Report from the Joint OAR-NMFS Modeling Uncertainty Workshop

Jason Link, Doran Mason, Terra Lederhouse, Sarah Gaichas, Troy Hartley, Jim Ianelli, Richard Methot, Charles Stock, Craig Stow, Howard Townsend

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

NOAA Technical Memorandum NMFS-F/SPO-153
July 2015

Report from the Joint OAR-NMFS Modeling Uncertainty Workshop

Jason Link, Doran Mason, Terra Lederhouse, Sarah Gaichas, Troy Hartley, Jim Ianelli, Richard Methot, Charles Stock, Craig Stow, Howard Townsend

NOAA Technical Memorandum NMFS-F/SPO-153
July 2015

U.S. Department of Commerce
Penny Pritzker, Secretary

National Oceanic and Atmospheric Administration
Kathryn D. Sullivan, Ph.D., Administrator

National Marine Fisheries Service
Eileen Sobeck, Assistant Administrator for Fisheries

Recommended citation:

Link, J.S., D. Mason, T. Lederhouse, S. Gaichas, T. Hartley, J. Ianelli, R. Methot, C. Stock, C. Stow, and H. Townsend. 2015. Report from the Joint OAR-NMFS Modeling Uncertainty Workshop. U.S. Dept. of Commer., NOAA. NOAA Technical Memorandum NMFS-F/SPO-153, 31 p.

Copies of this report may be obtained from:

Kenric Osgood, Ph.D.
NMFS Office of Science and Technology
National Oceanic and Atmospheric Administration
1315 East-West Highway SSMC3/12555
Silver Spring, MD 20910

Or online at:
http://spo.nmfs.noaa.gov/tm/

Executive Summary

The National Oceanic and Atmospheric Administration's (NOAA's) Office of Oceanic and Atmospheric Research (OAR) and National Marine Fisheries Service (NMFS) held a joint workshop April 13–15, 2015, at NOAA's Great Lakes Environmental Research Laboratory in Ann Arbor, Michigan, to examine methods and means of addressing, reviewing, and presenting uncertainty in ecosystem and living marine resource models and assessments. This workshop explored the range of practices used within OAR, NMFS, and other organizations in the modeling community to deal with uncertainty, developed a set of best practices, and identified recommendations to better address model output uncertainty and improve model skill.

A broad range of NOAA's living marine resource and ecosystem modeling approaches were discussed. Examples from these and related disciplines show a healthy and robust gradient of models along multiple dimensions of complexity. As such, continued model development remains an important research activity and should include evaluations of model uncertainty. Best approaches to address uncertainties depend on the type of model and application. In this context, the workshop focused on model skill evaluation, noting the need for and feasibility of multiple measures.

Participants identified 14 best practices at the workshop that will serve ongoing and future efforts to address model uncertainty. A simple "cheat sheet" on matching approaches with specific types of uncertainty across model applications was suggested. This could lead to a template of standard reporting for living marine resource and ecosystem assessment outputs. The benefits of having a common usage of quantitative information to explore model skill would provide the public with a more consistent communication of scientific results and management implications. Some of the key practices that emerged included the use of multi-model inference, adoption of management strategy evaluations, and improved use of communication tools. These formed the basis of eight recommendations specific to model efforts. Overall, cross-disciplinary, cross-organizational meetings like this to coordinate and advance modeling efforts were considered useful and should continue. Next steps recommended are:

1) Seek NOAA leadership support for a full range of quantitative modeling efforts across the spectrum of complexity and disciplinary emphasis in support of living marine resource management mandates.
2) Establish routine and regular venues for the NOAA modeling community to meet and interact.
3) Support and advance cross-line-office and cross-disciplinary (including social science) coordination on model development.

Acknowledgements

The Workshop Steering Committee, on behalf of all participants, would like to thank the local host of the workshop, the Great Lakes Environmental Research Laboratory. We especially thank Doran Mason, Barbara Gerych, Patricia Sousa, and Tanya Mathis for their efforts in local logistics and travel planning. We acknowledge the efforts of the invited speakers, and thank them for their participation. Finally, we thank both the NMFS and OAR Offices of the Assistant Administrator for partial funding support allowing invitational speakers to participate in the workshop.

Table of Contents

Introduction

Models of living marine resources (LMRs) and their ecosystems provide a means to integrate and synthesize a broad range of temporal and spatial data and provide the foundation for resource management decisions. Living marine resource and ecosystem modeling is typically performed through assessments that represent the operational implementation of ecosystem science. Ecosystem science provides fundamental understanding of marine organisms, their habitats, and how they interact with each other—all in a changing environment. Assessments leverage this fundamental science to provide specific advice needed by managers, regulators, and the affected public. Given the complexity of marine ecosystems, the performance of relevant assessments and models is affected by their inherent uncertainty and skill.

From multiple mandates NMFS is responsible for providing scientifically based management advice for approximately 450 managed fish stocks and more than 200 protected and marine mammal species, 2,000 habitats, 200 aquaculture permits, and 10 large marine ecosystems. OAR provides tools, process studies, and data streams that contribute to these modeling approaches and assessments. NOAA's National Ocean Service (NOS) provides tools and physical and ecological models to support safe navigation and coastal zone management. These NOAA line offices have a solid history of developing models together.

Models to conduct LMR and ecosystem assessments have inherent uncertainty, as all are based on incomplete data and are necessarily simplifications of natural processes at work. Such assessment models are continually being updated and improved in order to provide the best scientific advice to inform resource management, and part of this process is to evaluate how these models address uncertainty. This workshop was convened to further discuss and explore advancements in the characterization, reduction, and communication of uncertainty, with particular emphasis on exploring best practices for addressing model uncertainty in an LMR context. The deliberations at this workshop capitalized on ongoing interactions among OAR and NMFS modeling communities, and were intentionally broadened to include other NOAA line offices and partners from other disciplines to ensure salient practices were considered. The results reported herein represent state-of-the-art approaches used in addressing uncertainty in LMR and ecosystem modeling, identify areas for advancement in addressing model uncertainty, and note key recommendations for future modeling efforts.

This report is structured around four workshop terms of reference that guided discussions. An additional section on the human dimension of uncertainty, including the perception of uncertainty, is also included in this discussion.

Terms of Reference

1) Outline and summarize main types of assessment models used for LMR and ecosystem assessments.
2) Characterize, define, and discuss means to address various types of model and assessment uncertainty.
3) Characterize, define, and evaluate model skill.
4) Capture best practices for addressing uncertainty in modeling and assessment; next steps from the workshop.

Types of models used for living marine resource and ecosystem assessments

Representatives from across NOAA and other participating organizations provided a brief primer on the types of models used for LMR or ecosystem assessments to help guide workshop discussions. These have been catalogued elsewhere (Quinn and Deriso 1999, Cloyd et al. 2007, Plaganyi 2007, Shertzer et al. 2008, Townsend et al. 2008, Leonardi et al. 2009, NOAA 2010 generally; Townsend et al. 2014, Deroba et al. 2015; LMR and ecosystem models specifically) and are not repeated here. The presentations demonstrated a robust gradient of models along multiple dimensions of complexity (Figure 1).

The workshop discussions noted that these models can be used and applied to a variety of issues. For example, in some cases stocks of interest may have very limited data, yet resource managers are required to determine annual catch limits as a reference point for determination of overfishing status. In such cases various data limited modeling methods could be employed. At an intermediate level, nearly 150 of the major fish stocks are managed according to reference points and catch limits determined from age-structured demographic models of the harvested population. At the other extreme, trade-offs among fishery and protected resources, all under different climate change scenarios and given multiple objectives to consider, might require evaluation for a management body to prioritize its decisions. In that case some of the more comprehensive ecosystem models would be employed as data constraints allow. Many additional examples and applications exist in between these scenarios.

Workshop participants recognized that, although there are appropriate uses of distinct model classes for different issues to be modeled (e.g., Townsend et al. 2008), there remains uncertainty across all types of models. For example, some models used for tactical fisheries management (e.g., setting stock-wide annual catch limits or quotas) will overlap any spatial dynamics and suffer from process uncertainty due to their simplifications. Ecosystem models used for strategic advice on management strategies (e.g., exploring options across a range of objectives and relevant performance measures) may have multiple functional forms and suffer from structural uncertainty. Some types of uncertainty may be more important to consider than others given the type of model or application. The various types of modeling uncertainty are further discussed later in this report.

The degree to which these models provide credible advice for use by the full range of NOAA's partners largely depends on how well the modeling community presents, communicates, and addresses the associated uncertainties.

