

**CHAIR'S SUMMARY OF
PROGRAM REVIEW OF STOCK ASSESSMENT PROCESS
NORTHEAST FISHERIES SCIENCE CENTER
WOODS HOLE, MA
MAY 19-23, 2014**

I. Review Panel Members

John Armor (Chair)	NOAA National Ocean Service
Mark Dickey-Collas	International Council for Exploration of the Sea
Patricia Livingston	NOAA/NMFS Alaska Fisheries Science Center
Gunnar Stefansson	University of Iceland
Jon Helge Vølstad	Institute of Marine Research, Norway
Stephen Walsh	Department of Fisheries and Oceans, Canada

II. Background and Overview of Meeting

From May 19-23, 2014, the National Oceanic and Atmospheric Administration's Northeast Fisheries Science Center (Center) convened a panel to examine and evaluate the Center's fishery stock assessment program pursuant to the Magnuson-Stevens Act (2006) and comparable international agreements. The review focused on the overall program of assessment modeling, approach, review process and communication. Panelists received most presentations and other background documentation prior to arriving in Woods Hole. Throughout the week, Center scientists provided the six panelists with presentations and information on these topics to inform the recommendations. Presentations also provided the scientists' perceptions of the perceived strengths, challenges, and opportunities for solutions on each topic.

The following is the panel chair's summary of the other five panelists' individual reports. This summary includes observations and recommendations that two or more panelists had in common, however, they should not be considered consensus recommendations of the group. General overarching observations and recommendations are provided first, followed by observations and recommendations for each term of reference. The material provided in this summary is very cursory by design. Details and rationale behind observations and recommendations can be found in the panelists' individual reports.

III. General Observations & Recommendations

The Center's leadership and scientists do a tremendous job dealing with the extraordinary complexity of their tasks: performing assessments for more than 60 species and providing scientific input for 18 fishery management plans, two regional fishery management councils, and one interstate fishery commission. In addition, there are numerous *ad hoc* requests for information and follow-ups from GARFO, councils, commission, external partners and headquarters.

The complexity of the Center's task and the regulatory environment, however, seems to have given rise to an overly complicated stock assessment process that seems to (in some cases) undermine the quality and timing of the advice provided to fishery management authorities.

IV. Major Observations & Recommendations by TOR

TOR 1: Does the Center apply a suitable scientific/technical approach to fishery stock assessment modeling?

Panelists found the Center's approach generally suitable and that Center scientists are working to improve the approach. However, panelists noted:

- The overly complex review process leading up to the quota advice not only overburdens the assessment scientists, but also adds a possible substantial error to assessment process since the ABC may be based on 2-year old fishery-independent and fishery-dependent survey data.
- Data processing issues centered on the often late arrival of catch data from federal and coastal state fisheries results in large and unacceptable lag times for completing the assessment process.
- Staff are being severely burdened by the number of stocks that have to be assessed annually and the frequency demanded by Councils and Commission, including too many terms of references and ad hoc requests.
- There is also very little time to investigate new modelling techniques, e.g. SURBA, Stock Synthesis, recent developments in statistical assessment models (SAMs), etc., and little opportunities to work on improving existing models.
- The NOAA Fisheries toolbox provides many advantages (particularly as a tool for rapid responses to additional analysis requests in the review process and communicating with stakeholders) but also creates some limitations (particularly the complexity resulting from multiple programming languages, the data led introductory interface and the resulting issues with keeping the system up to date).

Panelists recommended:

- The "assessment bucks" concept may have some promise to help achieve some balance in staff focus and energy.
- Refocusing staff time on increasing knowledge necessary for tackling critical research questions, such as changes in natural mortality, catchability, ecosystem interactions, and climate impacts, may yield improvements to stock assessment approaches and results.
- Staff development should consist of training in new stock assessment methods, participation in international meetings and working groups such as ICES, NAFO and NASCO, etc. to keep abreast in new developments in data collection, surveys and modelling, and promotion of scientific writing for primary publications.

TOR 2: Is the assessment process efficient, effective and clearly described, including terms of reference for assessment reports?

In general, panelists found the assessment process to be highly complex and burdensome. Specific observations included:

- The assessment process was described and varied depending on the FMP and region.
- The process for developing terms of reference for the assessment reviews was described and seemed fraught with the potential to add any number of items for consideration in the assessment, with varying priority.
- The process is over engineered, inefficient, and is not necessarily within the control of the NEFSC.

Panelists recommended:

- Consider ways to reduce terms of reference or to move some to a research only track and only brought forward when ready for inclusion.
- Investigate and institute a better delivery system for fishery dependent data collection and transfer using automation to minimize time lags in arrival times.
- The assessment process itself needs to be more streamlined to improve delivery.

TOR 3: Does the Center, in conjunction with other entities such as the Councils' Scientific and Statistical Committees, have an adequate peer review process?

In general, panelists found the peer review process to be very thorough and appropriate, however, there is some room for streamlining to enhance timeliness of scientific advice. Some specific observations included:

- Attempts to maintain stability or improve timeliness may well be needed but care must be taken that this is not done at the expense of the independence of the reviews.
- The peer review process is adequate but there are many complexities introduced because of the lengthy timing of the SAW/SARC process, the different reviewers involved in the review of a particular assessment across time in the SAW/SARC, the two different SSCs with different control rules and criteria for acceptance of assessments.
- The SAW/SARC process tends to be somewhat disjoint because the CIE experts come and go and may make recommendations that are more in line with practices in other countries.

Panelists recommended:

- Consider either keeping at least one CIE member at consecutive reviews, or have a chair from SSC.
- Revising TORs of the contract to the CIE for obtaining reviewers may be needed. It should be (and perhaps already is recognized to be) incumbent on the SSC chair to ensure there is good communication of the results and recommendations from the previous SAW/SARC and keeping reviewers on track with considering the primary TORs and what is needed for ABC/OFL setting.
- Efforts should continue to communicate which TORs are important, provide each SAW/SARC panel with sufficient history of the previous peer review, what was worked on and what was not worked on and why (i.e., low priority or inadequate data, etc.) to give context to these independent but perhaps naïve in the US ACL setting process reviewers of what is trying to be accomplished. Separating immediate TORs for ACL setting from medium term to longer term TORS for strategic improvement of the assessment should be part of the communication process, if it is not already done.

TOR 4: Does the Center work effectively internally and in coordination with the NEFMC, MAFMC, ASMFC, and GARFO to accomplish needed assessments according to a set of priorities? Considerations should include program structure, staffing, funding, and the stock assessment prioritization process.

In general, panelists found priority-setting could be improved and better coordinated to improve the overall stock assessment process at the Center. Some specific observations included:

- There was little evidence that the Northeast Regional Coordinating Council (NRCC) has led to effective prioritization.
- There is also some indication that staff time may be taken up with a large number of the science review meetings. Consideration should be given when assigning staff to reduce back to back assessments, and prioritizing who needs to attend various review meetings.
- Trying to ensure that staff have time after the completion of a SAW/SARC to work on finalizing reports and publishing results should be considered to the extent possible.

Panelists recommended:

- Consider whether non-stock assessment personnel participate in PDTs and free up stock assessment scientist time. NEFSC research can be brought to bear to assist in necessary Council analyses but it may not be necessary for the stock assessment scientist to sit on the PDT or actually write the analytical document. Depending on the analysis, Council or Regional office staff should have the lead.
- Consider staffing levels for each SAW working group and how many of these each stock assessment scientist participates in per year. Consideration could be given to reducing that load by including only the most necessary stock assessment scientist in each. Ways to reduce the allocation of stock assessment time to non-essential tasks should part of the assessment bucks concept.
- Apply a systematic approach for addressing problems in the stock assessment process. There is a need to have one body set the priorities in frequency and time-lines for assessments (benchmark vs updates) using a defined set of criteria.

TOR 5: Does the Center achieve adequate assessment accomplishments relative to mandates particularly with respect to the number of Fishery Management Plan (FMP) species assessed?

Panelists generally thought the Center's stock assessments were meeting the demands of managers, however, many felt some streamlining was needed to sustain the current level of service. Some specific observations included:

- The number of benchmark assessments has declined over time but the total of number of assessments has not.
- Timeliness of data inputs to the stock assessment process is hindering the quality and timing of advice.
- Improvements to the fishery dependent data availability problem along with the schedule mandated by the peer review and GARFO process are the most essential for getting the most timely data into stock assessments for ABC/OFL setting.

TOR 6: Does the Center have an effective process in place for taking ecosystem and climate change factors into consideration in the stock assessment process?

In general, panelists found the Center's efforts to take climate change and ecosystem impacts into stock assessments are laudable but that improvements could certainly be made. Some specific observations included:

- With regards to the ecosystem approach to fisheries management, the initial steps taken by NEFSC are encouraging.
- Incorporation of ecosystem considerations is currently most appropriate for strategic advice, rather than operational annual advice (called tactical by some at the center). Limit reference points should reflect the reality of the ecosystem dynamics.
- Challenges relate to research time to develop modeling approach, and cuts in ecosystem monitoring surveys.

The panel recommended:

- Stabilizing or increasing funding for ecosystem surveys.
- Secure time for stock assessment and ecosystem modelers to work together on methods development.

TOR 7: Does the Center adequately engage stakeholders in the stock assessment process and communicate assessment-related results, needs, and research to them effectively?

In general, panelists felt the Center does a good job communicating with stakeholders and others about the stock assessment process. Some specific observations included:

- The NEFSC puts out an impressive array of documents and communicates face-to-face through the public meetings set up for the formal SAW/SARC and SSC interactions surrounding the ABC/OFL setting process.
- Communication is not the same as dialogue.
- The entire assessment process is very transparent and there are no obvious problems in this regard.

Panelists recommended:

- Members of NRCC should take an active role in increasing communication within the assessment process and outside to the general public.
- Center staff should be given ample opportunities to publish results from stock assessment model development and critical research areas to enhance the scientific integrity of the Center.
- The one exception to transparency seems to be that the SARC has a closed-door writing session at the end of each meeting. This is a subject of some controversy and there have been calls for opening this session. Interestingly there (a) seems to be little reason for having this writing session closed and (b) there would be little benefit to transparency to opening it. Opening the session should be considered, if only to make the entire process transparent.

TOR 8: Are there opportunities for improving stock assessments and the stock assessment process?

