_____

Center for Independent Experts (CIE) Reviewer's Independent Report on the 2013 and 2014 assessments of Gulf of Mexico gag grouper (*Mycteroperca microlepsis*) and greater amberjack (*Seriola dumerili*).

_____

Prepared by

Neil Klaer

Prepared for

Center for Independent Experts

Review Meeting
Miami, Florida
24-27 February 2014

Contents

Executive Summary

The SEDAR 33 Review Workshop (RW) for Gulf of Mexico gag (*Mycteroperca microlepsis*) and greater amberjack (*Seriola dumerili*) met in Miami, Florida, from 24-27 February 2014. The meeting was chaired by Dr. Sean Powers from the Gulf of Mexico Fishery Management Council Scientific and Statistical Committee. The review panel (hereafter called the Panel) was composed of three scientists affiliated with the Center for Independent Experts: Dr. Mike Armstrong, Dr. Anders Nielsen, and Dr. Neil Klaer. Two additional reviewers, Benjamin Blount and Greg Stunz from the GMFMC SSC, completed the Panel. The meeting format included presentations mixed with questions and open discussion, and was open to the public. Activities of the reviewers were shared during the meeting. A draft summary report of the proceedings was developed in the final days of the meeting.

Findings by term of reference

1. Evaluate the data used in the assessment

   In general, the input data and methods used to process them for inclusion in the assessment were adequate and appropriate.

   For both gag and greater amberjack the data and assessment teams are commended for their work in compiling and evaluating the wide range of data and parameters used in the assessment.

   The gag data are reliable in the sense that deficiencies and uncertainties in the data have been explored and their effect on the assessment shown through sensitivity analyses.

   The nature of the data for greater amberjack and its fisheries means that data quality is patchy and in places insufficient to support the estimation of the many selectivity and other parameters in the stock synthesis model as presented. Further development is needed to identify a simpler model formulation appropriate to the information content of the available data.

2. Evaluate the methods used to assess the stock, taking into account the available data.

   The Panel agreed that stock synthesis (SS) provides an assessment method appropriate for application to the diverse sets of input data available for both gag and greater amberjack. A CASAL assessment provided a good cross-verification of assessment results for gag, and an ASPIC assessment provided continuity with the previous assessment for greater amberjack, while no longer providing the preferred assessment method.

   For gag, good fits to the data and the ability of SS to find a stable solution indicate that the data are sufficient to support the additional complexities in the model structure around selectivity and retention that are necessitated by the series of changes in management measures.

The Panel considered that further work after the RW was required to improve the base SS model for greater amberjack. Indicators to improve are the jitter convergence and the likelihood profiles – these should better indicate that the model converges to stable solutions. A number of methods were suggested to assist with model improvement.

3. Evaluate the assessment findings.

As a final model was unavailable for greater amberjack during the RW, assessment findings (and also projections) are only discussed for gag.

Whether the gag stock is overfished depends on the base model configuration, with the male-female combined SSB version showing a worse current condition than the female SSB only version.

The stock is not undergoing overfishing. A combination of management interventions of lower IFQs on the commercial fishery and size and season limits on the recreational fishery appear to have successfully lowered recent F values to below F_{msy} and also F_{spr30} for the agreed base model and most sensitivities. Overfishing measures for some sensitivities requested by the Panel were unavailable at the time of writing – particularly for steepness 0.7.

4. Evaluate the stock projections

For gag the analytical approach was appropriate and provides an acceptable basis for management advice.

Projection results that followed Panel recommendations for conveying uncertainty for both gag and greater amberjack were unavailable at the time of writing.

5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.

Uncertainty in estimated base model parameters was evaluated both from the inverse hessian matrix and also via a parametric bootstrap procedure. This within-model uncertainty was also propagated into projections using the bootstrap procedure. However, between-model uncertainty is often greater in magnitude. Plausible alternative models are therefore often used to better indicate the true uncertainty in the assessment results. Such plausible alternative models for gag assessments could be across key fixed assumptions such as values for M and steepness, different assumptions for discard mortality, different weightings applied to abundance indices, whether to include males in the SSB and whether to include a red tide event. The Panel recommended that the uncertainty in sex structure of the spawning biomass and steepness be carried into the management advice for the current gag assessments, and at least steepness for greater amberjack to better reflect the true range of uncertainty in assessment results.

6. Consider the research recommendations provided by the Assessment workshop and make any additional recommendations or prioritizations warranted.