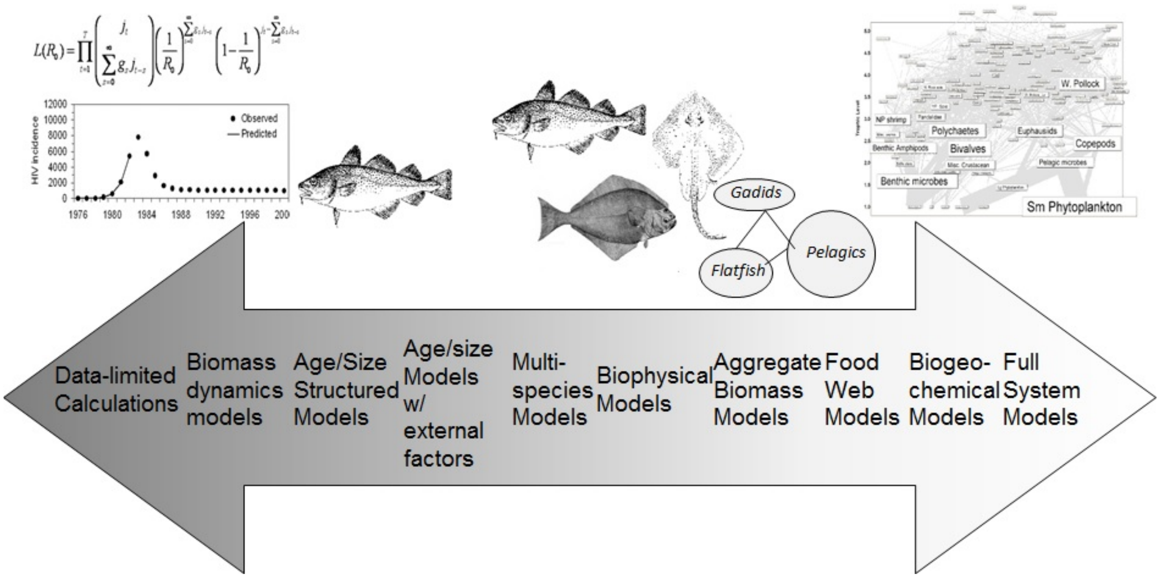

Figure 1. The gradient of modeling approaches used in living marine resource and ecosystem assessments.

Types of and methods for addressing uncertainty

There are many types and facets of model uncertainty. There have been several attempts to classify uncertainty in a living marine resource management context (Hilborn 1987, Morgan and Henrion 1990, Ferson and Ginzburg 1996, Francis and Shotton 1997, Anderson 1998, Charles 1998, Patterson et al. 2001, Link et al. 2002, Regan et al. 2002, Harwood and Stokes 2003, Peterman 2004, Mangel, 2006, Hill et al. 2007a, Townsend et al. 2008, Pine et al. 2009, McElhany et al. 2010, Link et al. 2012), which are generally recognized in other fields of complex systems modeling (e.g., Hawkins and Sutton 2009, NRC 2012). The uncertainties associated with any type of modeling are neither uniform nor insurmountable, but can be broken down into subcomponents. The workshop participants generally acknowledged the set of uncertainty types from Peterman (2004), as revised in Link et al. (2012; see Figure 2), as a useful rubric for the discussion. These include:

- Natural variability (or process uncertainty).
- Observation error (or measurement or estimation uncertainty).
- Structural complexity (or model uncertainty).
- Communication uncertainty.
- Objective uncertainty (lack of clarity on goals and objectives; this is often included with outcome uncertainty).
- Outcome uncertainty (or management performance uncertainty).

The workshop participants agreed that, for purposes of discussion, these six types of uncertainty were useful to consider and that they generally captured the literature on the topic.

An important observation from the workshop is that model uncertainty is influenced by, and comprised of, more than just statistical precision, probability distribution functions, etc., or even model accuracy, although all are important parts of elucidating model uncertainty. In particular, observational error, and associated statistics, tends to be much higher in LMR models compared to more physical models. The many types of uncertainty that affect model uncertainty could instead be termed "model-associated uncertainty." The term "model uncertainty" is used in this report to describe both model uncertainty and model-associated uncertainty.

As noted above, there are multiple dimensions to model uncertainty. Parts of these considerations emerge more prominently as different standards across different disciplinary uses or types of model application are noted. The type of model application, particularly whether the model outputs would be used in a regulatory versus informative context, was a repeated point of discussion (NRC, 2007). For example, regulatory outputs that materially stop an activity have different standards than those that provide general information to inform decisions. Considerations regarding the spatial scale, time frame (both extent and resolution), taxonomic resolution and range, or whether the model output is used for a forecast, nowcast, or hindcast all merit attention when discussing model uncertainty, as well as how best to handle the types of model uncertainty and the level of rigor at which such uncertainty should be addressed (c.f. Table 1). There was a clear sense of the need for standards on the rigor of model uncertainty treatment, conditional on the different types of application.

One element that arose from workshop discussions was that consistency in treatment of uncertainty can be as important as rigor, absolute determination, and accuracy when it comes to dealing with how model uncertainty is perceived and used by stakeholders. For example, the magnitude of fishery management buffer is expected to be scaled according to the degree of scientific uncertainty, so the magnitude of the buffer is linked to the perception of the absolute scale of uncertainty. The performance of the buffer is more important than getting the absolute scale of uncertainty correct. Similarly, consistency in terminology was also recognized as important and simple to convey but is an often overlooked element of model uncertainty. It was noted that communication uncertainty occurs even between different groups of modelers from related but subtly distinct disciplines.

There are many recognized methods for addressing model uncertainty (Link et al. 2010, NRC 2012, Townsend et al. 2014). These methods include approaches such as sensitivity analysis, risk analysis, improved data or model output visualization, management strategy evaluation (MSE), and multi-model inference (MMI, or model ensembles). The concept of MMI was repeatedly noted (e.g., Thorpe et al. 2015) as an approach that merits further consideration, as it can address multiple types of uncertainty. The same is also true for MSE (e.g. Punt et al. 2014). Kinlan et al. (in Review) and Zipkin et al. (2014) provide examples of additional statistical approaches to define and portray uncertainty (as applied spatially in species distribution and abundance maps).

The MMI approach has been employed in other disciplines (e.g., weather prediction—especially hurricane prediction, climate science, and social sciences; Barnston et al. 2012, Demuth et al. 2012, IPCC 2013, Kirtman et al. 2014); doing so has helped to advance those fields of study and improve model utility. An easily recognizable example is the "cone of uncertainty" representing potential hurricane trajectories as estimated by an ensemble of models (Gall et al. 2013). MMI is also recommended for LMR and ecosystem assessment contexts (Townsend et al. 2014), with a few examples emerging (e.g., Gaardmark et al. 2013, Ianelli et al. 2015). The objective for MMI is to improve marine resource management through a more accurate assessment of the uncertainty in marine resource predictions. Further attention is warranted on how to combine multiple models with different structures, parameters, analytical engines, underlying functional forms and theoretical assumptions. Certainly there are biases in any model given these considerations, yet employing MMI is a means to address these factors. We tend to treat each LMR assessment as a final model with outputs, albeit even with some measure of statistical error, but do not formally consider other models as an ensemble. Even one model with different assumptions and parameterizations (e.g., sensitivity analysis of one model as a limited form of MMI) does not always formally have results and formulations routinely reported on. Thus, MMI affords an opportunity to address some of the issues leading to different elements of model uncertainty. There would need to be minimum performance standards for inclusion of models in any such multi-model ensemble and clear protocols for how to assign weightings to different models in an ensemble if the models are to be averaged in any way. Typically any such weightings are based on model skill, but there are many nuances to skill definition (see below) and also other factors to consider (Burnham and Anderson 2003). Ensemble modeling approaches in other disciplines are based primarily on averaging methods. These methods include: simple means, means with individual bias corrections, means with collective bias corrections, regularization, and Bayesian Model Averaging (Townsend 2014) and so on, with an exploration of best practices from other disciplines (e.g., climate, flooding, hurricanes, etc.; Cornuelle et al. 2014, Townsend et al. 2014) in the form of MMI case studies was noted as something clearly worth pursuing. A particular challenge will be in communicating the results of MMI into a regulatory environment in which there is a requirement for clear documentation of the scientific basis for LMR management limits. If a range of model types was executed, perhaps extra caution would be needed when single-species predictions and those from other models diverge.

The MSE approach also affords an opportunity to address many types of model uncertainty (de la Mare 1986, Smith et al. 1999, Punt et al. 2014). The primary purpose of MSE is to objectively identify trade-offs in achieving multiple and different management objectives. Part of that is the exploration of different configurations of how the ecosystem and LMR dynamics actually occur, so as to develop robust strategies under a range of conditions and assumptions. MSEs have been applied in stock-focused LMR assessment contexts (Punt and Donovan 2007, A'mar et al. 2009), and there is increasing ability to execute such approaches for multiple species models (Ianelli et al. 2015) and full ecosystem models (Fulton et al. 2011a, b). The value of this approach is that it can couple the full range of the NOAA modeling efforts by using more complicated, inclusive models as "operating" models against which the performance of "estimation" models and management systems such as those nearer the left of Fig. 1 can be tested. For instance, the performance of time-varying random process error in single stock assessment models can be tested against ecosystem models capable of generating realizations of the true processes occurring in nature. This approach has strong potential to simulate and test a range of direct model uncertainties, but also explicitly deals with a lot of model-associated uncertainties like communication and outcome uncertainties.