In general, panelists had multiple suggestions for improving the Center's stock assessment process. Refer to panelists' individual reports for details. A few of the suggestions included:

- The system needs to keep its legitimacy and yet needs to be radically streamlined. Researchers need some space to explore ideas. The center should consider the greater use of management strategy evaluation as a tool.
- The management systems in the NEFMC and MAFMC areas are evolving and the science is evolving with it. The challenges will be communicating the highest priority issues to deal with and making the resources and institutional linkages necessary to implement those.
- Many of those issues are outside the control of the NEFSC. GARFO, NEFMC, MAFMC, and ASMFC should be working together with the NEFSC to solve those problems.
- The Centre needs to avail of (or obtain) expertise in gear technology to address issues of calibration and selectivity; 2) apply management strategy evaluation (MSE) to stock assessment issues such as retrospective problem and decisions on catch advice; and 3) use funding from the set aside (RSA) programs to tackle immediate, critical research issues such as an independent estimate of natural mortality, movement and stock size for Georges Bank yellowtail and cod, e.g., with a re-introduction of tagging programs.
- Priorities must be set on terms of reference and research proposals coming out of the assessment process. The natural place for these to be discussed is at the NRCC meetings.
- The full process, from time of data collection through implementation, needs to be shortened. The assessment part is a minor part of the full process.
- One of the biggest potentials for improvement in the stock assessment likely relates to improving the quality of the fishery-dependent information. The current system of data-collection is largely ad-hoc (refer to the 2013 program review), which can result in variable bias of unknown magnitude.

Stock Assessment Science Program Review- Northeast Fisheries Science Center, 2014

Reviewer: 1

May 22, 2014

Background and General Comments

The NEFSC conducts stock assessments under numerous FMPs for two regional fishery management councils and a fishery commission. There are numerous challenges with respect to dealing with the complexities that result from the number of FMPs, organizations, data availability and timing, and the peer review process. The NEFSC has talented, dedicated stock assessment scientists and also long time series of information to use to inform stock assessments. Last year's program review was on the fishery dependent and independent data components that feed into stock assessment. Acting on the recommendations from that review is essential for improving the stock assessment enterprise. Some of those recommendations may get repeated below to emphasize those that are seen as bottlenecks or key to the process. Additional challenges involve the high scrutiny placed on stocks with poor status that may constrain fishing opportunities.

Reviewers were asked to organize comments according to eight themes that define the stock assessment program and provide advice to the Center on the direction and quality of the stock assessment science program and suggest areas of improvement.

THEME 1: Does the Center apply a suitable scientific/technical approach to fishery stock assessment modeling?

The NEFSC employs a variety of modeling approaches ranging from age and length based approaches for more data rich stocks to index methods for more data poor stocks. There are 61 stocks and 63 models. The diversity of modeling approaches can be viewed as both as a strength and a weakness. Having models tailored for the stock's particular life history and fishery characteristics and type of data available is important. In addition, there are assessments that do incorporate ecosystem factors. There is more movement towards statistical based modeling approaches, which are more commonly used in West coast and Alaska fishery stock assessments. The case studies provided indicated appropriate model choice given data availability for particular stocks.

There are challenges involved in maintaining and improving individual models such as staff time available to develop and test new model features. Also, the NMFS Toolbox interface would need to be modified to accommodate the new model, if it was one in the Toolbox. The use of the Toolbox to easily make changes during review meetings was seen as a transparent, documentable way of showing changes and when changes were made, which may be important in the environment of NEFSC stock assessment. Center scientists are making contributions to improving the suite of modeling approaches that go into the National Stock Assessment Toolbox.

It does appear that integrated modeling packages such as SS3 have been considered for stock assessments and one assessment at NEFSC does rely on the SS3 framework. When given sufficient time to work on modeling improvements, it appears those do get accomplished. The appropriateness of the analytical framework for assessments does appear to have been tested and show that they are using the appropriate

framework though there are indications that improvements are desired for many assessments. Not all improvements may be high priority, however, and that should be considered in this process. Using a model update only approach for some stocks could be an alternative to consider to help in stock assessment throughput.

THEME 2: Is the assessment process efficient, effective, and clearly described, including terms of reference for assessment reports?

The assessment process was described and was varied depending on the FMP and region. The process for developing TORs for the assessment reviews was described and seemed fraught with potential to add any number of items for consideration in the assessment, with varying priority among TORs. There was much discussion of the need to reduce TORs or to move some to a research only track and only brought forward when ready for inclusion. Also, the 2011 document on “A New Process for Assessment of Managed Fishery Resources off the Northeastern United States” was presented as a process that has been vetted through the NRCC as a way to deliver operational assessments, define a research track, and TORs for each of these along with defining responsibilities.

High priority should be placed on implementing the recommendations of this document. Some are being done but it appears that not all have been. In particular, using a common reporting format for assessments seems not to have been implemented. Presenting this document’s recommendations again to the respective SSCs would be important. With respect to implementing the stock assessment format template, it seems that there was some desire for authors to use a common method such as Latek for inputting figures, etc. However, that should not be seen as a requirement for moving forward. Some of the recommendations in the “New Process” document are even more important to implement, and involve the protocols for remand, re-examination, addressing error or new information into improved stock assessments, and developing a rational schedule for operational assessments. These will make significant progress in making the stock assessment process more efficient.

Stock assessment authors sometimes appear to be driven to try numerous modeling approaches to try to “fix” a problem that ultimately may not be solved through modeling but are rather an indication that there is incomplete knowledge of a process or inadequate data. Although stock assessment scientists are being responsive to demands being placed on them in this regard, there needs to be a process for identifying when modeling is not the answer to the assessment problem and ability to move those issues to the research track and brought back when sufficiently mature for consideration. It should be recognized that not every assessment needs to be brought up to the highest tier. The NMFS survey prioritization tool that will be employed regionally should be helpful in deciding where to allocate scarce assessment resources.

One very troubling aspect of the stock assessment process was the sheer amount of time and effort that assessment scientists must spend on processing the fishery dependent data. This is highly inefficient and is a definite bottleneck. Extremely high priority should be placed on making the recommended improvements to the timeliness and quality of fishery dependent data system from last year’s fishery dependent and independent data review. The need for an integrated catch accounting system that is available in a centralized place is critical not only for timely fishery management but also for timely fish stock assessments. It should not be incumbent upon individual analysts to spend months vetting these data sources but should be a responsibility that GARFO and ASMFC should be working together to accomplish. Fishery dependent data systems in other regions should be examined to see if there are ways

to implement an integrated system as quickly and efficiently as possible even in the light of changing management, which is a given in fisheries management. The NMFS Alaska Region has an efficient system and also a method for estimating rates spatially (Catch in Areas Database). There may also be ways to improve the timeliness of fishery dependent data used in stock assessments by methods for estimating the remainder of the year catch instead of waiting for the full year of terminal catch to be available. If there are ways to deal with some cross cutting TORs by one person such as doing survey data updates, bycatch estimation, that should be considered. Ways to incorporate the most recent data, not just fishery dependent data, should be considered.

THEME 3: Does the Center, in conjunction with other entities such as the Council's Scientific and Statistical Committees, have an adequate peer review process?

The peer review process is adequate but there are many complexities introduced because of the lengthy timing of the SAW/SARC process, the different reviewers involved in the review of a particular assessment across time in the SAW/SARC, the two different SSCs with different control rules and criteria for acceptance of assessments. Consideration should be given to having some stability in at least one of the SAW/SARC reviewers from one benchmark assessment to the next, although the practice of having an SSC member chair the SAW/SARC may also help with that process. Revising TORs of the contract to the CIE for obtaining reviewers may be needed. It should be (and perhaps already is recognized to be) incumbent on the SSC chair to ensure there is good communication of the results and recommendations from the previous SAW/SARC and keeping reviewers on track with considering the primary TORs and what is needed for ABC/OFL setting. Upon completion of the SAW/SARC process, the SSC can play a role in helping advise stock assessment scientists on the importance of dealing with research recommendations that come out of the SAW/SARC and what may or may not be valuable to address in the next SAW/SARC or sent on to the research track. Those recommendations should similarly be reviewed and discussed within the NEFSC and communications back to the SSC should be made to alert them and the Council about what can realistically be dealt with given the data or resource availability. The research track would be the best place to deal with vetting outside ideas for stock assessment improvement that come from industry consultants and other entities. These should not be dealt with during the ABC/OFL setting process.

It appears that there may be scheduling problems with getting SSC review of numerous annual updates. Perhaps an SSC member from the NEFMC and MAFMC can attend other region's SSC meetings such as NPFMC (December SSC meeting) to see how the annual updates can be reviewed and received. Operationally, consolidating some FMPs for species with similar fishing years could help in reducing complexity and make the process more efficient. Definition of a benchmark assessment requiring SAW/SARC review seems to be very strict and consideration should be given to whether the SAW/SARC process should be used more for strategic review of the assessment models, data inputs, and research enterprise supporting these outside of the annual ABC/OFL setting process.

THEME 4: Does the Center work effectively internally and in coordination with the NEFMC, MAFMC, ASMFC, and GARFO to accomplish needed assessment according to a set of priorities? Considerations should include program structure, staffing, funding, and the stock assessment prioritization process.

The NRCC process seems to be useful for discussing timing conflicts and working out priorities between regions. Although the Councils think the current peer review process is rigorous and desirable, there is also conflict between that goal and the ability to schedule the SAW/SARC and SSC reviews in a fashion that meets all the goals for desired timeliness for ABC/OFL setting. Ultimately, it appears the GARFO regulatory deadlines are also a large driver in terms of the scheduling and also add a significant amount of time at the end of the review process to account for time needed to get the recommendations implemented. This is puzzling given the ability of AKRO to implement catch recommendations so much more quickly.

Councils feel that the NEFSC is responsive to their needs given the constraints of staffing and scheduling. There seems to be support across all the entities for the idea of doing operational updates and streamlining the data update process to free up time for improving key assessments. This needs to be implemented. There is also some indication that staff time may be taken up with a large number of the science review meetings and consideration should be given to assignments for staff to reduce back to back assessments and who needs to attend various review meetings. Trying to ensure staff have time after the completion of a SAW/SARC to work on finalizing reports and publishing results should be considered to the extent possible. Can non-stock assessment personnel participate in PDTs and free up stock assessment scientist time? NEFSC research can be brought to bear to assist in necessary Council analyses but it may not be necessary for the stock assessment scientist to sit on the PDT or actually write the analytical document. Depending on the analysis, Council or Regional office staff should have the lead. It isn't clear how many stock assessment scientists are on each SAW working group and how many of these each stock assessment scientist participates in per year. Consideration could be given to reducing that load by including only the most necessary stock assessment scientist in each. Ways to reduce the allocation of stock assessment time to non essential tasks should part of the assessment bucks concept. This is not just a stock assessment branch issue but one that NEFSC, GARFO, and the Councils should consider in their priorities and assignments for accomplishing priorities.

THEME 5: Does the Center achieve adequate assessment accomplishments relative to mandates particularly with respect to the number of Fishery Management Plan species assessed?

