Modeling Approaches in Support of Ecosystem-Based Fishery Management at the Northeast Fisheries Science Center, Woods Hole MA

External Independent Peer Review

by

Gunnar Stefansson

Prepared for the Center for Independent Experts

April 2011

Executive Summary

An independent peer review panel met with several participants at a workshop at the Northeast Fisheries Science Center (NEFSC) in Woods Hole, March 29-31 2011 where approaches to ecosystem modeling were presented. The models comprise several of the best currently available and the fundamental data used with these models are data collected by the NEFSC, also appropriate for this purpose.

Overall, the work conducted by the group is exceptional in breadth and in fact the number of approaches, developed or tested, is quite unusual. When developing models it is indeed important to consider several approaches as the group has done, not only to avoid bias and errors but also to see the data and ecosystem through different looking glasses to gain new insights. Having done this, however, it becomes important to put some limitations on how to move forward: The number of approaches to be used in the future needs to be restricted to fewer, more select models. These models need to be developed in greater depth than before.

Members of the NEFSC ecosystem modeling group are well connected to other parts of the NEFSC, to various stakeholders and to international colleagues. This is an important aspect of staying at the forefront of development to make the approaches scientifically sound, yet applicable to the tasks at hand, such as providing timely advice for an ecosystem approach to fisheries management or other issues regarding the ecosystem.

Given the complexity of the food web one would not *a priori* expect reasonable results from the simplest models, which might for example simply consider pairs of species. This is in stark contrast with Arcto-Boreal systems where one can commonly isolate and focus on interactions between only two species at a time. Instead more complex models are needed.

The immediate issue facing the ecosystem group at NEFSC is, however, not simply one of a choice of model or modeling environment. Rather, the immediate issue is to find or participate in a venue for deciding on what kind of management should or is likely to ensue as a result of the decisions to move towards ecosystem based management. That venue will inevitably include a dialogue with stakeholders. This dialogue will define how ecosystem based (fishery) management should proceed and that again will determine which models are needed, some of which are already in the ecosystem modeling toolbox at NEFSC.

There is clearly a need to increase the number of individuals directly involved in developing models, as the current number is too low to be able to both develop and use the highly complex models that will inevitably be needed.

Background

This review is based principally on information (material and presentations) made available at a meeting held at the Northeast Fisheries Science Center (NEFSC) in Woods Hole, March 29-31 2011. The material is listed in Appendix 1 and the formal statement of work is given in Appendix 2 (including terms of reference in App. 2.2 and tentative meeting agenda in App. 2.3). The meeting was conducted with presentations on the topics given in App. 2.3 along with considerable and useful discussions among the participants (App. 3) on most topics.

This reviewer has a background in fisheries management and advice based mainly on statistical stock assessment methods, both single- and multispecies models. The review necessarily reflects this background to some extent.

The formal project description for this review (from Appendix 2) is as follows: *The purpose of this review is to evaluate the appropriateness and performance characteristics of community-level and ecosystem models employed at NEFSC as operating models in support of the development of Ecosystem-Based Fishery Management (EBFM) strategies for the Northeast U.S. Continental Shelf. NMFS strongly endorsed the concept of Ecosystem-Based Management and the related need for the development of Integrated Ecosystem Assessment in support of EBFM. Although this review is directed at efforts in the NEFSC, the findings will be more broadly applicable throughout the agency.*

A paper, "the overview paper[1]", provides an overview of the entire ecosystem modeling enterprise at the NEFSC. As such it is the single most important reference for this review. Most presentations at the March meeting corresponded to expanding on a part of that document. A "white paper", on ecosystem-based fishery management, has also been prepared by the New England Scientific and Statistical Committee (SSC)[2]. These two documents give a clear description of the importance of taking ecosystem considerations into account in general, e.g. referring to national policy in a White House Executive Order as well as recommendations by the Commission on Ocean Policy.

[1] Link *et al.* (2011) – to be referred to as "the overview paper" throughout.

[2] See NEFMC SSC, 2010 in the list of documents.

Similar issues were raised by Dr. F. Almeida in his introduction to the workshop, in particular that although the single species assessments may be very good as such they do not generally take ecosystem considerations into account nor are they a natural part of an ecosystem approach to fishery management (EAFM).

Much of this review refers to the "fish-part" of the various models. Of course there are many ocean uses other than fishing and the ecosystem group is likely to become tasked with questions related to more general and conflicting uses of the marine ecosystem, i.e. questions relating to general ecosystem based management (EBM). This has happened in the past through a variety of questions put to the group but such requests can only increase in numbers as society steadily invents more novel ways of using the marine environment, whether for windmill farms, subsurface cables or other uses. In these cases there will be requests for advice, usually spatially oriented and usually across the entire ecosystem. When ecosystem models are under discussion, one must bear in mind not just fisheries but also these potential applications, some of which have extended far beyond traditional fishery science.

This move to requests for advice on ocean policy outside traditional single-species issues is global and the number and diversity of such requests will only increase. Responding to these requests requires a battery of models and analyses, of the types which have been addressed within the ecosystem group at the NEFSC as well as some which will require completely new approaches.

Model uses vary in general and multispecies or ecosystem models are no exception. One particular use of such a model may be to function as an operating model for testing EBFM **strategies**. These contrast with models for **tactics** that form a part of the typically annual assessment-to-advice cycle. The types of models available for consideration as operating models alone form quite a large class, ranging from holistic models which attempt to describe the entire system (i.e. starting from hydrography, energy transfer and/or microbial activity up to fisheries harvests, economic yield and/or employment rates as the case may be) to models which only take into account the population dynamics of a single species (possibly slightly modified to take the species' role as predator or prey into account). At these two extremes are Atlantis and the extended (single species) stock assessment (ESAM). In between is a plethora of models that may consider only a few fish species in detail (typical of Gadget models[3]) or one may model a

[3] For a background to models with Gadget, see Stefansson and Palsson (1998); Begley and Howell (2004); Begley (2004); Taylor et al. (2007) and Taylor (2011).

large number of fish species but only take a few important processes into account (typically predation mortality as with MSVPA)[4].

An exercise in modeling may arise simply as a research proposal out of academic interest or it may arise as a response to requests from stakeholders. In the context of this review it is important to realize that the main purpose of the present ecosystem modeling undertakings at the NEFSC is to form a basis for advice to stakeholders[5]. Discussions during the meeting, issues raised within the overview paper and further discussions with the users/developers of the ecosystem models have made this clear and hence the terms "*appropriateness and performance characteristics*" in the above project description will be interpreted in that light.

One way to classify modeling approaches is by how rigorously they incorporate data. An exploratory tool may, in principle, be developed without any data. A simple mathematical model may thus shed some light on questions on general interest. Usually this does not meet the needs of any stakeholder, who want (quantitative) directions on how to proceed on a given topic. A next step is therefore to make sure that the model behaves at least roughly like the ecosystem (or part thereof) that it is intended to mimic, and this may be done by simple eyeballing to verify overall trends predicted by the model. Statistically speaking, however, one should set up formal criteria (e.g. likelihoods) that describe the data and how well the model can predict the available data. Most single species stock assessment models continue this process to the bitter end. However many multispecies/ecosystem models really only take a cursory glance at data but rely on hand-tuning parameters to obtain what the investigator interprets as a decent or "good-enough" description, "tuning" or fit to the data. This may be appropriate when describing in general how a marine protected area (MPA) may work or how sets of ecosystem indicators may reflect the state of the resources, i.e. to give informed advice on strategy.[6] It is unlikely, however, that such broad-brush model fits

---

[4] References to all models discussed in this report, except Gadget, can be found in Link et al., 2011.

[5] See Link et al., 2011 and NEFMC SSC, 2010.

[6] See Stefansson and Rosenberg (2005, 2006) for simple MPA models developed only with a cursory glance at reality and Fulton et al. (2005) for fairly generic tests of ecosystem indicators, in both cases using specific ecosystems only as guideposts.

will be adequate descriptions of any given system and it is therefore also unlikely that they will be found to be an appropriate basis for tactical decisions. In spite of their problems as tactical models, they may be very useful as operating models for testing harvest strategies, i.e. a model reflecting roughly the ecosystem in question will be a potential candidate as an operating model. Nonetheless, unless these models describe the data sets reasonably well they will always be subject to criticism and may not be accepted as e.g. a primary operating model.

These are of course only some of the reasons why multispecies models have not generally replaced single species models. Rather, single species models have been fitted *ad nauseam* to data [7] and then augmented by incorporating simple ecosystem considerations.[8] This pretty much ensures that the earlier single-species fit to the data is maintained while taking the (apparently) most important ecosystem considerations into account.

Needless to say, members of the NEFSC ecosystem group are fully aware of these issues and many of them were discussed during the meeting. They are (re-) stated here to provide background for several of the methods-specific comments that appear below.

[7] Where we nonetheless need to acknowledge that such fits are often less adequate than desired, *cf* common retrospective patterns, trends in catchability etc.

[8] In a general sense, whether technical interactions, effects of predation on predator growth or prey mortality, effect of temperature on the stock-recruit curve or on catchability, *etc*.

Summary of Findings

General

From presentations to the panel, it is clear that the ecosystem group sees its mandate in the light of stakeholder requirements, with notable reference to the US Commission on Ocean Policy, the National Ocean Policy, the NOAA Strategic Plan and the SSC white papers on ecosystem approaches to fisheries.[9] It is not at all clear, however, how these official guidelines can be used to drive the research within the ecosystem group. The guidelines in many cases first need to be interpreted and implemented as management action. This is by its nature a chicken-and-egg problem: one can not implement an ecosystem approach to fisheries (EAF) without the science, which can not take direction without knowing where the EAF is going.

Given the number of tasks undertaken, the Ecosystem Assessment Program of the NEFSC (or "ecosystem group", below) at Woods Hole contains only a small number of permanent "model-oriented" employees. In addition to these the group does have several support staff, with a particular emphasis on access to databases. In spite of the low numbers, the group has managed to convincingly evaluate a wide range of models, most for the purpose of providing EAF directions. Further, this development is done in the appropriate international context. Thus, publications go through the normal peer-review process, contact is maintained to international professional groups (one of the chairs of the ICES Multispecies Assessment Working Group is within the ecosystem group) and contact is maintained with stakeholders through participation in the appropriate Science and Statistical Committees (SSCs). The last connection is particularly important since it links the ecosystem group with stakeholders in a natural manner as seen below.

There is clearly a need to increase the number of individuals directly involved in developing models: The modeling subgroup simply needs to include more individuals with a background in modeling (i.e. to have simultaneous expertise in applied mathematics, statistics, computer programming and database access).

[9] See references in the overview paper, Link et al. 2011.

Very many different models have been set up and tested by the ecosystem group. This is important up to a point, not only to verify results but also to capture different ecosystem aspect or processes not all captured within a single model or modeling framework. On the other hand one must at some stage focus on the way forward and select a reduced set of models satisfying a few specific criteria:

- They should describe the system "adequately" by being able to predict measurements, i.e. they should *relate to data*.
- They should relate to quantities or issues of interest to the stakeholders.