   A number of research recommendations already proposed were endorsed, and some new ones proposed.

   I have outlined a view that directing the SEDAR process towards closely associated species or complexes may have advantages.

7. Provide guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.

   I have made a small number of additional specific recommendations.

1 Introduction

1.1 Background

The SEDAR 33 Review Workshop (RW) for Gulf of Mexico gag (*Mycteroperca microlepsis*) and greater amberjack (*Seriola dumerili*) met in Miami, Florida, from Monday, February 24, through Thursday, February 27, 2014. The meeting was chaired by Dr. Sean Powers from the Gulf of Mexico Fishery Management Council Scientific and Statistical Committee (GMFMC SSC). The review panel was composed of three scientists affiliated with the Center for Independent Experts: Dr. Mike Armstrong, Dr. Anders Nielsen, and Dr. Neil Klaer. Two additional reviewers, Benjamin Blount and Greg Stunz from the GMFMC SSC completed the Panel.

Reports from the SEDAR 33 Assessment Workshop and Data Workshop as well as all associated background documents were made available via a secure webserver to the Panel on 4 February prior to the review meeting. During the meeting, all documents were available electronically via the same webserver, and notes and presentations were uploaded as they became available.

The meeting format included presentations by the assessment teams mixed with questions and open discussion. The Panel participated in the review of each term of reference. The meeting was open to the public although no public comments or questions were received by the Panel.

1.2 Review Activities

A brief description of presentations, Panel requests and responses are given in the summary report. Activities of the reviewers were shared during the meeting. It was a requirement that a first draft of the summary report be produced during the Review Workshop. Initial drafting of the report against the Terms of Reference (TORs) was divided among the reviewers and I drafted the text for TOR3 on the Stock Synthesis (SS) assessment for gag and SS and ASPIC assessments for greater amberjack. Draft text for the summary report was compiled with the assistance of the chair on Wednesday and Thursday of the meeting, with further compilation and editing in the two weeks following the meeting.

2 Review of gag and greater amberjack assessments

2.1 Terms of reference

The Panel considered the assessments in light of the terms of reference provided as follows:

1. Evaluate the data used in the assessment, addressing the following:
   a. Are data decisions made by the Assessment Workshop sound and robust?
   b. Are data uncertainties acknowledged, reported, and within normal or expected levels?
   c. Are data applied properly within the assessment model?
   d. Are input data series reliable and sufficient to support the assessment approach and findings?
2. Evaluate the methods used to assess the stock, taking into account the available data.
   a. Are methods scientifically sound and robust?
   b. Are assessment models configured properly and used consistent with standard practices?
   c. Are the methods appropriate for the available data?
3. Evaluate the assessment findings with respect to the following:
   a. Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?
   b. Is the stock overfished? What information helps you reach this conclusion?
   c. Is the stock undergoing overfishing? What information helps you reach this conclusion?
   d. Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?
   e. Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?
4. Evaluate the stock projections, addressing the following:
   a. Are the methods consistent with accepted practices and available data?
   b. Are the methods appropriate for the assessment model and outputs?
   c. Are the results informative and robust, and useful to support inferences of probable future conditions?
   d. Are key uncertainties acknowledged, discussed, and reflected in the projection results?
5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.
   • Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods
   • Ensure that the implications of uncertainty in technical conclusions are clearly stated.

6. Consider the research recommendations provided by the Assessment workshop and make any additional recommendations or prioritizations warranted.
   - Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.
   - Provide recommendations on possible ways to improve the SEDAR process.
7. Provide guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.

2.2 Findings by term of reference

2.2.1 TOR1 Evaluate the data used in the assessment.

In general, the input data and methods used to process them for inclusion in the assessment were adequate and appropriate. I agree with the points made in the summary report.

a. Are data decisions made by the Assessment Workshop sound and robust?

The SEDAR process allows good consideration of the properties of available data and the ability of those data to inform stock assessment as the task of a dedicated Data Workshop (DW). Terms of reference for the following Assessment Workshop (AW) explicitly include addressing known uncertainties in the input data. Both for gag and greater amberjack, the processes were similar, with considerable work at the DW to characterize available data, and at the AW to account for uncertainties in the assessment input data.

b. Are data uncertainties acknowledged, reported, and within normal or expected levels?

Both gag and greater amberjack have similar uncertainties from the viewpoint of available catch data. Both have uncertain catch history because of catch records combined with other species, significant levels of discarding that has changed throughout history often in association with changes in management measures, and significant catch from both commercial and recreational sectors.