The number of benchmark assessments has declined over time but the total of number of assessments has not. Timeliness of data in the stock assessment has already been mentioned previously and this is a feature of the fishery dependent data availability problem already discussed along with the schedule mandated by the peer review and GARFO process. Improvements in those processes are the most essential for getting the most timely data into stock assessments for ABC/OFL setting.

THEME 6: Does the Center have an effective process in place for taking ecosystem and climate change factors into consideration in the stock assessment process?

There are clear examples of the incorporation of ecosystem and climate change factors into the stock assessment process. The NEFSC has done a good job at working on this and is thinking at both the tactical and strategic levels of how to make this happen. Obviously, there is lots of discussion around the US and worldwide about the ways to incorporate multispecies and ecosystem models into the fishery management process. It appears that both Councils are thinking about this and there is good communication between ecosystem scientists and the stock assessment enterprise about ways to move forward.

The information shown on climate effects on changes in spatial distribution and overlap of species and biological parameters are important considerations for Councils and thinking strategically about fishery management adaptation under climate change.

THEME 7: Does the Center adequately engage stakeholders in the stock assessment process and communicate assessment-related results, needs, and research to them effectively?

The NEFSC puts out an impressive array of documents and communicates face-to-face through the public meetings set up for the formal SAW/SARC and SSC interactions surrounding the ABC/OFL setting process. The SAW/SARC process tends to be somewhat disjoint because the CIE experts come and go and may make recommendations that are more in line with practices in other countries. Efforts should continue to communicate which TORs are important, provide each SAW/SARC panel with sufficient history of the previous peer review, what was worked on and what was not worked on and why (i.e., low priority or inadequate data, etc.) to give context to these independent but perhaps naïve in the US ACL setting process reviewers of what is trying to be accomplished. Separating immediate TORs for ACL setting from medium term to longer term TORS for strategic improvement of the assessment should be part of the communication process, if it is not already done.

Having an SSC member chair the SAW/SARC process can help in this process but there may need to be more communication of NEFSC with the SSC outside of the immediate SAW/SARC process to arrive at a common understanding of the standards for accepting models and material that each SSC hopes to receive in order to make ABC/OFL determinations from the SAW/SARC advice. It is not clear if the SSCs or Councils on the east coast arrange for workshops with scientists to discuss how to move forward in these areas but that is a common practice in the Alaska region. The SSC in that region looks to the Science Center to help advance the science and propose potential scientific avenues to move forward in areas needing attention such as stock structure, dealing with recruitment, etc. These are done in the form of both workshops and working groups that advance white papers for SSC discussion. It appears on the east coast that the SSCs do more of the initiating of working groups and white papers but involving center scientists in these interactions and working groups would be most helpful to make sure that there is a common understanding of the science issues and ability to advance them to meet Council needs.

It is important for NEFSC stock assessment scientists to participate in regional, national, and international working groups to communicate and advance stock assessment science. It appears there is participation of NEFSC staff in these areas. Stock assessment scientists also need to communicate their science advancements through the peer reviewed literature and it isn't clear how much that has been happening recently.

THEME 8: Are there opportunities for improving stock assessments and the stock assessment process?

There are always needs for improving stock assessments through improvements in the data inputs and in the models based on changing inputs and understanding of the processes influencing stock dynamics. The challenge is to identify the highest priority improvements and the resources to complete those. It is clear that improvements in the fishery dependent data stream need to be made and have already been commented earlier in this document. There are also research issues that could be dealt with by NEFSC survey and gear scientists, such as those dealing with catchability. Obtaining estimates of catchability

may be very useful for stocks where there is uncertainty about both catchability and natural mortality. Field studies to estimate catchability are more tractable than those to estimate natural mortality. It isn't clear what the process is for communicating and prioritizing research to improve stock assessments within the Center but the new science planning process should further facilitate that. AFSC scientists currently identify cross cutting research issues and jointly work on research activity plans that can address those. A recent example is an identified need to improve the research on fish maturity, which has been an ad hoc area of research improvement in Alaska until now. AFSC scientists held a workshop that involved stock assessment scientists needing the information, and survey and observer program personnel who would obtain and process samples. They are in the process of identifying current resources available to accomplish what they view are the highest priority items and also what additional resources are needed. This will go into the AFSC's science planning process for the next year to see if there are additional resources to accomplish this. Like most other centers, however, the AFSC will likely need to find additional resources through preparing proposals to compete for funds through regional or national RFPs. But it will be well positioned to do so because there has been a collaborative effort among stock assessment scientists and other research programs at the center.

Other research efforts may take the form of identifying modeling strategies or approaches for catch advice. These have been well articulated at this program review and the NEFSC appears to be making progress on many of these issues, such as incorporating climate and ecosystem factors into individual assessments and working on multispecies and ecosystem approaches. The management systems in the NEFMC and MAFMC areas are evolving and the science is evolving with it. The challenges will be communicating the highest priority issues to deal with and making the resources and institutional linkages necessary to implement those. As mentioned throughout this report, many of those issues are outside the control of the NEFSC. GARFO, NEFMC, MAFMC, and ASMFC should be working together with the NEFSC to solve those problems.

NOAA Fisheries Science Centre Stock Assessment Science Program Reviews

NEFSC fishery stock assessment programme

Review Report- 2 30 May 2014

This report reflects the personal views of the author, and was written independent of the other panel members. It leads with text describing the general impression and then reflects on the 8 terms of reference specific to the NEFSC review and finishes with personal recommendations. It will undoubtedly contain errors with regards to the scientific and management structures as a result of the complexity in the system and ignorance of the reviewer.

The review is written under the assumption that best practice for the provision of fisheries scientific advice is through a participatory process that is transparent, accountable and based on the best available science at the time; and that it is communicated in an effective manner. This assumption is not dissimilar to the objectives of NS2. Stock assessments are a tool for the provision of fisheries management advice. The stock assessment must be fit for purpose.

The preparation and cooperation by the NEFSC staff for the review was greatly appreciated.

General Impressions

The NEFSC was and still is one of the preeminent operational stock assessment centres in the world. While focus is often drawn to the troublesome New England groundfish assessment and management challenges, the large majority of the stock assessments carried out by the centre provide the effective evidence base for fisheries management advice across the region (invertebrates, pelagic and coastal fish, Mid-Atlantic fisheries). The researchers at NESFC work extremely hard and experience a very high work load. The review highlighted a need to reduce the operational “crank-the –handle” work carried out by the staff to allow the development of approaches to resolve current and future challenges to fisheries management in the North East Atlantic region. This was termed “research track” by the centre management. The research track or development work is crucially needed to enable innovative and responsive solutions to be found for assessment and management problems.

The centre is wonderfully rich in data. This is a huge resource. The data led approach however seeps into the approaches used by the stock assessors. The precision of the assessments, or uncertainty in the advice is provided based on uncertainty from data or model fits. Reduced attention appears to be given to uncertainty in structural knowledge, or stock assessment model assumptions. Science is about increasing knowledge through the testing of ideas against the evidence. The evidence in this case is the data. I feel that the centre needs to give increased attention to the development of ideas. For this, the researchers need space to think (including space to read and write). This search for the best fit to the data, contrasting to the best provision of knowledge to the management issue leads to the “best fit” approach to stock assessments. Globally the “best fit” approach is being challenged, often because it

fails to explore uncertainty due to model assumptions. The recent World Conference on Stock Assessment Methods and Deroba et al (in press)¹ highlight the issues. The researchers at the centre should be aware of the limitations of the “best fit” approach. The NOAA Fisheries toolbox provides many advantages but also creates some limitations. It leads the user into a data-led approach to management challenges, rather than a conceptual one.

Many of the strains in the NEFSC stock assessment system appear to be due to the demands of the regulatory management system and the peer review process. Much of these appear to be the product of regional norms rather than national requirements. To an outsider, and despite the work of the NRCC, there appears little prioritisation of stock assessment or research needs. At the current programme review, the regional Fisheries Management Councils and Commission stated that they wanted more assessments, more often and delivered quicker. When questioned, they described a system of competition for NEFSC resources and research which made prioritisation difficult. This high demand from the fisheries management system needs to be addressed as it appears non-sustainable. The demand from the management side, made me question the lack of management strategy evaluation (MSE)² used by the group at NESFC. I understand the concern expressed by NEFSC researchers that some forms of MSE could blur the boundaries between the provision of scientific evidence for decision making and entering the environment of making policy decisions. However experience from Australia, South Africa, the International Whaling Commission and Europe shows that when handled carefully MSE provides powerful and robust tools for exploring research and exploitation challenges and provides a mechanism for the development of participatory processes with partners and stakeholders. Many of the management challenges and suggested changes to the stock assessments could be explored using MSE. These MSEs could occur prior to the investment of resources or to test the likely effectiveness of proposed adaptations or management actions. MSE should be used to engage with stakeholders and partners to explore potential management options, and likely consequences of scientific developments.

The effort to populate and maintain the fisheries management plans (FMP) is huge. I could not see the rationale for the number of FMP in the NEFMC area as the system was currently constructed. There are contradictions across the plans (see figure). The NEFMC concept of a fisheries management plan was slightly alien to me. My lived experience is that many of the problems in mixed- groundfish fisheries cannot be solved with single species approaches. Thus I would expect FMPs to be fleet oriented and relate to the organisms that these fleets catch or impact, rather than be a collection of single species targets within a management area. This later approach works well when fisheries target and catch single species in an area, but not when the fisheries catch are mixed. I know that the FMP are legal agreements, but the definition of FMP in the Magnuson-Stevens Act and NS1 appears to allow for fleet based approaches.

Publishing scientific developments and results is a crucial part of a stock assessors work. Apart from the standard reporting, this is best done through peer review publications. The process of publishing in the

¹ Deroba et al. in press. *Simulation testing the robustness of stock assessment models to error: some results from the ICES Strategic Initiative on Stock Assessment Methods*. *ICES Journal of Marine Science* doi:10.1093/icesjms/fst237

² Or in a broader sense management procedure evaluations

primary literature ensures that the researchers are up to date with global methods and it provides extra scrutiny to the methods developed and conclusions reached (extra peer review). Whilst the current system requires traceable and auditable stock assessments, I doubt that the provision of large stock assessment documents with varying structure is the best use of stock assessors' time. I know that the NEFSC strategic plan has proposed approaches to resolve the challenge. Preparation for the extensive peer review process takes a large commitment by NEFSC researchers.



Conflicts in NEFMC FMP. (Slide 11 from Gaichas et al presentation, 22 May 2014).

