Chair's Summary of Program Review of Stock Assessment Process
NOAA–NMFS Southeast Fisheries Science Center
Miami, Florida
8–10 July 2014

Review Panel Members

- Michael Hansen, USGS – Great Lakes Science Center, Chair
- Robert Atlas, NOAA – Atlantic Oceanographic and Meteorological Laboratory, Reviewer
- Ewen Bell, UK – Center for Environment, Fisheries & Aquaculture Science, Reviewer
- Joseph Hightower, USGS – North Carolina Cooperative Fish & Wildlife Research Unit, Reviewer
- William Karp, NMFS – Northeast Fisheries Science Center, Reviewer

Background and Overview of Meeting

The stock assessment process conducted by the NOAA Southeast Fisheries Science Center (SEFSC) was reviewed at the Mayfair Hotel in Miami, Florida, during 8–10 July 2014. The review was convened as an annual review of science programs at National Marine Fisheries Service (NMFS) Science Centers (including associated laboratories) and the Office of Science and Technology (ST) to: (1) evaluate quality, relevance, and performance of science and research conducted in NMFS Science Centers and associated laboratories; and (2) strategically position Science Centers and ST in planning future science and research. As defined in the Terms of Reference (TOR) for the review, the objective of the review was to examine and evaluate the SEFSC fishery stock assessment program that is conducted pursuant to the Magnuson-Stevens Act (2006) and comparable international agreements. Stock assessments apply mathematical and statistical models to data collected from living resources and their associated fisheries to provide scientific advice on current and future status of managed resources. Fishery, survey, and biological data available for stock assessments were reviewed in 2013, so the present review focused on the process of developing stock assessments from these data, including modeling approaches, review process, and communication of advice. The review was not intended as an in-depth review of any particular stock assessment, but rather, as a review of the body of assessments conducted in response to available data and management needs. During the review, the panel considered materials provided by the Center to comment on seven assessment themes related to the NMFS marine stock assessment program in the southeast. Each theme and operational discussion item was led off by a short presentation, followed by in-depth discussion among panelists and designated discussants (SEFSC, SEDAR, and management council representatives). Each panel member, including the chair, summarized their comments about seven themes in individual summary reports (attached below the Chair's report). The Chair also summarized comments that emerged from multiple panel members (immediately below).

General Observations and Recommendations

Stock assessment scientists at the SEFSC complete an extraordinary number of assessments for a very large number of high-valued species. Clearly, the requested number of stock assessments exceeds total capacity of the work force of lead scientists. The Center Director is therefore challenged to balance a need to produce the largest number of assessments possible within work force limits (i.e. the number of stock assessment scientists), while also ensuring the work force of assessment scientists is allowed time for professional development (e.g. publishing scholarly articles and attending professional conferences). The latter need for career development has been increasingly sacrificed in favor of the former need to complete the largest possible number of stock assessments, because of attrition in the work force, which is at least partly induced by the intensive work load of lead scientists, and lags in replacing scientists who leave, which is largely a function of the agency's personnel management system. A need for stock assessments is not expected to dissipate in the future, which argues for stream-lining the stock assessment process. Similarly, stock assessment scientists must be allowed to develop their careers by allowing them time to publish scholarly articles and to attend scientific conferences, which argues for changing work-force management in the Center to explicitly reserve time for professional development. Specific issues related to thematic areas of the review are summarized below, which collectively support this general observation and overall recommendation.

Summary of Panel Member's Major Observations and Recommendations

- **Theme 1: Science and Technical Approaches**
    - Observation – The approach used by stock assessment scientists at the SEFSC is state-of-the-art, and relies on a lead scientist to develop a model that is appropriate for the available data, but also meets management needs for fishery advice. Clearly, data limits model complexity, whereas managers may desire a model that is more complex than is warranted by the available data. A consequence of the overall approach is that any assessment model has an appearance of a "one off" model that is highly unique for the available data and the specific fishery management need. Another consequence of the overall approach is that managers may

desire a model that is more complex than warranted by available data, because a more complex model is often believed to be better than a less complex model. This tension between quality and quantity of the available data and a desire by managers and modelers to seek the most complex model possible can lead assessment scientist to seek a more complex model than is reasonable or necessary.

- Recommendation #1 – Invest in data collection that is commensurate with management system needs, so data needed for the most appropriate stock assessment model are of sufficient quality and quantity (e.g. age compositions for catch-age models; fishery independent surveys).
- Recommendation #2 – Develop a written standard operating protocol to formalize the process of model development in response to: (1) quality and quantity of available data; and (2) needs of the management system.
- Recommendation #3 – Ensure that uncertainty of each data stream is propagated through model development, so management advice correctly reflects model uncertainty, to ensure managers are fully aware of the level of risk associated with their management actions.

- **Theme 2: Assessment Process**
  - Observation – The assessment process aims to secure assessments of high priority stocks by prioritizing stock assessment needs in relation to availability of lead scientists. Overall, the process is sound and seems to produce assessments of stocks that are of highest priority. However, the process sometimes seeks a level of assessment that may not be reasonable or necessary, as when a "Benchmark" assessment is requested to seek a better answer than an earlier "Benchmark" assessment.
    - Recommendation #1 – Account for lead scientist time in hours, rather than "slots", to provide a better match between available effort and requested effort.
    - Recommendation #2 – Shift emphasis from "Benchmark" assessments to "Update" assessments, to ensure the most appropriate level of assessment is applied.
    - Recommendation #3 – Thoroughly review the SEDAR process every 10 years, to ensure the process is both meeting management needs and appropriately using SEFSC scientist effort.

- **Theme 3: Peer Review Process**
  - Observation – The peer review process used for stock assessments ensures an objective review of the process and products (i.e. the gold standard in science). Overall, the peer review process is sound. However, over-emphasis on "Benchmark" assessments and extensive documentation for assessments challenges the process and peer reviewers.
    - Recommendation #1 – Develop a standard operating procedure for methods commonly used in stock assessments (e.g. perhaps through a working group).
    - Recommendation #2 – Peer-review standard operating procedures using standard review methods, such as the CIE.
    - Recommendation #3 – Exempt subsequent stock assessments that rely on standard methods from peer review, and require peer review only of alterations in standard methods.

- **Theme 4: Communication**
  - Observation – Stock assessments are documented through extensive open-access publication of documents that elaborately describe all details of each assessment, from data compilation and manipulation through model selection and development. This extensive documentation is redundant among assessments when standard methods are used and lacks transparency by being too technically dense for many stakeholders and cooperators.
    - Recommendation #1 – Simplify documentation of stock assessments by referring to standard methods wherever possible, which should be posted in the same internet location, along with peer review comments and findings.
    - Recommendation #2 – Summarize departures from standard methods and previous assessments early in the documentation, to facilitate reader understanding of departures from standard operating procedures or protocols.
    - Recommendation #3 – Add a concise summary of each assessment that is understandable to stakeholders and cooperators, preferably to be written by a communication specialist or by a lead scientist who is trained in public outreach and communication.

- **Theme 5: Research Opportunities**

- Observation – Stock assessment scientists produce high-quality state-of-the-art assessments that are highly responsive to cooperator needs. However, scientists are allocated too little time for research to both improve stock assessments and to advance their careers, a consequence of which is turnover among stock assessment scientists that is higher than expected or desired.
    - Recommendation #1 – Reserve a portion of each lead scientist's annual FTE for research into improvement of stock assessment methodology and to personal research.
    - Recommendation #2 – Ensure scientists are current in their knowledge by supporting their attendance at one conference per year.

- **Theme 6: Ecosystem Considerations and Next-Generation Assessments**
    - Observation – The SEFSC is actively engaged in developing models that integrate multiple species and the surrounding environment. However, this effort was funded externally, which suggests the effort was not high enough in priority for base resources to be allocated. Further, next-generation assessments should include management strategy evaluations to test performance of the overall system.
        - Recommendation #1 – Provide base support for the GOM IEA 3-year plan by allocating a portion of lead scientist's time to this effort.
        - Recommendation #2 – Incorporate environmental variables within assessment models to increase precision of management advice, if variables are predicable (i.e. decadal oscillations).
        - Recommendation #3 – Undertake a management strategy evaluation of the stock assessment process and associated fishery management system.

- **Theme 7: Organization, Priorities, and Accomplishments**
    - Observation – See descriptions of Themes 1–6, for background related to these recommendations, which cross over all preceding themes.
        - Recommendation #1 – Adopt the national model for using the most appropriate assessment model for each stock selected by SEDAR.
        - Recommendation #2 – When allocating scientist effort to stock assessments, reserve a portion of each scientist's time for professional development
        - Recommendation #3 – Support scientist's attendance at national and international conferences, to ensure scientist knowledge is state-of-the-art and to promote career development.

Conclusions – The stock assessment program at the SEFSC benefits from a highly skilled and extraordinarily dedicated workforce whose throughput of stock assessments serves fishery managers and stakeholders very well across a broad region of ocean resources. Throughput of stock assessments would benefit from streamlining the process by: (1) shifting more assessments from benchmarks to updates; (2) standardizing methods that overlap among assessments; and (3) shortening stock assessment reports by using standard operating protocols. Workforce management would benefit from: (1) treating lead stock assessment scientists in units of hours, rather than slots; and (2) reserving a portion of each lead scientist's FTE for career development (publishing scholarly articles and attending conferences). Communication of stock assessment findings would benefit from: (1) summarizing key changes or innovations at the front end of documentation; and (2) adding a concise layperson summary to each assessment.

Reviewer Report on Program Review of Stock Assessment Process
NOAA–NMFS Southeast Fisheries Science Center
Miami, Florida
8–10 July 2014

Background

A review of the stock assessment process conducted by the NOAA Southeast Fisheries Science Center (SEFSC) was undertaken at the Mayfair Hotel in Miami, Florida, during 8–10 July 2014. The review was convened as an annual review of science programs at National Marine Fisheries Service (NMFS) Science Centers (including associated laboratories) and the Office of Science and Technology (ST) to: (1) evaluate quality, relevance, and performance of science and research conducted in NMFS Science Centers and associated laboratories; and (2) strategically position Science Centers and ST in planning future science and research. As defined in the Terms of Reference (TOR) for the review, the objective of the review was to examine and evaluate the SEFSC fishery stock assessment program that is conducted pursuant to the Magnuson-Stevens Act (2006) and comparable international agreements. Stock assessments apply mathematical and statistical models to data collected from living resources and their associated fisheries to provide scientific advice on current and future status of managed resources. Fishery, survey, and biological data available for stock assessments were reviewed in 2013, so the present review focused on the process of developing stock assessments from these data, including modeling approaches, review process, and communication of advice. The review was not intended as an in-depth review of any particular stock assessment, but rather, as a review of the body of assessments conducted in response to available data and management needs. During the review, the panel considered materials provided by the Center to comment on seven assessment themes related to the NMFS marine stock assessment program in the southeast. Each theme and operational discussion item was led off by a short presentation by a facilitator, followed by in-depth discussion among panelists and designated personnel who were assigned as discussants. Discussants included key SEFSC, SEDAR, and management council representatives. Each panel member summarized their comments about the seven themes in individual summary reports. The Chair also summarized comments that emerged from multiple panel members as General Observations and Recommendations (immediately below).

