Hendrik Tolman	Hendrik.Tolman@noaa.gov	NWS NCEP

Appendix C.

Steering Committee (with alternates noted)

Kenric Osgood	Kenric.Osgood@noaa.gov	ST
Howard Townsend	Howard.Townsend@noaa.gov	NCBO
Jason Link	Jason.Link@noaa.gov	NEFSC
Kerim Aydin	Kerim.Aydin@noaa.gov	AFSC
(Sarah Gaichas)		
George Watters	George.Watters@noaa.gov	SWFSC
(John Field)		
Phil Levin	Phil.Levin@noaa.gov	NWFSC
(Chris Harvey, Isaac Kaplan)		
Michael Schirripa	Michael.Schirripa@noaa.gov	SEFSC
Don Kobayashi	Donald.Kobayashi@noaa.gov	PIFSC

Appendix D.

Breakout Groups for each TOR
Trigger Questions

These questions were used to reinforce and further explore the TOR, augmenting the plenary discussions and presentations. They are presented here as a catalyst in case any particular Center would like to further examine any of these topics.

Session I.
Addresses TOR 1

Are the main classes/types of models still appropriate for your Center? [MRM (multispecies and ESAMs), bulk biomass (network and aggregate) and full system (ecosystem and biophysical) models]
Which ones are you still using or plan to use, and in what context?
What are the one to two top unknowns in each Center's EM activities?
What are the major sources of uncertainty in those EMs?
How can those sources of uncertainty be identified?
If a Center is not doing EM and evaluating EM uncertainty, why not? What needs to happen to facilitate the execution of EMs?

Session II.
Addresses TOR 4

How important has it been-- or has it not yet happened-- for EMs to undergo peer review? Can you describe how that peer review (of EMs) process has worked? If not, what needs to take place to establish these review venues?
Regarding EM uncertainty, what are the key "checklists" that external reviewers need for reviewing EMs?
Are there appropriate standardized questions and model applications for the broad range of model classes? Specifically, how can we best facilitate the evaluation of model uncertainty in these reviews?
How important has it been for colleagues, stakeholders, and/or managers to informally see "under the hood" of EM 's as a means to build credibility?
How can we make these review criteria more widely available to reviewers?
How do these review criteria compare to other model (e.g. assessment) review contexts?
Can you discuss what are the venues that each Center is using to present and review EM? Are they appropriate/sufficient?

Session III.
Addresses TOR 1.4, 1.3

How easy is it to identify sources of ecosystem modeling uncertainty?
How can those sources of uncertainty, or can they, be mitigated? That is, how is each Center addressing EM Uncertainty?
Are there any best practices for addressing uncertainty in EMs?
Are there lessons/efforts from other regions/Centers in NMFS that are particularly effective?
Are there lessons/efforts from other disciplines/Offices in NOAA that are also effective and germane?
If so, please describe them and what you think makes them exemplary.
Are there any demonstrable and portable tools or approaches that can be used to address EM uncertainty across NMFS?

Session IV.
Addresses TOR 1.5, 2.2, 2.4

How uncertain are some of the EM results compared to SS (TS or PS) results?
What ecosystem models can produce precise BRPs and related outputs that are specific, precise and point estimates (tactical advice)?
What is the (expected and reasonable) variability about those estimates from different models classes?
What models can not produce precise BRPs, but rather accurate, directional (strategic) outputs?
What is the (expected and reasonable) variability about those estimates?
Are SS and EM model outputs being compared/complementary in each Center? If so, how?
If a model is to be used for LMR management advice, are there additional criteria that model must meet?

Session V.
Addresses TOR 4, 3

What are some case studies/examples where a more strategic approach would have helped?
What features about those cases are noteworthy?
What is the best place for strategic advice?
How is each Center engaging Stakeholders, particularly Councils, regarding Strategic Advice and Guidance from EM?
If your Center hasn't started engaging Stakeholders regarding EM, what needs to take place to do so?
What would you say are the next major steps for NMFS EM?

Appendix E.

Ecosystem Modeling Best Practices List

(Editor's Note: this text and section was adapted heavily from an FAO report (FAO 2008) and the prior NEMoW report (Appendix B in Townsend et al. 2008).

A best practice approach to ecosystem modeling must include specification, implementation, evaluation, reporting and review steps. Model scoping undertaken during model specification must include the iterative construction of conceptual models that are used to define the relevant subsystem to be modeled. Once this subsystem is identified, the final model representation must be defined based on the question being considered, available data, the important system features and the appropriate scales (regarding space, time, taxonomic and human impacts resolution) and process representations.