The NEFSC, NEFMC, MAFMC, ASMFC and GARFO need to develop mechanisms to engage and utilise any proposed alternative approaches to the stock assessments, i.e. approaches coming from outside NEFSC. A suggested mechanism should be participatory and not confrontational. The work load of the centre researchers is high, so I would encourage further partnerships with academics and other researchers to solve stock assessment challenges.

With regards to the ecosystem approach to fisheries management, the initial steps taken by NEFSC are encouraging. I welcomed the acceptance that incorporation of ecosystem considerations is currently most appropriate for strategic advice, rather than operational annual advice (called tactical by some at the centre). Limit reference points should reflect the reality of the ecosystem dynamics³. The stock assessors should accept that data sources come with varying precision. This varying precision does not negate the usefulness of the information, but may require the development of new techniques. The aim of including greater ecosystem knowledge should not just be the improvement of fisheries management advice, but also the increasing the robustness of the model assumptions. Including knowledge for knowledge sake is not appropriate, but a model that successfully uses process information to determine growth or condition in a time series would be more robust for projections into unknown space (temporal or spatial) that projecting recent empirical findings. The SARC system needs to find approaches that can allow greater cross-disciplinary review of ecosystem considerations.

³ I welcome the 2013 report "Evaluating the Effectiveness of Fish Stock Rebuilding Plans in the United States"

The political solution of the TRAC for Georges Bank cod and haddock appeared difficult to justify in terms of fisheries science (use of different methods in a subset of the area).

For future research needs and developments, I would have liked to have seen an NEFSC roadmap with proposed timelines for the delivery of new products, including the streamlining of data delivery, the incorporation of ecosystem approaches and the delivery of the 2011 new process for the assessment of managed fishery resources in the north east Atlantic region.

Terms of Reference

1) Does the Centre apply a suitable scientific/technical approach to fishery stock assessment modelling?

The centre uses a range of approaches, most crafted to the specific data set concerned. Most stock assessments function well as advisory tools. The centre should consider other length based, ecosystem based and data poor approaches within its tool box. The centre needs to find a mechanism to incorporate alternative assumptions into their modelling approaches.

2) Is the assessment process efficient, effective and clearly described, including terms of reference for assessment reports?

The process is over engineered, not efficient, but this is not necessarily within the control of the NEFSC.

3) Does the Centre, in conjunction with other entities such as the Councils' Scientific and Statistical Committees, have an adequate peer review process?

Generally the peer review process is extremely robust and perhaps overly burdensome. Improved peer review of innovations on climate change, spatial approach and incorporation of ecosystem dynamics needs to be addressed in the future. The current system is creating a "snowball" effect caused by researchers responding to reviewers concerns made at each review iteration.

4) Does the Centre work effectively internally and in coordination with the NEFMC, MAFMC, ASMFC, and GARFO to accomplish needed assessments according to a set of priorities? Considerations should include program structure, staffing, funding, and the stock assessment prioritization process.

There was little evidence that the NRCC has led to effective prioritisation. Within NEFSC there also appeared to be little structured prioritisation other than day to day needs. It was pointed out that a large proportion of NEFSC researchers time was spent dealing with "non-standard" Council or Commission requests.

5) Does the Centre achieve adequate assessment accomplishments relative to mandates particularly with respect to the number of Fishery Management Plan (FMP) species assessed?

I am unclear about the meaning of this ToR.

6) Does the Centre have an effective process in place for taking ecosystem and climate change factors into consideration in the stock assessment process?

A process is developing, and examples were shown. The lack of a committed budget line to ecosystem monitoring and development of the ecosystem approach is a challenge. Considering that this is seen as a policy priority, this lack of budgetary commitment should be seen as a major concern. The population

dynamics group and the ecosystem assessment must further integrate to allow more spatial tools to be developed and for the groups to catalyse each other.

7) Does the Centre adequately engage stakeholders in the stock assessment process and communicate assessment-related results, needs, and research to them effectively?

Engagement with managers, policy developers and stakeholder appears multi-layered. It appears to be based on the objective of transparent communication of decisions and process rather than a participatory dialogue. The best engagement is a dialogue based around a participatory process rather than a “communication” of results. In a participatory and transparent process, the engagement should be equitable. This requires stakeholders and researchers to work together on challenges. Clearly pre-define responsibilities prevent any blurring of the role of the provision of scientific knowledge and the roles of deciding or lobbying for specific management objectives.

8) Are there opportunities for improving stock assessments and the stock assessment process?

The system needs to keep its legitimacy and yet needs to be radically streamlined. Researchers need some space to explore ideas. The centre should increase the use of MSE as a tool for investigating management approaches, resourcing of scientific priorities and building dialogue with partners and stakeholders. The NEFSC should consider approaches for multi-model inference.

Recommendations

NEFSC leadership should:

- Create an NEFSC roadmap with proposed timelines for the integration of new methods and delivery of products, including the streamlining of data delivery, the incorporation of ecosystem approaches and the delivery of the 2011 process for the assessment of managed fishery resources.
- Challenge the existing over-burdensome peer review process, and management system.
- Challenge the NRCC to prioritise stock assessment and research and development needs.
- Implement the research track including the creation of space for conceptual development and mentoring for publishing in the peer reviewed literature.
- Encourage the Population Dynamics group to further think outside the single species box (e.g. fleet based approaches, spatial dynamics of fish populations, MSEs and exploration of management objectives for regional fisheries).
- Review whether the “best fit” approach to the provision of fisheries advice is most appropriate, being aware of the regulatory implications of such a review.
- Develop further mechanisms to increase operational interaction between the Ecosystem Assessments and Population Dynamics groups in NEFSC.
- Consider the operational and conceptual impact of the stock assessment tool box, suggest creating a new entrance page and including more approaches in the suite of models.

Reviewer report 3

1 Introduction

The meeting was provided with extensive documentation as well as presentations from scientists, NEFSC management and Council/Commission views. The structure and presentations were excellent but the complexity of the overall process of proving and implementing management advice implies that a one-week meeting is somewhat short for an extensive review.

This reviewer's report is split into a few sections, followed by a short summary, which links conclusions to the terms of reference.

There is some room for improvement on several fronts, and some of these can lead to a reduction in total processing times. It must be emphasized, however, that these improvements will generally not drastically change the total amount of time from data collection through implementation of the management decision: The total amount of time is to a very large extent determined by meetings which are neither determined by nor limited by Center staff time.

As a general comment, the overall performance of the Center seems to be very good, but the friction between the different partners/stakeholders makes the general environment very difficult. This is not unusual around the world, but the interactions are particularly complex here.

2 Assessment Process, Tasks and Timeliness

When any advisory process and implementation is considered as a whole, this naturally includes the analyses of data, developments or modifications of assessment models, running the models, writing reports and paper, presentations in-house and work at workshops, followed by political negotiations and eventual implementation. Different parts of this process vary in importance and duration on a case-by-case basis.

At the NEFSC the assessment scientists also have a suite of responsibilities, which range from attending SARC/SSC to PDT and SC meetings as well as other review meetings. These are all a part of the process from data through assessments to implementation.

In addition to the assessment process, again in an international context, "other work" is given to assessment scientists and there is considerable variation in how much this is.

In the assessment process under consideration here, the entire process is unusually complex and the resulting path from data to implementation is very long indeed. Parts of

the process can probably be shortened. For example, under Data Handling below, some possible methods are mentioned which might shorten the data processing steps and under Alternative Harvest Rules, some methods are discussed which might possibly alleviate some of the assessment issues. Unfortunately, this is not likely to have a major effect on the length of the process as a whole: The real bottleneck on time does not seem to be within the NEFSC but in the following regulatory/review/implementation process which takes much longer than the time from data becoming available through the assessment becoming available.

On timeliness, the panel was presented with the following example:

... catch data from 2013 that become available in May 2014 would be analyzed and reviewed at a November 2014 SARC, reviewed by the SSC in January 2015, considered by the Plan Development Team (PDT) and Council Committees throughout 2015 with Acceptable Biological Catch (ABCs) to be implemented by GARFO at the start of the 2016 fishing year in May 2016.

In this example, the December 2013 catch data has been through the complete assessment process by January 2015, i.e. in just over 12 months. This is long, but the remaining political process adds another 16 months to this and it is therefore important to consider the data-assessment-implementation process as a whole.

The general issue of timeliness is important: Having a long time period from an assessment to implementation has various implications, which can be very serious in certain scenarios. To mention an extreme case, consider a harvest control rule, which sets a quota as a proportion of a survey abundance index. In this case, there is no reason for lengthy annual discussions on methods: Upon completion of the survey, everything follows from a given method of computation. If the stock is likely to be in a poor state, then it would in most places be considered unacceptable to wait for two years for a decision: There would already be 2-3 new measurements of stock size before the fishery started and these could easily demonstrate that the stock was in better or worse shape than predicted. This also applies in the more complex cases of using an assessment, and even though completing and reviewing an assessment does take somewhat longer than just computing a simple index, it does not add years on top.

3 Development of Assessment Methods and the Fisheries Toolbox

For assessments to be reliable, correct and satisfy modern standards, the assessment scientist need to be up to date on recent research and ideally to be researchers, directly involved in model development or developers of surveys, sampling schemes etc. It is therefore crucial for any research center that their scientists are also active researchers.

The NOAA Fisheries Toolbox has many similarities to work in other locations. For comparisons, another toolbox will be mentioned here: GADGET (Globally applicable

Area-Disaggregated General Ecosystem Toolbox, at <http://www.hafro.is/gadget/>), which is open-source and written with a similar purpose in mind.

The Fisheries Toolbox has several important design features, such as disentangling model runs from the graphical user interface (GUI). This was also done in the case of GADGET (based on experience from previous attempts).

The development of a GUI as a part of an assessment toolbox is a somewhat contentious issue in the first place. On the one hand GUI development requires enormous resources, which might potentially be used on improving models, but on the other, the GUI facilitates teaching and presentation in a working group meeting. In this particular case, the GUI has clear benefits in terms of how it can be (and is) used within the advisory process, e.g. to immediately evaluate the effects of minor changes in assessment assumptions. Unfortunately one of the obvious drawbacks to the Fisheries Toolbox is the internal programming language used for the GUI, since Visual Basic is really quite outdated and non-portable.

Hopefully the more general and multi-platform R can be used more in the Fishery Toolbox in the future, moving away from difficulties in using specific features of specific versions of specific operating systems, and this is clearly the desire of many Center staff. This will also make it much easier for individual staff to add new features to the system, e.g. in the form of new plotting routines or diagnostics.

The approach taken here is thus in many ways the same as has been taken e.g. in some European projects (with GADGET used in several of those), which takes all data from files and outputs all results to files. As with the Center, developmental add-ons are underway to provide R commands for plots and controlling GADGET assessment runs.