General Observations and Recommendation

Stock assessment scientists at the SEFSC complete an extraordinary number of assessments for a very large number of high-valued species. Clearly, the requested number of stock assessments exceeds total capacity of the work force of lead scientists. The Center Director is therefore challenged to balance a need to produce the largest number of assessments possible within work force limits (i.e. the number of stock assessment scientists), while also ensuring the work force of assessment scientists is allowed time for professional development (e.g. by publishing scholarly articles and attending professional conferences). The latter need for career development has been increasingly sacrificed in favor of the former need to complete the largest possible number of stock assessments because of attrition in the work force, which is at least partly induced by the intensive work load of lead scientists, and lags in replacing scientists who leave, which is largely a function of the agency's personnel management system. A need for stock assessments is not expected to dissipate in the future, which argues for stream-lining the stock assessment process. Similarly, stock assessment scientists must be allowed to develop their careers by allowing them time to publish scholarly articles and to attend scientific conferences, which argues for changing work-force management in the Center to explicitly reserve time for professional development. Specific issues related to thematic areas of the review are summarized below, which collectively support this general observation and overall recommendation.

Key Specific Findings and Recommendations

- **Theme 1: Science and Technical Approaches**
    - Overall Approach – The Center employs a rigorous approach to stock assessments that aims to use a model appropriate for the available data. Within a hierarchy of models that extends from simple (e.g. surplus production) to complex (e.g. statistical catch-age), a lead stock assessment scientist integrates available data into the most appropriate model for the data. In general, a more complex model is used for stock assessments of species that are highly valued or heavily fished, and therefore, have more available data, than for species that are less valued or lightly fished, and therefore, have less available data.
        - Strengths – The overall approach is reasonable and should produce a stock assessment model that optimizes use of available data. Further, the overall approach should produce a stock assessment model that is state-of-the-art for the available data.
        - Challenges – The overall approach is reasonable, but lacks a formal written protocol to both guide model development and to communicate the standard approach to stakeholders. Consequently,

stock assessments have an appearance of "one off" uniqueness that may seem to be more art than science to naïve stakeholders or managers.
        - Recommendation – Develop a standard operating procedural manual that describes the overall approach, as a tool for guiding future stock assessments and for communicating the procedure to stakeholders and managers. Post the procedural manual on a publicly available website.
    - Classification of Stock Assessments – The Center is asked to produce stock assessments for a large number of species that range widely in available data from "data poor" to "data rich" species. In general, a more complex and integrated assessment can be derived for "data rich" species than for "data poor" species (as noted above). Further, the lead stock assessment scientist will likely be drawn to apply the most data intensive (i.e. complex) model to each species, because a more complex model would hopefully produce a more convincing estimate of stock status.
        - Strengths – The overall approach described above would hopefully produce a stock assessment model that is most appropriate for the available data, as described above. Further, the overall approach would generally maximize, rather than minimize, use of available data by seeking to use the most complex model possible for available data.
        - Challenges – The desire to maximize use of available data, and thereby, to produce a more convincing stock assessment using the most complex model possible fails to recognize that some fisheries may be adequately assessed with models of low complexity.
        - Recommendation – A decision rule should be developed that anticipates management need in addition to data availability, when selecting the most appropriate stock assessment model. Some fisheries can be adequately managed using simple stock assessment models that lead to simple fishery management rules, thereby foregoing unnecessarily complex data compilation and assessment model development.

- **Theme 2: Assessment Process**
    - Overall Approach – To ensure stock assessments are of good quality, reliable, and relevant to management needs, stock assessments are prioritized for inclusion in the Center's work plan using a coordinated process (SEDAR or ICCAT). The process is designed to be inclusive by allowing participation of all relevant interest groups and also transparent by posting all records in a publicly available internet location. The Center identifies the number of "slots" (approximately equivalent to a lead stock assessment scientist) that are available for leading the development of stock assessments, which are then assigned in priority order to requested stock assessments.
        - Strengths – The process seems likely to designate the highest priority needs for stock assessments to the species that are most in need of management advice, thereby ensuring relevancy of effort spent by the Center on stock assessments.
        - Challenges – The process allows managers to designate the level of a requested stock assessment (i.e. "benchmark" or "update") that differ greatly in the effort required for completion of the stock assessment by a lead scientist and supporting effort.
        - Recommendation – The process could be revised to match the level of stock assessment effort requested to the amount of time available for stock assessment biologists in the Center. This will require stock assessments to be defined in terms of hours required (e.g. "benchmark" = 1000 hours; "update" = 500 hours) and lead stock assessment "slots" to be replaced by "hours" or some other measure of lead stock assessment scientist time available in each year.
    - Priority Setting – Cooperators request stock assessments of species that are deemed to be most in need of management advice. Therefore, species that are prioritized to be most in need of stock assessments are, not surprisingly, the species that are most highly valued, most controversial, or potentially over-fished.
        - Strengths – The process assures that stock assessments are completed for species that are in need of management advice, or for species for which the original stock assessment was questionable or contentious with stakeholders.
        - Challenges – For some high-profile or contentious species, "benchmark" stock assessments may be requested in the hope of obtaining a different answer, because the original "benchmark" stock assessment produced unpopular (usually, overly restrictive) management advice.
        - Recommendation – The process could be revised to only allow benchmark stock assessments for species for which a stock assessment was not previously completed. This would thwart attempts by unhappy stakeholders to "shop" for a better answer.

- **Theme 3: Peer Review Process**

- Overall Process – The stock assessment process is subjected to rigorous peer review at all stages, as required by the national standard. The process meets the national standard for being transparent, inclusive, unbiased, independent, and not duplicative.
    - Strengths – Peer review of the stock assessment process ensures scientific quality of products, as for many other science processes, such as scientific publications, research funding programs, and scientist tenure and promotion. Peer review is widely viewed by stakeholders and cooperators as a standard approach to ensure integrity of scientific processes and products.
    - Challenges – The requirement for peer review adds time to any process and ensures rigorous application of the scientific method, but also ensures "benchmark" stock assessments cost much more than "update" stock assessments. Further, standard methods used for multiple assessments are peer reviewed for each assessment.
    - Recommendation – Reduce the number of "benchmark" stock assessments by shifting recurring assessments to an "update" of the previous assessment (see Theme 2). In addition, develop a standard operating protocol (SOP) for stock assessments (see Theme 1) that can be peer reviewed once, but not for each stock assessment that relies on the SOP thereafter.
- Sequence – Independent peers review stock assessment products after the Stock Assessment Workshop (which follows the Data Workshop) and the assessment has completed.
    - Strengths – Placement of peer review after the stock assessment has been developed should ensure that stock assessment products are of high scientific quality.
    - Challenges – Placement of peer review after the stock assessment has been developed prevents peer review from "advising" the peer review process until after, rather than before, extraordinary effort has been expended.
    - Recommendation – Consider moving peer review ahead of the stock assessment workshop, to advise stock assessment development toward a model that is most appropriate for available data. This would provide support for an assessment that aims for the most parsimonious approach, thereby reducing effort spent toward a more complicated approach than is either possible, based on available data, or necessary, based on management need.

- **Theme 4: Communication**
    - Stock Assessment Documentation – To achieve transparency, stock assessments are documented in great detail, including detailed recording of workshops, data streams, and stock assessment model structure. The number and extent of documents has grown significantly over time.
        - Strengths – The level of detail documenting each stock assessment ensures a level of detail that serves as a permanent record of all details related to the stock assessment.
        - Challenges – The level of detail documenting each stock assessment may not be necessary for data or methods in common to multiple assessments. Further, the level of detail documenting each peer review is challenging to peer reviewers, who are asked to review assessments that are described in extraordinarily large numbers of documents and pages.
        - Recommendation – Use of standard methods described in standard operating protocols (see Theme 1) would allow for fewer documents describing individual stock assessments.
    - Stock Assessment Complexity – To provide an appropriate level of detail for peer review, stock assessments must include enough detail to enable the assessment to be reproduced (i.e. reproducibility standard). The number and extent of documents is substantial, to serve as a reference for peer review.
        - Strengths – The level of detail documenting each stock assessment ensures a level of detail that can be reviewed by peers.
        - Challenges – Extensive details and complexity of stock assessments may not be understandable for all stakeholders and cooperators.
        - Recommendation – A simplified summary of each stock assessment would enable communication of key findings to stakeholders and cooperators.

- **Theme 5: Research Opportunities**
    - Workload Allocation – Each lead scientist is treated as a "slot" when planning and prioritizing assessments within the Center's work plan. This system enables scheduling of stock assessments to high priority species over species of lower priority (see Theme 2).
        - Strengths – The system for assigning lead scientists to assessments allows the highest priority assessments to be incorporated into the Center's work plan. Further, the Center has a remarkable record of completing assessments to support management council cooperators.

        - Challenges – Development of stock assessments consumes nearly all available time of scientists in the Center. Therefore, devotion of nearly all available time of lead scientists to stock assessments affords little time for research. A strong and recurring signal from participants of this review was related to this single issue!
        - Recommendation – The Center Director could define a "slot" to include a smaller fraction of a lead stock assessment time (< 1.0 FTE), to reserve a fraction of each scientist's time for research. The resulting effort toward research would enable innovation and improvement in stock assessments. To enable conversion of "slots" into time, time spent on assessments must be quantified.

- **Theme 6: Ecosystem Considerations and Next-Generation Assessments**
    - Ecosystem Approach – Incorporation of environmental variables has explained some of the previously unexplained variation in fishery harvest, recruitment variation, and distribution of some species of fish in the Gulf of Mexico. Clearly, such measured environmental variation can be used to understand variation in fishery parameters, but may also be useful for predicting future fishery states when the environmental variation follows a predictable pattern (e.g. decadal oscillations).
        - Strengths – Stock assessment models could be improved by incorporating environmental variables by increasing precision of stock assessments and thereby reducing uncertainty of management advice and associated risk of overfishing based on that management advice.
        - Challenges – Incorporating environmental variation in stock assessment models increases model complexity, which increases time to maintain data streams supporting stock assessment models (may compete with a need for less time spent on stock assessments in favor of more time needed for research). Unfortunately, randomly varying environmental variables cannot improve model predictions in the absence of a predictable linear or nonlinear pattern of variation in the variable.
        - Recommendations – Wherever possible, environmental variables should be tested as correlates of apparent changes in the temporal pattern of catchability or recruitment. Large-scale, long-term trends in ocean currents may be especially useful for both correlating to trends in catchability and recruitment, and by following a predictable nonlinear pattern of change through time.

- **Theme 7: Organization, Priorities, and Accomplishments**
    - Workload Balance – The number of stock assessments needed by cooperators is matched to the number of scientists available to lead assessments through a process of prioritization that ensures high-priority stock assessments are included in annual work-load planning at the Center. The number of assessments is not expected to decline in the future, according to the Regional Manager. Further, the number of lead scientists is not expected to grow beyond a few vacancies that are presently in the process of being filled.
        - Strengths – The process used to assign assessments to lead scientists would seem to ensure that the highest priority assessments are completed. Therefore, cooperators would seem to be served well by the present process.
        - Challenges – The process used to assign assessments to lead scientists does not seem to allow lead scientists time in their annual work plan for professional development, which leads to frustration, and ultimately, attrition and turnover.
        - Recommendations – The process used to assign assessments to lead scientists must explicitly reserve time for professional development in annual work-force planning. Continuation of the present process will perpetuate job frustration, and thereby perpetuate attrition and turnover.