An uncertainty not explicitly dealt with is that of unit stock, and therefore uncertainty in the catch level most relevant to that unit stock, whether they are smaller units within the currently defined boundary (basically inside the US EEZ within the GM), or more widely throughout the GM or into the Atlantic.

The gag stock for assessment purposes is assumed to be separate to the south Atlantic, and mostly in the GMFMC area with the bulk of the US catch taken in the W Florida region. There has been a suggestion that larvae are transported into the US stock region from Mexico and beyond, so it is possible that recruitment to the US gag stock may be dependent on fishery impacts outside of US management control. Additional biological work would be required to determine if this was an important consideration.

Greater amberjack are also assumed to form a single stock within the GMFMC area, with some evidence for sub-regionalization with less mixing between fish between the east and west of that range.

It is a common problem for various US stocks that catch histories from neighboring countries such as Mexico are unavailable. The collection of landing statistics from such places for potential straddling stocks would allow investigation of the possible important influences on those stocks.

Variability in catch associated with uncertainty in the estimation of that catch was examined and considered in sensitivity analyses. An improvement to that process is to explicitly include catch uncertainty within the assessment model. The Panel recommended that a level of uncertainty (starting with a suggested CV of 0.2) be assigned to the recreational catch estimates within the stock assessment.

For both gag and greater amberjack the natural mortality rate was estimated using Hoenig's method, and a Lorenzen procedure used to assign a declining age-specific value. These methods used early available fishery data, but a concern is that little is known about age or size composition prior to about 1980 for either species. Substantial catches were taken earlier, so the estimates of natural mortality were made from depleted stocks, which could cause the natural mortality rate to be overestimated. Upper and lower values for M were examined as sensitivities, and the range chosen seems appropriate.

The report cards for abundance indices filled by the DW provides the beginning for a process that would enable semi-quantitative expert judgment to decide what relative weighting should be given to those indices within a stock assessment. At present the outcome of the report cards are applied as an aid to the decision of whether or not to use an index in the stock assessment, and the relative weights among indices within the assessment are decided based on measurement error from the standardizations. I can only encourage the continued development of the report card process and improved methods for assigning relative weights among indices based on their relative ability to accurately reflect abundance.

It has become generally accepted that the effective sample size for length or age composition data is more related to the number of trips or fishing operations sampled, than the aggregate number of fish measured across sampling operations – particularly if large numbers of measurements derive from a small number of trips or operations. Often the number of trips or operations is not available for input for a stock assessment, but for gag and greater amberjack those figures are mostly available. The gag and greater amberjack assessments brought to the RW both used numbers of fish measured (capped following convention at 200) as the input sample sizes for composition data. The Panel recommended that trips or fishing operations are the preferred measure.

There are a large number of data sources for catches, discards, abundance indices and age/size composition for gag and greater amberjack. The development of a single document that describes each of these sources and summarizes important details in relation to sampling design, implementation,

sampling achievement and analysis has some value. Ideally, a data group would update the document periodically, and reference data sets would be archived electronically. Such a system would allow the precise specification of the procedures used to create reference data sets from the raw data, to help avoid future data loss or misuse, and to provide a principal data reference source for stock assessors (and also reviewers).

c. Are data applied properly within the assessment model?

The Panel considers that, based on the workshop reports and presentations at the Review Meeting, the data have been properly applied within the assessment model.

d. Are input data series reliable and sufficient to support the assessment approach and findings?

For gag the Panel concludes that the input data series are reliable and are sufficient to support the assessment approach and findings. The data and assessment teams are commended for their work in compiling and evaluating the wide range of data and parameters used in the assessment. The data are reliable in the sense that deficiencies and uncertainties in the data have been explored in detail and that assumptions and decisions made in compiling input parameters and data have been clearly presented and their effect on the assessment shown through sensitivity analyses. Good fits to the data and the ability of Stock Synthesis to find a stable solution also indicate that the data are sufficient to support the additional complexities in the model structure around selectivity and retention that are necessitated by the series of changes in minimum landing sizes and IFQs which affect size compositions of retained and discarded fish.

For greater amberjack the data and assessment teams are commended for their work in compiling and evaluating the wide range of data and parameters used in the assessment, but the nature of greater amberjack and its fisheries means that data quality is patchy and in places insufficient to support the estimation of the many selectivity and other parameters in the SS model as presented. Further development is needed to identify a simpler model formulation appropriate to the information content of the available data.