The Center appears to have been quite heavily involved in the development of the Toolbox. This development, along with the use of the Toolbox constitutes an important assessment accomplishment.

The Toolbox includes stock synthesis, which is one of the most elaborate statistical single-stock assessment tools on the planet (not developed at the Center).

The Center has developed several assessment methods well beyond the basic "canned" methods provided by commonly available tools. As always, such developments range from the use of a regression covariate through an entire new assessment method, such as length-based scallop assessment method.

An alternative approach, which is also used at the Center, is to have a basic tool, which can fit a range of models of varying complexity to any number of data sources. The approach implies that one can start with simple models and add complexity or data as needed for a particular stock, at each stage verifying fits and the need for complexity (at

least in principle). This can be done using tools such as ASAP⁴, generic ADMB⁵ and GADGET and Stock Synthesis (SS)⁶, but the extensibility of these varies quite a bit, some of them permitting specific extensions and others being completely open-ended. Notably, a staff member is one of two authors on the ASAP assessment methodology.

4 Data Handling

The Center collects, receives and uses data from a wide variety of sources. The incoming annual data need to be verified and appropriately processed for storage in relevant databases. It is important that these databases are appropriately designed and the preprocessing of the data be made as automatic as possible. These issues have been addressed by all fisheries centers around the world but not always very successfully.

Usually raw data are stored in the institutional database and any processing normally takes place outside those databases. It has been found extremely useful in some cases to implement "back-ends" which automatically extract and preprocess all data to be used in single- and multi-species assessments into a new "summary database", which is "mildly aggregated", only to the finest level of aggregation which will be used for any of the tools. If this is done carefully, then this is the step when all error-handling and filtering is done.

The next step in data handling includes aggregation to appropriate regions and time steps, as well as any kind of imputation or prorating or "borrowing" data from adjacent time cells or spatial cells. This step can be completely automated and be based only on metadata files, which describe the aggregation mechanisms from the summary database, as required for a particular assessment.

When this approach is taken one sees immediate benefits in terms of time spent by the users, particularly for new users who are given ready-made extraction scripts which completely automatically generate syntactically correct input files for the assessment programs.

Another virtue of going through formal and automated database schemes is that they can be formally tested using bootstrapping or similar methods. In at least one case this has been used to test the entire process from database of raw data through the assessment and reference points. In that example no difference was found between the automated methods and the nitty-gritty manual means of scrutinizing every data value and choices of how to combine age-length keys etc. Obviously such individual results may not carry over to other cases.

⁴ <http://nft.nefsc.noaa.gov/ASAP.html> or Legault and Restrepo (1998).

⁵ <http://admb-project.org/>

⁶ http://nft.nefsc.noaa.gov/Stock_Synthesis_3.htm

Many of the data sets used in assessments are externally provided and the Center has limited control over their timeliness. The NRCC may want to consider whether they, as a body, can assist to influence such processes.

Within the Center there is some room for improvement on exactly the issue of data handling prior to assessments. Work on such improvements should of course be a continuous part of the Center, but a long-term strategy is needed on the data-flow, from data sources into the databases and onwards, possibly through an intermediate database, finally to input files for assessment routines.

5 Assessment Issues, Updates and Alternative Harvest Rules

The panel was presented with a number of assessments, assessment methods, issues and solutions. It is abundantly clear that the assessments overall are of a high standard and the data are usually both appropriate and appropriately used.

Pollock assessments are an example of a particularly difficult assessment. Problems with pollock assessments are known in other regions and probably pertain to the biology and behavior of the species rather than specific problems with pollock in this region. It is therefore unlikely that modifications to or further development of assessment methods will yield a grand truth on the state of the stock. It is more likely that basic research is needed for new biological information or completely new data sources. Alternatively, one might want to consider management strategies, which involve simpler quantities rather than uncertain assessments (unfortunately it is not obvious what quantities could be used in the case of pollock).

Georges Bank yellowtail flounder assessments were presented and discussed several times during the meeting. Like the pollock assessment, these are also very difficult and currently there is no obvious single answer to what the best stock estimate should be. Again, it is not obvious that putting tremendous effort into assessments alone will give the results needed for traditional assessments to become an acceptable and reliable basis for management.

However, the 25% rule based on indices and catch is an example of a harvest control rule (HCR), which gets around the assessment issues. This is an excellent example of how one might proceed to alleviate a whole suite of problems (subject to appropriate tests).

Testing such rules is basically a management strategy evaluation (MSE), which is often a very useful method to move away from a stalemate in an assessment-implementation scenario. An MSE approach is always tailored to a specific situation and is one way to bring all the partners to the table.

The EGB cod and EGB haddock components appear to be assessed and managed somewhat inconsistently from the corresponding complete stocks and the TRAC-

derived allowable catches are simply subtracted from the total. There is a lot to be said for consistency.

Moving some assessments to "updates" is clearly a step forward in terms of reducing workload on the assessment team, and gives staff more time to focus on stocks where it is important to do more comprehensive assessments.

6 Process for Setting Priorities on Research Questions:

It seems to this reviewer that it is important to funnel as many *ad hoc* requests, pressures or other outside research requests into a formal scheme, to move towards longer-term research objectives instead of treadmill-style assessments with the same main issues recurring year after year.

SARC may recommend research but there is no obvious procedure for taking these into action. This contrasts ICES WGs which can e.g. send research questions, in the form of proposed terms of reference, up through the system, which might then set up new working groups on the topic. Councils do have research set-aside programs, but not such a formal channel. The center has internal procedures for multi-year strategic research planning and of course handles internal research proposals, but there is no formal procedure for including recommendations from the assessment process. Some sort of formal channel for setting research priorities based on recommendations coming out of the assessment process should be useful input for the NEFSC.

NRCC appears to be the obvious place to set priorities. Although it may be difficult since the NEFMC, ASMFC, and MAFMC compete for the "resources" of the NEFSC, that very fact also makes the NRCC the most appropriate place to figure out these priorities.

7 Transparency, communication and CIE Reviews

The Center has quite an amazing number of communication and educational roles, many of which were presented to the panel.

The entire assessment process is very transparent and there are no obvious problems in this regard.

The one exception to transparency seems to be that the SARC has a closed-door discussion session at the end of each meeting. This is a subject of some controversy and there have been calls for opening this session. Interestingly there (a) seems to be little reason for having this writing session closed and (b) there would be little benefit to transparency to opening it. Opening the session should be considered, if only to make the entire process transparent.

The CIE reviewers form an important part of the full assessment-regulatory process, but

there are several difficulties associated with having independent reviews. If all reviewers are to be independent of each other and of the institutes involved then there are obvious potential problems of stability: Sequential SARC/CIE review panels can potentially reach opposing conclusions, e.g. on whether to accept or reject assessments. Put bluntly, the selection of CIE reviewers may affect the conclusion.

The principle of having independent reviews is quite important and it avoids historical (and international) problems of a lock-in on procedures, which eventually become outdated. So care needs to be taken to maintain the independence (not to drop the "I" in "CIE").

There is always room for improvement, however, and one may think of many ways to modify the CIE involvement. For example, the above stability issues may be an indication that the ToR for SARC/CIE panels need to be sharpened, or specifying that at least one member (or an SSC chair) should continue to sit on consecutive review panels.

The SAW-SARC-SSC part of the process takes considerable time and the Center has certainly tried many approaches to limit this. One may want to consider limiting CIE involvement further, e.g. to only extensive strategic reviews. Changes such as modifications to selection patterns or natural mortality currently go through the review process, but one might argue that if a SARC with CIE input provides longer term strategic input, then that would include directions on what should be tested (i.e. a research direction), which then results in assessment method development which get implemented, used and further developed until a new review with CIE involvement takes place.

To this reviewer it does not seem particularly useful to get an outside review in order to modify a selection pattern. Such changes should be data driven.

Finally on the topic of communication, several projects around the world have placed an emphasis on developing management regimes from scratch in a cooperative manner. These approaches may not be led by management agencies. In extreme cases management regimes have only social input and no biological feedback. This contrasts other approaches, which may be completely dominated by biologists and modelers and/or political inputs.

These recent developments include implementations of management schemes in Australian waters, which have taken a multitude of different objectives into account, through current European research projects on co-creating management systems and decision support frameworks⁷. By including all interested parties in the creation of the entire system, the hope is that the usual antagonism will be less than that observed in systems which are mandated completed top-down.

⁷ MareFrame: <http://www.mareframe-fp7.org/>

These bottom-up approaches are no panacea and may be impossible to implement within a given legal framework, but as a long-term project it is almost certain that getting the partners around one table to discuss the entire process is beneficial in terms of reducing friction.

8 Human Resources and Funding

The various assessments and data handling are of the sort of quality expected from an institute with a good international reputation, done by a correspondingly qualified staff.

The Center is, however, extremely stretched on staff resources. This has a number of implications. In the short-term, there are potential implications in how it may not be possible to include a variety of potential improvements in an assessment. This becomes much more of an issue in the longer-term, however, since stretching the human resources will can easily lead to a brain-drain and steadily declining quality of assessments. It is imperative that the staff be given the opportunities to conduct research on assessment methods, including studies of alternative methods, development of new ones and

The panel was informed that 10-15% of the Center's funding was from competitive research funds. Although this might potentially be increased, it is not a likely source for major increase given that the Center can not receive NSF grants. These external schemes have funded some very important research projects including the HABCAM development.

Assuming that there are no immediate new sources for funding, the question remains as to how to best use the available human resources. The Center presented approaches, which had been or could be taken for improvement.

Although the focus above (and during the meeting) focused on the assessment-related issues and timing, it is important to note that various "additional requests" tax the human resources considerably. When staff time is considered, these additional requests also need to be accounted for.

As pointed out above, it is extremely important for assessment personnel to also be researchers, actively publishing research papers. The assessment scientists thus need to be given time for this purpose. This is not for academic reasons, but for personal development and simultaneously keeping assessment methods up-to-date. Most assessments automatically include some new development. Such developments must be evaluated for adequacy and their effects and these tests commonly lead to results, which can and should be published in scientific journals.

9 Ecosystem considerations

The ongoing work by the ecosystem group is very advanced and has already led to changes and new views in the assessment process, but it is “ongoing work”. With the current resources it will be difficult to extensively include ecosystem concerns into the assessment process. That may not be the most difficult part of including ecosystem effects, since the entire advisory process is very single-species oriented, including such concepts as biological reference points.

Some particular ecosystem/environmental effects have from time to time been included as extensions in single species assessments. These include the effects of consumption on prey mortality, the relationships between temperature, salinity or visibility and survey catchability and the effect on environment on the stock-recruitment relationship. Examples of each of these go back a few decades, and one can find examples not only of statistical significance but also of considerable importance in terms of effects on assessment results.