Summary of Group Discussion

- There are many types of model uncertainty.
- Not all uncertainty is direct model uncertainty. Much is model-associated uncertainty.
  - Careful use of terminology is warranted when discussing model uncertainty.
- Climate/weather forecasts are informational, whereas LMR forecasts/models are regulatory and directly impact how the resource will be used. That distinction may help differentiate which approaches of dealing with uncertainty, and level of rigor, are most appropriate.
- There are many methods for dealing with model uncertainty.
- Multiple model inference is an important area to explore to address multiple forms of uncertainty.
  - Models included in MMIs must meet certain requirements to be brought into the ensemble. Those standards would need to be developed by the modeling community.
  - One type of MMI inference is among models of similar complexity and scope, such as among the various models used for hurricane forecasting.
  - Another type of MMI is in the nesting of models of different complexity. For example, using large-scale ecosystem models to design fishery harvest strategies and then more tactical, simpler approaches to implement year-to-year implementation of that strategy.
  - An extended exploration and discussion of model weighting is warranted.
  - Differing philosophies and underlying assumptions behind each of the models in an ensemble can be beneficial.
- Management strategy evaluation is also an important area to explore to handle multiple forms of uncertainty.
  - MSEs that couple ecosystem and stock dynamics are highly desirable.
  - MSEs that also include socio-economics are even more desirable.
  - Formal evaluation of the trade-offs in different modeling assumptions and biases is usefully explored via MSEs.
  - Formal evaluation of the trade-offs in meeting management objectives is the main goal of MSEs, and should be adopted more frequently than it is.
- Resources for model development have been limited. There has been a struggle to develop single models, let alone multiple models.

Best practices

- Use MMI to bracket, include different perspectives on, compare mechanisms, and identify main sources of uncertainty. This is particularly important where predictions may qualitatively diverge.
- Use MSE to bracket, communicate, and explore consequences of uncertainty.
- Inform “epsilons” (i.e., statistical error) in simpler LMR assessment models by using more complicated models to help define epsilons and obtain covariance structure.
- Develop consistency in modeling protocols and approaches for handling uncertainty.
- Implement tighter (i.e., consistent, streamlined) reporting of model outputs into a standard format for LMR and Ecosystem assessments.

Recommendations

- Establish guidelines for uniform application of MMI. Convene experts to address levels of rigor necessary (and uncertainty tolerance) for different model types and model applications. An outline is presented in Table 1.
- Establish two to three pilot projects of MMI, similar to Gaardmark et al. 2013 and Ianelli et al. 2015.
- Increase widespread development of MSE capacity.

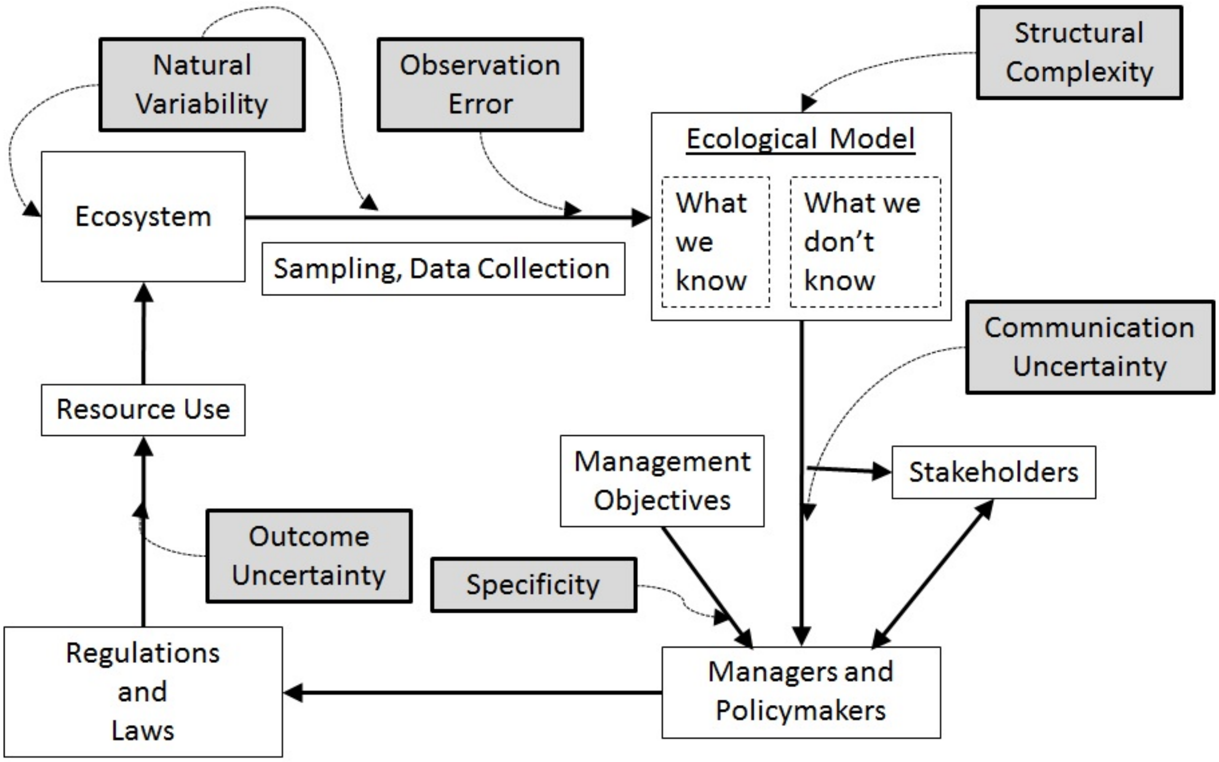

Figure 2. Main types of modeling uncertainty in a typical LMR management system (adapted from Peterman 2004; Link et al. 2012). These uncertainties are represented by the ellipses as part of an LMR management system, natural variability (or process uncertainty), observation error (or estimation uncertainty), structural complexity (or model uncertainty), communication uncertainty, objective uncertainty (lack of clarity on goals and objectives; this is often lumped in with outcome uncertainty), and outcome uncertainty (or management performance uncertainty).

Table 1. Example of a table to contrast factors contributing to model uncertainty, the type of model application, and recommended level of addressing both for each type of uncertainty. Such a table would need to be filled out by experts; an example is provided here to compare across heuristic, regulatory, and informational applications.

Application or Consideration/ Uncertainty	**Process Uncertainty**	**Estimation Uncertainty**	**Structural Uncertainty**	**Communication Uncertainty**	**Outcome (Objective) Uncertainty**
Tactical LMR & Ecosystem Mgt					
Quota, ACL, ABC, etc.					
Forecast					
Status Determination					
Include Ecosystem considerations					
Strategic LMR & Ecosystem Mgt					
Tradeoff Evaluation					
Compare Objectives					
Incorporate Climate Forecast					
Socio-economic outcomes					
Heuristic Use	Hi	Mod	Lo	Mod	Lo
Regulatory	Hi	Hi	Mod	Hi	Hi
Informational	Lo	Lo	Lo	HI	Mod
Spatial Extent					
Spatial Resolution					
Temporal Extent					
Temporal Resolution					
Taxonomic Extent					
Taxonomic Resolution					
Disciplinary focus					
Physio-chemical					
Bio-ecological					
Socio-economic					

Model skill evaluation

Generally when we assess skill we are asking: how well does the model represent the “truth” over a specified range of conditions? Model skill assessments seek to provide an objective evaluation of how well the model forecast (or nowcast) performs when compared to a reference set of conditions (typically, observations). Unfortunately, the "true conditions" against which models can be compared are not known nearly as well for LMR and ecosystem models as they are for climate and meteorological models. The tendency with LMR models has been to evaluate their performance according to how well the model compares to historical data. However, historical data can have many inadequacies in representing the past state of the resource, so skill in matching these data does not always equate with skill in forecasting future states of the resource (Legault 2009, Deroba 2014, Brooks and Legault in press). Another approach is to create an artificial simulated reality and then to test performance of a proposed approach against multiple realizations of that artificial reality, for example testing the performance of simple aggregated models against spatially structured realities (McGilliard et al. 2014).

Evaluation of model skill has multiple facets. Typically these revolve around very quantitative and statistical measures of either model fit to historical data or model forecast. Perceptions of skill are often subjective, even with quantitative measures thereof, and this invokes the issues associated with communication uncertainty. Systematically evaluating model skill and cogently presenting those results goes a long way to addressing many facets of model uncertainty (e.g., Lynch et al. 2009, Stow et al. 2009, Zhang et al. 2010, Adams and Higdon 2012).

Again, the main reason we develop these LMR and ecosystem assessment models is as decision support tools. Quantifying the uncertainty in the information provided by these models factors into the decision-making process. Given the complexity of most of these models, it is unlikely that we can execute a full uncertainty or sensitivity analysis for them, but given the emergence of novel methods, we can conduct rigorous skill assessments that reveal key uncertainties for quantities and temporal or spatial scales of interest (Stow et al. 2009, Olsen et al. in review).