These criteria will be discussed in general terms first and some will then be followed up on within the model classes. Data fitting in particular is discussed in some detail below, since this is an important aspect of obtaining reliable model fits and outputs.

Management strategies and their evaluations were discussed repeatedly during and between presentations at the meeting. A particular issue is how ecosystem concerns can be built into management strategies, how they can then be evaluated and whether such strategies can be implemented. Examples are discussed below, with reference to issues regarding implementation.

The panel was presented with several models and analyses where species had been grouped into e.g. "species groups" or "taxonomic groups" or "guilds". This is often very useful and important, whether for understanding total consumption by ecosystem components, energy transfer between groups or the potential for managing harvests from a group caught by a fleet. As always however, one must eventually evaluate whether the analyses should terminate (with scientific publications as appropriate) or be carried forward to become a part of the toolbox for answering stakeholder questions, whether in an EBM or EBFM context.

Relating to data

As noted above, some of the models developed by the group give outputs differing considerably from single species results, with specific examples given below. That is of course not an issue unless the output actually deviates considerably from what measurements imply (i.e. it is not an issue *per se* if models differ in things which can not be measured directly such as the natural mortality rate or absolute abundance in a given year), or can easily be questioned on general biological grounds. Whether these various issues are serious or not depends on the application. Importantly however, an

immediate problem with not being able to predict e.g. the overall shape or trend reflected in a survey time series reasonably well is the lack of credibility within stock assessment circles, which of course may affect credibility among a considerable portion of the stakeholders who will be familiar with survey results or single species stock assessments. Thus one may be faced with a credibility issue even if a model is only to be used as one of several operating models and even if the model is only being used for simple exploratory analyses. Such issues need to be thought of beforehand and tackled in some way.

In the setting of a management strategy evaluation one may alternatively simply want to specify different parameter sets which correspond to different biological assumptions in cases when these come under obvious debate. This is common practice anyway when developing operating models, as these have to cover a wide range of plausible assumptions (see also section on databases below).

No matter which way is taken forward with a given multispecies model, there is a need to make sure that it either explains data series reasonably well or that there is some other mechanism to make sure that the properties of the actual data series can be captured for the purpose of the task at hand.[10] Whether the task is an important MSE or "simply" uncovering the effects of sharks it is always important to try to ensure that the output is not merely a result of discrepancies between the model and some important features of the data.

In cases when there are no formal measures of the goodness of fit to data available, but only outputs from different models without indications of which fit "best", there is an obvious problem of comparing the models. The NEFSC ecosystem group has addressed this in part through a process of inter-model comparisons. This is needed in any case, when there is a multitude of models to choose from, regardless of the fitting methods. Having the models first formally verified by comparisons to actual data merely reduces the likelihood of some models performing inconsistently with the observations.

---

[10] A single issue may be solved e.g. by adding appropriately correlated variation to a predicted recruitment series, outside the model (if that is the issue). Alternatively one may simply want to acknowledge the fact that e.g. an appropriate fitting method is needed and either modify the approach accordingly or abandon the model as the case may be.

While it is clear that formal fitting to data is difficult when nonlinear models have parameter sets which number in the hundreds or even thousands, it is more important than ever to make sure that approaches to determining these parameters is objective and can be repeated. Traditional methods for determining values for large sets of parameters range used to be *ad-hoc*, such as those originally used for making MSVPA consistent with stomach content data but many current ADMB models may include many more parameters than the earlier MSVPA models: The fact that the ADMB models can be fitted "properly" is a consequence of improved methods and improved computer performance.

One alternative approach includes transforming the data (using some model) in order to have an information source comparable with output from the model in question. This was e.g. done with MULTSPEC where stomach content data were first converted to consumption estimates and subsequently the model fitting mechanism compared the internal estimates of consumption to these values.[11]

An extreme version of this approach, which pretty much ignores the original data, is to fit the parameters of one model to reflect the outputs from another.[12]

These methods basically correspond to fitting the model to a function of the data and in general this is not a recommended procedure since important properties of the data may be lost. In terms of fitting methods, this is similar to going back to the now-outdated methods of transforming proportions using a logistic transform and then fitting a straight line to the data. With the advent of the generalized linear model these transformations have largely been abandoned, the whole point of the GLM exercise being to adapt the model to fit to the original data, not *vice versa*.

Given the problems of fitting complex models to large data sets one may nevertheless be forced to resort to such less-than-optimal methods. At a minimum, the approach should be made objective and repeatable. Before resorting to such methods, however, it is important to at least consider alternative formal model-fitting approaches since they have advanced considerably in the past few decades. For example, Bayesian models are used to describe fMRI data, which make both fisheries data and models minute in comparison.

---

[11] See e.g. Tjelmeland and Bogstad (1989, 1998); Bogstad et al. (1992).

[12] Pope (1989) used this approach to obtain parameters in a (simple) multispecies production model based on output from a more complex model.

For multispecies fisheries models specifically methods have been developed to use formal model fitting to stomach content data instead of transforming to consumption and to use formal methods to verify consistency of the model fits.[13] These particular methods were developed for simpler ecosystems and may not be plausible approaches to the intertwined NE US ecosystems but one should at least investigate whether current approaches to parameter estimation can be improved.

The problem of fitting overly complex models to data, using chronically inadequate computers, is a very old one. Recent methods typically use several phases of model fitting, estimating parameters one group at a time, possibly only using a portion of the data at each stage. At the extreme this corresponds to picking up some estimates from the literature (e.g. consumption estimates obtained from stomach content data and not modified further), but a more deliberate approach is usually preferred. This is pretty much the method used in age-old methods for age-disaggregating a length distribution[14] and is routine in both ADMB and Gadget[15] work.

In addition to data-driven methods to evaluate model quality, there is a need to compare in some manner widely different models (*inter*-model comparisons). This field has been developed considerably within climate models, in some cases using Bayesian approaches to compare an array of different models.[16] One can not avoid noting, however, that if neither of two models fit data well then it is very hard to state why they differ. In general, if two models fit well to abundance series and are in accordance with

---

[13] See e.g. Stefansson and Palsson (1998) for a Gadget-approach to modeling, Taylor (2003) and Taylor (2011) for model-fits to single/multi-species models and Stefansson (1998) for model diagnostics.

[14] Users of the MIX program of MacDonald and Pitcher (1979) typically first estimated proportions, then mean lengths followed by standard deviations, repeating the process as needed.

[15] Within the Gadget environment models are typically developed parameters estimated for each stock component separately, see e.g. Taylor (2011). Once an adequate model fit is obtained for each component these are combined and linked to more data (e.g. stomach content data). Even then parameters may be estimated in phases at each stage.

[16] See e.g. Berliner and Kim (2008) for an example.

stomach content data and catches, then one does not expect a major discrepancy between them. Thus improved fitting to data should reduce the need for inter-model comparisons.

Relating to stakeholders

For a modeling exercise to have relevance in the context of an institute such as the NEFSC it must usually provide some input to discussions with stakeholders. If this is not the case then the models are usually considered an academic exercise and downgraded in importance accordingly. Given the limits on resources this approach is often appropriate (though sometimes short-sighted). The modeling exercise may be more than just a model of course, particularly in the context of a management strategy evaluation (MSE).

Some of the models presented to the panel can be considered candidates for use in a management strategy evaluation. The devil is in the detail, however, and it is nowhere nearly enough just to have decent underlying models. In particular, the MSE is largely irrelevant unless it is done in tight collaboration with stakeholders and uses a strategy which can be implemented. Put bluntly, an evaluation of irrelevant strategies is a largely irrelevant exercise, though it may be of some theoretical interest.

An important aspect of the ecosystem group is the participation of two group members in Science and Statistical Committees, that forum being appropriate to advise stakeholders on e.g. management strategies.

In the single-species framework management strategies usually aim to limit the fishing mortality for the species and use some biomass level as a target or lower bound. In an ecosystem context these individual values may not make much sense, usually because the single species BMSY values may add up to more than the ecosystem has seen before.[17]

[17] See earlier work by the ICES Multispecies Assessment Working Group, which led to revised natural mortality vectors being used in many ICES single species assessments, or Walters *et al.* 2005 for a more recent reference.

Although it may seem imminently sensible to go ahead and suggest alternatives to the single species harvest control rules (HCR), this can really only be done within a forum that is set up by stakeholders with a mandate to evaluate and propose such rules.

For example, one might easily envisage a 2-tier system for demersal fish, where an overall ecosystem quota is first set, followed by allocating this to species. Experience in many countries indicates that this is really only feasible as a part of a dialogue with stakeholders. Otherwise it is highly likely that the evaluation of the new ecosystem harvest control rule becomes an academic exercise of little interest to decision makers. Thus the appropriate way forward is for a body such as a SSC, which already has the appropriate mandate of suggesting a new HCR, to set terms of reference and nominate a subgroup to formulate or evaluate revised ecosystem-based harvest control rules. Such terms of reference (or the process as a whole) will specify rules that have the potential to be implemented and are not just theoretical.

In light of the somewhat uncertain procedure it is not clear when additional input will be required from economists or social scientists. On the one hand for an "ecosystem MSY" it **is** clear that it makes little sense to simply add up tons of mussels, lobster, mackerel and cod without at least multiplying by the unit price. On the other hand simple economics (first-hand value) can fairly easily be added to an ordinary multispecies model. In terms of the EBFM a much more pressing need is to start the dialogue with stakeholders in order to define just what general kind of framework is needed.

Notes on databases

The ecosystem group appears to have good access to NEFSC databases and the expertise to extract data from these as required to feed the models. Naturally this is an essential part of any modeling exercise.

Some modeling environments (e.g. Atlantis) can be very data-hungry and it is a major task to set up the input files for such models. Some analyses may be based on data from satellites, trawl surveys, acoustic surveys, hydrographic measurements, biological samples from catches, surface sightings or price information. These data sets are on different time scales and variable spatial resolution implying that care must be taken at all stages of analysis.

It is common practice to set up input files manually for programs such as Atlantis. It follows that it is a major hassle to revisit the input data. For example converting input data from 1cm groups to 5cm groups may require many man-months of work in some

instances. In principle this should be attainable by merely changing the extraction command from the original database, but it rarely is.

The modeling group may want to consider generic alternatives by defining a single interface to all the databases. Such an interface is typically a new database but could be a collection of specific views into the present systems. Once the interface is defined the next step is not to extract data for a given model but to write extraction routines (whether in perl, python or php) to give input data for the modeling environment in a syntactically correct format. At the same time the extraction routines output the metadata which the modeling environment needs, be they spatial scales, information on length groupings etc. If this is done correctly then a lot of time can be saved later if different spatial, temporal or other scales are to be evaluated. If the new databases are generic (and not hard-linked into NEFSC databases) then the extraction routines can be made applicable to other areas as well.[18]

These databases may in principle be used to generate different data combinations to evaluate models using bootstrapping (or the jackknife). The issue involves setting up a resampling procedure which avoids the usual problems with the *intra*-haul correlation, but in a much more general sense since all the different data sources collected in a small spatial-temporal cell tend to be correlated. Although such methods have only been tested for a few cases, initial results are quite promising.[19]

Common issues of the general "ecosystem" type are the various gear conflicts. In many (most) cases these are at a finer resolution than that used in most models. Thus, although commonly posed by management, these questions usually fall outside the range which can be answered by the models. Rather than try to operate models on the

---

[18] This approach was used in the EU-funded dst^2 project for setting up input files for Gadget, see http://www.hafro.is/dst2/ or Taylor (2003); Kupca (2005, 2004a); Kupca and Sandbeck (2003). Given a standardized data base, the same extraction routines have been used to extract data for Gadget models for various areas in the North-East Atlantic including Icelandic waters, the Irish Sea, North Sea as well as the Bay of Biscay and Tyrrhenian Sea.