One fundamental problem with these analyses remains. In statistical circles this is called hunting for significance with a shotgun, i.e. searching through potential explanatory variables until something significant is found. This aspect is quite important due to the sheer number of potential environmental variables in any given case study. In the examples shown to the panel, this is unlikely to be an issue (the variables are well justified), but needs to be considered whenever stock assessment scientists are considering multiple environmental variables for inclusion into an assessment.

The use of an environmental/ecological variable to explain natural mortality (M) or catchability (q) implies the potential for disentangling the effects of M and q in the assessment. Normally these two quantities are completely confounded but the inclusion of explanatory variables might make it possible to estimate both. This is well established, when natural mortality is estimated through consumption and q is estimated freely, but the unusual potential here is to force q to have a link through an explanatory variable (thermal habitat change) and estimate M freely. As always, care is needed in how this is done and various criteria can be used to verify just what model is most appropriate.

Finally, the panel was presented with an operating model for simulating an ecosystem to test several simpler multi species approaches. The operating model will initially be Hydra, possibly followed by Atlantis. This is an example of state-of-the art research, which, if the assessment team is included, has tremendous potential in terms of moving assessment methods forward and towards ecosystem-based assessments.

The ecosystem operating model should in principle be able to test almost any of the assessment models used by the Center. As such it can become an extremely important research tool providing important feedback into the assessment process in the longer term. It is important to ensure that the operating model can output data, which can be automatically uploaded into the other models and in this context it should be useful to consider some of the database issues raised above.

10 Summary

The following section is a summary of the above, with recommendations arranged by term of reference.

Issues pertaining to ToR1 (appropriate models) are discussed above in sections 3 and 5. In short, the approach is suitable and the Center is involved in furthering these approaches to improve them. The main problem is one of allocation of resources to research on methods (including self-improvement).

Recommendations:

- Staff should be allocated time for research, publications and similar activities.

Issues pertaining to ToR2 (assessment process efficiency) are discussed above in section 2. In short, the full process is extremely long, but only a minor part of that is due to the time taken by the Center. There are a few potential areas where data handling or assessment time might possibly be reduced, but that would not have major impact on the total length of time to implementation, given the current process. The fact that up to over 2 years pass from data availability through implementation is really quite serious, but this can, apparently, only be drastically changed by shortening parts of the process over which the Center has no control. Modifications to the data handling schemes are likely to improve many aspects of the process however. Similarly, moving species to "update" status can have a considerable effect.

Recommendations:

- Generic methods of data handling across species, areas and data sources should be considered, with the aim of increasing automation, consistency and timeliness.
- Considerations of timeliness should consider the entire process from data collection through implementation.

Issues pertaining to ToR3 (peer review) are discussed above in section 7. In short, the review process appears quite adequate. Attempts to maintain stability or improve timeliness may well be needed but care must be taken that this is not done at the expense of the independence of the reviews.

Recommendations:

- Considerations should be given to either keep at least one CIE member at consecutive reviews, or have a chair from SSC.

Issues pertaining to ToR4 (efficiency and funding) are discussed above in section 8

Recommendations:

- The Center's population dynamics branch does seem to work very effectively
- Research priorities need to be set in a more structured manner.

Issues pertaining to ToR5 (assessment accomplishments) are discussed above in sections 2, 3, 6 and 8. In short, the Center has a fair list of assessment accomplishments, ranging from new methods of assessments (notable scallops) to moves to updates instead of extensive assessments in an attempt to improve

timeliness. Similarly, proposals to replace full-blown assessments on extremely difficult stocks by simpler harvest control rules are potentially major improvements, but of course they need to take appropriate input from stakeholders. To move further on the various issues, priorities need to be set as to where to put research effort.

Recommendations:

- The NRCC should consider research recommendations coming out of the assessment process and assign priorities, which the Center can subsequently use for guidance.

Issues pertaining to ToR6 (ecosystem and climate effects) are discussed above in section 9. In short, excellent work is being conducted on ecosystem models and several attempts have been made at including ecosystem/environmental considerations into single-species assessments.

Recommendations:

- The various processes of extending single species models to incorporate environmental/ecosystem effects should be continued.
- The process of setting up an ecosystem operating model should be a high priority within the ecosystem group.

Issues pertaining to ToR7 (stakeholder engagement and communication) are discussed above in section 7. In short, transparency is extremely important as are research publications. In a few other areas research has been conducted and even implementations exist, where bottom-up approaches have been used to redefine the entire fisheries management system. Whether the friction inherent between partners/stakeholders can be alleviated in the present setting using such bottom-up approaches is not clear, but there is enough at stake that this should be investigated.

Recommendations:

- Whenever possible the assessment process should be transparent, even including writing sessions if this leads to reduced tension.
- Developments of methods (assessments, surveys etc) can and should be published.
- Methods for "co-creating" management strategies should be investigated with the longer-term goal of reducing friction between the partners.

Issues pertaining to ToR8 (improving the assessment and the process) are discussed above in sections 2 and 5. Needless to say, any assessment procedure (or human-defined process) can be improved. One example of potential improvement is the 25% rule for yellowtail. A full study of such rules would constitute a Management Strategy Evaluation (MSE).

Recommendations (see also under ToR1):

- Priorities must be set on terms of reference and research proposals coming out of the assessment process. The natural place for these to be discussed is at the NRCC meetings.
- The full process, from time of data collection through implementation, needs to be shortened. The assessment part is a minor part of the full process.
- Simplifying methods such as the 25% rule for yellowtail should be extensively

studied and evaluated.

NOAA Northeast Fisheries Science Center (NEFSC)

Stock Assessment Science Program Review

Reviewer 4

The objective for this review was to examine and evaluate each Center's fishery stock assessment program that is conducted pursuant to the Magnuson-Stevens Act (2006) and comparable international agreements. Panel was provided with sufficient materials (document plus ppt presentations) to comment on 8 themes that define the stock assessment program.

Overview

Many of these ToRs are related and posed a problem in commenting on them separately. As a consequence there is some repetition, and several of my observations could easily fit under more than one ToR.

1. Does the Center apply a suitable scientific/technical approach to fishery stock assessment modeling?

- Observations: The Centre has a strong stock assessment modelling capacity and apply best scientific practices in modeling to achieve their goals.
- Strengths: 1) Many of the staff are internationally recognized from their publications and participation at Canadian, NAFO, NASCO and ICES fora, and demonstrate commitment to scientific integrity in their work; 2) There are long time series for annual multiple surveys and landings; competent aging and biological groups; and a good port and observer program; and 3) the use of NOAA Fisheries Toolbox programs is a mainstay to many of the current stock assessments.
- Challenges: 1) Data processing issues centered on the often late arrival of catch data from federal and coastal state fisheries which results in large and unacceptable lag times for completing the assessment process; 2) staff are severely burdened by the number of stocks that have to be assessed annually and the frequency demanded by Councils and Commission, including too many Terms of References and ad hoc requests; and there is also very little time to investigate new assessment models or modelling approaches, e.g. SURBA, Stock Synthesis, etc., and little opportunities to work on improving existing models.

Recommendations to address the issue: 1) The Center should reinstate its Methods Working Group and expand it to include experts from the other groups in the Population Dynamics Group to discuss commonality in approaches to assessment modelling which could reduce individual

modeling workloads, e.g., projections methods. Here also staff would be discussing and sharing new innovations in current models and the results of investigations into new models e.g. SURBA and other similar survey based assessment models, and framework approaches such as integrated analysis (IA), e.g. stock synthesis, MULTIFAN-CL, etc.; 2) the Centre needs to find a mechanism to reduce the assessment workload of staff generated by Councils, Commission and the Centre; and 3) the Centre needs to refocus (>20%) staff time on increasing biological knowledge necessary for tackling critical research questions such as changes in M, catchability, movements of stocks, the retrospective patterns that are occurring in many stocks, and improvements to stock assessment approaches. Staff development should consist of training in new stock assessment methods, participation in international meetings and working groups such as ICES, NAFO and NASCO, etc. to keep abreast in new developments in data collection, surveys and modelling, and promotion of scientific writing for primary publications.

2. Is the assessment process efficient, effective and clearly described, including terms of reference for assessment reports?

- **Observations:** The current stock assessment cycle of fulfilling the need of the Councils and Commission, additional assessment demands following the SAW/SARC/SCC process and numerous ad hoc request from the regulatory bodies has resulted in no downtime. Time for current assessments are compromise to deal with many of these extra demands. More importantly, the stock assessment results to set ABCs may have catch data that can be up to 3 years old because of delays in data arrival and is unacceptable for giving management advice on current stock status and projections. The whole lengthy assessment process and the extended 2 tier review process of SARC and SSC, the Council/Commission process, the management plan formulation and the public comment process periods is too long and needs to be streamlined. All of this can undermine the quality of the science. No Panel member reported this happening in their respective countries.
- **Strengths:** The Centre produces high quality assessments, and the NOAA Toolbox plays a large role in delivering the assessment products. The staff try to balance their assessment workloads with additional major demands.
- **Challenges:** 1) the ever increasing work load of the Centre staff does not permit them time to deal assessment related issues such as catchability, dome-shaped and/or time-varying selectivity, M, stock productivity, and characterization uncertainty. Many of these are common problems across many species. Staff feel they are on a treadmill with no time to make improvements to the models or the science; and 2) sensitivity tests on various parameter estimations are carried out, but there minimal mention of risk analysis being performed so I am not fully aware of the extent of its usage.

Recommendations to address the issue: 1) investigate and institute a better delivery system for fishery dependent data collection and transfer using automation to minimize time lags in arrival

times at the Centre. Also the trip identifier issue needs to be resolved quickly. It was noted that ICES reduces the complexity of the science to meet delivery of data, analysis, and implementations within the same year; 2) The assessment process itself needs to be more streamlined to improve delivery. Chief among this is limiting the number of ToRs coming in from all levels of the assessment process. This should be undertaken by the NRCC who should screen and prioritize the ToR timelines.

3. Does the Center, in conjunction with other entities such as the Councils' Scientific and Statistical Committees, have an adequate peer review process?

- Observations: the peer review process is extensive and thorough with sufficient transparency, and takes into consideration qualifications of experts, balance of perspective, conflict of interest, and independence. This holds true for SAW/SARC, TRAC, operational and update assessments. The Centre does not seem to have established protocols for considering and including input from scientists not on the agency assessment team. The hiring of scientists by industry to conduct independent modelling is starting to become commonplace and the Centre has to develop these protocols.
- Strengths: the peer review process is extensive and thorough with sufficient transparency
- Challenges: Securing qualified experts for reviews both internally within the system and outside agencies such as CIE.