There are many measures of model skill. For example, measures of skill can include various types of correlation coefficients between model predictions and observations, measures of departure from accuracy in terms of central tendency and trend, measures of bias, modeling efficiency, and many others (Allen and Somerfield 2009, Stow et al. 2009). There are also some semi-quantitative and even qualitative evaluations of model skill, but these are not typically recommended unless there is very limited observation data with which to compare the model. All these measures inform different aspects of model skill. The key point from the discussions, which reinforces emerging conclusions in the literature, is that model skill is evaluated relative to a specific set of quantities of interest (Adams and Higdon 2012), and that multiple measures of skill are needed to fully evaluate model performance (Stow et al. 2009).

Concerns of extrapolation outside the range of observations (which is a common issue for climate change applications) were also discussed. Statistical or empirical approaches, which often lack mechanistic reality, frequently perform as well as specified complex 3D models (e.g., for El Nino, Barnston et al. 2012). If a complex model does not do any better than a long-term average, it would have poor skill. However, an empirical model may not perform well when predicting outside of the range of data. Conversely, a complex mechanistic model that does not perform any better than an empirical model within the range of observations could potentially capture the underlying dynamics of the modeled system and generate more reasonable extrapolations outside the range of observations. How to move beyond data assimilation techniques and evaluations of model skill in limited data situations or situations beyond the extent of collected data remains an important area of research. In particular for fisheries and ecosystems, most of our data come from the past 30 to 50 years and represent ecosystems in transition for lightly fished systems to more extensive levels of harvest to which the systems have not yet fully adjusted. Thus the full range of potential compensations may not yet have been observed. In the LMR and ecosystem assessment context, the ability to hindcast is an important aspect of model skill that also merits continued attention, although hindcast skill does not necessarily imply forecast skill (Hastie et al. 2009, Fienen and Plant 2015).

One method of using MSEs is to develop test datasets for testing the performance of simpler models. In physical systems that have more direct observations of the system, one can use model skill against observational datasets. Absolute model skill is easier to evaluate when there are direct measures of the quantities of interest from the modeled system, as in physical systems such as weather or hurricane forecasting (e.g., Hamill and Juras 2006). This approach is not available for the vast majority of living marine resource models. Data simulation would be necessary for LMR model skill evaluations. Thus, using complex models with more interacting parts that can be used to generate test environments for simpler models seems a viable approach.

A repeated observation at the workshop was that the application of model skill criteria in the published literature is inconsistent. There is no clear standard or protocol generally for this broad class of models, certainly not for any LMR or ecosystem modeling. There appears to be a gradient in maturity of skill assessment methods (e.g., tides perhaps most mature, harmful algal bloom forecasts needing more work; Hess et al. 2003, Zhang et al. 2010), and certainly the criteria of what is acceptable skill will differ depending upon on the application of the model and the quantities of interest (Adams and Higdon 2012). Developing basic criteria for model performance would be helpful to lend consistency to review processes, especially as new and more complex models are developed for use in management contexts. Further, performance criteria could be used to determine which models might best be assembled into an ensemble for management decision-making.

Beyond identifying direct model issues and diagnostics, model skill helps address other model-associated facets of uncertainty. For instance, model skill helps communicate to users, stakeholders, and resource managers how well the model works. It can also provide feedback to data collection and monitoring programs.

Summary of Group Discussion

- Model skill is the performance of particular model fit and forecast in comparison to a reference set of information.
- Model skill is important to convey the confidence of model performance.
    - Communicating model skill simultaneously addresses many facets of model uncertainty.
- Different modeling approaches have unique strengths and complementary values. To use models as an ensemble suite, there need to be some common skill metrics.
- There are many measures of model skill.
- Multiple measures are necessary to fully characterize model skill.

- Testing models with test-bed dataset (often generated by more complex models) is an increasingly viable approach.
- Standards for model skill should consider that:
  - There are several "best practices" publications, but not regarding skill assessment.
  - They are likely different across model applications.
  - Some minimal performance standards for use in LMR and ecosystem assessments are warranted.
- Continuing to collect data with which to validate models is obvious, but important.
  - Feedback to monitoring systems should be made more explicit.

Best practices

- Utilize multiple metrics of quantitative model skill.
- Adopt Verification, Validation, and Uncertainty Quantification (VVUQ; NRC 2012) standards.
- Utilize appropriate diagnostic and visualization tools when reporting model skill.
- Communicate regularly among models and observing systems.

Recommendations

- Document model skill for all LMR and ecosystem assessment models.
- Establish guidelines to determine Minimal Performance Standards of skill criteria for the different levels of model application.

Human Dimensions

Many challenges with model uncertainty have little to do with the direct quantitative facets of the model. Rather, they stem from the human perceptions of uncertainty, how familiar users and stakeholders are with the model, and how communication about the model, its outputs, and methods contribute to that perception (Spiegelhalter et al. 2011). Specifically, communication transmits information, data, and findings, but *how* communication is done also enables or inhibits the building of trust, credibility, and legitimacy in the science, models, and the scientific and modeling process. In other words, this human domain of model uncertainty has limited responsiveness to increases in Mohn's rho, R2, MEF, RMSE. Rather, it has almost everything to do with how well users, stakeholders, managers, and other partners perceive the model and the modeler and how we might communicate about the model, Mohn's rho, R2, MEF, or RMSE (Fulton et al. 2011b).

These types of considerations are not typically the priority consideration among natural scientists executing these modeling exercises. Fortunately this workshop had active participation by several social scientists expert in human perception and cognition. There are several Social Science fields that study decision-making under conditions of uncertainty. These include: decision sciences, risk perception (psychometrics), behavioral economics, communication science, and public policy, among others (e.g., Kahneman et al. 1982, Rosenberg and Restrepo

1994, Slovic 2000, Eggert and Martinsson 2004, Fletcher 2005, Kompas et al. 2008). Learning key messages from those fields will continue to be invaluable for the modeling community.

How do people, particularly non-experts, perceive and respond to uncertainty? All people have and will apply their cognitive biases or heuristics. For instance, people tend to reject information that does not support their pre-existing values and world view, which are defined in large part by their social context (e.g., their groups, professional norms, cultural norms, etc.). Attempting to address these biases directly through education has shown that the "deficit model" is not effective—i.e., educating the public or stakeholders to better understand science (science literacy) will not lead to fewer controversies or disagreements over science, models, and uncertainty. People yearn for predictability, can be uncomfortable with ambiguity, and can go to extraordinary cognitive efforts to seek patterns in information and establish predictability. People tend to jump to conclusions, rapidly binning new information with familiar stereotypes and categories, such that first impressions are very important and often hard to overcome. Context also matters in how different audiences perceive uncertainty. For example, the general public can interpret the term "uncertainty" as ignorance and the term "error" as mistake (c.f. Table 2; Johnson and Slovic 1995). Or resource managers, fishermen, NGOs, and other stakeholders who might be more knowledgeable about the modeling may respond to scientific information, uncertainties, etc. in different ways when they are sitting at a council table in a political, public setting than they might in the hallway over a coffee break, in a non-public, less decision-making context. Trust in and credibility of the messenger impacts how people perceive the information and uncertainty. The source, context, and presentation of the information matter.

How can LMR and ecosystem modelers effectively communicate about uncertainty given these inherent challenges and known biases? Recognizing that different audiences have different perspectives, will make different assumptions, and communicate differently is a key first step. Accounting for these known cognitive biases and group norms is one way to more effectively communicate. For example, developing narrative stories appropriate for a given culture helps effective communication. It is clear that the modeling community would be wise to develop protocols with principles for reporting these various sources of uncertainty, particularly such that frameworks are established to allow our partners and the public to easily understand highly technical information.

It is critical to identify and characterize the range and diversity of audience segments (e.g., even within the LMR-associated community), as multiple sectors will have different social norms and values, different preferences for receiving information, different foundations of knowledge, etc. Tailoring messages and communication tools for each audience segment is therefore prudent. While no single set of answers will work in all situations and for all stakeholders, the climate change communication science community has developed tips on how to convert scientific terminology into more common language (Table 3; Somerville and Hassol 2011, Johnson 2012, Fiske and Dupree 2014).

Since trust and credibility of the messenger is important, there are strategies for increasing trust and credibility in the scientist. For example, listening well and asking follow-up, clarifying questions of stakeholders illustrate that a scientist is genuinely interested and respects the opinion and knowledge of the stakeholder. A person is more likely to be given respect and credibility if he shows respect and credibility to others.

Further, context matters in many ways in human interactions and that is clearly true among fishermen, scientists, managers, NGOs and others in LMR management. In a fishery management council meeting the stakes are high, the allocation of fishery quota and viability of livelihoods are made, and stakeholder interests can be in competition with one another. Communication can take very different forms at the council table versus in the hallway during a coffee break. Changing the context can provide a more constructive, safer, less threatening setting for listening and learning about science and models. For example, fishery management council training programs, stakeholder training programs (like the Marine Resource Education Project), collaborative fisheries research, etc. provide venues for the more iterative dialogue that helps overcome biases and builds greater trust and credibility (Hartley and Robertson 2009, Hartley and Glass 2010). Fisheries-specific tactics have been developed to increase openness and transparency in the scientific and modeling process, which in turn increases credibility of the outputs and provides venues to promote engagement between scientists and stakeholders that build trust between the two communities.