Table E.1 shows best practices for modeling. These are not benchmarks but rather are a set of practices that should guide thinking as to the importance of different model attributes and suggested approaches for handling each of these. These practices should be followed to the extent possible. This list summarizes some of the key attributes to be considered in model development and suggests the current best practice for handling each of these, noting that this may not be practically achievable in most circumstances.

Table E.1. List of recommended best practices for modeling ecosystems.

Consideration in Model Development	**Best practice approach**
Setting up a model	
How many species or groups?	Aggregate based on shared characteristics of the species and omit the least important to keep food web tractable
Include age, size or stage structure of the species of interest?	Include if there are major shifts over the course of the life history of (harvested) species of interest
Include spatial structure?	Include where there are major shifts in the location of the species of interest over the course of its life history
Include seasonal and temporal structure?	Where there are large differences in the seasonal dynamics in species movement or production
Defining boundary conditions	Basing boundaries on biological/geological rather than anthropogenic considerations such as national boundaries.
Is fishery harvesting more than one stock of a particular species?	Model needs to distinguish such different stocks when the harvesting practice is such as might impact these stocks to different extents; this may necessitate spatially structured models
Distinguish different fleets?	Important in the context of provision of advice at the tactical level, if for the same mass of catch, they make substantially different impacts on target and bycatch species or on the habitat.
Explicitly represent primary productivity and nutrient cycling	May only be necessary when bottom-up forces or lower trophic levels are of key concern. Inclusion of these processes can be highly informative for some strategic modeling exercises.
How to model recruitment?	Recruitment may be included either as an emergent property or as a derived relationship (which should not be based on uncritical correlation studies of recruitment and environmental parameters). Recruitment variability is likely to be important for tactical and risk analyses, but is not a strict requirement in many strategic models.
How to model movement?	This involves testing sensitivity to a range of movement hypotheses. Where possible, best practice involves parameterizing movement matrices by fitting to these data. If decision rules are used to drive movement, attention should be focused as to whether the resultant changes in distribution are sensible.
Explicitly consider fleet dynamics?	Important to consider if substantial changes to the spatial distribution of fishing may result from, for example, the declaration of an MPA. The population model must include of spatial component in these circumstances, and it may be necessary to develop a model of the manner in

	which fishing effort patterns will change in response.
How much detail in representing predator-prey interactions?	Represent as bi-directional unless it can be strongly demonstrated that it is adequate to include a one-way interaction only in which the predator ration is fixed and changes in prey abundance have no effect on predator populations.
Which functional response?	Test sensitivity and robustness to alternative functional relationships.
Include environmental forcing?	Only if it is an absolute requirement for capturing system dynamics. When it is included there must be some means of generating future forcing for use in predictions and closed loop simulations and a clear understanding of probable mechanism
Other anthropogenic forcing?	Their influence on shallow coastal and estuarine systems should be considered in conceptual models and if they are found to be significant pressures on the system then they should be empirically included (e.g. simply as a forcing term) in any strategic models and management strategy evaluations for the system.
Alternative stable states?	Strategic models in particular need to ensure forecasting the consequences of environmental change, contain the capacity (e.g. functions) which allow for phase shifts, either directly (in accordance with past observations), or as an emergent property of the functions in the model. Even if such a functional form is used, it must be recognized that, until a threshold is crossed in the data, it may not be possible to parameterize the threshold point: uncertainty reporting should evaluate possible thresholds either on a theoretical or empirical basis.
Dealing with uncertainty	
Model structure uncertainty	Identify alternative qualitative hypotheses for all of the processes considered likely to have an important impact on the model outputs and then formulate these hypotheses mathematically (or as the values for parameters of a general relationship).
Implementation uncertainty	Implementation failures introduce biases in fishery data which will impact assessment and tactical models. It also creates biases in the expected impacts of simulated management measures within an MSE. Implementation uncertainty needs to be linked to consideration of fleet dynamics and is largely driven by, and must be included in, economic considerations.
Other process error considerations	Other process error, arising from natural variation in model parameters, needs to be included in projections, whether they be strategic or tactical, when that variation contributes