Recommendations to address the issue: 1) A further step to streamline the SAW/SARC assessment process timing is to eliminate one of the 2 peer review groups by merging responsibilities of the SSC into the SARC; and 2) operational assessment which involve the application of existing models with update data may be a quick answer to handle the increasing number of assessments per year efficiently, while reducing the overall process time especially the peer review process (only SSC involved). However it must be recognized that this approach would severely limited improvement to existing models or development of new models or assessment frameworks

4. Does the Center work effectively internally and in coordination with the NEFMC, MAFMC, ASMFC, and GARFO to accomplish needed assessments according to a set of priorities? Considerations should include program structure, staffing, funding, and the stock assessment prioritization process? AND 5. Does the Center achieve adequate assessment accomplishments relative to mandates particularly with respect to the number of Fishery Management Plan (FMP) species assessed?

- Observations: 1) The Centre is approaching the breaking point in exceeding intellectual capacity to carry out its mandate due to the increase analytical and review demands from GARFO, Councils, Commission, TRAC and headquarters. These demands are outstripping the resources of the Centre, and Centre management needs to make all internal and external partners aware of this situation; and 2) it was voiced that industry wants more benchmarks because they perceive benchmark assessments will lead to increase quotas so there is pressure to have more – this perception has to change.
- Strengths: Talented but overworked staff are able to meet most of the regional expectations in terms of quantity and quality in both data poor and full assessments.
- Challenges: Some protocols seem to exist within the stock assessment process to prioritize need, frequency and appropriate level of stock assessments but there does not seem to be one decision making body to govern the process.

Recommendations to address the issue: Apply a systematic approach for addressing problems in stock assessment process. There is a need to have one body set the priorities in frequency and time-lines for assessments (benchmark vs updates) using a defined set of criteria. In addition the huge demand for ad hoc requests from all players in the assessment process during the year also needs to be prioritize. That body should include the Centre and the Councils and Commission chairs and should be an extension of the duties of the NRCC.

6. Does the Center have an effective process in place for taking ecosystem and climate change factors into consideration in the stock assessment process?

- Observations: The Center does consider ecosystem and environmental factors affecting fish stocks and their assessments. There are some recent examples of using ecosystem data to improve assessment performance in yellowtail flounder, butterfish, and herring. In addition a collaborative multispecies assessment project between Centre ecosystem and assessment scientists for Georges Bank has just begun which could, at a later date, address stock spatial issues, migration/movement, and other technical aspects of the assessment.

- Strengths: 1) The Centre has an outstanding history of ecosystem monitoring, data collection, and research; and extensive food habits database in the Food Web Dynamics program. 2) Several scientist in these groups are both nationally and internationally known.
- Challenges: The integration of ecosystem information into stock assessment is complex and challenging. This process is being hampered by budget cuts to the ecosystem observation surveys and the overcapacity in stock assessment workload. Scrutiny of ecosystem results in the stock assessment by the peer-review process may be biased or glossed over due to inexperience of experts in this combine research field.

Recommendations to address. 1) at the bare minimum ecosystem funding should be stabilized or increased; and 2) more integration within the Centre of analytical stock assessment and ecosystem experts to address common approaches and methods.

7. Does the Center adequately engage stakeholders in the stock assessment process and communicate assessment-related results, needs, and research to them effectively?

- Observations: The Centre has an elaborated communication setup to effectively communicate results to clients and the general public
- Strengths: Talented staff with good resources; well organised process sanctioned by NRCC; Council members can attend assessment peer reviews (SARC and SSC)
- Challenges: there is problem how to deal with public perception of external comments on assessment results; in addition SSC reviews may conflict with assessment results and generate negative perception in assessment integrity.

Recommendations to address the issue: Members of NRCC should take an active role in increasing communication within the assessment process and outside to the general public. Centre staff should be given ample opportunities to publish results from stock assessment model development and critical research areas to enhance the scientific integrity of the Centre.

8. Are there opportunities for improving stock assessments and the stock assessment process?

- Observations: There are several research areas that need to be address such as calibration of survey catchability (fixing Q will improve estimation of M) and selectivity (dome shape or not) in survey and fishery data; investigating underlying causes of retrospective

patterns; incorporating environmental data into stock assessment and characterization of uncertainty.

- Strengths: talented staff that have the intellectual capacity to adapt to new challenges.
- Challenges: problems with cleanup and timely delivery of fishery dependent data; freeing up of staff time away from stock assessment workload to work on these neglected research areas.

Recommendations to address the issue: 1) The Centre needs to avail of (or obtain) expertise in gear technology to address issues of calibration and selectivity; 2) apply management strategy evaluation (MSE) to stock assessment issues such as retrospective problem and decisions on catch advice; and 3) use funding from the set aside (RSA) programs to tackle immediate, critical research issues such as an independent estimate of natural mortality, movement and stock size for Georges Bank yellowtail and cod; e.g. with a re-introduction of tagging programs.

NEFSC Program Review May 19-22, 2014

NOAA Northeast Fisheries Science Center 2014 Stock Assessment Science Program Review

Reviewer 5

Background

The National Marine Fisheries Service's Northeast Fisheries Science Center (NEFSC) in Woods Hole, Massachusetts, conducted an external review from May 19 to 23, 2014, to examine and evaluate the NEFSC fishery stock assessment program. The focus was on the overall process including data preparation, assessment modeling, peer review process and communication. A review of the scientific data gathering and data management procedures related to fishery stock assessments conducted by the in the NEFSC was conducted in August 2013.

All background materials were provided to the Panel electronically through the Center (or Office of Science and Technology) website well in advance of the review. All presentations were provided to the Panel, through the website, at the beginning of the review. A booklet folder with copy of all presentations was provided to the Panel.

During informational presentations each day there were specific intervals included for public comment. Stakeholders were invited to participate as observers and to comment during the daily public comment sessions. At the close of the review, the Panel and Center Directorate met to discuss the results of the review in closed session.

Some general comments and impressions

The peer-review process was very well organized, with the only critique being some overlap between TORs which will be reflected in the comments to each TOR. The quality of the scientific background material and the presentations and the very professional organization of the review meeting show that NEFSC have a highly skilled, motivated, and dedicated scientific and technical staff. It is also apparent that the NEFSC staff has a very high workload, while still delivering high quality science. The Population Dynamics Branch conducts assessment of around 60 species caught in commercial and recreational fisheries which is an incredible achievement given the number of staff. The data preparation, analysis, and the comprehensive review process that is part of the assessment process, seems to be so demanding that NEFSC will not have sufficient time to develop and implement solutions to some of the challenges they have identified related to climate change and changes in the ecosystem. The overly complex review process leading up to the quota advice not only overburdens the assessment scientists, but also may add substantial error to the assessment process since the ABC and quota may in actuality be based on 2-year old scientific survey and catch monitoring data. The timing and also quality requirements for each stock assessment would ideally be prioritized to carve out more time for research and development. Improved allocation of NEFSC staff time between routine assessments and research for methods development has potential for substantial advancements in the scientific basis for the fisheries management advice. In particular, it is important that time be freed up to make capacity for the planned re-designing of the fishery-dependent data collection system, and for the development of assessment methods that better can account for factors related to climate and ecosystem changes. The NEFSC has tremendous data and expertise in food web modeling, and research teams that combines expertise in ecosystem modeling and assessment modeling (which requires that time be freed for assessment scientists to participate) could facilitate much greater contribution to stock assessments. Providing time for such cutting-edge modeling efforts can also aid in retaining key staff that might otherwise be lost due to burnout.

Comments on the 8 questions in the terms of reference

- 1) Does the Center apply a suitable scientific/technical approach to fishery stock assessment modeling?

The Northeast Fisheries Science Center (NEFSC) employs suitable peer-reviewed stock assessments models. Models included in the NOAA Fisheries Toolbox which are used for the

assessment of a majority of stocks are peer-reviewed and accepted in the scientific community. One advantage with the Toolbox as identified by the NEFSC staff is that it allows rapid update runs during the assessment process, for example in response to requests during the review-process. Also, it is easier for new staff to get up to speed with assessment modeling, especially since the format for input data is standardized. However, a downside of the GUI interface developed in Visual Basic is that it can be time-consuming to implement new models, or model updates, so the models actually used in the assessments may not be the best available in some cases. The frequent updates in the stock-synthesis model are an example where the current version in the Toolbox is not up-to-date. It is also well recognized by the NEFSC that the toolbox has limitations for incorporating ecosystem considerations in the stock assessment process. Also, it could be beneficial to expand the suite of statistical assessment models available through the Toolbox to include for example the ICES SAM model (Anders Nielsen, DTU Aqua) and other recent developments. I am confident that the very strong modelers at NEFSC would have developed more cutting-edge methods if they were provided time and resources. For example, the Toolbox currently available does not include any modeling tools that can account for multiple species interactions, and none of the models are spatially explicit. NEFSC has already developed several assessment modeling approaches in addition to the Toolbox to deal with stocks where the Toolbox models are not suitable. The Toolbox is not used or is not applicable for ca. 25% of stocks, including scallops. The scallop assessment for example is based on the CASA (Catch at Size Analysis) model, with overall biomass and recruitment information obtained from several surveys. More recently, the NEFSC scallop survey has evolved into a combined dredge and optical survey (HabCam) which provides abundance data of exceptional quality. The stock assessment for scallop is performing very well, and has allowed catch levels resulting in fishing mortality close to reference points.

According to the presentations and material provided for the review it appears that diagnostics of the model performance of at least the age-based assessment models focus on retrospective patterns. Diagnostics related to the quality of input data presented appeared to be limited, which makes it more difficult to assess the reasons for poor model performance. In some cases it seems like a fairly large effort is spent on trying fixing retrospective patterns, when perhaps the real problem is model assumptions or inadequate input data from the surveys and biological sampling of catches. For age-based assessments, tables of # aged fish per species were presented as an indication of the information available for estimating abundance indices or catch at age. It is recommended that NEFSC consider using some of the data quality diagnostic currently being developed in ICES (e.g., ICES WKPICS). Because of the multi-stage sampling, where primary sampling units (PSUs) typically are trips, the number of fish aged is not very useful for evaluating if there is sufficient data to estimate catch-at-age adequately. The effective sample size for estimating catch-at-age is largely driven by the number of primary sampling units (PSUs). Even if a large number of fish is aged, the precision in estimated catch-at-age may be very poor if the biological samples for length and age compositions are collected from only a few

vessels and trips in the observer and port sampling programs. Also, the precision of abundance indices by age and length are driven by the number of trawl stations, and much less by the number of fish measured for length and age. If assessments are conducted for stocks that occupy relatively small spatial areas, then the number of trawl haul, and the number of trips sampled for length and age may be quite small, resulting in poor precision in key input data to the stock assessment models. I therefore recommend that diagnostics to go along with age-based assessments include number of trawl hauls (in the fishery-independent surveys) and the number of vessels and trips with age-samples by gears and area for the fishery-dependent data. The separate stock assessment of cod and haddock for portions of the Georges Bank (EGB cod, and EGB haddock), as part of the Transboundary Resources Assessment Committee (TRAC), is an example where a thorough evaluation of the data-support (e.g., number of trips where catches were sub-sampled for age) likely would reveal that catch-at-age cannot be estimated with adequate precision. Also, for a small geographic area, only a limited number of trawl stations in the fishery-independent survey would provide data on numbers-at-age, resulting in poor precision.