[19] This kind of bootstrapping has been done for entire tagging experiments in Hannesson et al. (2008) but Kupca (2004b) developed a formal method for bootstrapping from a data base system and this was tested and used in Taylor et al. (2011) (or see http://arXiv.org/abs/0807.3677)

very fine scales required to describe gear interactions it is often more appropriate to base responses on catch data (logbooks) or survey information, possibly using geographical information systems (but see below).

A note on geographical information systems

It has become customary to use a geographical information system (GIS) to layer different pieces of information and even to obtain plots of species distributions. Unfortunately there is no guarantee that these systems will use appropriate (statistical) methods to provide such plots and the methods may not even be documented. It is usually more appropriate to obtain such plots using a mapping sub-package within a statistical package – the plots themselves can then be outputs of a GAM, GLM or combinations of multiple such models.[20]

[20] Typical examples can be seen in Stefansson (1996) or Stefansson and Palsson (1997) but alternative zero-inflated distributions are also common.

Findings specific to Terms of Reference (ToR)

1. Evaluation, findings and recommendations of overall modeling strategy

A. ***Summarize evaluations, findings and recommendations of overall community-ecosystem level modeling strategy in practice for the NEUS LME system***

Only a few bullet points will be given here, with more detail given on each topic below.

- The overall strategy has led to understanding of many important issues in the ecosystem under investigation.
- Data analyses and models clearly demonstrate the skills of the members of the ecosystem group.
- Spatial areas have been defined in order to have "relevant" areas e.g. for geographic management. General ecosystem questions are most likely to have a spatial component. "Space" is therefore a particular component and models incorporating this should be given a particular emphasis.
- The modeling strategy has tested many different approaches. At this time it is appropriate to reduce the total number of models and approaches, moving focus to fewer models and trying to emphasize those most likely to be beneficial to answering stakeholder queries, yet not completely omitting those analyses likely to further enhance understanding of this complex ecosystem.

B. ***Determine whether the science reviewed is considered to be the best scientific information available.***

As documented in several places in this report, the individuals are seen to be active in the local and global scientific community and publish their results in internationally peer-reviewed journals – as is expected from a group at the forefront of model development and applications. The models and analyses used and developed are certainly at the forefront worldwide. As always, these models can be enhanced and suggestions on how to move the models towards stakeholder interest, increased reliability and increased credibility are given below.

C. ***Determine if the intended uses of overall community-ecosystem level modeling that have been identified as priorities for the NEUS LME are being executed in accordance with global best practices.***

The New England Scientific and Statistical Committee (SSC) has produced a white paper on ecosystem-based fishery management. This white paper is the outcome of work initiated with a 2009 workshop "with over 60 participants providing a cross section of Council members, scientists, managers, invited experts, NGOs and fishing industry involved in EBFM.[21] This type of dialogue forms an integral part of the decision making process which ultimately ends up in an implementation of the EBFM. This SSC will need to be involved in taking this work one step further, namely to map out the actual types of harvest control rules or management strategies which are implementable and in accordance with the guidelines in the white paper. In many international fora this is not done by groups of scientists but by special groups that have a mandate to suggest strategies. The panel has been informed that the SSC has such a mandate and is therefore ideally suited to suggest e.g. that a body such as (or with ties to) the ecosystem group evaluate different types of management strategies with an EBFM focus.

[21] See the white paper, NEFMC SSC. 2010

Several of the models developed by the ecosystem group have the potential to become a part of such work.

The models considered and presented to the panel have been developed, set up, tested and verified using fairly standard methods, much along the lines recommended as "best practice" international standards.[22] Overall, the breadth of models that have been used is overwhelming. Details on each model type, along with recommendations, are given below.

The way each type of model applies to EBM or EBFM varies considerably. Some are seen to be highly applicable to e.g. becoming an operating procedure but this does not apply to all of them.

D. *Provide recommendations for further improvements.*

See recommendations at the end of this document (p. 30).

2. Evaluation of strengths and weaknesses, and recommendations for specific methodologies

A. *Energy Transfer Models (Fogarty)- Production Potential Models*

Results from production potential and network models were demonstrated to the panel. In the case of production potential models uncertainty estimates were also demonstrated. These types of models are of considerable interest for a number of reasons, one of which is to have some idea about the total potential production from the system and another being how that is likely to change in response to predicted global warming. Apart from such one-off questions, these models are however, unlikely to answer many questions of interest to stakeholders.

Interestingly, several results were given where an overall fishing mortality multiplier was applied to the entire collection of fisheries to demonstrate how yields did not reduce much by decreasing fishing mortality from an overall FMSY, but the proportion of collapsed stocks declined considerably.

Suggestions have been put forward that these methods may be used as a part of

[22] But note that the FAO-defined "best practice" does involve formal fitting to data: "Fitting to data is best practice, and this requires careful specification of likelihoods ", *cf* p. 59 of FAO Technical Guidelines for Responsible Fisheries, ISSN 1020-5292, Suppl. 2, Add. 1 at http://www.fao.org/docrep/011/i0151e/i0151e00.htm.

management. For example one might design a 2-tier approach to setting quotas where some sort of overall quota is initially set at the ecosystem level, subsequently followed by a disaggregation to the species level. The exact implementation can be envisioned in a variety of ways – the overall ecosystem level could be based on an overall fishing mortality rate or could be based on the current or new production relative to some baseline or total production (the "f-ratio").

Several concepts have emerged from these investigations, many useful but some require more work. For example it is not clear how stable the numbers in the f-ratio are but this is crucial if such a number is to be used.

The idea of estimating total production potential using energy transfers, one way or another, is certainly a process which is useful in principle and is something that has to be done at least once for a given system. It is not as clear whether such work can or should form a basis for fisheries management in the long or short term, other than as a reference.

Investigations using an overall fishing mortality have been done as part of other work, some under this heading of production potential. This topic needs to be given considerably more attention.

One could of course simply investigate the effect of using a single fishing mortality across species. In practice this would at best apply within a complex such as groundfish, since it is quite far-fetched that one would ever want to use the same fishing mortality for a pelagic species and a groundfish species. Since the species will have age- (or length-) dependent selection patterns, the fishing mortality is not strictly constant across components of the ecosystem anyway.

A more appropriate method would therefore be to use an underlying fishing mortality vector and scale this by some number *k*. The scaling factor *k* then replaces the constant fishing mortality in the analyses. The underlying vector of fishing mortalities needs to be "sensibly" chosen. One approach is to set it to the FMSY values for each species separately.[23] When modeling without interactions the FMSY values will most naturally be chosen as the single-species FMSY values. However in the present, multispecies setting, the most natural approach would be to first find the vector of fishing mortalities that maximizes the total yield (possibly economic yield), i.e. the FMSY vector in the multispecies sense. This vector forms a natural

[23] This is done by Walters et al. 2005.

vector for scaling the results.

The problem with any of these methods is that it may not be feasible for the fleets to target the species in this manner, i.e. the ratios of fishing mortalities induced by the fleets may be such that the vector of overall FMSY levels can not be attained. Likely examples of this scenario include cases where a vulnerable species is a bycatch with a less vulnerable species (e.g. Atlantic halibut with plaice or most skates with most demersal (teleost) fish).

It follows that to make these kinds of analyses credible one must first evaluate the vector of fishing mortalities per species as actually induced by the combined fleets. Scaling this vector (e.g. to unit length) provides a standard vector of (relative) fishing mortalities, which has been achieved in the real world. The annual variation in these vectors provides a measure of what kind of deviations can be reasonably implemented. The ecosystem is harvested by many fleets so the partial fishing mortalities inflicted by each fleet form components which can in principle be combined in any manner by placing limits on individual fleets. If the ecosystem FMSY vector is within such ranges of plausible values (done separately for the different components such as demersal *vs* pelagic) then it can be used as mentioned above.

Otherwise the exercise would appear merely academic since it can probably not be implemented.

It follows that to make these kinds of analyses credible one must first evaluate the vector of fishing mortalities per species as actually induced by the combined fleets. Scaling this vector (e.g. to unit length) provides a standard vector of (relative) fishing mortalities, which has been achieved in the real world. The annual variation in these vectors provides a measure of what kind of deviations can be reasonably implemented. The ecosystem is harvested by many fleets so the partial fishing mortalities inflicted by each fleet form components which can in principle be combined in any manner by placing limits on individual fleets. If the ecosystem FMSY vector is within such ranges of plausible values (done separately for the different components such as demersal *vs* pelagic) then it can be used as mentioned above. Otherwise the exercise would appear merely academic since it can probably not be implemented.

Finally, and most importantly, although the idea of managing with respect to overall ecosystem measures is very appealing, it can only be done as a part of a dialogue with stakeholders. Basically, the concepts need to be introduced at meetings that

have the mandate to decide or recommend harvest control rules and feedback from such meetings needs to be used to further develop the concepts.

B. *Energy Transfer Models (Link)- Network Models*

Most of the comments on the Energy Transfer Models in (A) above also apply here but are not repeated.

The network models have been used for understanding at several levels. One such is to obtain estimates of transfer efficiencies between trophic levels. The intention is to use these in other models.

Considerable attention has been given to "tune" the different models, thus e.g. obtaining fairly consistent results in Ecopath and Econetwrk.

This reviewer is not an expert in Ecopath, Ecosim, Ecospace or Econetwrk, but one particular general comment is in order: The single species stock assessment methods (whether extended to account for some ecosystem concerns or not) have been extensively developed to take uncertainty into account. Model diagnostics are routinely used to evaluate whether the models appropriately fit to the data sets (and often these tests fail even in the final assessments). Basically, the development over the past 3 decades has been away from *ad-hoc* fitting methods to formal nonlinear statistical models in most cases. It is hard to envisage that these models will be scrapped in favor of other models that do not have these statistical underpinnings – even if the new ones are more likely to capture important ecosystem concerns. It is not lost on this reviewer that it is easier said than done to enforce objective statistical methods on the very high dimensional multispecies models under consideration. It is merely hard to see *ad-hoc* data fitting methods or outputs with no uncertainty estimates replacing current (extended single species) methods – and as mentioned above, it is FAO-defined best practice to fit formally to data.