	substantially to uncertainty in the model outcomes.
What features to include in closed loop simulations?	As many as are feasible to parameterize for addressing the question at hand.
Should the model be fit to data?	Fitting to data is best practice, and this requires careful specification of likelihoods.
Taking account of parameter uncertainty	Include clear statements about uncertainties in model parameters; Bayesian methods and bootstrapping are considered best practice for quantifying parameter uncertainties in extended single-species models and MRMs; Improving current practices requires 1) that there is an explicit accounting of the number of parameters that are being estimated and fixed, 2) qualitative estimates of the uncertainty in every parameter, and 3) **sensitivity analyses.**
Taking account of parameter uncertainty for mass balance / static models	To develop and fully document a formal data pedigree (quality ranking), and if possible include error ranges for estimates, with input from data providers as to potential biases. Sensitivity analyses may be conducted using available routines. For dynamic models, best practices is to fit to as much data as possible using appropriate likelihood structures, while being clear about both potential biases arising from fixing parameters, as well as fully reporting error ranges resulting from freeing parameters. In case of fixing parameters, additional sensitivity analyses (e.g. resampling, Monte Carlo routines) should be used to assess model sensitivity to the assumptions. An important component is using results of sensitivity analyses to guide future data collections and the continuation of critical time series.
Use and outputs	
Should code be freely available?	Documentation and source code must be freely available to allow for review and understanding of the model. Using existing models can be of great help in learning, but careful thought is required when using a pre-existing model so that the tool is not misused.
Social and economic outputs	Have economic experts collaborating with fisheries ecologists when designing a model implementation of economic factors.
Ease of modularization	Best is object-oriented design

Appendix F.

National Standards for Ecosystem Modeling in a Living Marine Resource Management Context

From the initial NEMoW (Townsend et al. 2008), here we reiterate a set of standards as guidance for use of EM and to serve additionally as review criteria (Table F.1).

Table F.1. Proposed national standards for ecosystem modeling.

1. Adequate Documentation

2. Clearly Stated Objectives

3. Peer Review

4. Best Practice Use

5. Uncertainty Characterized

---

1. Adequate Documentation

It is proposed that all Ecosystem Models should be fully documented. This would include descriptions of input data and parameterization (Table F.2), model structure and equations; major modeling "tweaks" and tips to allow for functionality and execution of the model; model assumptions; specific model implementation and/or application to a particular system; and key diagnostics. It was noted that NOAA technical memoranda, Center reference documents, and webpages should be more fully utilized for this purpose.

Table F.2. A proposed template for documenting model input data and parameters. The rows would be the state variables, parameters, or similar input properties to the model, which would then have major properties (columns) described.

	Input Value	Units	Description	Type (State Variable, constant, etc.)	Origin (Including Species From Which Parameter Was Derived)	Multiple Measures?	Timeframe for Derivation of Value	Type of Review	Reliability/ Confidence	References
Parameter 1										
Parameter 2										
Parameter 3										
Parameter 4										
Parameter 5										
etc.										

2. Clearly Stated Objectives

It is proposed that all Ecosystem Modeling activities have clearly stated objectives. Although a seemingly obvious consideration, often the purpose of a modeling exercise is not forthrightly stated, leading to confusion about intent and application of results. Particularly for those EMs used in a LMR management context, the objectives should be clearly stated. Additionally, if EMs are not to be used in a particular context, these limitations should be clearly identified if they are for primarily research/heuristic purposes.

Additionally, there were several common issues of why EM is invoked. These are listed in Table F.3. We also note the generic model classes (Plaganyi 2007) with model types in that table. Using the rows and columns, we provide a list of recommended model uses for particular applications. These are to help guide that the generic model classes are appropriately applied to the major, common issues facing LMR issues for which EM is invoked.

We want to be clear; this table is not intended to be too prescriptive or to limit innovation. Rather it provides guidance on established approaches for common sets of issues such that a novel use would need a strong justification to be used outside of the recommended objective-model mapping given.

Table F.3. Major model classes as typically applied to common objectives of model use.

Model Classes (Plaganyi 2007) / Major Topics			*Extended SS Assessment Models*	*Minimal Realistic Models*			*Dynamic Systems Models*				*Whole Ecosystem Models*
	Generic Model Types / Common Issues & Objectives	Single Species	Single Species w / add-ons	Multi-species	Aggregate Biomass	Food Web	Habitat	Biophysical	Biogeochemical	Bioeconomic	Full System
Technological Interactions	technological interactions		x	x	x	x				x	x
	bycatch	x	x	x	x	x				x	x
Trophic / Ecological Interactions	protected species and species of interest	x	x	x	x	x		x		x	x
	commercial fishing on forage species		x	x	x	x				x	x
	trade-offs among predators being targeted		x	x	x	x					x
	predation of targeted species		x	x	x	x		x			x
Physical / Climate Drivers	effects of fishing on habitats						x		x		x
	habitat effects on stocks		x	x	x	x	x	x			x
	climate		x	x	x	x	x	x			x
	cumulative effects					x	x	x			x
	toxins / bioaccumulation					x			x		x
Spatial Features	MPA efficacy, structure, placement						x	x		x	x
	range shifts		x				x	x			x
	habitat restoration strategies						x		x		x
System Considerations	invasive species										
	"ecosystem health"—sustainability, resilience				x	x					x
	ecosystem status				x	x		x			x
	biodiversity				x	x					x
	underlying system carrying capacity				x	x	x	x			x
	regime shifts							x			