Conclusions – The stock assessment program at the SEFSC benefits from a highly skilled and extraordinarily dedicated workforce whose throughput of stock assessments serves fishery managers and stakeholders very well across a broad region of ocean resources. Throughput of stock assessments would benefit from streamlining the process by: (1) shifting more assessments from benchmarks to updates; (2) standardizing methods that overlap among assessments; and (3) shortening stock assessment reports by using standard operating protocols. Workforce management would benefit from: (1) treating lead stock assessment scientists in units of hours, rather than slots; and (2) reserving a portion of each lead scientist's FTE for career development (publishing scholarly articles and attending conferences). Communication of stock assessment findings would benefit from: (1) summarizing key changes or innovations at the front end of documentation; and (2) adding a concise layperson summary to each assessment.

Reviewer Report on Program Review of Stock Assessment Process
NOAA–NMFS Southeast Fisheries Science Center
Miami, Florida
8–10 July 2014
Background

General Observations and Recommendations: The Stock Assessments that are performed, are in response to both the Magnuson-Stevens Sustainable Fisheries Act and the Atlantic Tuna Conservation Act. These assessments are extremely intensive, and are difficult to perform due to the limited data that is available. I am extremely impressed with the leadership provided by the SEFSC Center Director Dr. Bonnie Ponwith, and her division directors and branch chiefs, as well as by the dedication and motivation of the staff scientists at SEFSC. All of the above personnel are highly motivated to preserve the important fisheries that they are responsible for, while minimizing the impact on commercial and recreational fishermen. The methodologies that they apply are state of the art and the work involved is extensive.

The presentations given by stock assessment scientists clearly demonstrate a mastery of the biological and statistical knowledge required to conduct or contribute to the required Data, Assessment, and Review phases of the SEDAR process for stock assessments. Every presenter clearly communicated a strong commitment to transparency in all aspects of data analysis and conclusions. I also commend the presenters for reflecting on the current SEDAR process and identifying areas that would benefit from evaluation and restructuring, and for proposing specific suggestions for improvement. This demonstrates a continued effort to improve and refine a process to produce the best possible science to inform fisheries management decisions, and personal dedication as public servants.

My most important recommendations are to streamline the process wherever possible, go to the Proposed National Approach to set priorities for assessments, invest in people, and to perform research in combination with partners to further validate and improve their models and to improve the representation of uncertainties in the assessments.

Key (Specific) Findings and Recommendations (as reviewer has comments on)

- **Scientific and technical approaches**
    - Observations
        - Strengths: The approaches reflect the current state of the art.
        - Challenges: It is very complex and time consuming.
    - Recommendations to address issue: Streamline where possible.

- **Assessment process**
  - Observations
    - Strengths: Utilizes the current state of the art and is very comprehensive.

    - Challenges: The SEDAR process that determines the number of stock assessments and updates targeted for each year, and for prioritizing targeted species, results in a stock assessment workforce that is overwhelmed and unable to dedicate time to researching new models and understanding to advance stock assessment science. This also leads to reduced workforce morale and higher than expected staff turnover. The current SEDAR process requires a considerably greater number of benchmark assessments versus update assessments and is another factor contributing to an overwhelmed workforce. The benchmark assessments are appropriately more time consuming, requiring a greater investment of staff time and effort. The number of workshop webinars is excessive.

  - Recommendations to address issue: Reduce and possibly limit both the number and duration of workshop webinars conducted during the assessment process. I recommend that the Center reconsider methods used to determine how many "slots" are to begin a new assessment each year. Creating specific, planned opportunities for staff to alternate between months dedicated to an assessment, and months available for research could yield a more productive workforce and greater staff retention. The SEDAR process should consider an approach to prioritize when certain stock assessments need to be updated. SEDAR also needs to clearly and carefully define thresholds that would trigger a benchmark assessment versus an update. Consideration should be given to types of new indices, understanding, or model updates that would be considered a significant change and thus provide compelling reason to dedicate the increased resources necessary to conduct a benchmark assessment, as opposed to an update. The assessment process could potentially be improved by testing and validating fisheries models using an approach similar to that used for Observing System Simulation Experiments (OSSEs). This could provide a means to test various models used to estimate future fisheries stock populations and demonstrate their relative uncertainties in a controlled setting. This could increase the relative confidence in the resulting model projections for stock populations and predicted changes. I recommend that the SEFSC explore the possibility for doing this with NOAA AOML.

- **Peer review process**
  - Observations
    - Strengths: Very comprehensive.
    - Challenges: Too demanding on available reviewers.
  - Recommendations to address issue: Streamlining the documents would make this less labor intensive.
- **Communication**
  - Observations
    - Strengths: Well documented, transparent, publicly accessible.
    - Challenges: While the SEDAR process clearly meets high standards of transparency, the final stages of the assessment process falls short in not effectively or clearly communicating a summary of the assessment for stakeholders. Leaving lay audiences and non-scientific stakeholders to dig through hundreds of pages of analysis likely leads to greater confusion and distrust of results and a lengthy process that is not well understood. Effective communication and stakeholder buy-in of assessment results is vital and should be considered just as essential to the other more technical components of the assessment process.
  - Recommendations to address issue: Streamline assessment documents. Focus on rationale and basis for critical decisions. Include an executive summary for each assessment. Prepare a simplified plain language summary for stakeholders and the general public. This should be prepared by or in conjunction with a communication specialist. Explore increased cooperation with Sea Grant. Look for ways to communicate and increase awareness of the stock assessment process in the stakeholder community, leveraging or continuing the successful MREP approach. Look for opportunities to communicate the success of assessment-driven management decisions to stakeholders, for example where a stock was effectively rebuilt or sustained.
- **Research opportunities**
  - Observations
    - Strengths: Development of improved models and indices, biophysical modeling, and incorporation of environmental data and satellite observations are major strengths.

        - Challenges: Additional research is needed to improve assessment process and policy advice. The current assessment load is too large and there is not enough time for research. This affects morale, retention and recruitment adversely.

    - Recommendations to address issue: Current level of research needs to be expanded, through options such as hiring of more personnel, streamlining the stock assessment process, and increased partnering. Research should be considered a priority that is not diminished if the demand for additional stock assessments increases. Requirements for promotion should be realigned to reflect the work being performed by stock assessment scientists. Avenues to enhance ownership of stock assessments should be explored through interactions with other NMFS centers and NOAA line offices.

- **Ecosystem considerations and next-generation assessments**
    - Observations
        - Strengths: Excellent progress through very effective collaborations. This activity is proactive and represents substantial forward thinking. It takes into account environmental factors that affect productivity and recruitment, and has potential to result in improved assessments, by explicitly accounting for physical effects. This contributes to an improved relationship with stakeholders.

        - Challenges: Resources are limited and must be borrowed. Historic data is missing.

    - Recommendations to address issue: Follow the GOM IEA 3 year plan, and continue to work closely with NOAA AOML and other partners. Development of an ecosystem based fisheries management plan should be the highest priority. An Ecosystem team if formed should be cross-line office. It appears that the addition of ecosystem indices have the potential to improve the accuracy of assessments, but consideration of trading-off or removing other indices should be considered so that the overall process is not further overwhelmed. (See comments on modeling for suggestions on how to evaluate which indices would be most important to include.)

- **Organization, priorities, and accomplishments**
    - Observation

        - Strengths: Outstanding leadership at all levels of SEFSC. Excellent partnerships. Stock status overall has improved.
        - Challenges: High turnover of staff. Diverse locations. Prioritizing assessments. Personnel (and monetary) costs substantial for each SEDAR assessment.

    - Recommendations to address issue: Adopt the national approach to determine level and frequency of assessment. Simplify assessments where possible. Reduce length and number of documents where possible. Increase investment in people, both staff and constituents. Travel of NMFS SEFSC scientists to scientific conferences should be encouraged and supported.

Conclusions: See General observations and recommendations (above)

Reviewer Report on Program Review of Stock Assessment Process
NOAA–NMFS Southeast Fisheries Science Center
Miami, Florida
8–10 July 2014
Background

General Observations and Recommendation

Key (Specific) Findings and Recommendations (as reviewer has comments on)

- **Scientific and technical approaches**
    - The approach taken to stock assessment by the South Eastern Fisheries Science Center appears to be thorough, investing considerable resource in ensuring that both data sources and modelling approaches are fit for purpose and as compliant with world class standards as possible.
    - The use of well documented and reviewed assessment software is appreciated in that it means that one, highly technical aspect of the review process is already covered. This does not mean, however, that the search for stock assessment methodologies is complete. The development of more integrated ecosystem approaches is important and should deliver alternative approaches to the field in the long term, however there is a pressing need to develop novel assessment techniques for the data limited situations which are prevalent in the SEFSC jurisdiction.
    - The range of potential data sources (and often their fragmentation) means that data compilation exercises are considerably more complex than in other fishery assessment arenas. I understand that the previous program review on data issues concentrated on this area and that there are moves to house a centralized database for several of the data streams which should some way to streamlining the whole assessment process.
    - The high level of recreational activity and the high discarding ratio means that estimates of current catches is quite uncertain and it is entirely appropriate that the assessment methodologies used do not treat catches as being exactly known (in the majority of cases). Avenues for increasing the precision of estimates for these two fishery components should continue to be actively explored.
    - In the absence of an absolute abundance index from survey series, the model estimations of absolute stock abundance will be scaled by the landings and hence any bias in the reporting of catches will influence estimates of abundance, reference points and subsequently the various catch limit levels. Should the magnitude of any bias change through time then this will have consequences for the perception of stock status. Reference points are typically influenced by long term stock dynamics and therefore will be less influenced by shifts in bias levels in the terminal years when compared to the stock biomass levels. Whilst this is true for any stock assessment the impact of such biases will be more acutely felt where the management system is actively seeking to achieve given biomass levels.
    - The use of abundance indices in stock assessment models as linear predictors of stock abundance has at its base the statistical assumption that every fish in the stock system has an equal probability of capture by the gear. This assumption is always violated but the degree to which this violation might be influential varies according to the capture methods employed. There was discussion regarding how changes in the behavior of fishers in relation to management action would affect the commercial catch rate indices and efforts are made by the SEFSC staff to account for such changes in their standardization approaches. There appears to be the implicit assumption that fishery independent data (i.e. scientifically designed survey) will be comparatively free from potential bias, however the behavioral traits of the target species will influence their catchability. In particular gears which rely upon attraction are likely to be influenced by behavioral interactions within and between species at the sample site. Pot/trap gears rely upon the behavioral response of individuals to enter the trap and the presence of other individuals can influence the choice as to whether to enter the trap and can experience saturation. Catch rates from hook and line/long-line fishery methods may also be influenced by the abundance of other species in the area by either out-competing at the hook, or by displacement from the immediate vicinity. The potential for such influences upon the survey indices should be explored.