On the same note, the use of species groups, whether called guilds or taxonomic groups, is useful for understanding many concepts but it is not quite clear how these can be translated into management. Possible exceptions may be cases such as a demersal fishery in a given area. If a considerable portion of the catch taken by that fleet consists of species that form a "guild" then one might see a way forward in providing "ecosystem advice" for that species group in that particular area, for example by overall effort control for that particular fleet.

C. *Aggregate Production Models (Link/Fogarty)*

A class of aggregate production models was presented where, instead of fitting to a population, fitting was done to a group of species. Although this is in many ways different from fitting production models based on individual species, there are enough commonalities that this is discussed with the multispecies production models below (D).

D. *Multispecies Production Models (Gamble/Fogarty)*

The MS-PROD model, as presented, is a "multispecies extension of the Schaefer model". The various production models (items C and D) have interesting aspects and potential use in several ways. They are, however, bulk biomass models, which implies certain limitations on their usefulness in terms of advice to stakeholders.

As a start it should be noted that, just as the single species Schaefer model is a very useful tool, so can the multispecies versions be useful. In particular a well-designed quadratic model like this is very useful to quickly investigate the effects of harvest control rules and to get a first idea about likely directions when the certain components change.

One concern is reflected in the description of how parameters were set: "r's were **informed** by using 2 x F_{MSY} as a starting point, and modified as necessary" and similarly for Ks, competition interaction strengths, spatial overlap etc. What this means, basically, is that no formal method was used to estimate any parameters in these models. Further, the multispecies production models contain parameters describing "competition" as well as "availability" of prey to predator. *A priori* one would expect these parameters to be confounded and it is not at all clear whether their use can be justified from a model-fitting point of view (*i.e.* whether in fact using all of these parameters merely adds noise).

Further, these models have no explicit spatial components, arguably the most important component of models to be used for ecosystem based assessments or modeling.

There have been many uses of multispecies bulk models, some theoretical and some with applications. At least one of these appears to have been fitted in an objective manner by fitting to output from an age-disaggregated multispecies

prediction model (MSFOR, for the North Sea).[24]

In addition to these issues, which arise from this debatable parameterization and method of assigning values to parameters, other issues arise when considering some of the simulations undertaken. For example, it is not clear to this reviewer that it can be completely logical to (1) taking the single-species F_{MSY} as an (even initial) basis to give r or (2) setting interaction terms to zero without re-"fitting" the remaining parameters. In general, one would expect the same effects here as in multiple regression where parameters need to be modified as others are dropped from or inserted into the model. For example if the initial parameters were indeed obtained from a formal (nonlinear regression) fit to a (very large) data set, then one would always re-estimate the parameters after inserting any assumption on certain interactions being zero etc.

These models are not spatially disaggregated so they are not likely to be amenable to EBM. Further, at present they probably do not provide a description of the fish part of the system adequate to make them credible to stock assessors or stakeholders so it is not clear whether they can be used directly for EBFM. It is therefore recommended that they be downgraded in terms of modeling emphasis and relegated to become a less-used part of the toolbox, e.g. only for occasional studies such as the effects of serious overfishing a certain species group, and then only with comparisons to other models.

In spite of this, there are several examples of aggregate biomass production models that have found their way into being a part of a management strategy evaluation. Such cases typically involve systems where age readings are hard to come by and one would normally not think of a biomass production model as being optimal for a data-rich scenario such as the area off the New England coast.

E. *Full System Models (Link/Gamble)- ATLANTIS*

The Atlantis model, in the form set up by the group, has considerable potential since this modeling environment can encompass almost any aspect of the ecosystem handled by any other ecosystem model. Naturally, this does not come without problems. In particular Atlantis does not incorporate any statistical estimation

[24] See Pope (1989), but note that this ignores the stock-recruitment curve and also here there are issues with the method of fitting, although it appears to be a formal (objective) fitting mechanism.

method but is adjusted to reflect more-or-less the overall behavior of the ecosystem in question. This is quite adequate for obtaining an overall view of a system and for answering a variety of what-if questions. If Atlantis is to be used as an operating model various error terms need to be added to the output and this can in principle be done externally. Atlantis has in fact been used for evaluating a range of very different management methods where these approaches have been developed and described.[25]

The implementation of Atlantis presented to the panel is a promising first attempt at an Atlantis model for the region. The model is general enough in principle to answer a wide range of EBM-based questions.[26]

On the other hand it must also be noted that there are some problems with the Atlantis approach and these will be serious in certain scenarios. The primary issue (from this reviewer's point of view) is the complete lack of a method to objectively estimate parameters, evaluate goodness of fit, compare different models quantitatively or obtain uncertainty estimates for parameters. The problem stems from the run times, i.e. it is infeasible to attach numerical estimation routines to a computer program as unwieldy as Atlantis.

This computer program does not automatically provide sums of squares or other measures of goodness-of-fit (and in particular survey indices etc are not even input into Atlantis). It would certainly be possible to automate some of the „tuning" process, e.g. by selecting a set of important parameters and important time series and computing sums of squared deviations outside Atlantis, followed by a battery of runs to obtain a better fit. Given the run times however, this would need to be run on a computer cluster, even for only a handful of parameters – and the entire set of parameters will probably not be estimated in this manner in a single human lifetime.

In spite of these problems, if a model of this order of complexity is to be completed to any level of confidence (even just to be used as a primary operating model) then some such fitting process needs to be invoked. Given the excessive run times, that can only reasonably be done by running different Atlantis parameterizations simultaneously on a large computer cluster and evaluate fits to some of the more

[25] See e.g. Fulton et al. (2005) for a complex example.

[26] For the narrower EBFM questions, MSVPA would go a long way as might Gadget but the choice is not at all clear.

important parameters. However, if using Atlantis is to be taken to the logical conclusion with the associated costs in man-hours, then purchasing or getting access to a 50 or 500-core computer cluster may not be the biggest worry.

As explained in the background section, however, these issues do not preclude Atlantis from being a very useful tool in the toolbox. In particular one can see Atlantis as one of the tools for answering fairly general spatial questions. Similarly, Atlantis can in principle be an operating model for evaluation of management strategies (or one of many such models).

Thus, even without any fitting procedure, Atlantis is clearly a candidate for an operating model for management strategy evaluations and for simple what-if analyses. Further, since this model is designed to include more detail than any other, it can be used as a basis to generate data for pretty much any other model under consideration and thus explore consistency between other estimation methods.

One typical problem regarding "realism" or "credibility" is that recruitment bursts are generally not picked up (certainly not in the implementation presented to the panel) since annual recruitment deviations from e.g. an underlying Ricker curve are not estimated. Deviations from a stock-recruit curve would need to be implemented in some manner if e.g. the typical variation in haddock recruitment is to be captured. In principle this could be kick-started by using as input the observed time series of recruitment from single species assessments (with the obvious caveats related to trying to maintain consistent natural mortality assumptions and thus scaling of stock size between the single species and multispecies versions). Since observed (apparent) autocorrelation in recruitment deviations are not uncommon these stocks can easily have "red noise" which is not captured (*cf*. the gadoid outburst).

This contrasts a minimally realistic model (MRM) where one would typically select a handful of important factors (whether species or processes) and make sure that these are modeled so that they appropriately reflect the underlying data sources. This is not just the case for extended single species models but also the suggested method with some multispecies models.[27]

In the case of using Atlantis as an operating model one would presumably (more naturally) split the process of obtaining a credible result into the two parts of (1) obtaining a "reasonable" overall description and (2) generating data with properties

[27] See for example the 2-area, 2-species Gadget example in Taylor 2011.

similar to those measured.

Atlantis appears to be the most promising generally applicable model with the suite of models tested by the group, for general EBM applications and EBFM in particular (though alternatives exist).[28]

This begs the question of what is needed to conduct serious modeling work using an environment such as Atlantis. The same applies to MSVPA, Gadget or similar modeling environments. In order to maintain any of these environments it is important to have a user, which has one of them as a primary responsibility. If there is only one individual then that user needs to be able to take on all tasks from extracting data through modifying the computer code and implementing appropriate statistical estimation methods. Quite commonly this implies more than one individual. It is important that the person who is mainly involved in model runs also has these as a primary responsibility, i.e. does not get swamped in other activities. Experience based on Gadget work seems to imply that one needs to continually have 2-3 staff members working mainly tasks related to that single model. One of these persons needs to be a "user" who actually applies the model to the ecosystem in question.

F. *Other models (Fogarty/Link)*

A variety of "other" models and analysis were presented to the group. All of these are useful, but usually only in a rather limited way. For example, an analysis indicating that (many components of) the ecosystem may behave in a "nonlinear" manner (so as not to be predictable) raises many more questions than it answers. Within a single species framework this kind of effect (almost only) arises if the right-hand limb of the stock-recruit curve is too steep and this is almost never seen from data. If this kind of effect is estimated to be common across a system, then there must be other causes and the questions raised range from doubt ("is the methodology biased in some manner?") to queries on what can give rise to these effects (species interactions, migrations, fisher behavior or combinations thereof). If the analysis does not answer these questions then it is probably best relegated to

[28] In terms of EBFM alone one could presumably use Gadget to get the same results but with somewhat more objective data fitting, more piecemeal modeling and possibly a less daunting overall modeling exercise. Gadget does come with its own set of problems and this is not intended to indicate that Gadget should be adopted.

the status of a scientific curiosity, to be analyzed as such with a low priority. Methods designed to either directly answer stakeholder concerns or provide explanations of ecosystem behavior should take precedence.

Extensions of single species models have been developed in the "usual" way, taking natural mortality from MSVPA into single species assessments, including environmental effects into various processes etc. In other regions many such analyses are conducted by an ecosystem group, others as a part of a stock assessment process or management strategy evaluation.[29] These extensions will not go away: Even with improved ecosystem models there will still be a need to investigate simpler single species models, which may better capture the essential dynamics of the species in questions – yet taking into account the main interactions through simple methods.

The various analyses (single-species or otherwise) have clearly implied that there are many important "ecosystem effects".[30] Many of these will need to be taken into account when management strategy evaluations are undertaken. At a minimum one will need to consider how not including these will affect outputs.

Conclusions

The group has demonstrated a capability to set up state-of-the art models, test them and implement them – for a very wide range of models. The primary need at this stage is to reduce the number of models and go from the "breadth" to "depth", i.e. to select only a few models (which should be clearly applicable to providing advice relevant to stakeholders), but develop these in sufficient detail to provide more confidence in each one.

As an example of why this is important one can mention an application presented at the workshop. This particular application demonstrated the effect of reducing dogfish in the system, implemented both in Atlantis and MS-PROD. The inter-model

[29] Such as the ICES Multispecies Assessment Working Group on the one hand and species-oriented working groups on the other, whether in the North Sea, Barents Sea or off Iceland.

[30] In the general sense of technical interactions, effects of predation on predator growth or prey mortality, effect of temperature on the stock-recruit curve or on catchability.

comparison in this case demonstrated that some of the more important effects were estimated to be comparable in the two models, thus giving confidence in the results. The flip side of this is that without the comparison one could not have had confidence in either result: While it is true that it is generally important to compare model output across models, it is also true that such comparisons are essential when one lacks trust in either model! In the simplest case of linear regression with data clearly on the line there is no need to doubt a prediction made within the observed range of x values. The present situation is at the other extreme, when one is extrapolating outside the range of the data using models, which do not fit or explain the data well.