Table 2. Procedures for addressing uncertainty in fisheries management deliberations. From Hartley in press and Dankel et al. 2012.

Technique:	**Description:**
Pedigree analysis	Multi-criteria, qualitative characterization of the origins and status of information and data.
Uncertainty matrix	Classification method where a panel of experts numerically rate the nature and scale of the uncertainty on several defined parameters
Extended peer-review	Involving multiple disciplines and stakeholder perspectives on a peer-review panel
Incorporating and respecting traditional or local knowledge	Participatory approaches and incorporate traditional or local knowledge, e.g., Q-Method is based upon the conceptual framework of factor analysis, seeking correlations between variables. The Q-Method is concerned with individuals' viewpoints, seeking shared views or correlations across a sample of individuals and clarification on points of agreement and disagreement.
Participatory modeling	Facilitated, structured dialogue about uncertainty and the quality of the state of knowledge among scientists and stakeholders to enhance scientific understanding.
Collaborative research	Joint development, design and implementation of scientific monitoring or research activities. Collaborative at all stages of the scientific process from developing questions, through design and implementation, to communicating the findings and results.

Additional tactics being used in LMR management includes Management Strategy Evaluation, and a range of Visualization tools (e.g., Kemp and Meaden 2002, Mayer et al., 2002, Holland 2010).

While effectively communicating complex scientific and technical information to non-experts can be daunting and requires patience and targeted effort, it is essential to develop the credibility in the message and the messenger, build trust, and establish legitimacy of the modeling and management process in the eyes of the managers, fishermen, NGOs, and other stakeholders in LMR management. There are a wide range of tools and approaches to assist with communicating model uncertainty, and the modeling community would be wise to increasingly adopt these tools and enter partnerships with the social science disciplines that develop them.

Summary of Group Discussion

- Considerable uncertainty surrounding LMR and ecosystem models arises from the way humans perceive and communicate about risk and uncertainty.
- There are many biases and factors that influence perception when communicating about model uncertainty.
    - There is a challenging balance and some disagreement among social scientists on the impact of focusing on uncertainty to varying degrees. Communicating about uncertainty in an open and transparent manner builds trust and credibility in the messenger, which in time will engender trust and credibility in the model. But before the messenger is accepted as trustworthy, the discussion of scientific or model uncertainty may be perceived by some as ignorance of the issue.
    - Uncertainty can be in the eyes of the beholder—e.g., a scientist may think talking about uncertainty provides a deeper understanding of the natural phenomenon, whereas a distrustful fisherman may interpret it as the scientist not knowing what they are talking about. Yet, one needs to talk about uncertainty to be open, transparent, and build trust with the fishermen.
- The modeling community should consider whether we should use our best storyteller or our best technical expert in key communication contexts. This decision should be driven by the target audience.
- Knowing one’s audience is the paramount rule (what information do they need, how do they make decisions, what are their concerns, etc.?). There are readily available mechanisms that can help assess particular audiences of interest.
- Addressing/communicating uncertainty is essential, and fortunately there are many extant tools to assist in doing so and social scientists and practitioners (e.g., Sea Grant) willing to partner on advancing effective communication.

Table 3. Examples of different perspectives on technical terminology by scientists and the public. From Somerville and Hassol 2011.

Terms that have different meanings for scientists and the public		
Scientific term	**Public meaning**	**Better choice**
enhance	improve	intensify, increase
aerosol	spray can	tiny atmospheric particle
positive trend	good trend	upward trend
positive feedback	good response, praise	vicious cycle, self-reinforcing cycle
theory	hunch, speculation	scientific understanding
uncertainty	ignorance	range
error	mistake, wrong, incorrect	difference from exact true number
manipulation	illicit tampering	scientific data processing
scheme	devious plot	systematic plan

Best practices

- Know your audience.
    - First consider the audience and the ways that it best receives information.
    - Maintain consistency in message but mechanism of delivery may be quite different based on the audience.
- Tell stories.
- Build trust by taking advantage of as many face-to-face interactions as is feasible—both those we create and those ad hoc opportunities we attend.
    - Putting a face to the message improves public trust, regardless of who is delivering the message.
    - Seek connections, commonalities.
- Practice with family/friends who are not experts in the field.
- Focus on aspects that people care most about—what makes this relevant to someone, why do they have a stake in the outcome?

Recommendations

- Explore communication training options for the modeling community.
- Establish venues for further interaction among communications and cognitive experts with the modeling community.
- Codify protocols for effective reporting on various sources of model output and uncertainty.

Summary and Conclusions

There are many similarities between NMFS and OAR modeling efforts, including LMR assessments, water quality modeling, habitat modeling, and ecosystem assessments. Other partners, including NOS modelers, also provided highly comparable perspectives. This workshop capitalized on those commonalities and leveraged a significant body of work to explore model uncertainty. As the need for LMR and ecosystem models continues to grow, ensuring these models are credibly received and utilized remains an important task.

The suite of best practices identified at the workshop will well serve ongoing and future efforts to handle model uncertainty. A key outcome of this workshop includes a catalogue of best practices to address model uncertainty that describes methods to improve NOAA's LMR and ecosystem modeling enterprise.

The enhanced networking and cross-disciplinary perspectives that arose during workshop deliberations is an important outcome to recognize. It was clear that continued coordination on common issues is warranted. This was a rare meeting in which social scientists were fully engaged and integrated in the planning and discussions, which emphasized the need to continue to engage social scientists to better communicate about model uncertainty. Additional outcomes include collaborations, future/enhanced/ongoing innovations, research, and operational development that will continue to pay dividends in the future.

The workshop made clear that, despite the many lessons learned, no "silver bullet" exists to address uncertainty across all dimensions of the various models and types of uncertainty. The suite of best practices noted here will go a long way to address model uncertainty. Even something as simple as a "cheat sheet" of when to use certain approaches to address specific types of uncertainty across model applications, a "template" of standard output reporting for LMR and ecosystem assessment outputs (e.g., ICES 2010), and common usage of quantitative information to explore model skill would be beneficial.

Key observations centered on the use of multi-model inference, management strategy evaluation, and improved use of communication tools. Those items form the bulk of our best practices and recommendations. It bears repeating that Management Strategy Evaluation (MSE) and Multi-model Inference (MMI) present significant utility to the modeling community.

We summarize the best practices and recommendations from this workshop below and note that cross-disciplinary, cross-organizational meetings like this workshop to coordinate and advance modeling efforts are very beneficial and should continue.

Summary of Best practices

Uncertainty

- Use MMI to bracket, include different perspectives on, compare mechanisms, and identify main sources of uncertainty.

- Use MSE to bracket, communicate, and explore consequences of uncertainty.
- Inform "epsilons" (i.e., statistical error) in simpler LMR assessment models by using more complicated models to help define epsilons and get covariance structure.
- Develop consistency in modeling protocols and approaches for uncertainty handling.
- Implement consistent, streamlined reporting of model outputs into a standard format for LMR and ecosystem assessments.

Model Skill

- Use multiple metrics of quantitative model skill.
- Adopt VVUQ (NRC 2012) standards.
- Use appropriate diagnostic and visualization tools when reporting model skill.
- Communicate regularly among modelers and observing systems.

Communication

- Know your audience.
- Tell stories.
- Build trust by adopting as many face-to-face interactions as is feasible.
- Practice with family/friends who are not experts in the field.
- Focus on aspects that people care most about.

Summary of Recommendations for Addressing Modeling Uncertainty

- Establish guidelines for uniform application of MMI.
- Establish 2 to 3 pilot projects of MMI.
- Increase widespread development of MSE capacity.
- Document model skill for all LMR and ecosystem assessment models.
- Establish guidelines to determine Minimal Performance Standards of skill criteria for the different levels of model application.
- Explore communication training options for the modeling community.
- Establish venues for further interaction among communications and cognitive experts with the modeling community.
- Codify protocols for effective reporting on various sources of model output and uncertainty.

Next steps: Major Recommendations to NOAA leadership

1. Seek NOAA leadership support for a full range of quantitative modeling efforts across a spectrum of complexity and disciplinary emphasis in support of living marine resource management mandates.
2. Establish routine and regular venues for the NOAA modeling community to meet and interact.
3. Support and advance cross-line-office and cross-disciplinary (including social science) coordination on this issue.

References

Adams, M., and Higdon, D. 2012. Assessing the reliability of complex models: mathematical and statistical foundations of verification, validation, and uncertainty quantification. The National Academies Press, Washington DC. 1 p. http://www. nap. edu/catalog.php

Allen, J. I., and Somerfield, P. J. 2009. A multivariate approach to model skill assessment. Journal of Marine Systems.

A'mar, Z.T., A.E. Punt, and M.W. Dorn. 2009. The evaluation of two management strategies for the Gulf of Alaska walley Pollock fishery under climate change. ICES Journal of Marine Science 66(7): 1614-1632.