- Age based assessment is usually seen as the "gold standard" for fisheries assessments as they offer the potential for the most accurate estimates of current stock size and exploitation rate. The data requirements for age-based assessment are vastly more than for simpler models and given the often complex and fragmented nature of the data inputs it is not clear that moving to an age-based platform will necessarily be more useful to management. Indeed even the current attempts to use age or length disaggregated assessment techniques may be stretching the limits of what is possible given the data quality. Discussions during the last session made it clear that the existing otolith collection program is at capacity and consequently there is no scope for an expansion of the routine age based program. With this in mind a critical evaluation should be undertaken to ensure that age or length disaggregated assessments are only conducted when the benefits of such an approach outweigh the potential risks given the data limitations. The SESFC should take all care to ensure that assessment terms of reference are in keeping with the available data.
- Some reference was made during the presentations to the issue of Natural Mortality (M) and the reviewer acknowledges the difficulty in arriving at accurate or realistic values of this parameter, but it is one of the cornerstones of the assessment process and can have a stronger influence on reference point estimation than the estimate of terminal population size. Reference points based upon the virgin stock status (e.g. spawner per recruit levels) are particularly susceptible to the choice of M, I have experience of stocks being classed as significantly under-exploited or significantly over-exploited depending upon which value of M is chosen from a range of published values! There are several ways of tackling this: 1) investing in research into values of M, 2) incorporating uncertainty in M into the assessment process 3) moving away from reference points based upon virgin stock states.
- Whilst there is evidently a significant body of work undertaken in respect to capturing uncertainty within the data and modelling processes, it is not clear if this fully feeds through to the final product (e.g. the small buffer between the OFL and ABC for GOM Gag Grouper).
- **Recommendations to address issues**
  - Ensure sufficient research time is available to staff for the encouragement of novel assessment technique creation.
  - Fishery independent surveys have the potential to offer indices of abundance with the least level of bias. These surveys should be defended from rationalization and expanded wherever possible to be as inclusive of stock range/species composition as possible.
  - Further research and analysis on the potential for behavioral interaction around survey gear to ensure the index is as appropriate as possible for use as a linear predictor of stock abundance.
  - Develop programs to obtain better estimates of natural mortality and incorporate uncertainty around this crucial factor in the stock assessment/reference point estimation process.

- **Assessment process**
  - The process of delivering stock assessment products to the Council SSCs was well laid out for the panel and it was clear that there are several issues with the process which cause the process to be sub-optimal for all parties.
  - The SEDAR program appears to have expanded beyond its original brief as the place for the thorough review of particularly difficult and potentially contentious stock assessments. The current practice of putting a majority of assessments through the SEDAR process would seem to be missing the point.
  - Potential for the number of annual stock assessments is limited by availability of staff. There are 248 stocks under FMPs of which we were informed that 107 could undergo some form of stock assessment. With 20 potential assessment leads this represents a formidable task as being able to conduct a meaningful assessment means that the assessment lead needs to understand and have recall on the intricacies of the data streams, biology and ecology of each stock they are assigned to. This means that each assessment will have a significant "spin-up" time whilst the assessor reacquaints themselves of the details.
  - There are a number of ways in which the system seems to be imploding under its own weight including how the assessments are commissioned and the steps taken to deliver assessments to the SSCs.
  - **Commissioning**.

- The high level of input by fishery manager into the scheduling and prioritisation of stocks assessments causes issues not only for the SEFSC but potentially for the managers as well. There has been an understandable tendency to focus on the key species for fishery managers, but this is open to misuse in that assessments could be repeatedly requested until one is found that delivers a desirable outcome for management subsequent to which an assessment may not be requested for many years in which time the true stock status may have changed substantially
- There would appear to be a desire for benchmark assessments over the use of update assessments, particularly on high-profile stocks. Benchmarks as currently undertaken represent a particularly burdensome task and the number of benchmarks requested is suffocating the system.
- A benchmark should deliver the most contemporary and appropriate methodology to assess a given stock and the frequency of new data streams appearing or advances in biological understanding is such that a further benchmark is unlikely to deliver a significant shift in assessment methodology within a short time frame. The use of update assessment procedures would appear to be more appropriate to deliver contemporary estimates of stock status unless there was a significant body of evidence to indicate that the previous benchmark process was no longer suitable.
- The purpose of benchmark processes in the ICES arena is to establish the data streams, determine biological parameter inputs and decide upon an assessment approach (i.e. create a recipe for an assessment). The recipe is then used to generate the next scheduled assessment. This process allows the optimal scientific approach to be developed independently from consequences of the management advice which would result. Benchmark products (the recipe) are then used for the subsequent assessments with only minor deviations from the process tolerated (e.g. missing years from a particular survey). More substantial revisions to a data series or change in model structure may require an "inter-benchmark" process in which scientists make specific proposals and analyses which are peer-reviewed before acceptance. Such "inter-benchmark" processes are often undertaken by correspondence. Finally where major revisions have occurred then a full benchmark process is scheduled but this would not typically happen within a 5 year window from the previous benchmark. The decision of which category any proposed amendments fall within is at the discretion of the professional staff within ICES.
- By sticking to a moderately ridged system of update assessments interspersed with periodic benchmarks, managers would benefit as they would be able to have more frequently revised stock statuses for a wider number of stocks (as the update process should be vastly quicker and easier for SESFC staff). SESFC staff would benefit from having fewer benchmarks to prepare for and be able to devote a more appropriate amount of time to research targeted at delivering improved scientific understanding for future benchmarks.

- **Conduct**
  - The division into three separate processes has a certain amount of logic to it, however it places a considerable constraint upon the way of working in that there is little or no opportunity to use common approaches to cover multiple stocks and explore mixed fishery interactions.
  - Having all meeting fully open to the public does satisfy the openness mandate of the SEDAR process but it does mean that meetings can become large and cumbersome. It is unlikely that stakeholders have interest throughout the full duration of the meetings and it could be beneficial to have a set portion of the meeting where public consultation is undertaken. Such sessions could be partly presentation of the work undertaken to date and part input from stakeholders.
  - The dispersed geography of the area does place significant limitations upon physical meetings and I understand the use of video-conferencing / webinars. From the description of the webinar program it would appear that these are often being scheduled too frequently and

for too long. They are also of questionable use when feedback and interaction from stakeholders is minimal, especially given that they require significant preparation time from SEFSC staff. The use of webinars should be carefully scheduled at key points to disseminate progress.

- Documentation of the process is important to demonstrate the science behind the process and the decisions necessary to move the process forward, however the current volume of documentation provided presents a significant challenge to the SEFSC staff and to those interested in the assessments. This topic is dealt with in the communications section of my report.

- The SEFSC have proposed a new approach to the SEDAR process which should deliver benefits to both managers and the SEFSC staff.
  - **Methods working group.** This group would be responsible for appraising stock assessment methodologies and data analysis tools. This has the great advantage that many stock assessments have similar issues and it save going over the same ground each time a new benchmark is commissioned. It also enables SESFC staff to see more across the spectrum rather than being narrowly focused on their own stocks and this can only help develop a more integrated approach to assessment and advice. The work of this group could significantly streamline the benchmark process.
  - **Stakeholder Advisory Panels.** This group would establish better routes for communication and dialogue between science and stakeholders which can only be a good thing!
  - **Reducing the frequency of Benchmark Assessments.** As discussed above, the role of the benchmark should only be to revise a process when there are significant developments in data or science.
  - **Moving the balance of review responsibility from CIE to SSC.** My comments on this are in the "Peer review" section of the report.
- **Recommendations to address issues**
  - The SESFC needs to be more assertive in setting out what is possible in terms of assessment workload. This could include setting a general rule for the minimum amount of time permissible for the conducting of benchmarks and a limit on the number of benchmarks commissioned per council per year. All other assessments would be updates and again these would be limited in number.
  - The proposed national approach to the prioritization and frequency of stock assessments looks to be an eminently sensible approach.
  - The adoption of the plans for the Methods Working Group and Stakeholder Advisory Panel is encouraged.
  - Provided that the Methods WG system is operational, stock-specific benchmark meetings would be condensed to a single point to undertake compilation of data and assessment model selection. Peer review is probably best undertaken at this stage.
  - To help identify the number of assessments possible under a revised SEDAR process the Center needs to have a better understanding of the true cost in terms of finance and person hours that assessments take.
  - Monitor staff time on assessments and the various tasks associated with them. Whilst I appreciate that collecting such data represents an additional administrative burden upon staff it needn't be onerous (5 minutes per week completing some form of on-line database sheet).

- **Peer review process**
  - Peer review serves two purposes
    - ensuring that the approach taken is the most appropriate and

        - Checking that the approach has been applied properly
    - It is not clear that the current review process is able to fully deliver on these two factors simultaneously. The sheer volume of data inputs and data working group deliberations means that it is difficult for CIE review to properly appraise whether the appropriate methodologies have been applied to the data. This would be vastly simplified if the methods were pre-approved by a Methods Working group.
    - In the current scheduling, the review being placed at the end of the process is really best placed to ensure that due process has been followed and that the selected methods have been applied correctly. Given the protracted period of time the current SEDAR process takes, it is simply too late for reviewers to identify fundamental flaws in assumptions and methodologies. It would be better to have independent reviewers be involved with the benchmarking process as it develops to offer advice and scrutiny. CIE input is probably best around the assessment group stage to bring in outside ideas. Final review (i.e. has the method been applied properly for the current assessment) could be the job of the SSC.
    - As stated previously the proposals for the creation of a Methods Working Group is sound. It should avoid duplication of methodology review by the benchmark groups. The benchmark groups can then be condensed to a single compilation and assessment meeting which will considerably aid the job of those tasked with reviewing the data and methodologies selected by the benchmark.
    - I am not wholly sold on the idea of using desk based CIE reviewers as additional inputs to the benchmark process. This would have to occur after the benchmark meeting had occurred and is likely to deliver less real benefit to the system than having them embedded within the benchmark meeting itself. As was commented in the discussions, there is only so much one can gain from reading papers and the face-to-face discussions of physical meetings are invaluable. It partly depends upon what benefit the SEFSC wish to gain from the use of CIE participants. I would contend that they could be used to gain additional input /alternative viewpoints to the process just as usefully as reviewers.
    - If the suggested changes are made to the SEDAR process then the majority of assessments conducted within any one year will be update assessments. The review process for this is considerably less than for a benchmark process and it would be entirely appropriate that the SSC undertake the role of ensuring that the methodologies defined by the benchmark process have been followed.
    - **Recommendations to address issue**
        - Provided the Methods Working Group system is adopted, peer review of the benchmark system will be greatly easier and therefore more likely to be constructive.
        - Embedding (truly) external reviewers/experts in a single benchmark meeting is likely to deliver the greatest benefit of their experience.
        - Final review of the assessment products be conducted by the SSC with a mandate to ensure that the processes identified as most appropriate by the benchmark have been correctly carried out.

- **Communication**
    - There is a universal need for improved communication between science and stakeholders, these issues to not only apply to SEFSC!
    - Transparency has two components, openness and inclusiveness. The current process of publishing all materials is open, but not inclusive as the contents are not accessible to all. To those less steeped in assessment science they are likely to be baffling and simply presenting those documents might cause some tension. To determine which documents are used by people it should be possible to use web statistics to see which documents have been downloaded and how many times however I suspect that the vast majority of document accesses will be by SEDAR staff and reviewers.
    - For benchmark assessments a summary document explaining key pieces of data / analysis may well be sufficient for the majority of stakeholders.