The need for inter-model comparisons should be considerably reduced when more effort is put into making each model better match the observations.

Most of the above has placed an emphasis on the fisheries part of the ecosystem but it seems reasonable to assume that general EBM-related questions will come forth at an increasing pace in the future, even more so than EBFM-related questions. Most of the EBFM issues will almost certainly need to take spatial issues into account and it is not at all unlikely in the future that these will involve issues such as considering overall effort targeted at a system or overall harvests. In principle many of these can be handled using management strategy evaluations using models such as Atlantis as an operating model (but noting the incredible attention which need to be given to detail). The EBM issues are wider-ranging, will likely also mainly be of a spatial nature and it is not clear that any single toolbox will be generally useful for answering such questions. Familiarity with all available databases and capabilities in data analysis will be very important, however: Log-books, satellite information and other data sources of widely variable nature have and will need to be analyzed to answer these questions. Thus the primary capability needed may not be as much ecosystem modeling as database extractions and statistical analyses.

Several of the ideas presented in different documents relating to changes in fishery management, i.e. moving from species-directed management towards area-based management, need to be considered in detail. Given that there is a current management scheme it is not clear, however, just how such changes can be implemented. At the very least they will need to be developed within appropriate fora, i.e. in dialogue with stakeholders. Even the ideas for such changes need to originate in dialogue since otherwise one is unlikely to select appropriate models to cover the various aspects that will crop up in subsequent debates.

Recommendations

- Regarding fisheries, as a matter of priority, a dialogue should be set up (probably through the two SSCs) to advance discussions with (fisheries) stakeholders on how ecosystem issues can best be taken into account (in accordance with the various mandates relating to EBFM) with the intent of bringing discussions to the stage where revised and implementable management procedures can be formalized for evaluation purposes.
- The dialogues will define more clearly the types of toolboxes needed but at present it would seem that spatially disaggregated models are the most likely candidates. The focus of these discussions needs to be on what general forms of revised management strategies are feasible in terms of implementation and in accordance with how management needs to be moving towards the EBFM.
- Regarding other uses of ocean resources (EBM) it may not be possible to develop toolboxes to answer generic stakeholder questions. Development of skills in database extractions, spatial analyses and statistical modeling will always be important however and this should be continued.
- Among the models that have been developed and tested by the group, Atlantis appears to have the greatest potential as an operating model for management strategy evaluations in an ecosystem context. Other candidates are not obvious for this task (but see the text).
- The number of models in use and development should now be reduced and depth rather than breadth should be the priority.
- More modelers with simultaneous expertise in statistics and computer modeling should be added to the group.
- A general move should be made to always attempt to incorporate data in models in a statistical (and objective) manner (only).
- Non-spatial (aggregated) biomass production models, time-series analyses and other models which are not seen to be clearly linked to spatial issues or stakeholder questions should be downgraded in terms of modeling emphasis and relegated to become a less-used part of the toolbox.
- Methods and models designed to either directly answer stakeholder concerns or provide actual explanations of ecosystem behavior should take precedence.

Finally, as a critique of process, this reviewer would like to note that this particular review consists of reviewing more modeling approaches than is usually done during a review process. Future reviews need to concentrate on “smaller” questions such as choosing an operating model OR evaluating the implementation and value of energy transfer models OR evaluating the Atlantis approach as implemented at the NEFSC OR how one can suggest harvest control rules in the multispecies setting. Any one of these would be more than enough for one review.

References

Other references cited in the review but not listed here can be found in the appendices.

J. Begley. Gadget User Manual. Technical Report 120, Marine Research Institute, 2004.

J. Begley and D. Howell. An overview of Gadget, the Globally applicable Area–Disaggregated General Ecosystem Toolbox, 2004. ICES (Int. Counc. Explor. Sea) Document C.M. 2004/FF:13.

L.M. Berliner and Y. Kim. Bayesian design and analysis for superensemble-based climate forecasting. Journal of Climate, 21(9):1891–1910, 2008. ISSN 1520-0442.

B. Bogstad, S. Tjelmeland, T. Tjelda, and O. Ulltang. Description of a multispecies model for the Barents Sea (multspec) and a study of its sensitivity to assumptions on food preferences and stock sizes of minke whales and harp seals. Technical report, SC/44/O 9, 1992.

E.A. Fulton, A.D.M. Smith, and A.E. Punt. Which ecological indicators can robustly detect effects of fishing? ICES Journal of Marine Science, 62(3):540, 2005.

S. Hannesson, A. Jakobsdottir, J. Begley, L. Taylor, and G. Stefansson. On the use of tagging data in statistical multispecies multi-area models of marine populations. ICES Journal of Marine Science, 65(9):1762–1772, DEC 2008. ISSN 1054-3139. doi: 10.1093/icesjms/fsn132.

V. Kupca. A standardized database for fisheries data, 2004a. ICES (Int. Counc. Explor. Sea) Document C.M. 2004/FF:15.

V. Kupca. Standardised fisheries database. In Final Report: dst2: Development of structurally detailed statistically testable models of marine populations, volume 118 of Technical Report, pages 1–57. Marine Research Institute, Reykjavik, 2005.

V. Kupca. Bootstrapping of areas. In Final Report: dst2: Development of structurally detailed statistically testable models of marine populations, volume 118 of Technical Report. Marine Research Institute, Reykjavik, 2004b.

V. Kupca and P. Sandbeck. dst2 Datawarehouse structure and data import. In dst2: Development of structurally detailed statistically testable models of marine populations, volume 98 of Technical Report, pages 37–46. Marine Research Institute, Reykjavik, 2003.

J. S. Link, R. J. Gamble, and M. J. Fogarty. An overview of the nefsc ecosystem modeling enterprise for the northeast us shelf large marine ecosystem: toward ecosystem-based management., 2011. MS for NEFSC Ecosystem Modeling Review, March 2011.

P. D. MacDonald and T. J. Pitcher. Age–groups from size frequency data: a versatile and efficient method of analyzing distribution mixtures. J. Fish. Res. Board. Can., 36:987–1001, 1979.

J.G. Pope. Simple representations of the yield surface., 1989. Working Paper submitted to the ICES Multispecies Assessments Working Group.

G. Stefansson. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES Journal of Marine Science, 53(3):577, 1996.

G. Stefansson. Comparing different information sources in a multispecies context. In F. Funk, TJ Quinn II, J. Heifetz, JN Ianelli, JE Powers, JF Schweigert, PJ Sullivan, and CI Zhang, editors, Fishery Stock Assessment Models: Proceedings of the international symposium; Anchorage 1997, 15th Lowell Wakefield Fisheries Symposium, pages 741–758, 1998.

G. Stefansson and O.K. Palsson. Statistical evaluation and modelling of the stomach contents of Icelandic cod (Gadus morhua). Canadian Journal of Fisheries and Aquatic Sciences, 54 (1):169–181, 1997.

G. Stefansson and OK Palsson. A framework for multispecies modelling of Arcto-boreal systems. Reviews in Fish Biology and Fisheries, 8(1):101–104, 1998.

G. Stefansson and A.A. Rosenberg. Combining control measures for more effective management of fisheries under uncertainty: quotas, effort limitation and protected areas. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1453):133, 2005.

G. Stefansson and AA Rosenberg. Designing marine protected areas for migrating fish stocks. Journal of Fish Biology, 69(Suppl C):66–78, 2006.

L. Taylor. Datawarehouse for Icelandic waters. In dst2: Development of structurally detailed statistically testable models o f marine populations, volume 98 of Technical Report, pages 64–75. Marine Research Institute, Reykjavik, 2003.

L. Taylor, J. Begley, V. Kupca, and G. Stefansson. A simple implementation of the statistical modelling framework Gadget for cod in Icelandic waters. African Journal of Marine Science, 29(2):223–245, AUG 2007. ISSN 1814-232X. doi: 10.2989/AJMS.2007.29.2.7.190.

L. A. Taylor. Gadget models of cod–shrimp interactions in icelandic waters. Technical Report RH-03-2011, Science Institute, University of Iceland, 2011.

L. A. Taylor, V. Trenkel, V. Kupca, and G. Stefansson. A bootstrap method for highly

disparate data sets used with statistical multispecies models. Technical Report RH-02-2011, Science Institute, University of Iceland, 2011.

S. Tjelmeland and B. Bogstad. MULTSPEC-a review of a multispecies modelling project for the Barents Sea. Fisheries Research, 37(1):127–142, 1998.

S. Tjelmeland and B. Bogstad. MULTSPEC: The manual. Institute of Marine Research, Bergen, 1989.

C.J. Walters, V. Christensen, S.J. Martell, and J.F. Kitchell. Possible ecosystem impacts of applying MSY policies from single-species assessment. ICES Journal of Marine Science: Journal du Conseil, 62(3):558, 2005. ISSN 1054-3139.

Appendix 1: Bibliography of materials provided for review

NEFSC Ecosystem Modeling Review
Background Readings

Modeling Overviews

Link, J.S., Bundy, A., Overholtz, W.J., Shackell, N., Manderson, J., Duplisea, D., Hare, J., Koen-Alonso, M. & Friedland, K. 2011. Northwest Atlantic Ecosystem Based Management of Fisheries. *in* Belgranno, A. and Fowler, C.W. (eds.) Ecosystem Based Management for Fisheries: Linking Patterns to Policy. pp. 32-112.

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

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.

ICES. 2009. Report of the Working Group on Multispecies Assessment Methods (WGSAM), 5–9 October 2009, ICES Headquarters, Copenhagen. ICES CM 2009/RMC:10. 117 pp.

ICES. 2008. Report of the Working Group on Multispecies Assessment Methods (WGSAM), 6–10 October 2008, ICES Headquarters, Copenhagen. ICES CM 2008/RMC:06. 113 pp.

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 NEMoW (National Ecosystem Modeling Workshop). NOAA Technical Memorandum NMFS-F/SPO-87. pp. 93.

ICES. 2007. Report of the Working Group on Multispecies Assessment Methods (WGSAM), 15–19 October 2007, San Sebastian, Spain. ICES CM 2007/RMC:08. 134 pp.

Whipple, S., Link, J.S. , L.P. Garrison & Fogarty, M.J. 2000. Models of predation and fishing mortality in aquatic ecosystems. *Fish. and Fisheries*. 1:22-40.

Hollowed, A.B., Bax, N., Beamish, R., Collie, J., Fogarty, M., Livingston, P., Pope, J. & Rice, J.C. 2000. Are multispecies models an improvement on single-species models for measuring fishing impacts on marine ecosystems? ICES Journal of Marine Science, 57:707-719.