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.

Barnston, A.G., M.K. Tippett, M.L. L'Heureux, S. Li, D.G. DeWitt. 2012. Skill of Real-Time Seasonal ENSO Model Predictions during 2002-11: Is Our Capability Increasing? Bull. Amer. Meteor. Soc., 93, 631-651.

Brooks, E.N. and C.M. Legault. (in press). Retrospective forecasting - evaluating performance of stock projections for New England groundfish stocks. Canadian Journal of Fisheries and Aquatic Sciences.

Burnham, K.P. and D.R. Anderson. 2003. Model Selection and Multimodel Inference. Springer.

Charles, A.T. 1998. Living with uncertainty in fisheries: analytical methods, management priorities and the Canadian groundfishery experience. Fisheries Research 37: 37-50.

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-+.

Cloyd, E.T., A.P. Leonardi, D.L. Scheuer, and E.J. Turner. 2007. Establishing National Ocean Service priorities for estuarine, coastal, and ocean modeling: capabilities, gaps, and preliminary prioritization factors. NOAA Technical Memorandum NOS-NCCOS 57.

Cornuelle, B., J. Hansen, B. Kirtman, S. Sandgathe, and S. Warren. 2014. Issues and Challenges with Using Ensemble-Based Prediction to Probe the Weather–Climate Interface. Bull. Amer. Meteor. Soc., 95, ES213–ES215.

Dankel, D. J., R. Aps, G. Padda, C. Rockmann, J. P. van der Sluijs, D. C. Wilson and P. Degnbol. 2012. Advice under uncertainty in the marine system. *ICES Journal of Marine Science* 69(1):3-7.

de la Mare, W. K. 1986. Simulation studies on management procedures. Reports of the International Whaling Commission, 36: 429 – 450.

Demuth, J.L., R.E. Morss, B.H. Morrow, and J.K. Lazo. 2012. Creation and Communication of Hurricane Risk Information. Bull. Amer. Meteor. Soc. 93: 1133–1145.

Deroba, J.J. 2014. Evaluating the consequences of adjusting fish stock assessment estimates of biomass for retrospective patterns using Mohn’s rho. North American Journal of Fisheries Management 34(2): 380-390. DOI: 10.1080/02755947.2014.882452

Deroba, J.J., D.S. Butterworth, R.D. Methot, J.A.A. De Oliveira, C. Fernandez, A. Nielsen, S.X. Cadrin, M. Dickey-Collas, C.M. Legault, J. Ianelli, J.L. Valero, C.L. Needle, J.M. O’Malley, Y-J. Chang, G.G. Thompson, C. Canales, D.P.Swain, D.C.M. Miller, N.T. Hintzen, M. Bertignac, L. Ibaibarriaga, A. Silva, A. Murta, L.T. Kell, C.L. de Moor, A.M. Parma, C.M. Dichmont, V.R. Restrepo, Y. Ye, E. Jardim, P.D. Spencer, D.H. Hanselman, J. Blaylock, M. Mood, and P.-J. F. Hulson. 2015. Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. ICES Journal of Marine Science. 72(1), 19–30. DOI:10.1093/icesjms/fst237.

Eggert, H., and P. Martinsson. 2004. Are commercial fishers risk-lovers? *Land Economics* 80: 550-560.

FAO (Food and Agriculture Organization of the United Nations). 2008. Fisheries Management. 2. The ecosystem approach to fisheries. 2.1 Best practices in ecosystem modeling 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).

Ferson, S. and L.R. Ginzburg. 1996. Different methods are needed to propagate ignorance and variability. Reliability Engineering and System Safety 54: 133-144.

Fienen, M.N., and N.G. Plant. 2015. A cross-validation package driving Netica with python. Environmental Modeling and Software, 63: 14–23. 10 pp

Fiske, S.T. and C. Dupree. 2014. Gaining trust as well as respect in communicating to motivated audiences about science topics. *Proceedings of the National Academy of Sciences*. 111(Supplement_4):13593-13597.

Francis, R.I.C.C., and R. Shotton. 1997. ‘Risk’ in fisheries management: a review. Canadian Journal of Fisheries and Aquatic Sciences 54: 1699e1715.

Fletcher, W.J. 2005. The application of qualitative risk assessment methodology to prioritize issues for fisheries management. *ICES Journal of Marine Science* 62 1576-1587.

Fulton, E.A., J.S. Link, I. Kaplan., M. Savina-Rolland, P. Johnson., C. Ainsworth, P. Horne, R. Gorton, R.J. Gamble, A.D.M. Smith, D.C. Smith. 2011a. Lessons in modeling and management of marine ecosystems: The Atlantis experience. *Fish and Fisheries* 12:171-188.

Fulton, E.A., A.D.M. Smith, D.C. Smith, I.E. van Putten. 2011b. Human behaviour: the key source of uncertainty in fisheries management. *Fish and Fisheries* 12:2-17.

Gaardmark, A., M. Lindegren, S. Neuenfeldt, T. Blenckner, O. Heikinheimo, B. Muller-Karulis, S. Niiranen, J.T. Tomczak, E. Aro, A. Wikstrom, C. Mollman. 2013. Biological ensemble modeling to evaluate potential futures of living marine resources. *Ecological Applications* 23:742-754.

Gall, R., J. Franklin, F. Marks, E.N. Rappaport, and F. Toepfer. 2013. The Hurricane Forecast Improvement Project. Bull. Amer. Meteor. Soc., 94, 329–343.

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

Hamill, T. M., and Juras, J. 2006. Measuring forecast skill: is it real skill or is it the varying climatology? Quarterly Journal of the Royal Meteorological Society, 132: 2905–2923. Wiley Online Library. 19 pp.

Hartley, T.W. (in press) In the Eye of the Beholder: Scientific Uncertainty and Information Flow in Coastal and Marine Governance Networks. In B.H. MacDonald (ed) *Does Information Matter? A critical question for the future of coastal zone management.* Taylor & Francis.

Hartley, T.W., and C. Glass. 2010. Science-to-management pathways in US Atlantic herring management: using governance network structure and function to track information flow and potential influence. ICES Journal of Marine Science. 67(6):1154-1163.

Hartley, T.W. and R.A. Robertson. 2009. Stakeholder Collaboration in Fisheries Research: Integrating Knowledge Among Fishing Leaders and Science Partners in Northern New England. Society and Natural Resources. 22(1):42-55.

Harwood, J., and K. Stokes. 2003. Coping with uncertainty in ecological advice: lessons from fisheries. Trends in Ecology and Evolution 18(12): 617-622.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. The Elements of Statistical Learning. Springer-Verlag. 763 pages.

Hawkins, E., and R. Sutton. 2009. The Potential to Narrow Uncertainty in Regional Climate Predictions. *Bull. Amer. Meteor. Soc.* 90: 1095–1107.

Hess, K.W., T.F. Gross, R.A. Schmalz, J.G.W. Kelley, F. Aikman, E. Wei, M.S. Vincent. 2003. NOS Standards for Evaluating Operational Nowcast and Forecast Hydrodynamic Model Systems. NOAA Technical Report NOS CS 17. 48 pp.

Hilborn, R. 1987. Living with Uncertainty in Resource Management. North American Journal of Fisheries Management 7(1): 1-5.

Hill, S.L., G.M. Watters, A.E. Punt, M.K. McAllister, C. Le Quere, J. Turner. 2007a. Model uncertainty in the ecosystem approach to fisheries. Fish and Fisheries 8: 315–336.

Holland, D.S., 2010. Management Strategy Evaluation and Management Procedures: Tools for Rebuilding and Sustaining Fisheries. OECD Food, Agriculture and Fisheries Working Papers, No. 25. OECD Publishing. http://www.oecd.org/greengrowth/fisheries/45497984.pdf

IPCC. 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Ianelli, J., K.K. Holsman, A.E. Punt, K. Aydin. 2015. Multi-model inference for incorporating trophic and climate uncertainty into stock assessments. Deep Sea Research Part II: Topical Studies in Oceanography. http://dx.doi.org/10.1016/j.dsr2.2015.04.002

ICES. 2010. Report of the Working Group on Multispecies Assessment Methods (WGSAM), 4-8 October 2010, San Sebastian, Spain. ICES CM 2010/SSGSUE:05. 95 pp.

Johnson, B.B. 2012. Climate Change Communication: A Provocative Inquiry into Motives, Meanings, and Means. Risk Analysis. 32(6):973-991.

Johnson, B.B. and P. Slovic. 1995. Presenting uncertainty in health risk assessment: initial studies of its effects on risk perception and trust. Risk Analysis 15(4): 485-494.

Kahneman, D., P. Slovic, and A. Tversky. 1982. *Judgment under uncertainty: Heuristics and biases.* New York NY: Cambridge University Press.

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-+.

Kemp, Z. and G. Meaden. 2002. Visualization for fisheries management from a spatiotemporal perspective. ICES Journal of Marine Science. 59(1):190-202.