- There is a vast literature depository created by each benchmark in terms of research and working papers. It is acknowledged that this is a burdensome task, however it is important for future stock assessment scientists to have access to the research and findings of these exercises especially as a large proportion of these analyses never make it to peer reviewed publication. In order to make these working documents as useful as possible a brief lay-person description (i.e. executive summary) could be the first page, followed by description of data, methods and findings *sufficient* for future generations to follow.
- **Recommendations to address issue**
- The full scientific documentation is important for the scientists to follow what has been done, but the style and content could be significantly pared back unless the working document could be used as the basis for a peer-reviewed journal article.
- Short descriptions of stock assessments in plain language would be of help. This is probably best facilitated by the use of dedicated communication experts to help draft such documents.
- MREP is evidently working and should continue, including funding of fishers.
- A training program could be developed for scientists to assist them in delivering to a wide range of audiences. This is likely to be most beneficial if it combines some training from professional experts in verbal communication alongside the experiences of senior staff in dealing with the specific stakeholder groups unique to the SEFSC. Learning by trial and error is not ideal (I know from experience), one bad interaction with industry can take several good interactions to repair!
- Use the Stakeholder Advisory Panel to define what style and level of documentation is required to enable the widest possible access.
- Use the services of communication experts to assist the processes of public outreach in disseminating assessment and research results.

- **Research opportunities**
  - The current assessment system places considerable demands upon the SESFC staff. There appears to be a widespread feeling of despondency regarding control over the workload and scientific direction that individuals can take. Against this there is an obvious need for the Center to deliver products of sufficient quality to meet the needs of the fishery management councils however the ability of the Center to deliver is wholly dependent upon the availability of staff. There is a clear need to find the middle ground between Centre needs and fostering a motivated workforce.
  - Low morale was mentioned and seems to be a function of a) promotion prospects, b) lack of recognition for their efforts, c) being trapped in a closed cycle and d) divorced from the original interests which brought them into the field.
  - If the difficulties in conducting original research are widely appreciated outside the organization this could cause difficulty in attracting the optimal caliber of staff to existing vacancies.
  - The ability to conduct original research and produce peer-reviewed journal articles is seen as a pre-requisite to maintain their sense of being scientists rather than assessment technicians. However the ability to conduct and understand rigorous stock assessment should not be viewed as a purely technical exercise and does require a high level of training and insight. Mechanisms should be sought to recognize this and ensure that those involved in stock assessment know that their work is just as valuable to the Center as those focusing on research. For many people their feeling of self-worth in relation to their employment is inextricably linked to their title and position on the pay-scale so wherever possible, the demands of stock assessment should be taken into account when appraising staff for promotion although I realize there are Federal limits to this.
  - As mentioned in previous sections, an accurate picture of how long different stock assessment tasks take is vital to understand where the bottle-necks are and how tasks could be re-structured/reallocated to free up time for individual researchers.

- The disparity between the time allocated to "research" compared to what is considered desirable is not inconsiderate and the suggested value of ~20% of staff time that needs to be "found" is not unreasonable as a minimum average. This is not going to be an easy task. Moving towards the suggested frameworks of the Methods group, limited benchmarks and more updates should free up some time but it is not clear how much time this will potentially deliver.
- There is also a desire of the Management Councils to see more throughput of assessments which is at odds with the desire of scientific staff. Managing the expectations of the Councils is a key role for senior management. Again, an improved understanding the resource requirements for benchmark/update assessments will be a vital tool for the Centre management.
- The Centre staff are clearly highly professional and dedicated and have garnered a reputation for getting assessments out "whatever the cost". Whilst this is admirable it is almost certainly contributing to the current work-load. I also suspect that there is a tendency to try and "gold-plate" the outputs (as evidenced by the volume of writing in the working documents). Management should work with the scientific staff to determine when a product has reached an appropriate quality for the task in hand. As with so many comments and observations in this review, these issues are not unique to the Centre and are problems that panelists are tackling in their own institutes (mine included).
- Specifying set proportions of individual staff time for research would seem like an obvious first step. Getting people to stick to them will be another matter. For protracted assessment time-frames (e.g. benchmarks), other tasks are likely to get squeezed out unless staff are particularly disciplined. Having update assessments with a condensed time frame should allow a "clear run" for research efforts. Resolving these issues will require dialogue between scientists and managers. Adopting a more project-based work system in which staff are allocated specific time windows to the various tasks may help. I acknowledge that within a scientific enquiry framework, projected time per task rarely meets reality but it could help people be more rationale in deciding when a task is sufficiently complete.
- **Recommendations to address issue**
  - Determine the staff time required for assessment purposes making sure that Fishery Management Councils understand how much time they can allocate in any one year.
  - Delineate a proportion of staff time to research and tasks other than assessment with the management structure to follow it up.
  - Determine approaches for appropriate recognition of assessment work.

- **Ecosystem considerations and next-generation assessments**
  - Setting appropriate harvest levels across the board for long term sustainability is the ultimate, aspiration of fishery managers, but the short term decisions of setting the next year's quota often over-ride the longer term look.
  - Developing ecosystem based modelling approaches for fishery management is an important task but is very much a long term process and there will be considerable challenges in maintaining momentum and demonstrating utility to funding bodies and managers. The interjections to assessments already made by some of the research strands will help maintain the interest but such efforts need to continue.
  - Staff involved in the fisheries assessment part of ecosystem approaches should not be divorced from the current reality of single species approaches and all avenues for cross-over should be sought. As a minimum, I would suggest that ecosystem based researchers should have exposure to the full working of at least one assessment round, preferably a benchmark.
  - The development of the Ecosystem Approach will by definition require the integration of many different skill sets and there are obvious challenges to combining the expertise whilst allowing staff to continue to develop. My personal view point is that it would seem most appropriate to have an umbrella group of scientists sitting in different specialisms who can pool together to combine their skills. Such a group would likely require a small core of facilitators who can manage the various experts. The alternative of a new Ecosystem Group would be more prone to "silo-mentality" (from

both within the team and from those teams outside) and ultimately have a less productive outreach across the specialisms. Ultimately it is a management decision about how best to structure the work to balance budgets and needs, there is no perfect model for this!

- Irrespective of how staffing organization is developed, the Ecosystem Approach to fisheries is a big program which requires big thinkers, lots of data and significant resourcing. It is unlikely that "salami-slicing" budget from current areas will deliver the kind of resources required for this long term venture so new sources of funding should be sought. It also requires that those involved with an Ecosystem Approach program (and I would expect this to be a considerable number) have the time to make significant contributions. This in turn means that the assessment burden would be required to be spread amongst a wider pool of scientists.
- Ultimately the Ecosystem Approach will necessitate a shift in thinking and approach of fishery managers and scientists need to be continually (but subtly) appraising fishery managers of how approaches and developments are helping inform their decisions.

- **Organization, priorities, and acccomplishments**
  - Many of my points regarding the actions of the Center in relation to current demands have already been addressed in the above sections.
  - The Center is blessed with a hardworking and dedicated staff who continually rise to the challenge of meeting the large workload, however this cannot be sustained in the face of increased assessment demands and the stress is evident.
  - With 248 stocks under FMPs, 51 of which have been assessed over the past 10 years (some of the multiple times) and 56 further potential assessments, the current system is clearly inadequate for the production assessments as demanded by the Fishery Management Council in response to legislation.
  - With the filling of current vacancies, the projected assessment throughput can only increase modestly under the current system.
  - Even with a more streamlined system, once the balance of research & development requirements has been redressed, the assessment potential of the Center will be limited compared to the potential demand.
  - The main restrictions to further assessment throughput have been identified by the Center, (data availability, over-use of the SEDAR benchmark process, scheduling control and staff availability). The suggestions of the Center staff, along with further inputs from the review panel should, if promptly and sufficiently enacted, help to redress the balance of capacity vs expectance.

Reviewer Report on Program Review of Stock Assessment Process
NOAA–NMFS Southeast Fisheries Science Center
Miami, Florida
8–10 July 2014

Background: Panelists were tasked with an evaluation of the SEFSC's fishery stock assessment program and were asked to consider modeling approaches, the review process, uncertainty of results, communication of advice, usefulness to managers and explicit consideration of environmental factors. Evaluations were based on reference material made available prior to the review and presentations made by SEFSC staff and colleagues from the Councils and other participants in the management process. My comments below are not based on an up-to-date knowledge of stock assessment or NMFS policies and procedures. They are my attempt to suggest some alternatives that might improve the quality or timeliness of assessments and management targets. I did not restrict my comments based on what I thought might be feasible with regard to budgets, policies, or politics.

General Observations and Recommendation: The SEFSC has a strong team of assessment scientists, using up-to-date methods and developing new approaches to meet management needs and in attempts to account for data limitations. Other Center scientists (in programs not reviewed here) provide underlying biological data, monitor commercial and recreational fisheries, and generate fishery-independent data that are essential for reliable assessment. As detailed below, the primary issues related to stock assessment are: (1) poor data quality for many species, (2) assessment models that are not always consistent with the type and quality of data available, (3) management needs that are not consistent with the type and quality of data available, and (4) assessment workloads that provide very limited time for research or publishing.

Key Findings and Recommendations

- **Scientific and technical approaches**

  Does the Center apply a suitable scientific/technical approach to fishery stock assessment modeling? What is the suitability of the stock assessment models employed, taking into account the constraints imposed by the available data?

  Center scientists use an array of analytical approaches, from data-poor methods based on landings or size data only to data-intensive methods such as a catch-at-age model requiring landings, size and age composition, as well as survey indices. Scientists consider the available data when selecting models but in some cases the management process provides an underlying pressure to use models more complex than warranted by the data, in order to provide stock status or fishing mortality targets/levels. The Magnuson-Stevens Fisheries Conservation and Management Act (MSA; see National Standards) requires estimates of acceptable biological catch (ABC) and Annual Catch Limit (ACL). The difference between ABC and ACL is supposed to account for the risk of overfishing, which depends on uncertainty, which is difficult to estimate in general but especially for data poor stocks. If fairly characterized for data-poor stocks, there would be a large difference between ABC and ACL. These reference points and measures of uncertainty may be reasonable for data-rich stocks and can clearly protect against overfishing, but are poorly suited to many of the southeastern stocks (especially those in the Caribbean). Very simple assessment and management options are needed for these data-poor or data-absent situations.

  Center scientists are familiar with the full spectrum of assessment models in use in the U.S. and in some cases develop new approaches. However, because of the constant demand for assessments, there appears to be very little time for this type of research or to do simulation studies to assess the performance of existing methods. This might be less critical in other regions with strong data (e.g. long set of catch-at-age and survey data, for standard utilitization of a catch-at-age model) but are very important here since Center staff are trying to address management needs with

marginal or inadequate data. Providing time for research (developing methods for data-poor fisheries) would help to address Council needs but would also be a welcome break from the assessment treadmill. This is important to support career growth and to reduce burnout.