Link, J.S. and Bundy, A. (In Press). Ecosystem Modeling in the Gulf of Maine Region: Towards an Ecosystem Approach to Fisheries. *AFS Gulf of Maine Symposium Proceedings*

Link, J.S., A. Bundy, W.J. Overholtz, N. Shackell, J. Manderson, D. Duplisea, J. Hare, M. Koen-Alonso, K.D. Friedland. (In Press). Northwest Atlantic Ecosystem-Based Fisheries Management. *Fish and Fisheries*

ESAMs & Influence on BRPs (MRMs 1)

Tyrrell, M.C., Link, J.S. & Moustahfid, H. 2011. The importance of including predation in some fish population models: implications for biological reference points. *Fisheries Research* 108-1-8.

NEFSC. 2011. 51st Northeast Regional Stock Assessment Workshop (51st SAW) Assessment Report. Northeast Fisheries Science Center Reference Document 11-02; 856 p.

NEFSC. 2011. 51st Northeast Regional Stock Assessment Workshop (51st SAW): Assessment Summary Report. Northeast Fisheries Science Center Reference Document 11-01, part A, B, C.

Debroba, J., G. Shepherd, F. Grégoire, J. Nieland, and J. Link. 2010. Stock Assessment of Atlantic Mackerel in the Northwest Atlantic - 2009. TRAC Reference Document - 2010/01. 59pp.

DFO. 2010. Proceedings of the Transboundary Resources Assessment Committee (TRAC) Spiny Dogfish Review. Proceedings 2010/01. 57 pp.

NEFSC. 2010. 50th Northeast Regional Stock Assessment Workshop (50th SAW) Assessment Report. Northeast Fisheries Science Center Reference Document 10-17, part A, C.

NEFSC. 2010. 49th Northeast Regional Stock Assessment Workshop (49th SAW) Assessment Report. Northeast Fisheries Science Center Reference Document 10-03. Appendix B1.

Link, J.S. & J.S. Idoine. 2009. Predator Consumption Estimates of the northern shrimp *Pandalus borealis*, with Implications for Estimates of Population Biomass in the Gulf of Maine. *N. Amer. J. Fish. Manag.* 29:1567-1583.

Moustahfid, H., Tyrrell, M.C. & Link, J.S. 2009. Accounting explicitly for predation mortality in surplus production models: an application to longfin inshore squid. *N. Amer. J. Fish. Manag.* 29:1555-1566.

Moustahfid, H., Link, J.S., Overholtz, W.J. & Tyrell, M.C. 2009. The advantage of explicitly incorporating predation mortality into age-structured stock assessment models: an application for Northwest Atlantic mackerel. *ICES J. Mar. Sci.* 66: 445-454.

Fogarty, M. L. Incze, K. Hayhoe, D.Mountain and J. Manning. 2008. Potential Climate Change Impacts on Atlantic Cod (Gadus morhua) off the Northeastern United States. Mitig. Adapt. Strat. Glob.Change. 13:453-466.

Overholtz, W.J., Jacobson, L.D. & Link, J.S. 2008. Developing an ecosystem approach for assessment advice and biological reference points for the Gulf of Maine-Georges Bank herring complex: adding the impact of predation mortality. *N. Amer. J. Fish. Manag.* 28:247-257.

Link, J.S. & Sosebee, K. 2008. Estimates and implications of Skate Consumption in the northeastern US continental shelf ecosystem. *N. Amer. J. Fish. Manag.* 28:649-662.

NEFSC. 2007. Assessment Report (45th SARC/SAW). Section A.10. [TOR 6]. Northeast Fisheries Science Center Reference Document, 07-16. pp 13-138.

NEFSC. 2007. Assessment Report (44th SARC/SAW). Section B.8. [TOR 6]. Northeast Fisheries Science Center Reference Document, 07-10. pp 332-344, 504-547.

Tyrrell, M.C., Link J.S., Moustahfid, H. & Smith, B.E. 2007. The dynamic role of pollock (Pollachius virens) as a predator in the Northeast US Atlantic ecosystem: a multi-decadal perspective. *J. Northwest Atl. Fish. Sci.* 38:53-65.

Overholtz, W.J. & Link, J.S. 2007. Consumption impacts by marine mammals, fish, and seabirds on the Gulf of Maine-Georges Bank Atlantic Herring (*Clupea harengus*) complex during 1977-2002. *ICES J. Mar. Sci.* 64:83-96.

Overholtz, W.J., Link, J.S. & Suslowicz, L.E. 1999. Consumption and harvest of pelagic fishes in the Gulf of Maine-Georges Bank ecosystem: Implications for fishery management. Proceedings of the 16th Lowell Wakefield Fisheries Symposium- Ecosystem Considerations in Fisheries Management. AK-SG-99-01: 163-186.

Multispecies Models (MRMs 2)

Link, J.S., B.A. Megrey, T.J. Miller, T. Essington, J. Boldt, A. Bundy, E. Moksness, K.F. Drinkwater & R. I. Perry. 2010. Comparative Analysis of Marine Ecosystems: International Production Modelling Workshop. *Biol. Lett.* doi: 10.1098/rsbl.2010.0526

Garrison, L.P., Link, J.S., Cieri, M., Kilduff, P., Sharov, A., Vaughan, D., Muffley, B., Mahmoudi, B., & Latour, R. 2010. An Expansion of the MSVPA Approach for Quantifying Predator-Prey Interactions In Exploited Fish Communities. *ICES J. Mar. Sci.* 67: 856-870.

Gamble, R.J. & Link, J.S. 2009. Analyzing the Tradeoffs Among Ecological and Fishing Effects on an Example Fish Community: a Multispecies (Fisheries) Production Model. *Ecol. Model.* 220:2570-2582.

Tyrrell, M.C., Link, J.S., Moustahfid, H. & Overholtz, W.J. 2008. Evaluating the effect of predation mortality on forage species population dynamics in the Northwest Atlantic continental shelf ecosystem: an application using multispecies virtual population analysis. *ICES J. Mar. Sci.* 65:1689-1700.

Hall, S.J., Collie, J.S., Duplisea, D.E., Jennings, S., Bravington, M. & Link, J. 2006. A length-based multi-species model for evaluating community responses to fishing. *Can. J. Fish. Aquat. Sci.* 63:1344-1359.

NEFSC 2006. 42nd Northeast Regional Stock Assessment Workshop (42nd SAW). Northeast Fisheries Science Center Reference Document, 06-09b, 308 pp.

Garrison, L. and Link, J. 2004. An Expanded Multispecies Virtual Population Analysis Approach (MSVPAX) to Evaluate Predator-Prey Interactions in Exploited Fish Ecosystems. VERSION 1.1. Users Manual and Model Description. Atlantic States Marine Fisheries Commission, Wahsington, D.C.

White, G.C., Kline, L.L., Garrison, L.P. & Link, J.S. 2003. Expansion of multispecies modeling assessment approach for management of coastal fisheries for Atlantic menhaden, bluefish, striped bass and weakfish. Annual Report to the NOAA Chesapeake Bay Office.

Food Web & Network Models

Link, J.S. 2010. Adding Rigor to Ecological Network Models by Evaluating a Set of Pre-balance Diagnostics: A Plea for PREBAL. *Ecol. Model.* 221:1582-1593.

Byron, C., D. Bengtson, J. Link, B. Costa-Pierce, R. Rheault, D. Beutel. and D. Alves. 2010. Modeling Carrying Capacity for Bivalve Aquaculture Using Mass-Balance Modeling and Stakeholder Collaboration. ICES CM 2010/J:03, 31p.

Overholtz, W.J. & Link, J.S. 2009. A simulation model to explore the response of the Gulf of Maine food web to large scale environmental and ecological changes. *Ecol. Model.*220:2491-2502.

Gaichas, S., Skaret, G., Falk-Petersen, J., Link, J.S., Overholtz, W., Megrey, B.A., Gjoesaeter, H., Stockhausen, W., Dommasnes, A. & Aydin, K. 2009. A comparison of community and trophic structure in five marine ecosystems based on energy budgets and system metrics. *Prog. Oceanogr.* 81:47-62.

Link, J.S., Col, L., Guida, V., Dow, D., O'Reilly, J., Green, J., Overholtz, W., Palka, D., Legault, C., Vitaliano, J., Griswold, C., Fogarty, M. & Friedland, K. 2009. Response of Balanced Network Models to Large-Scale Perturbation: Implications for Evaluating the Role of Small Pelagics in the Gulf of Maine. *Ecol. Model.* 220: 351-369.

Link, J., Overholtz, W., O'Reilly, J., Green, J., Dow, D., Palka, D., Legault, C., Vitaliano, J., Guida, V., Fogarty, M., Brodziak, J., Methratta, E., Stockhausen, W., Col, L., Waring, G., & Griswold, C. 2008. An Overview of EMAX: The Northeast U.S. Continental Shelf Ecological Network. *J. Mar Sys.* 74:453-474.

Link, J.S., O'Reilly, J., Dow, D., Fogarty, M., Vitaliano, J., Legault, C., Overholtz, W., Green, J., Palka, D., Guida, V. & Brodziak, J. 2008. Energy Flow on Georges Bank revisited: the energy modeling and analysis eXercise (EMAX) in historical context. *J. Northwest Atl. Fish*. *Sci.* 39:83-101.

Steele, J.H., Collie, J.S., Bisagni, J.J., Gifford, D.J., Fogarty, M.J., Link, J.S., Sullivan, B.K., Sieracki, M.K., Beet, A.R., Mountain, D.G., Durbin, E.G., Palka, D. & Stockhausen, W.T. 2007. Balancing end-to-end budgets of the Georges Bank ecosystem. *Prog. Oceanogr*. 74:423-448.

Link, J.S., Griswold, C.A. Methratta, E.M. & Gunnard, J. (eds). 2006. Documentation for the Energy Modeling and Analysis eXercise (EMAX). Northeast Fisheries Science Center Reference Document, 06-15. 166 pp.

Steele, J., Bisagni, J., Collie, J., Fogarty, M., Gifford , M., Link, J., Sieracki, M., Sullivan, B., & Beet, A. 2005. Constructing end-to-end budgets for the Georges Bank ecosystem. ICES CM 2005/M08 16 pp.

Link, J., W. Overholtz, J. O'Reilly, J. Green, E. Methratta, D. Dow, D. Palka, C. Legault, S. Edwards, G. Waring, W. Stockhausen, D. Mountain, J. Vitaliano, V. Guida, J. Kane, J. Jossi, M. Fogarty, J. Brodziak, C. Griswold, C. McCandless, N. Kohler, S. Fromm, L. Col, T. Smith, C. MacKenzie, and R. Goldberg. 2005. An Overview of EMAX: The Northeast U.S. Continental Shelf Ecological Network. ICES CM 2005/L02 40 pp.

Link, J.S., Stockhausen, W.T. & Methratta, E.T. 2005. Food web theory in marine ecosystems. pp. 98-113 *in* Belgrano, A., Scharler U.M., Dunne, J. & Ulanowicz, R.E. (eds.) Aquatic Food Webs: an Ecosystem Approach. Oxford Univ. Press, Oxford.

Link, J.S. 2004. A general model of selectivity for fish feeding: a rank proportion algorithm. *Trans. Amer. Fish. Soc.* 133:655-673.