Kinlan, B.P., M. Poti, A.F. Drohan, D.B. Packer, D.S. Dorfman, and M.S. Nizinski. 2015. Predictive modeling of suitable habitat for deep-sea corals offshore of the northeast United States. *Deep Sea Research Part I: Oceanographic Research Papers.* In review.

Kirtman, B.P., D. Min, J.M. Infanti, J.L. Kinter, III, D.A. Paolino, Q. Zhang, H. van den Dool, S. Saha, M.P. Mendez, E. Becker, P. Peng, P. Tripp, J. Huang, D.G. DeWitt, M.K. Tippett, A.G. Barnston, S. Li, A. Rosati, S.D. Schubert, M. Rienecker, M. Suarez, Z.E. Li, J. Marshak, Y.-K. Lim, J. Tribbia, K. Pegion, W.J. Merryfield, B. Denis, and E.F. Wood. 2014. The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction. Bull. Amer. Meteor. Soc., 95, 585–601.

Kompas, T., T.N. Che, and R.Q. Grafton. 2008. Fisheries instrument choice under uncertainty. Land Economics 84: 652-666.

Legault, C.M. (Chair). 2009. Report of the Retrospective Working Group, January 14-16, 2008, Woods Hole, Massachusetts. US Dept Commerce, Northeast Fish Sci Cent Ref Doc. 09-01; 30 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://www.nefsc.noaa.gov/nefsc/publications/

Leonardi, A.P., K. Boyd, L.D. McCullough. 2009. A Summary of NOAA's Environmental Modeling Activities. NOAA White Paper.

Link, J.S., J.K.T. Brodziak, S.F. Edwards, W.J. Overholtz, D. Mountain, J.W. Jossi, T.D. Smith, and M.J. Fogarty. 2002. Marine ecosystem assessment in a fisheries management context. Canadian Journal of Fisheries and Aquatic Sciences 59(9):1429-1440.

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

Link, J.S., T.F. Ihde, C.J. Harvey, S.K. Gaichas, J.C. Field, J.K.T. Brodziak, H.M. Townsend, R.M. Peterman. 2012. Dealing with uncertainty in ecosystem models: The paradox of use for living marine resource management. Progress in Oceanography 102: 102-114.

Lynch, D.R., D.J. McGillicuddy, F.E. Werner (editors). 2009. Skill assessment for coupled biological/physical models of marine systems. Journal of Marine Systems, Volume 76, Issues 1–2, Pages 1-250.

Mangel, M. 2006. Commentary: accounting for uncertainty in marine reserve design. Ecology Letters 31: L07303.

Mayer, L, Y. Li, and G. Melvin. 2002. 3D Visualization for Pelagic Fisheries Research and Assessment. ICES Journal of Marine Science. 59(1): 216-225.

McElhany, P., E.A. Steel, K. Avery, N. Yoder, C. Busack, and B. Thompson. 2010. Dealing with uncertainty in ecosystem models: lessons from a complex salmon model. Ecological Applications 20: 465-482.

McGilliard, C.R., A.E. Punt, R.D. Methot, Jr., and R. Hilborn. 2015. Accounting for marine reserves using spatial stock assessments. Canadian Journal of Fisheries and Aquatic Sciences 72(2): 262-280. http://dx.doi.org/10.1139/cjfas-2013-0364

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

Morgan, M.G. and M. Henrion. 1990. Uncertainty; A guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge, England, 332 pp.

NOAA. 2010. Environmental Modeling Strategic Plan: The Backbone for NOAA's Predictions and Assessments. NOAA White Paper.

National Research Council (NRC). 2007. Models in Environmental Regulatory Decision Making. Committee on Models in the Regulatory Decision Process, National Research Council. 286 pp.

NRC. 2012. Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification. Washington, D.C.: National Academies Press. 131 pp.

Olsen, E., Fay, G., Gaichas, S., Gamble, R., Lucey, S., Link, J.S. (Submitted ms, internal review, June 2015). Ecosystem model skill assessment: Yes we can! PLOS Biology.

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)

Patterson, K., R. Cook, C. Darby, S. Gavaris, L. Kell, P. Lewy, B. Mesnil, A. Punt, V. Restrepo, D.W. Skagen, G. Stefansson. 2001. Estimating uncertainty in fish stock assessment and forecasting. Fish and Fisheries 2(2): 125-157.

Peterman, R.M. 2004. Possible solutions to some challenges facing fisheries scientists and managers. ICES Journal of Marine Science 61: 1331–1343.

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.

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.

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

Punt, A.E. and G.P. Donovan. 2007. Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. ICES Journal of Marine Science 64: 603-612.

Punt, D.S. Butterworth, C.L. de Moor, J.A.A. De Oliveira, and M. Haddon. 2014. Management strategy evaluation: best practices. Fish and Fisheries. doi: 10.1111/faf.12104

Regan, H.M., M. Colyvan, and M.A. Burgman. 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. Ecological Applications 12(2): 618-628.

Rosenberg, A. and V. Restrepo. 1994. Uncertainty and risk evaluation in stock assessment advice for US marine fisheries. Canadian Journal of Aquatic Science 51: 2715-2720.

Shertzer, K.W., M.H. Prager, D.S. Vaughan, and E.H. Williams. Fishery Models. Pages 1582–1593 *in* Sven Erik Jørgensen and Brian D. Fath (Editor-in-Chief), Population Dynamics. Vol. 2 of Encyclopedia of Ecology, 5 vols. Oxford: Elsevier.

Slovic, P. 2000. *The perception of risk* London UK: Earthscan Publications, Ltd.

Smith, A.D.M., Sainsbury, K.J. and Stevens, R.A. 1999. Implementing effective fisheries-management systems –management strategy evaluation and the Australian partnership approach. ICES Journal of Marine Science 56: 967–979.

Somerville, R.C.J. and S.J. Hassol. 2011. Communicating the science of climate change. Physics Today 64(10): 48–53.

Spiegelhalter, D., M. Pearson, and I. Short. 2011. Visualizing uncertainty about the future. Science 333(6048): 1393-1400.

Stow, C.A., J. Solliff, D.J. McGillicuddy, S.C. Doney, I. Allen, M.A.M. Friedricks, K.A. Rose, P. Wallhead. 2009. Skill assessment for coupled biological/physical models of marine systems. Journal of Marine Systems 76(1-2): 4-15.

Thorpe, R.B., W.J.F. Le Quesne, F. Luxford, J.S. Collie, S. Jennings. 2015. Evaluation and management implications of uncertainty in a multispecies size-structured model of population and community responses to fishing. Methods in Ecology and Evolution 6: 49–58.

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, 80 pp.

Townsend, H.M., C. Harvey, K. Y. Aydin, R. Gamble, A. Grüss, P. S. Levin, J. S. Link, K. E. Osgood, J. Polovina, M. J. Schirripa, B. Wells (editors). 2014. Report of the 3rd National Ecosystem Modeling Workshop (NEMoW 3): Mingling Models for Marine Resource

Management – Multiple Model Inference. National Marine Fisheries Service, NOAA Technical Memorandum NMFS-F/SPO-149, 93 pp.

Zhang, A ., K. W. Hess and F. Aikman III (2010) User-based skill assessment techniques for operational hydrodynamic forecast systems, Journal of Operational Oceanography, 3:2, 11-24 online at: http://dx.doi.org/10.1080/1755876X.2010.11020114

Zipkin, E.F., B.P. Kinlan, A. Sussman, D. Rypkema, M. Wimer, and A.F. O'Connell. 2015. Statistical guidelines for assessing marine avian hotspots and coldspots: a case study on wind energy development in the U.S. Atlantic Ocean. *Biological Conservation*. http://dx.doi.org/10.1016/j.biocon.2015.06.035

Appendix A – Agenda

Day 1 – Monday April 13			Presenter or discussion leader/facilitator
12:00 -12:10	Plenary 1	Welcome, Logistics	Doran Mason
12:10-12:30		Introduction, Background Layout plans for workshop	Jason Link
12:30-2:30		**ToR 1: Types of assessment models used for LMR and ecosystem assessments**	**ToR 1 Discussion Leader: Jason Link**
12:30		• Primer on types of models used in NMFS ecosystem and living marine resource assessments	• Rick Methot and Howard Townsend
1:15		• Brief overview of models used in OAR	• Doran Mason
2:00		Discussion on models and assessments	**Jason Link**
2:30-2:45		**Coffee break**	
2:45-4:00, cont. day 2	Plenary 2	**ToR 2: Methods for addressing uncertainty**	**ToR 2 Discussion Leader: Jessie Carman**
2:45		• Types of modeling uncertainty	• Jim Ianelli and Sarah Gaichas
3:30		• Forecasting uncertainty in the physical environment: approaches and implementation	• Gregg Jacobs
3:45		• Forecasting impacts of Asian carp on the Lake Erie food web; application of expert solicitation to estimate parameter uncertainty	• Hongyan Zhang
4:00-5:00	Breakout 1	Discuss methods for addressing uncertainty in ecosystem and LMR models and assessments	**Group 1: Sarah Gaichas and Jim Ianelli** **Group 2: Howard Townsend and Rick Methot**
5:00		**Adjourn Day 1**	
6:30	Dinner	**Grizzly Peak Brewing Company**, 120 West Washington Street, downtown Ann Arbor. **Meet in the hotel lobby at 6 pm** to carpool.	