- 2) Is the assessment process efficient, effective and clearly described, including terms of reference for assessment reports?

The assessment process leading to the catch advice appears to be inefficient due to problems with data-streams, excessive reporting demands, and an overly complex and time-consuming meeting schedule related to the peer review process. The timing of availability of raw input data for stock assessments are scheduled to be available by early April, but in practice timing is variable so data may be received up to a month behind schedule in some years. This is incompatible with a rigid schedule of assessment reviews. The handling of fisheries-dependent data to prepare the input data to stock assessments takes significant time. The assessments are generally completed within 2-4 months of receiving the input data. With the time added to complete the following review process leading up to the ABC and catch quota advice the whole assessment process can take more than a year. The proposed streamlining of assessment reporting has the potential to significantly reduce workloads in the long-run. Also, NEFSC staff suggested that the number/amount and type of information requests from CIE reviewers be restricted. In the current system, there are situations where many requests (which could entail a lot of extra work for NEFSC scientists) likely provide information of limited value for determining if the assessments are adequate.

The stocks assessed by NEFSC are managed under fisheries management plans (FMPs) from the Northeast and Mid-Atlantic Management Councils and the Atlantic States Fisheries Management Commission. The assessments of fish stocks in the offshore US waters of the northwest Atlantic are subject to a very thorough scientific peer-review process through the Northeast Regional Stock Assessment Workshop (SAW). The SAW protocol is used to prepare and review

assessments. Assessments are prepared by SAW working groups (federally led assessments) or Atlantic States Maine Fisheries Commission (ASMFC) technical assessment committees (state led assessments) and reviewed by an independent panel of stock assessment experts called the Stock Assessment Review Committee (SARC), which includes reviewers from the Center of Independent Experts (CIE). On top of this, the Statistical and Scientific Committee conducts reviews of the reviews before coming up with the quota recommendation. This Stock Assessment Review layer is so cumbersome that it can compromise the assessment process. Due to the significant delay in the determination of ABC that is the basis for the quota, and the time to put catch quotas into regulations by GARFO (including the EIS process) may result in the final quota being based on 2 year old data. This is clearly risky, since it essentially means that the quota is based on a 2-year forward projection that adds significant uncertainty on top of the uncertainty in the stock assessment results. In addition, the enormous workload on NEFSC assessment scientists is at the expense of time for research and innovation to develop assessment models that are better suited to deal with the ecosystem approach to fisheries management.

It is critical that the assessment process be more streamlined and timely, and that the workload for NEFSC assessment associated with routine production assessments be significantly reduced to allow for more time for research. The current situation is not sustainable and there is a clear risk that NEFSC can lose highly talented staff if it is not resolved quickly. One of the proposed solutions that benchmark only is required when stock assessments are supported by new data or new model development, which I agree with. The current process seems like over-kill.

- 3) Does the Center, in conjunction with other entities such as the Councils' Scientific and Statistical Committees, have an adequate peer review process?

The high focus on the review process related to stock assessments conducted by NEFSC is commendable. In addition to the very thorough reviews conducted through SAW/SARC and the SSC as part of the routine process, there have been several comprehensive review meetings since 2000. In particular, the Groundfish Assessment Review Meetings (GARM): 2001:GARM I—Benchmark all stocks; 2005: GARM II—Update all stocks; 2008 GARM III Benchmark for all 19 stocks, and the Data Poor Workshop 2009.

The inclusion of CIE in the review process is very important, and demonstrates that NOAA has a strong commitment to move towards best available science, and not just best available NMFS science. This very strong focus on independent peer-reviews of stock assessments and assessment methods, and the transparency (e.g., the SSC open meetings) is lacking in Europe. While this thorough and open review process sets a standard internationally, and has many advantages, the burden on staff doing stock assessments is too high. The current multi-layer review process is not sustainable given current staffing levels. The tentative processes to reduce number of peer reviews by allowing for operational assessments where a certain level of changes can be implemented without a full review cycle is promising. There appears to be some duplication of peer-reviews in SARC and SSC. A reduction in meetings can maybe be achieved through a merging or better coordination of the SARC and SSC, a stricter prioritization of TOR for reviews, and by focusing on the identification of essential research questions that can realistically be addresses within available funding.

Some additional observations:

It appears that there is no organized way of presenting all the recommendations coming from the peer reviews, and responses to these recommendations over time. It could be beneficial to have a SharePoint website where research requests are presented for each review, along with the follow-up research to address the requests.

It was also noted that there are no defined standards to determine when an assessment should be rejected or accepted, and this can be particularly problematic when there are many review panels operating independently. The proposed process of conducting operational assessments for a number of stocks, and then deal with a more limited number of stock assessments where major advances are needed through a research track assessments, is promising.

It was also pointed out that CIE reviewers may not be familiar with complex management system in Atlantic States, which can result in misconceptions. Some limitations on the type of requests in the CIE review process could likely reduce the work-load on NEFSC staff without jeopardizing the independence of the review.

- 4) Does the Center work effectively internally and in coordination with the NEFMC, MAFMC, ASMFC, and GARFO to accomplish needed assessments according to a set of

priorities? Considerations should include program structure, staffing, funding, and the stock assessment prioritization process.

The current situation where the population dynamics branch is overburdened by the stock assessment process suggests that the NEFSC, MAFMC, ASMFC, and GARFO have not managed to set priorities that sufficiently considers NEFSC staffing level and the need to secure time for professional development. Ability to innovate is not fully at NEFSC met because of continued overload. It appears that there is little room for NEFSC to prioritize because of mandates driven by the Councils, the ASMF, and the Magnuson Stevens act. The many demands from the NEFMC, MAFMC, and ASMFC do not seem to fully take into account the availability of staff and funding at NEFSC to address all the requests. There is clearly a need for prioritization. The assessment buck idea presented by Paul Rago is one possible way of improving the stock assessment prioritization process. In particular, it appears that a large number of requests from MAFMC, ASMFC, and GARFO throughout the year are not related to stock assessments. Clearly, there is a need to coordinate and prioritize the requests from all these stakeholders so that the required work can be achieved without over-burdening NEFSC.

- 5) Does the Center achieve adequate assessment accomplishments relative to mandates particularly with respect to the number of Fishery Management Plan (FMP) species assessed?

Given the available staffing and funding, the answer is clearly yes. Given the complexity of the management system, and in particular the fact that many stocks falls under more than one FMP, and the very large number of research requests it is remarkable that the population dynamics branch manages to deliver the stock assessments requested every year.

- 6) Does the Center have an effective process in place for taking ecosystem and climate change factors into consideration in the stock assessment process?

The current high reliance on the Toolbox, combined with very limited time for research for the assessment scientists, is likely to be a bottleneck for the process of developing new modeling approaches that can incorporate ecosystem and climate change. Research teams with ecosystem and stock assessment modelers have the potential to make significant advances in the stock assessment process that are not yet fully realized because of the heavy workload required for stock assessments. The NEFSC is in an exceptional position with a very strong ecosystem assessment research team, a very strong team of assessment scientists, and exceptional time series of data to support the incorporation of climate and ecosystem factors in the assessment process. The NEFSC has a formal process to incorporate ecosystem approach. Challenges relate to research time to develop modeling approach, and cuts in ecosystem observational surveys.

- 7) Does the Center adequately engage stakeholders in the stock assessment process and communicate assessment-related results, needs, and research to them effectively?

My impression is that NEFSC has a strong system for engaging stakeholders, and that assessment-related results are communicated effectively. When presenting information about uncertainty to the fishery industry and other stakeholders it may be useful to separate uncertainty related to random sampling errors (precision) versus bias that can results from poor coverage of the stocks in fisheries-dependent surveys for example, or from poor model formulation. It is important to communicate that it cost money to improve accuracy of stock assessments. Emphasizing what sources of uncertainty that is quantified (known) versus uncertainty than cannot be quantified is important. This becomes particularly important for defending the level of monitoring required to achieve adequate stock assessments. I think the issue of uncertainty is particularly important related to the spatial resolution of stock-assessments and demands for fisheries-management. Quota allocation at finer spatial scales, and for métiers defined by area and gear-combinations, will have to take into account the data-support for the stock-assessments. I believe that the higher demand for assessment results for small spatial areas, and time periods, is incompatible with the level of sampling effort in the fishery-dependent and fishery-dependent surveys. The reason, again, is that the effective sample sizes for estimating abundance-indices and catch at length or age is driven by the number of PSUs sampled, which usually are quite few for species that occupies small areas, or for fisheries that concentrates in smaller areas. The assessment of George bank cod and haddock, where separate estimates are required for subareas of the Georges Bank, is an example where uncertainty in the stock assessment results would be substantially increased because of the reduction in effective sample sizes of the input data.

As part of the outreach, it is also recommended that NEFSC place more emphasis on scientific publications related to stock assessment methods, which is important for recognition in the scientific community.

- 8) Are there opportunities for improving stock assessments and the stock assessment process?

This is partly addressed in TORs 1,2, 6. I would also like to add that there could be improvements made by moving away from VPA type assessment models. More research on statistical assessment models that fully can account for sampling errors (including clustering effects) in fishery-dependent and fishery-independent survey data, could lead to improvements especially for age-based assessments. One of the biggest potentials for improvement in the stock assessment likely relates to improving the quality of the fishery-dependent information. The current system of data-collection is largely ad-hoc (ref. 2013 Program review), which can result in variable bias of unknown magnitude.

The stock assessment process as discussed above under several TORs has many rooms for improvements - in particular a more efficient peer review process that can improve timeliness and thus improve the basis for quota advice. Also, a streamlining of the reporting requirements, as for example the proposed standardization of the stock assessment reports. Use of Sharepoint for large number of figures and graphs that are not essential to include in the main report could be useful. Also, a better system to track research recommendations in the peer-revie process, and

the advancement on solving these, would be beneficial. This could perhaps also be done through a Sharepoint site.