Link, J.S. 2002. Does food web theory work for marine ecosystems? *Mar. Ecol. Prog. Ser.* 230:1-9.

Link, J. 1999. (Re)Constructing Food Webs and Managing Fisheries. Proceedings of the 16th Lowell Wakefield Fisheries Symposium- Ecosystem Considerations in Fisheries Management. AK-SG-99-01: 571-588.

Link, J. & Keen, R. 1999. A model of salmonid planktivory: Field test of a mechanistic approach to size-selection. *Ecol. Model.* 117:269-283.

Byron, C., J. Link, D. Bengston, B. Costa-Pierce. (in press). Modeling Carrying Capacity of Shellfish Aquaculture in Highly Flushed Temperate Lagoons. *Aquaculture.*

Byron, C., J. Link, D. Bengston, B. Costa-Pierce. (in press). Calculating Carrying Capacity of Shellfish Aquaculture Using Mass-Balance Modeling: Narragansett Bay, Rhode Island. *Ecol. Model.*

Aggregate (Production) Models

NEFSC 2008. Assessment of 19 Northeast Groundfish Stocks through 2007 Report of the 3rd Groundfish Assessment Review Meeting (GARM III), Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 4-8, 2008. Section 2.1. Northeast Fisheries Science Center Reference Document, 08-15. pp. 855-865.

Overholtz, W.J., J.S. Link, M.J. Fogarty, L.Col, and C. Legault. 2008. US Northeast Shelf LME Biomass, target biological reference points for fish and worldwide cross-system comparisons. GARM WP 3.1 GARM-III-BRP Meeting

Link, J.S., W.J. Overholtz, C. Legault, L. Col, M.J. Fogarty. 2008. Energy budget contextualization of fish biomasses at B_MSY. GARM WP 3.2 GARM-III-BRP Meeting

Overholtz, W.J., M.F. Fogarty, J.S. Link, C. Legault, and L. Col. 2008. Estimates of aggregate surplus production for the GARM and other stock groups for the US Northeast Shelf LME. GARM WP 3.3 GARM-III-BRP Meeting

Link, J.S., R. Gamble, W.J. Overholtz, C. Legault, L. Col, and M.J. Fogarty. 2008. An Aggregate and MS Production Model: A Simulator Tool GARM WP 3.4 GARM-III-BRP Meeting

Fogarty, M.J., W.J. Overholtz, and J. Link. 2008. Fishery Production Potential of the Northeast Continental Shelf of the United States. GARM WP 3.5 GARM-III-BRP Meeting

Overholtz, W.J., J.S. Link, M. Fogarty, L. Col, and C. Legault. 2008. Target Biological Reference Points, Worldwide Cross System Comparisons, and Aggregate Production Model Results for GARM Stocks. GARM WP 2.1 GARM-III-Final Meeting

Link, J.S. 2003. A Model of Aggregate Biomass Tradeoffs. ICES Annual Science Conference. Theme Session on Reference Point Approaches to Management within the Precautionary Approach. ICES CM 2003/Y08 28 pp.

Overholtz, W.J., M.J. Fogarty and J.S. Link. (In review). Using aggregate surplus production models to assess the overall production potential of the demersal fish resources of the Northeast U.S. Shelf Large Marine Ecosystem.

Full System Models

Link, J.S., E.A. Fulton, and R.J. Gamble. 2010. The Northeast US Application of ATLANTIS: A full system model exploring marine ecosystem dynamics in a living marine resource management context. *Progress in Oceanography.* 87:214-234.

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. (In Press). Lessons in modelling and management of marine ecosystems: The Atlantis experience. *Fish and Fisheries*

Link, J.S., Gamble, R.J., and Fulton, E.A. (In press). NEUS – ATLANTIS: Construction, Calibration and Application of an Ecosystem Model with Ecological Interactions, Physiographic Conditions, and Fleet Behavior. NOAA Tech. Memo. NEFSC

Patrick, W.S., P. Spencer, J. Link, O. Ormseth, J. Cope, J. Field, D. Kobayashi, T. Gedamke, E. Cortés, K. Bigelow, W. Overholtz, & P. Lawson. 2010. Assessing the vulnerability of U.S. fisheries to becoming overfished or undergoing overfishing through the use of productivity and susceptibility indices. *Fish. Bull.* 108:305-322.

Shin, Y.-J., L.J. Shannon, A. Bundy, M. Coll, J.L. Blanchard, M.-F. Borges, J. Cotter, P.M. Cury, J.J. Heymans, D. Jouffre, J.S. Link, C. Möllmann, H. Ojaveer, K.O.M. Abdallahi, M.-J. Rochet, & D. Yemane. 2010. Using indicators for evaluating, comparing and communicating the ecological status of exploited marine ecosystems. Part 2: Setting the scene. *ICES J. Mar. Sci.* 67: 692-716.

Shin, Y.-J., A. Bundy, L.J. Shannon, M. Simier, M. Coll, E.A. Fulton, J.S. Link, D. Jouffre, H. Ojaveer, S. Mackinson, J.J. Heymans, T. Raid. 2010. Can simple be useful and reliable? Using ecological indicators for representing and comparing the states of marine ecosystems. *ICES J. Mar. Sci.* 67: 717-731.

Blanchard, J.L., Coll, M., Cotter, J., Link, J., Trenkel, V., Vergnon, R., Yemane, D., & Shin, Y-J. 2010. Trend analysis of indicators: a comparison of recent changes in the status of marine ecosystems around the world. *ICES J. Mar. Sci.* 67: 732-744.

Link, J.S., Yemane, D., Shannon, L.J., Coll, M., Shin, Y-J., Hill, L., Borges, M.F., Bundy, A. & Aydin, K. 2010. Relating Marine Ecosystem Indicators to Fishing and Environmental Drivers: An Elucidation of Contrasting Responses. *ICES J. Mar. Sci.* 67: 787-795.

Coll, M., Shannon, L.J., Yemane, D., Link, J., Ojaveer, H., & Shin, Y-J. 2010. Ranking ecological relative status of exploited ecosystems using multiple indicators. *ICES J. Mar. Sci.* 67: 769-786.

Ecosystem Assessment Program. 2009. Ecosystem Status Report For the Northeast U.S. Continental Shelf Large Marine Ecosystem. Northeast Fisheries Science Center Reference Document, 09-11. 34 p.

Patrick, W.S., P. Spencer, O. Ormseth, J. Cope , J. Field, D. Kobayashi, T. Gedamke, E. Cortés, K. Bigelow, W. Overholtz, J. Link, and P. Lawson. 2009. Use of productivity and susceptibility indices to determine the vulnerability of a stock, with example applications to six U.S. fisheries. NOAA Tech. Memo. NMFS-F/SPO-101, 90p.

Pranovi, F. & Link, J.S. 2009. Ecosystem exploitation and trophodynamic indicators: a comparison between the Northern Adriatic Sea and Southern New England. *Prog. Oceanogr.* 81:149-164.

Methratta, E.T. & Link, J.S. 2006. Evaluation of Quantitative Indicators for Marine Fish Communities. *Ecol. Indicators* 6:575-588.

Link, J. S. 2006. Northeastern United States. Pp. 65 in Kruse, G.H., Livingston, P., Overland, J.E., Jamieson, G.S., McKinell, S. and Perry, R.I. (eds). Report of the PICES/NPRB Workshop on Integration of Ecological Indicators of the North Pacific with Emphasis on the Bering Sea. PICES Scientific Report No. 33.

Link, J.S. 2005. Translation of Ecosystem Indicators into Decision Criteria. *ICES J. Mar. Sci.* 62:569-576.

Link, J.S. & Brodziak, J.K.T. 2003. Discussion Group Report for Defining Ecosystem Overfishing. Proceedings of the 7th National Stock Assessment Workshop: (Re)Building Sustainable Fisheries and Marine Ecosystems. NOAA Tech. Memo. NMFS-F/SPO-62. pp. 29-32.

Link, J. & Brodziak, J. (eds.). 2002. Report on the Status of the NE US Continental Shelf Ecosystem. NEFSC Ecosystem Status Working Group. Northeast Fisheries Science Center Ref. Doc. 02-11, 245 pp.

Link, J.S., Brodziak, J.K.T., Edwards, S.F., Overholtz, W.J., Mountain, D., Jossi, J.W., Smith, T.D. & Fogarty, M.J. 2002. Marine Ecosystem Assessment in a Fisheries Management Context. Can. J. Fish. Aquat. Sci. 59:1429-1440.

Link, J.S., Brodziak, J.K.T., Edwards, S.F., Overholtz, W.J., Mountain, D., Jossi, J.W., Smith , T.D. & Fogarty, M.J. 2001. Ecosystem status in the Northeast United States Continental Shelf Ecosystem: Integration, synthesis, trends and meaning of ecosystem metrics. Or Getting to the brass tacks of ecosystem based fishery management. ICES Annual Science Conference. Theme Session on the Use and Information Content of Ecosystem Metrics and Reference Points. ICES CM 2001/T10 41 pp.

Murawski, S.A. 2000. Definitions of overfishing from an ecosystem perspective. ICES J. Marine Science 57:649-528.

Overholtz, W.O., Link, J.S. & Murawski, S.A. 2000. Working group report: ecosystem indicators. Proceedings of the 6th National Stock Assessment Workshop: Incorporating Ecosystem Considerations into Stock Assessments and Management Advice. NOAA Tech. Memo. NMFS-F/SPO-46. pp. 51-53.

IEAs, MSE and associated context

ICES. 2010. Report of the Working Group on the Northwest Atlantic Regional Sea (WGNARS), 20–22 April 2010, Woods Hole, USA. ICES CM 2010/SSGRSP:03. 63 pp.

NEFMC SSC. 2010. White Paper On Ecosystem – Based Fishery Management For New England Fishery Management Council. Newburyport, MA. 25 pp.

ICES. 2009. Report of the Working Group on Holistic Assessments of Regional Marine Ecosystems (WGHAME), 12-16 October 2009, ICES Headquarters, Copenhagen. ICES CM 2009/RMC:13. 76 pp.

Levin P.S., M.J. Fogarty, S.A. Murawski, and D. Fluharty. 2009. Integrated Ecosystem Assessments: Developing the Scientific Basis for Ecosystem-Based Management of the Ocean. PLoS Biol 7(1): e1000014. doi:10.1371/journal.pbio.1000014

Levin, P.S., M.J. Fogarty, G.C. Matlock, and M. Ernst. 2008. Integrated ecosystem assessments. U.S. Dept. Commer.,NOAA Tech. Memo. NMFS-NWFSC-92, 20 p.

NAFO. 2008. Report of the NAFO SC Working Group on Ecosystem Approach to Fisheries Management (WGEAFM). SCS08-24. 19 pp. Dartmouth, NS.

NAFO. 2008. Report of the WGEAFM Meeting May 26-30, 2008. NAFO SCS 08/10, Serial No. N5511, 70 p. Dartmouth, NS.