Day 2 – Tuesday April 14			
8:30	Plenary 2	Recap previous day. Layout plans for day 2	**Jason Link**
		ToR 2 continued	**Jessie Carman**
8:45		Breakout groups report on methods for addressing uncertainty (ToR 2)	**Group 1: Sarah Gaichas and Jim Ianelli** **Group 2: Howard Townsend and Rick Methot**
9:30		Group discussion on breakout group findings, identify best practices	**Jessie Carman**
10:00-10:15	**Coffee break**		
10:15-3:45	Plenary 3	**ToR 3: Model skill evaluation**	**ToR 3 Discussion Leader: Charlie Stock**
10:15		• Overview of model skill evaluation: A Reasonable Fit to Data	• Craig Stow
10:45		• Skill assessment and estimating forecast uncertainty in National Ocean Service physical and ecological models	• John G.W. Kelley
11:00		• Hierarchy of model skills and evaluation	• Yan Jiao
11:15		• Skill assessment of the Atlantis Ecosystem model	• Erik Olsen
11:30		General discussion on model skill evaluation, especially in relation to characterizing model uncertainty	**Charlie Stock**
12:00-1:00	**Lunch break**		
1:00-2:00	Breakout 2	Discuss approaches for model skill evaluation	**Group 1: Sarah Gaichas and Jim Ianelli** **Group 2: Howard Townsend and Rick Methot**
2:00-3:00	Plenary 3	Breakout groups report on model skill evaluation (ToR3)	
3:00-3:15	**Coffee break**		
3:15-3:45	Plenary 3	Group discussion on breakout group findings, identify best practices	**Charlie Stock**
3:45-4:45	Plenary 4	**Human dimensions of uncertainty**: • How do people, particularly non-experts, perceive and respond to uncertainty? • How do we effectively communicate about uncertainty and models in order to enable constructive public and management deliberations?	**Troy Hartley** • Sol Hart • Karen Akerlof
4:45 -5:00		Group discussion on human dimensions, identify major best practices	**Troy Hartley**
5:00	**Adjourn Day 2**		
Dinner on your own. Explore Ann Arbor!			

Day 3 – Wednesday April 15			
8:30	Plenary 5	Recap previous day. Layout plans for day 3.	**Jason Link**
8:45-12:00		**Capturing best practices and next steps (ToR 4)**	**ToR 4 Discussion Leader: Jason Link**
8:45		Discuss how best to tie together concepts of addressing model and assessment uncertainty, evaluating model skill, and communication/presentation thereof	
9:15	Breakout 3	Identify major lessons learned for addressing model and assessment uncertainty, evaluating model skill, and communication/presentation	**Group 1: Sarah Gaichas and Jim Ianelli** **Group 2: Howard Townsend and Rick Methot**
10:15-10:30		**Coffee break**	
10:30	Plenary 5	Develop recommendations and best practices for addressing model uncertainty for use in ecosystem and LMR assessments; next steps (ToR 4)	**Jason Link**
12:00		**Adjourn Day 3**	

Appendix B – Participants List

Invited Speakers	E-mail	Center/Office/Affiliation	Duty Station
Karen Akerlof	kakerlof@gmu.edu	George Mason Univ.	Fairfax, VA
Sarah Gaichas	sarah.gaichas@noaa.gov	NMFS/NEFSC	Woods Hole, MA
Sol Hart	solhart@umich.edu	University of Michigan	Ann Arbor, MI
Troy Hartley	thartley@vims.edu	Virginia Institute of Marine Science	Gloucester Point, VA
Gregg Jacobs	gregg.jacobs@nrlssc.navy.mil	Naval Research Lab	Stennis, MS
Yan Jiao	yjiao@vt.edu	Virginia Tech	Blacksburg, VA
John G.W. Kelley	john.kelley@noaa.gov	NOS/OCS	Durham, NH
Doran Mason	doran.mason@noaa.gov	OAR/GLERL	Ann Arbor, MI
Richard Methot	richard.methot@noaa.gov	NMFS/AA	Seattle, WA
Erik Olsen	erik.olsen@noaa.gov	Institute of Marine Research	Bergen, Norway
Craig Stow	craig.stow@noaa.gov	OAR/GLERL	Ann Arbor, MI
Howard Townsend	howard.townsend@noaa.gov	NMFS/HC	Oxford, MD
Hongyan Zhang	hongyan.zhang@noaa.gov	Univ. of Michigan Cooperative Institute for Limnology and Ecosystems Research	Ann Arbor, MI
Additional participants			
Eric Anderson	eric.j.anderson@noaa.gov	OAR/GLERL	Ann Arbor, MI
Jessie Carman	jessie.carman@noaa.gov	OAR/OWAQ	Silver Spring, MD
Martin Dorn	martin.dorn@noaa.gov	NMFS/AFSC	Seattle, WA
Dana Hanselman	Dana.Hanselman@noaa.gov	NMFS/AFSC	Juneau, AK
Al Hermann	albert.j.hermann@noaa.gov	OAR/PMEL	Seattle, WA
Jim Ianelli	jim.ianelli@noaa.gov	NMFS/AFSC	Seattle, WA
Terra Lederhouse	terra.lederhouse@noaa.gov	NMFS/HC	Silver Spring, MD
Chris Legault	chris.legault@noaa.gov	NMFS/NEFSC	Woods Hole, MA
Jason Link	jason.link@noaa.gov	NMFS/AA	Woods Hole, MA
Patrick Lynch	patrick.lynch@noaa.gov	NMFS/ST	Silver Spring, MD
Mark Monaco	mark.monaco@noaa.gov	NOS/NCCOS	Silver Spring, MD
Ed Rutherford	ed.rutherford@noaa.gov	OAR/GLERL	Ann Arbor, MI
Charlie Stock	charles.stock@noaa.gov	OAR/GFDL	Princeton, NJ
Fernando Gonzalez Taboada	ftaboada@princeton.edu	OAR/GFDL	Princeton, NJ
Mariska Weijerman	mariska.weijerman@noaa.gov	NMFS/PIFSC	Honolulu, HI

Appendix C – Steering Committee

Jessie Carman
Sarah Gaichas
Troy Hartley
Chris Hayes
Jim Ianelli
Terra Lederhouse
Jason Link
Doran Mason
Richard Methot
Howard Townsend

Appendix D – Discussion of trigger and breakout questions for ToRs and Human Dimensions sessions.

TOR 1: Models and assessments

- What distinguishes purely a modeling effort from an assessment effort?
- What constraints are there in doing applied LMR & Ecosystem assessments?
- How do we handle the challenge of meeting our mandates yet maintaining space for innovation?
- What can NMFS learn from OAR modeling efforts?
- What can OAR learn from NMFS modeling efforts?
- How do the modeling communities handle uncertainty?
- Are there common LMR and ecosystem modeling efforts we could better harmonize?

TOR 2: Addressing uncertainty

Breakout Qs:

- Can we readily reference, and agree upon, definitions and types of model uncertainty?
- Are there common means for addressing uncertainty?
  - If so, can we list the top 3-5 approaches?
  - Can we reference them? Can we document such best practices?
- How do we determine which approach is most appropriate under given conditions?
- What is the particular role of Multi-model inference in handling uncertainty?

Best practices identified:

- Any common themes
- What 2-3 things can we highlight as best practices for addressing model uncertainty?

TOR 3: Model skill and uncertainty

- Why is model skill important to characterize?
- Can we provide consistent and common measures on this?

- What is the role of qualitative measures or perceptions of model skill?
- Does an evaluation of model skill assist with how model uncertainty is handled?

Breakout Qs:

- Can we readily reference, and agree upon, a definition of model skill?
- Are there common means for evaluating model skill?
    - If so, can we list the top 3-5 approaches?
    - Can we reference them? Can we document such best practices?
- In what ways does model skill help address uncertainty?
- What levels of model skill are most appropriate under given circumstances?
- Best practices identified:
- Any common themes
- What 2-3 things can we highlight as best practices for addressing model skill?

Human Dimensions:

- How do people, particularly non-experts, perceive and respond to uncertainty?
- How do we effectively communicate about uncertainty and models in order to enable constructive public and management deliberations?
- How do we achieve decisions, with suitable stakeholder buy-in, in the face of uncertainty?

TOR 3 + Human Dimensions:

- What 2-3 things can we highlight as best practices for incorporating how we communicate model uncertainty?
- What are key lessons about the role of understanding how well models are perceived to work as that relates to how well models will be used?
- How do we handle and address perceptions regarding use of multiple, potentially competing, models?

TOR 4: Summarize best practices/lessons learned

- Are there main lessons learned that we can categorically employ to address model uncertainty?
- Are there 1 or 2 particular methods that will address and communicate model uncertainty and skill well?
- If you were the NOAA Administrator, what 1 modeling or model-related effort would you prioritize and fund to ensure our modeling is best used?