It may not be feasible given national requirements related to the MSA and overfishing standards, but it would seem appropriate to have alternative management needs and strategies for data-poor situations. The current system forces decisions that do not appear to be scientifically sound; for example, using length data and a whole set of questionable assumptions to generate Z estimates, then subtracting an assumed M to "estimate" fishing mortality. This assessment model should be evaluated through simulation, using realistic assumptions (i.e., that recruitment is not constant, that growth does vary from year to year). If performance of that model degrades under realistic assumptions about the species and the quality of the data, then other simpler management approaches will be needed.

One strong recommendation would be to put increased resources into developing fishery-independent survey methods, especially for the Caribbean stocks. For these stocks where fishery monitoring is difficult or infeasible, assessment and management should be based on surveys. It would seem desirable to develop habitat-based (stratified) surveys and to focus on estimating absolute abundance. There was brief mention of a habitat-based approach in a trap survey (Puerto Rico?). Such approaches could provide a strong, immediate basis for evaluating the impact of fishing. Even for data-moderate or data-rich stocks in the South Atlantic and Gulf of Mexico, reliance on "adjusted" data to obtain fishery-dependent indices seems ill-advised. It is difficult to measure effective fishing effort under the best of circumstances, and in a system with multiple regulation changes, it seems likely to produce biases that will be unknown in magnitude and direction.

In summary, the problem here is not in the assessment models or capabilities of the scientists, but strictly data quality and quantity. It cannot be solved by newer, more complicated models, but rather by obtaining consistent, long-term funding for fishery-independent surveys. Adequate survey data will improve the reliability of assessments and eliminate the need to generate fishery-dependent indices of unknown reliability.

- **Assessment process**

What is the efficacy of the assessment process regarding clarity of terms of reference, transparency to stakeholders, throughput, documentation, and reproducibility of results? What is the efficacy of the assessment process from an SSC perspective?

The SEDAR process provides a detailed, well-thought-out structure for stock assessments. There is plenty of opportunity for stakeholder involvement, although one concern is that the process requires too much time for effective participation of stakeholders or SSC members. A more streamlined process with a shorter primary assessment document would help. Details about the assessment can be put in appendices or covered through standardization of methods.

One recommendation would be to include more information about alternative models. This seems particularly useful for data-poor situations where model choice can substantially affect the results. It was mentioned that the stock synthesis model can be used to provide model fits of increasing complexity, from an age-structured surplus production model (with only two parameters) to more complex models as additional data sources are added. The change in results and uncertainty with additional data (and added model complexity) could be a routine part of stock assessments, rather than immediately moving to the most complex model supported by the data (or going beyond that, in some cases). There was also discussion about how to compare among models. This is a difficult topic and a useful area for staff research time.

Another recommendation is to carry forward the full level of uncertainty in stock assessments. This is difficult to do because many assumptions and decisions are made during the process. For example, when the natural mortality rates by age are treated as known constants based

on a Lorenzen curve, the assessment results become conditional on that (strong) assumption. When data are summarized outside the assessment model (e.g., converting length to age), that reduces the apparent uncertainty of the assessment results. Given the current emphasis on probability-based decision making, it is important to provide a realistic picture of the true level of uncertainty. If that causes assessment results to be disregarded, then that provides guidance for the future in terms of either reducing uncertainty (getting better data) or changing the assessment and management process for that species (e.g. going to a simpler management approach that does not require precise assessment results). It was mentioned in one presentation that the GOM SSC was investigating simplified control rules. That seems like a very productive direction, given that many assessments rely on marginal data and uncertainty seems to always be underrepresented.

It would also be useful to carry out an analysis similar to that described for PFMC assessments by Ralston. That analysis was a look back at assessment parameters, and how much they varied among assessments (as a better measure of uncertainty compared to internal measures). This would provide a firmer basis for adjustments between ABC and ACL and would be informative to the SSC and Council about the stability and uncertainty of assessment results. It should be done for stocks in all categories (data-rich, data-moderate, data-poor) because the cases considered by Ralston were probably all in the data-rich category. If it is difficult to find a common parameter among past assessments, this could be something to standardize moving forward. For example, for South Atlantic (SA) red snapper, it could be a standard policy to always present table of mean Fs for ages 2-5, and compare estimates 5+ years prior from the series of assessments. This is similar to a retrospective analysis (although done one year at a time) but would include more of the changes in model structure, software or assumptions that can vary among assessments.

On a related note, one difficult issue is that we never know the right answer. Unlike meteorologists who are proven right or wrong every day, we never know the level of error. Are there ways to validate assessment models? A retrospective analysis is somewhat useful, but that is mostly a measure of internal consistency. One possible way would be to carry out a tagging study and predict the number of tags that would be returned by each fishery sector in the year ahead. Including that information in an assessment provides a clear measure of performance, which is currently lacking.

On a related note, some of these species seem highly amenable to a tagging study. My experience has been that tagging studies of freshwater, estuarine and (accessible) marine species can provide detailed information about fishing and natural mortality. Our recent tagging studies have shown, for multiple species, strongly seasonal fishing and natural mortality. These estimates of M are stock- and year-specific, rather than relying on meta-analyses (of estimates of unknown quality) that provide proxies that are a simple function of fish size or other life history features. In some cases, we have contracted with guides to distribute tags over the full range of the stock. In all cases, we use high-reward tags and double-tagging to address non-reporting and tag loss. Tagging studies provide immediate results, and do not depend on landings or tenuous analyses of fishing effort. Tagging should be quite feasible for some of the shallow reef fishes or sharks and would provide direct estimates of F and M. Telemetry methods have also been used for some east coast and Gulf migratory species. The current generation of transmitters lasts for 2+ years for larger fish, and a series of releases can be used to obtain precise seasonal information on total mortality (see Rudd et al. 2014 as an example).

- **Peer review process**
  What is the adequacy of the Center and SEDAR assessment peer review process, taking into consideration the participation of other entities such as the Council's Scientific and Statistical Committees?

  Stock assessments produced by Center scientists receive substantial review. In addition to the SEDAR Review Workshop that includes outside experts (CIE), assessments are reviewed by the

SSC. Presentations indicated that a shift was underway to have an increasing role of the SSC in reviews. An advantage of that shift is that SSC reviewers are very familiar with the data limitations and management needs of that Council. One concern about the current process is that decisions made at the Assessment Workshop are revisited and frequently changed at the Review Workshop or during SSC review. There are many subjective decisions made in an assessment, and every knowledgeable reviewer might have made a slightly different set of decisions. To keep the process moving forward without excessive second-guessing, it might help to develop a standard protocol for reviewing versus changing assessments. For example, the protocol could outline characteristics of a suitable base model versus sensitivity runs, and which things might warrant new runs in a Review Workshop versus things to be examined in the next assessment cycle. There is little point in spending a lot of time choosing (and writing up) a base model and alternate sensitivity runs in an Assessment Workshop if those decisions are routinely overturned at subsequent reviews.

- **Communication**
  Does the assessment programs adequately communicate to the councils, state commissions, and headquarters their methods and results? Does the assessment program adequately communicate to NMFS headquarters its research and needs?

  Center scientists seem to do a reasonable job in presenting results to councils and commissions, and documents are available to the public through the SEDAR website. As noted in the presentations, one thing that is not well documented (or at least easy to find) is the set of decisions or calculations that take the SEDAR assessment result and produce management targets (e.g. ABC or ACL). One obvious area for improvement would be to provide a short summary of each assessment (on the SEDAR site and that Council web site). For example, the GOM gag stock assessment report is 609 pages, but there is no Executive Summary. A one page summary for each assessed species on a web site would be useful as well as as a summary at the start of the SEDAR report.

- **Research opportunities**
  Are there opportunities for improving stock assessments and the stock assessment process? What are there avenues for improving stock assessments and the stock assessment process?

  Center scientists have made some significant contributions within the area of stock assessment models and software (e.g. Prager's ASPIC). Some of these have resulted in software additions to the NOAA Fisheries Toolbox (NFT) and highly cited journal articles (e.g. Prager 1994 regarding ASPIC; 109 journal citations). Others are useful new developments that have not been added to the NFT or published (Beaufort Assessment Model, Williams and Shertzer), but should be if some research time can be made available. The BAM approach was not discussed in any detail, but it seems particularly noteworthy because it is a linked set of routines in R and ADMB for carrying out the assessment and producing report elements (e.g. tables, graphs) that would aid in putting together an assessment. Carving out time to make this widely available (within and beyond the Center) would be valuable for the stock assessment community as well as a good opportunity for professional growth. It also has the potential for streamlining the assessment process if adopted or mimicked Center-wide.

  Center researchers has also shown environmental factors that affect stock dynamics such as the work linking gag grouper mortality to red tides. These studies can reduce bias in stock assessments (e.g., compared to assuming constant natural mortality), increase precision, and increase our understanding of biological factors that regulate these populations (versus fishing impacts). Again these research contributions require a reallocation of staff time to research and publishing. These are not "ivory tower" studies that are unlinked to stock assessment, but applied studies that will improve the assessment and ultimately management.

The rate of publishing in the primary literature is low for PhD scientists but that is fairly typical for scientists within the stock assessment world (NOAA and otherwise). Nevertheless, having sufficient time for research, publishing and attending scientific meetings is important for career advancement and preventing burnout/excessive turnover. It may be helpful to map out the time required for benchmark and update assessments, and to sketch out staff availability for assessments. This would allow adminstrators to set aside (and protect) research time. For example, each assessment scientist might block out one month per year for research (not the same month!), and that could be factored into the planning for the number of assessments and their timing. The demand for assessments is unlimited, and there is no way for assessment scientists to "catch up", so adminstrators will need to plan for and protect that time.

- **Ecosystem considerations and next-generation assessments**
  How important are ecosystem considerations and next-generation assessments for improving the science used in management of managed fishery species in the southeastern United States?

The overview presentation showed several impressive examples of the potential for incorporating environmental factors as covariates. These are straightforward improvements to current assessment models, but will not often happen without a "set-aside" of analyst time for research. Other larger scale plans for multispecies complex or multisector models will require larger allocations of research time. These models may be useful for improving management, but only if the products are consistent with management needs (e.g. the detailed, fishery specific benchmarks that are being used).

Also mentioned within that presentation was increased use of management strategy evaluations. These have proven useful in other agencies and regions within NMFS, and can allow for better decision-making about types and intensity of monitoring (e.g. fishery-dependent vs independent), types of assessment models, harvest policies, etc. The difficult part is to develop a sufficiently realistic simulation of the whole process, in order to provide real insights about what has potential. This is again a research activity that will require time away from assessments, but should provide important insights about monitoring, assessment and management.

- **Organization, priorities, and acccomplishments**
  How well is the SEFSC organized to maximize stock assessment throughput and quality based on best available data for a given stock?

Evidence was presented that stock status, on average, is gradually improving, so current approaches are working. Changing stock status is a slow process that relies not only on the quality of the data and assessment but on managers to set appropriate regulations. It is good to examine those scores as a sort of feedback about how well the process is working.