Appendix 2: CIE Statement of Work for Gunnar Stefansson

External Independent Peer Review by the Center for Independent Experts

Review of Modeling Approaches in Support of Ecosystem-Based Fishery Management

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description: The purpose of this review is to evaluate the appropriateness and performance characteristics of community-level and ecosystem models employed at NEFSC as operating models in support of the development of Ecosystem-Based Fishery Management (EBFM) strategies for the Northeast U.S. Continental Shelf. NMFS strongly endorsed the concept of Ecosystem-Based Management and the related need for the development of Integrated Ecosystem Assessment in support of EBFM. Although this review is directed at efforts in the NEFSC, the findings will be more broadly applicable throughout the agency. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein.

CIE reviewers shall have working knowledge and recent experience in the application of community-level and ecosystem models for EBFM. The CIE reviewers shall have expertise with a broad spectrum of complexity and mechanistic detail from static energy flow models to detailed simulation models, and familiarity with the ATLANTIS model is desirable. Our objective is to employ multi-model inference to assess options for EBFM. We are particularly interested in the question of tradeoffs between model complexity and predictive skill in meeting the needs for scientific advice in support of EBFM. Operating models lie at the heart of the development of Integrated Ecosystem Assessments (IEAs). IEAs have been strongly advocated at the agency level as the principle vehicle for developing and evaluating scientific advice in support of EBFM. It is essential that a rigorous review of modeling activities be undertaken to meet this need.

CIE reviewers shall have experience in different approaches to modeling exploited marine ecosystems. The approaches currently employed in this region include mass balance energy flow models, aggregate-species production models with implicit consideration of species interactions, multispecies production models with explicit consideration of interspecific interactions, state-space multispecies models, multispecies delay-difference models, and the ATLANTIS modeling framework. Reviewers shall have direct experience in model development with EBFM application.

Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Woods Hole, Massachusetts during 29-31 March 2011.

Statement of Tasks: Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/sponsor.html).

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

The reviewers will be supplied with a review document describing ongoing modeling efforts at NEFSC in support of ecosystem-based fishery management:

Community and Ecosystem Models in Support of Ecosystem-Based Fishery Management for the Northeast U.S. Continental Shelf. Projected length 125-150 pp maximum

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

The CIE peer reviewers will provide a critical evaluation of the community-level and ecosystem modeling conducted at NEFSC in support of EBFM. The adequacy of the overall modeling framework to meet the needs of EBFM in this region will be assessed and recommended changes to modeling strategies will be provided. The reviewers will contribute individual perspectives on the findings and recommendations for each ToRs. The panel Chair will be responsible for overall compilation of the report of the peer review and in the development of a summary statement of the adequacy of the modeling effort in relationship to the requirements for EBFM in this region.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Other Tasks – Contribution to Summary Report: Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting at the Northeast Fisheries Science Center in Woods Hole, Massachusetts during 29-31 March 2011.
3) At the Northeast Fisheries Science Center in Woods Hole, Massachusetts during 29-31 March 2011 as specified herein, conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than 14 April 2011, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and to Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

22 February 2011	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
15 March 2011	NMFS Project Contact sends the CIE Reviewers the pre-review documents
March 29-31 2011	Each reviewer participates and conducts an independent peer review during the panel review meeting
14 April 2011	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
28 April 2011	CIE submits CIE independent peer review reports to the COTR
5 May 2011	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: Requests to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on substitutions. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:

(1) each CIE report shall completed with the format and content in accordance with **Annex 1**,

(2) each CIE report shall address each ToR as specified in **Annex 2**,

(3) the CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Contracting Officer's Technical Representative (COTR)
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
shivlanim@bellsouth.net Phone: 305-383-4229

Roger W. Peretti, Executive Vice President
Northern Taiga Ventures, Inc. (NTVI)
22375 Broderick Drive, Suite 215, Sterling, VA 20166
RPerretti@ntvifederal.com Phone: 571-223-7717

Key Personnel:

NMFS Project Contact:

Michael Fogarty
Northeast Fisheries Science Center, 166 Water St. Woods Hole, MA, 02543
mfogarty@mercury.wh.whoi.edu Phone: 508-495-2352

Appendix 2.1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.
   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.
   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.
   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.
   e. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting

Appendix 2.2: Terms of Reference for the Peer Review – as presented to the Review Panel in advance

Review of Modeling Approaches in Support of Ecosystem-Based Fishery Management

1. Evaluation, findings and recommendations of overall community-ecosystem level modeling strategy
2. Evaluation of strengths and weaknesses, and recommendations of analytic methodologies
3. Evaluation and recommendations of model assumptions, estimates, and uncertainty
4. Evaluation, findings, and recommendations of result interpretation and conclusions
5. Determine whether the science reviewed is considered to be the best scientific information available.
6. Recommendations for further improvements
7. Brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations

Appendix 2.3: Terms of Reference for the Peer Review – as presented at the Meeting

Review of Modeling Approaches in Support of Ecosystem-Based Fishery Management

A. Overall Review- Synthesis & Summary
   a. Summarize evaluations, findings and recommendations of overall community-ecosystem level modeling strategy in practice for the NEUS LMR system
   b. Determine whether the science reviewed is considered to be the best scientific information available.
   c. Determine if the intended uses of overall community-ecosystem level modeling that have been identified as priorities for the NEUS LME are being executed in accordance with global best practices.
   d. Provide recommendations for further improvements.
   e. Provide brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations

B. Energy Transfer Models *(Fogarty)- Production Potential Models*
   i. Review and agree upon data requirements requisite for the model
   ii. Evaluate adequacy of input data as applied for the NEUS application of this model
   iii. Evaluate strengths and weaknesses of analytical methodologies
   iv. Evaluate if model structure, parameterization, calibration/tuning, validation/verification, and intended uses are adequately documented
   v. Evaluate strengths and weaknesses of assumptions, example estimations, and sources of uncertainties, especially with respect to known best practices in the field
   vi. Review types/levels of use for model outputs, especially with respect to adequacy of modeling relative to major topical issues
   vii. Recommendations

C. Energy Transfer Models *(Link)- Network Models*
   i. Review and agree upon data requirements requisite for the model
   ii. Evaluate adequacy of input data as applied for the NEUS application of this model
   iii. Evaluate strengths and weaknesses of analytical methodologies
   iv. Evaluate if model structure, parameterization, calibration/tuning, validation/verification, and intended uses are adequately documented
   v. Evaluate strengths and weaknesses of assumptions, example estimations, and sources of uncertainties, especially with respect to known best practices in the field

vi. Review types/levels of use for model outputs, especially with respect to adequacy of modeling relative to major topical issues
vii. Recommendations

D. Aggregate Production Models *(Link/Fogarty)*
i. Review and agree upon data requirements requisite for the model
ii. Evaluate adequacy of input data as applied for the NEUS application of this model
iii. Evaluate strengths and weaknesses of analytical methodologies
iv. Evaluate if model structure, parameterization, calibration/tuning, validation/verification, and intended uses are adequately documented
v. Evaluate strengths and weaknesses of assumptions, example estimations, and sources of uncertainties, especially with respect to known best practices in the field
vi. Evaluate levels, methods and ramifications for aggregation and compare to single species summaries
vii. Review types/levels of use for model outputs, especially with respect to adequacy of modeling relative to major topical issues
viii. Recommendations

E. Multispecies Production Models *(Gamble/Fogarty)*
i. Review and agree upon data requirements requisite for the model
ii. Evaluate adequacy of input data as applied for the NEUS application of this model
iii. Evaluate strengths and weaknesses of analytical methodologies
iv. Evaluate if model structure, parameterization, calibration/tuning, validation/verification, and intended uses are adequately documented
v. Evaluate strengths and weaknesses of assumptions, example estimations, and sources of uncertainties, especially with respect to known best practices in the field
vi. Review types/levels of use for model outputs, especially with respect to adequacy of modeling relative to major topical issues
vii. Recommendations

F. Full System Models *(Link/Gamble)- ATLANTIS*
i. Review and agree upon data requirements requisite for the model
ii. Evaluate adequacy of input data as applied for the NEUS application of this model
iii. Evaluate strengths and weaknesses of analytical methodologies
iv. Evaluate if model structure, parameterization, calibration/tuning, validation/verification, and intended uses are adequately documented

v. Evaluate strengths and weaknesses of assumptions, example estimations, and sources of uncertainties, especially with respect to known best practices in the field
vi. Evaluate levels, methods and ramifications for aggregation
vii. Review types/levels of use for model outputs, especially with respect to adequacy of modeling relative to major topical issues
viii. Recommendations

G. Other models *(Fogarty/Link)*
   i. Briefly review and comment upon other community and ecosystem models for the NEUS ecosystem. For each:
      1. Review simple summaries
      2. Evaluate examples of intended/extant uses
      3. Identify any gaps in model uses
   ii. Empirical multivariate time series
   iii. MRMs
      1. ESAMs
      2. Other MS models
   iv. Others
   v. Recommendations

Appendix 2.3: Tentative Agenda – as presented to the Review Panel in advance

Review of Modeling Approaches in Support of Ecosystem-Based Fishery Management

Northeast Fisheries Science Center, Woods Hole MA 02543

March 29-31 2011

March 29 2011

900 Welcome to Workshop and Overview of Objectives for the Review

930 Review of Overview Modeling Strategy and Philosophy for Multi-Model Inference

1030 Break

1100 Review of Data Sources in Support of Community-Ecosystem Modeling

1230 Lunch

1400 Review of Data Sources Continued

1530 Break

1600 Review of Community/Ecosystem Model Applications

1730 Adjourn

March 30 2011

0900 Review of Community/Ecosystem Model Applications Continued

1030 Break

1100 Review of Community-Ecosystem Modeling Applications Continued

1230 Lunch

1400 Review of Community-Ecosystem Modeling Applications Continued

1530 Break

1600 Discussion

1730 Adjourn

March 31 2011

0900 Discussion/Questions Continued

1030 Break

1100 Panel Deliberations

1230 Lunch

1400 Panel Deliberations Continued

1730 Adjourn

Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Panel members: Dr A. Smith, Dr G. Stefansson and Dr V. Christensen

Meeting participants in the Review of Modeling Approaches in Support of EBFM, 3/29 - 3/31/2011

1. Ingrid Biedson – Cornell
2. Tom Hoff – MAFMC
3. Wendy Gabriel – NEFSC
4. Kiersten Curti – NEFSC
5. Rich Bell – URI/NMFS
6. Anne Richards – NEFSC
7. Sean Lucey – NEFSC
8. Steve Sutton – NEFSC
9. Ron Schlitz – NEFSC
10. Burton Shank – NEFSC
11. Linda Deegan – MBL
12. Hui Liu – NEFSC
13. Rob Gamble – NEFSC
14. Tony Smith – CSIRO Australia
15. Villy Christensen - UBC
16. Gunnar Stefansson – University of Iceland
17. Frank Almeida – NEFSC
18. Jon Hare – NEFSC – Narragansett
19. Michael Jones – NEFSC
20. Kimberly Murray – NEFSC
21. David McElroy
22. Laurel Col – NEFSC
23. Devorah Hart – NEFSC
24. Mike Fogarty – NEFSC
25. Jason Link - NEFSC