Socioeconomic Drivers	economic issues	x			x	x				x	x
& Management	trade-offs among fleets			x						x	x
	determining reference points- systemic				x	x		x	x		x
	determining reference points- SS assessments	x	x	x							
	cumulative management effects				x	x				x	x

3. Peer Review

It is proposed that all NMFS EM used for LMR management be peer reviewed. This statement was viewed as imperative by the NEMoW participants. Although perhaps obvious, it merits stating outright.

This peer review would entail a review of the: model structure; model behavior & sensitivity analysis; software & code; and for a particular application a review of the: parameters & input data; calibration, validation & verification; and model outputs.

Such a peer review would be performed at several levels of model construction, with a preference for the model structure, behavior and software to be in the peer reviewed literature. Related to item #1 above, the particular application would also need to be documented in an appropriate venue. The model output as applied to a particular LMR management issue would also need to be reviewed by a panel of experts, *sensu* the CIE or some similar body.

4. Best Practice Use

It is proposed that all NMFS EM modeling efforts adopt a "best practices" approach (see Appendix E). This would effectively entail using the FAO (or variant thereof) "checklist" when initiating an EM application.

The discussion ranged widely on this topic, but there was consensus that this "Best Practices List" not be an absolute requirement but a set of guidelines of approaches to best address common modeling caveats. The converse of a "Best Practices List" was suggested as having a minimum standards of use, which is in effect another view of the checklist.

By ensuring that minimum EM standards are met, we mean to ensure that the data are sufficient for each generic model type and specific model (meet minimum requirements). We also mean to ensure that the data and model class are sufficient to specific issue being addressed (Item #3 above).

5. Uncertainty Characterized

It is proposed that each EM effort needs to explicitly characterize uncertainty. Although this is a large part of the best practices noted in #4 above, a clear, transparent set of statements of where a model may not perform adequately is needed. A transparent set of statements of what data or inputs or parameters are sensitive or highly variable, and how this might affect model behavior, is needed.

Although this point could be included in item #4 above, the NEMoW participants conveyed strong enough opinions on the matter to warrant it being noted as a separate item.

Characterization of model uncertainty would take into account structure, implementation, and parameter uncertainty as a key set of reported diagnostics. There should be a suite of standard, model-specific diagnostics in each EM application, but some form of uncertainty characterization would be mandatory for EM use in a LMR management context.

Appendix G.

Glossary of Abbreviations Used

AFSC: Alaska Fisheries Science Center

BRP: biological reference point

CBL: Chesapeake Biological Laboratory (a campus of the University of Maryland Center for Environmental Science)

CRC: Chesapeake Research Consortium, Inc.

CSIRO: The Commonwealth Scientific and Industrial Research Organisation (Australia)

EAF: ecosystem approach to fisheries

EM: ecosystem modeling (covering the full range from minimal realistic models, multispecies and extended stock assessment models, bulk biomass (network and aggregate) and full system (ecosystem and biophysical) models.

ESAMs: extended stock assessment models; term includes a wide variety of models from multispecies models to those that include some additional aspect of the environment (e.g., predator or prey, or climate variability)

FEP: fisheries ecosystem plan

IEA: Integrated ecosystem assessment

LMRs: living marine resources

MRM: minimal realistic models

MS: multi-species

MREP: Marine Research and Education Partnership

MSE: management strategy evaluation

NCBO: NOAA Chesapeake Bay Office

NEFSC: Northeast Fisheries Science Center

NEMoW: National Ecosystem Modeling Workshop

NMFS: National Marine Fisheries Service

NOAA: National Oceanic and Atmospheric Administration

NSAW: National Stock Assessment Workshop

NWFSC: Northwest Fisheries Science Center

NWS NCEP: National Weather Service - National Centers for Environmental Prediction

OAR GFDL: Office of Oceanic and Atmospheric Research - Geophysical Fluid Dynamics Laboratory

PIFSC: Pacific Islands Fisheries Science Center

PR: protected resources

SEFSC: Southeast Fisheries Science Center

SF: Office of Sustainable Fisheries

SFU: Simon Fraser University

ST: Office of Science and Technology

SWFSC: Southwest Fisheries Science Center

TOR: Terms of reference

UW: University of Washington

wrt: with respect to