One area for improvement is in setting assessment priorities. These are not internal Center decisions but nevertheless analyses done within the Center could help to influence the process. The proposed national standard appears to be a very useful template, and it would seem useful to move ahead and apply that to the set of candidate assessments facing the Center. Thinking through the various weighting factors and having a prioritized list would provide a stronger basis for decision making, or at least would clarify the extent to which current decisions are at odds with these (entirely reasonable) factors. For example, it would be valuable to see how assessment (as well as survey and fishery monitoring) effort compared to the magnitude of the various fisheries under Center purview. It would also seem valuable to develop a more formal policy regarding whether to do a benchmark assessment versus an update. This again is not a Center decision but it greatly affects workload and throughput. Currently there seems to be a poor understanding of what is to be gained from repeated benchmark assessments versus updates. There is also the appearance that calls for frequent assessments (of either type) are done in the hope of a different result, when in

reality not much changes from the addition of one or two years of additional data. Pushing at the Council level for a formal policy on assessment type could help in reducing the clearly excessive frequency of benchmark assessments.

Regarding time for research, administrators and assessment scientists could work together to identify a suitable fraction of time for research, then build that (and appropriate measures of research performance) into performance plans. If assessment scientists are going to be kept on the "assessment treadmill" without a set-aside of research time, then performance plans will need to reflect that, in order to be fairly judged on accomplishments.

One important future direction is to put much more effort into collecting more meaningful data. As mentioned above, new technologies such as acoustic telemetry or ROVs for absolute abundance estimation have the potential to greatly strenghen current assessments, or perhaps to replace current assessments based on traditional catch sampling that have little potential. The habitat-based trap study in the Caribbean and ROV work would seem to have very high potential for reef fish. These field studies also provide assessment scientists with opportunities to get on or in the water, gaining first-hand biological experience that would result in better assessments.

- **Conclusions**
  - Match assessment and management complexity to data quality. Spend time internally and with partners to categorize species based on a realistic assessment of data quantity and quality. This categorization should help in setting realistic expectations about assessment and management approaches
  - Invest in fishery-independent data, especially for data-poor situations
  - Better represent the true level of uncertainty in assessments and benchmarks
  - Find approaches for a true validation of assessment results
  - Consider alternative methods (e.g. tagging, visual surveys) that can replace or improve traditional catch-based assessments
  - Develop policies for assessment prioritization, review of assessments
  - Define and protect a suitable fraction of time for research

Reviewer Report on Program Review of Stock Assessment Process
NOAA–NMFS Southeast Fisheries Science Center
Miami, Florida
8–10 July 2014

Introduction

The Southeast Fisheries Science Center's (SEFSC's)2014 Program Review, focusing on the stock assessment process associated with meeting requirements under the Magnuson-Stevens Fishery Conservation and Management Act (MSA), took place in Miami, Florida, July 8-10, 2014. The review was well organized and thorough, and Center leadership and staff are to be commended for their preparation, organization, and thoroughness.

SEFSC is unique among the NOAA Fisheries Science Centers in the number and diversity of stock assessments that it carries out and in the number partners with whom it must work to provide assessment information and management advice. These include three regional fishery management councils (Southeast Atlantic, Caribbean, Gulf of Mexico), two interstate commissions (Atlantic, Gulf of Mexico), the NOAA Fisheries Highly Migratory Species Office (HMS) and the International Commission for the Conservation of Atlantic Tunas and (ICCAT). Meeting MSA requirements, in association with the Councils, has become especially challenging since the 2006 reauthorization of the act, which mandated setting of Annual Catch Limits (ACLs) for all managed stocks.

Even though it is faced with very substantial assessment demands, severe data limitations, and inadequate staffing in support of stock assessments, the Center is to be applauded for its ongoing accomplishments. The staff involved in conducting assessments are professional, highly qualified, and effective. Nevertheless it must be emphasized that too much is being asked of them and this is undermining their opportunities for research and professional development. These circumstances are already impacting morale and staff turnover and can be expected to undermine the integrity of the assessment process if not remedied.

To facilitate the regional stock assessment process, the SouthEast Data, Assessment, and Review process (SEDAR) was established in 2002. This enabled an effective partnership among the 8 entities involved in meeting MSA stock assessment requirements across the region. SEDAR has been very effective in establishing protocols and procedures, facilitating prioritization of stock assessments, and ensuring that assessments are adequately documented. However, over its 12-year history, it has not managed increasing demands very well, and procedural changes are necessary to improve efficiency, better constrain, focus and manage document production, and reduce burdens that are being placed on the stock assessment analysts.

While stock assessment methodology and the SEDAR process were emphasized during this review, attention was also directed towards the peer review process, opportunities for research, communication of assessment results, and ecosystem considerations. Observations and recommendations related to each of these topics are provided below.

- **Science and Technical Approach**

The panel received a thorough and comprehensive briefing regarding the scientific and technical approaches employed for assessing stocks and developing management advice.. Relevant presentations included the introductory general overview, background on the data used in the stock assessment process and related issues summarized from the 2013 Program Review, specific examples of stock assessments, and a very informative presentation of the entire process which included careful consideration of many of the data-related challenges, procedures for selecting and evaluating models used for assessment and provision of management advice, evaluation of model performance, and characterization of uncertainty. Together with documents provided to the reviewers, these characterized the breadth and depth of the stock assessment demands placed on the Center. In general, stock assessments performed throughout the Center appear to be carried out with a very high degree of professionalism, and methodology is appropriate with careful focus on important details and quality control. Methods selected for specific stocks are consistent with the biology of the stocks, fishery dynamics, and data limitations. Data limitations are of considerable concern for most assessments since long term, fishery dependent time series are lacking, and fishery dependent data are often inadequate or require careful, time consuming preprocessing which incorporate assumptions that may be difficult to substantiate. The overall lack of data necessitates the use of data poor methods for many stocks. For many fisheries, these data poor circumstances are likely to persist.

While the presentation of scientific and technical approaches on day two was thorough, it did not provide a basis for determining how well these approaches have actually been followed, on a stock-by-stock basis. However, the stock-specific assessment presentations on day one, which covered data rich and data poor stocks assessed under different authorities and from different regions with the Southeast, support thr perception that overall methodology is, in fact, appropriate.

Nevertheless, some overarching issues emerged during the presentations and ensuing discussion. These included concerns regarding characterization of uncertainty associated with different stages of the assessment process and propagation of uncertainty through the entire process, as well as the heavy emphasis placed on preprocessing fishery dependent data to develop CPUE based abundance indices. This reliance on fishery dependent data is necessary because useful time series of fishery independent data are almost entirely lacking. To quote directly from the presentation, these (landings reconstruction) approaches are time consuming and difficult to defend for many species although they can reduce possibilities for illogical assessment outcomes and stabilize estimation. Estimation of error also requires substantive and difficult to validate assumptions.

Even when survey time series are available, they are generally short and interruptions (missing years) can be problematic. The paucity of fishery independent data is a major limitation in the Center's stock assessment work.

Discard and bycatch data are also limited. Discard can be a major source of fishing mortality, and exceeds landings in some fisheries (especially recreational). Bycatch of small fish in the shrimp fisheries is also a major source of fishing mortality which has, traditionally, been very difficult to quantify due to very high levels of effort in the shrimp fisheries and low (or historically zero) levels of monitoring. In general, monitoring is inadequate and a high reliance is placed on self reporting even though this is known to be biased low. Furthermore, data on discard mortality are very limited. The relatively recent innovation of an electronic logbook which automatically logs, and transmits information on shrimp fishing effort has been highly successful and now facilitates effort-based bycatch estimation.

<u>Recommendations</u>

Even though assessment methodology does appear to be appropriate, consistency is lacking, especially among the different assessment groups within the Center. Many factors contribute to this inconsistency and it is certainly not appropriate to move towards a single assessment approach. The organizational structure, which maintains different reporting lines for the different assessment groups, likely contributes to this situation. Changes in organizational structure may be merited but occasional Center-wide assessment methods meetings may also be beneficial. Implementation of the stock synthesis approach for some GOM shrimp stocks has been successful even though it was challenging. This approach should be considered for the other shrimp stocks and implemented if possible.

Research should be directed towards improving methods for estimating uncertainty in each step of the assessment process and properly integrating uncertain within the overall assessment. Better characterization of assessment uncertainty is important for many reasons and will facilitate comprehensive analyses of the consequences of incorporating each data set, perhaps through management strategy evaluation. In particular, estimation of natural mortality and associated uncertainty is problematic for some stocks so research in this area should be prioritized.

Emphasis should be placed on improving estimates of discard and discard mortality. Broader implementation of the electronic logbook developed for the shrimp fisheries should be considered in this regard, together with other electronic reporting approaches. Working directly with commercial and recreational fishers to improve reporting is essential. Strategic use of observers to address specific bycatch/discard information needs should be considered. For example, by selecting target fisheries and bringing high levels of observer coverage to bear for relatively short time periods.

Data limitations are of serious concern in almost all fisheries. A concerted effort should be directed towards improving fishery independent data and moving away from the dependence on historic landings data and the associated preprocessing. Investments should include research to estimate survey catchability, and use of advanced observing technologies which facilitate absolute abundance estimation.

Some emphasis should be placed on improving size and age composition estimation while minimizing the need for reading otoliths and other aging structures. The Center is evidently engaged in this type of work and additional research would likely reduce sample processing costs and error associated with size and age composition estimation.

- **Assessment Process**

The SEDAR process was described, initially through an overview presentation on day one and in considerable detail on day two. This process has evolved over time to become a cooperative enterprise which involves the Southeast Fisheries Science Center, the Southeast Regional Office, the three regional fishery management councils (South Atlantic, Caribbean, Gulf); the two regional marine fisheries commissions (Atlantic and Gulf), and the NOAA Fisheries Highly Migratory Species Division. It is administered by the South Atlantic Fishery Management Council on behalf of the cooperators. Standard operating procedures for benchmark, standard and update assessments are documented and the roles and responsibilities of the participants are defined. The process works well in many ways but it is complex and labor intensive. This is to be expected given the number of cooperators (8), assessment teams (7), and management regimes. Transparency and thoroughness are emphasized but timeliness is not and this can be problematic. Since its inception in 2002, demands placed on the SEDAR process have increased and divergence in implementing SEDAR policies among cooperators has become problematic.

SEDAR provides an effective mechanism for prioritizing and scheduling stock assessments and defines procedures and requirements for three different levels of assessment (benchmark, standard, update). However, the process has become inefficient and administratively burdensome over time. While thoroughness and transparency have, appropriately, been emphasized, requirements for documentation and in-person meetings or webinars have grown to the point where it is no longer possible for many potential participants to engage in the process; thus transparency has been eroded. Many other procedural problems have also been identified, including unacceptable delays by some cooperators in providing data for assessments.

Demands for stock assessments have increased, especially following the establishment of ACL requirements in the 2006 reauthorization of MSA. Furthermore, cooperators tend to prefer benchmark assessments rather than standard or update assessments even when this is not merited and this places additional and sometimes unnecessary demands on the SEDAR process and the assessment teams. The SEDAR process requires the cooperator who proposes an assessments (generally a Fishery Management Council) to plays a lead role in drafting and approving assessment terms of reference. This sometimes results in unnecessary analytical demands being placed on the assessment team. Peer review requirements are clearly defined (see subsequent comments under "peer review" heading).

While one of the principles under which SEDAR was established was to ensure joint ownership of the process among the cooperators, stakeholders still tend to view the Center as the "owner" of assessments carried out by its staff, and negative feedback from unhappy stakeholders tends to be directed towards the Center. The SEDAR process does not apply to all stock assessments carried out in the region. In particular, different protocols apply to stocks assessed through ICCAT and this is appropriate given the nature of these stocks and their international governance. Shrimp assessments carried out by Center scientists are also not managed through SEDAR and it appears that some other exceptions apply, such as the recent SAFMC wreckfish assessment that was carried out by scientists retained by the fishing industry.

<u>Recommendations</u>

A thorough review and updating of SEDAR policies and procedures should be conducted every 10 years. Since it is now 12 years since SEDAR was established, the first review and update should be scheduled during 2014 or early 2015. The following issues and concerns should be considered:

- The need to ensure that data are provided to analysts in a timely manner, as agreed among the cooperators, should be emphasized
- Every effort should be made to allow the analysts to focus on analytical work and not be burdened by unnecessary documentation requirements or other procedural requirements
- Documentation requirements should be clearly defined with emphasis placed on brevity and clarity. Consistency of documentation across analytical teams should be strongly encouraged. A clear protocol should be established for documenting changes from previous assessments of the same stock.
- Executive summaries should be drafted for each assessment.
- Standardized formats for analytical documents and reports should be developed and enforced
- Procedures workshops should be continued
- The possibilities for third parties (academics, etc) to conduct SEDAR assessments under the same requirements that are placed on agency assessment teams should be considered.
- The Center should partner with the proposing cooperator in drafting and approving assessment TORs.

- All assessments other than those carried out under ICCAT protocols should fall within the purview of SEDAR. While this will increase the SEDAR workload, it will ensure consistency of peer review and documentation requirements as well as recognition of the resources required. Furthermore, stakeholders will perceive that a single set of protocols is applied uniformly.
- Criteria for prioritizing stock assessments, determining appropriate assessment frequency, and determining appropriate assessment level should be developed, documented and applied through SEDAR; these would likely be based on those currently being developed by NOAA. A policy for selecting the simplest assessment approach appropriate to meet management information needs should be embodied in this process.
- The idea of establishing a methods working group, and the potential for this WG to streamline the model selection process and reduce peer review requirements should be carefully considered.
- The peer review process, and the roles of the CIE and SSCs in this process should be reviewed and this review should consider alternatives to current practice. However, the CIE should continue to peer review benchmark assessments and desk reviews should be discouraged.
- Protocols should preclude SSCs from considering modeling changes during post review meetings and any assessment related analytical work carried out by or for the SSCs should be documented within the SEDAR archive

Three additional recommendations are offered:

- SEDAR should consider annual multispecies data workshops when appropriate
- Approaches for improving stakeholder involvement and trust should be implemented. The idea of a SEDAR advisory panel was raised during the review. This idea has merit but could be administratively burdensome. Including a SEDAR module within MREP may also be effective.
- Investment in software improvements should be made to streamline the SEDAR process, improve consistency and, potentially, to facilitate documentation.

- **Peer Review Process**

Peer review is essential to the stock assessment process. Peer review protocols and requirements are established under SEDAR policies and are, in part, discussed in the preceding section. However, this topic was also presented separately during the program review and will be addressed in this section of the reviewer's report.

The presenter asserted that a peer review process should be transparent, inclusive, unbiased, independent and not duplicative. These requirements are clearly articulated within the MSA National Standard 2 guidelines and implemented through the SEDAR process. The role of CIE as reviewers in the benchmark assessment process, and SSC as reviewers of assessment updates is consistent with this requirement.

Improving the efficiency of the overall assessment process and the peer review requirements was also discussed during this presentation. Certainly shifting towards more updates and fewer benchmarks would reduce overall assessment burdens as well as peer review requirements (or create "room" for additional assessments). Establishment of a SEDAR Methods Working Group could also reduce peer review requirements by eliminating duplicative reviews of the same methods, but establishment of this type of working group would also require drafting of clear guidelines for determining the point at which modifications to a previously-reviewed model require additional peer review. The idea of reducing the cost of CIE reviews by encouraging or requiring desk reviews seems unwise in most instances because discussion among CIE reviewers and between the reviewers and the analysts greatly enhances the peer review process.

- **Communication of Assessment Results and Data Needs**

The discussion under this theme focused on the effectiveness of communicating stock assessment results to the Councils and their SSCs, and to stakeholders and the public. Concerns regarding voluminous and inconsistent SEDAR documentation, and specific potential improvements are addressed in preceding sections. Consistency, brevity, clear documentation of changes from previous assessments, and the need for assessment executive summaries are of particular importance.

Specific technical communication is required by SSCs. This may require preparation of new reports and presentations describing assessments results. These are not currently archived on the SEDAR server but this should be remedied.

Assessment scientists generally brief the SSCs and communication appears to be satisfactory. Preparation of posters summarizing stock status and assessment results for display at council meetings may improve communication with members and attendees. Use of one-page summary stock status documents should also be considered.

Greater emphasis on communicating stock assessment methods and results to stakeholders should be encouraged. The Marine Resources Education Program (MREP) has been successful in this regard and the program should be continued and, if possible, expanded.

- **Research Opportunities**

Research to improve stock assessment methods and to improve stock assessments is an essential to the success of a stock assessment enterprise. This truism was articulated on several occasions during the review and was the focus of this theme. SEFSC has a long and very well respected history of noteworthy advances in this discipline and important contributions continue to appear in spite of the ever increasing demands for stock assessments. Analysis done more than a decade ago suggested that successful stock assessment scientists should spend about 30% of their time conducting research and about 40% of their time conducting stock assessments. Today, at SEFSC, much more time is spent conducting assessments, considerably less time is spend on research, and administrative requirements have increased to a substantial degree. Failure to invest sufficiently in this science undermines the ability of staff to improve their skills and bring innovation to their work; often prospects for advancement are enhanced through successful research and associated peer-review publications. Thus morale is also compromised. The integrity of the Center's scientific credentials is also eroded when this situation persists.

Steps have already been taken to leverage the Center's reduced capacity for stock assessment related research through collaborations with colleagues and students from other institutions and this has been effective to a limited degree. Since demands for stock assessments can only be expected to increase, the Center leadership must bring new resources to bear to support the stock assessment enterprise as well as improve efficiency as detailed elsewhere in this report. At SEFSC, as elsewhere in the NOAA Fisheries Science Enterprise, very difficult decisions must be made if to restore the balance between essential research and essential stock assessments.

- **Ecosystem Considerations and Next Generation Assessments**

Fish stocks do not exist in isolation, they are integral components of complex marine ecosystems and their dynamics are influenced, to varying degrees, by physical, chemical and biological conditions and by interactions with other species. The relative importance of fishing and environmental factors on stock dynamics has been and continues to be a topic of debate and active research. Under this theme, the panel was briefed regarding progress towards the Gulf of Mexico (GOM) Integrated Ecosystem Assessment, general approaches to merging ecosystem information into stock assessments, and the GOM Ecosystem Status Report. Three examples of research which has elucidated ecosystem influences on stock dynamics were also described; swordfish in the North Atlantic and Atlantic Multidecadal Oscillation, the influence of ocean currents on larval advection and resultant consequences for survival and recruitment, and red tide as an indicator for gag natural mortality.

The Center has made noteworthy progress in bringing ecosystem considerations to bear in stock assessments and in next generation assessments even though resources for this work are very limited. Most progress has occurred in the Gulf of Mexico and regional partners are beginning to recognize the importance of developing products that integrate ecosystem analyses into the SEDAR stock assessments.

Research and development of this capacity at SEFSC is strongly encouraged and ongoing interaction between stock assessment scientists and scientists involved in ecosystem research and assessment is essential to the success of this endeavor.

- **Organization, Priorities and Accomplishments**

Discussion under this theme focused on organization and staffing, accomplishments, and challenges and potential solutions. Information on the SEFSC organizational structure and the configuration of stock assessment teams had not been presented previously. The remainder of the presentation served as an eloquent review of many of the issues raised earlier in the week, along with constructive ideas for addressing some of the more compelling challenges.

The Center's stock assessment enterprise consists of 5 teams located in 4 different locations throughout the Southeast. Some teams are very small. Most individuals are located at the main laboratory, in Miami. Team responsibilities are defined by region and type of assessment (e.g., HMS, Gulf and Caribbean, South Atlantic, Sharks, Shrimp). Reporting lines differ and there is no common supervisory oversight of the entire enterprise. This structure is consistent with the overall organizational structure of SEFSC and has both strengths and weaknesses. Changes in organizational structure to consolidate stock assessment activities could improve consistency and the ability to be responsive to region-wide changes in stock assessment priorities.

The importance of bringing objective criteria to bear in setting stock assessment priorities as well as assessment levels and assessment frequencies has been raised earlier. Here the discussion included a thoughtful overview of the agency's draft prioritization scheme and this framework does, indeed, offer some very useful guidance. Tradeoffs between thoroughness, transparency and timeliness also demand greater consideration. What is actually necessary to support requirements for management advice? How can data quality and availability be improved? Here the need for improved surveys which employ advanced technologies and provide absolute abundance estimates was emphasized.

As discussed in preceding sections, staffing and workload problems persist and interact. Nevertheless, the Center has been effective in maintaining and, to a small degree, increasing stock assessment throughput. Furthermore, overall stock status, as indicated by FSSI and proportion of stocks not overfished and/or not subject to overfishing has improved. None of this obviates the urgent need to recognize the demands placed on staff and to develop strategies for relieving this pressure and encouraging and supporting more research and professional development. Needs for improving trust through communication and engagement with stakeholder, and for broader education and outreach were also detailed.

Conclusions

The SEFSC stock assessment enterprise excels technically and professionally in spite of major challenges associated with burgeoning demands (for stock assessment and other products). However, these increasing demands, as well as limitations associated with meeting assessment data requirements and an overburdened staff greatly constrain the Center's ability to meet throughput expectations. Furthermore the workload is compromising the ability of key staff to engage in research and professional development; this is resulting in morale issues and jeopardizes the integrity of the overall process.

Some specific recommendations for addressing scientific and technical concerns related to stock assessment and for a much needed reform of the SEDAR process are provided above. Several of these recommendations also relate to the peer review process. Some, especially those directed towards improving the SEDAR process, could be effected relatively quickly while others, such as those directed towards improvements in fishery independent monitoring, will be expensive and time consuming to implement. Concerns associated with excessive demands on staff, and increasing demands for stock assessment products are especially challenging to resolve but should be prioritized. Improving efficiency and managing expectations are both important in this regard. Provision of additional staff to support stock assessment would certainly be beneficial but this may not be possible due to budget limitations and the need to balance staffing requirements across the range of scientific activities within the Center.

In general, the research that is being carried out by stock assessment scientists at SEFSC is of very high quality and has resulted in noteworthy methodological improvements. Communication of research results is effective although emphasis on improving communication with stakeholders should be encouraged. The Center is to be commended for work that has been done in bringing ecosystem considerations to bear in stock assessments and towards broader, integrated ecosystem assessment in the Gulf of Mexico. Even though resources to support these activities are lacking, efforts should be directed towards maintaining and increasing this type of work.