2.2.2 TOR2 Evaluate the methods used to assess the stock, taking into account the available data.

The Panel agreed that SS provides an assessment method appropriate for application to the diverse sets of input data available for both gag and greater amberjack.

a. Are methods scientifically sound and robust?

For gag, stock assessments were completed using Stock Synthesis and CASAL. The primary assessment proposed for use for management purposes was SS. A CASAL model was configured as closely as possible to match with SS, while acknowledging some small differences in setup options in each framework,

making exact comparison impossible. Both the absolute spawning biomass series and the series of estimated annual recruitments were very similar, providing good evidence that the data leads to similar conclusions using either framework. This cross-framework replication provides model validation at a level not normally considered for most stock assessments, and is commendable.

An assessment using ASPIC was developed for greater amberjack. This differed from the most recently accepted greater amberjack assessment mostly in how the abundance indices were interpreted. An important feature of the index data available for greater amberjack is that the selectivity is widely different among those indices. For example, the size range of catches associated with the MRFSS index suggests a broad selection pattern, while the narrow and small range of catches associated with the headboat survey suggests a dome shaped pattern. A biomass dynamic model such as ASPIC is unable to allow selectivity differences across indices, with all applying equally to the total biomass. Experience with an age structured model for amberjack shows that the different trends shown by the indices can be more appropriately followed by expected values determined via appropriate size selection.

The AW concluded that factors that significantly change the size selectivity of an index (e.g. the implementation of a new retention size limit) affects the index catchability and therefore necessitate that the index be broken at that point. A long index showing potentially useful trend information such as MRFSS is therefore reduced to smaller segments that no longer show any significant trend. As management change is a common feature for the amberjack indices, an assessment that can deal with those changes in more parsimonious manner is preferred.

As a model that can deal with the issues outlined is available (SS), the Panel agreed that the ASPIC model provides continuity with previous assessments, but is no longer the preferred method for determination of stock status and management advice for greater amberjack.

The SS analytical approach was appropriate and provides an acceptable basis for management advice for both gag and greater amberjack, given the diversity of data inputs available including multiple abundance indices with different and time-varying selectivity/discarding as well as catch age and length composition data. I agree with the comments and recommendations in the summary report. The SS base cases for both gag and greater amberjack were modified during the review after examining sensitivity analyses and diagnostics requested by the Panel. It was possible to agree on a base model configuration for gag during the RW which was one of the sensitivities to the base already examined by the AP. Results from the modified base runs were to be compiled and made available in the weeks following the review for greater amberjack, and at the time of writing are not completed.

b. Are assessment models configured properly and used consistent with standard practices?

The gag assessment follows standard practices but also includes several novel features that the assessment team has captured in interesting and effective ways.

Gag are protogynous hermaphrodites (female at birth, then a proportion of the population transition into males) that has been included in the model as a fixed logistic function.

A red tide mortality event (an algal bloom, which had harmful effect on gag) occurred in 2005, and is well supported by most indices. The assessment team included this additional mortality in the model as an additional fleet, which only has positive effort in 2005. In this approach the level of mortality associated with the red tide was estimated within the model to best fit available data. The selectivity of the red tide fleet was set to apply equally to all ages except age zero. The implementation of the red tide event and evidence of improved model behavior would be of wide interest, and I encourage formal publication of the procedure.

The Panel considered that further work after the RW was required to improve the base SS model for greater amberjack. Indicators to improve are the jitter convergence and the likelihood profiles – these should better indicate that the model converges to stable solutions.

Ways to do this are:

- o Investigate correlations among model parameters (particularly for selectivity and retention), and either fix or provide informative priors for one or more values that have some supportable evidence. If supportable evidence is unavailable and the parameter has a strong influence on the results, then a range of alternative fixed values should be investigated.

- o For individual jitter starting points that resulted in different likelihood solutions, investigate which parameter estimates were affected that may also be fixed or provided with informative priors.

- o Examine the CVs of parameter estimates. If the CV is large and the value has little influence on results, then choose a fixed value.

- o Examine the time blocking of retention and selectivity for the fleets that converge at very high F values – consider very high to be values greater than 1.0, but preferably less than that. Consider adjusting the configuration of selectivity and retention of those fleets around the period of high F to see if the problem can be alleviated.

- o Investigate independent sources of information for plausible values for steepness for amberjack – e.g. Ram database, stock assessments of similar species to develop a plausible range of steepness values to set for

base, upper and lower cases to carry forward as sensitivities for reference points and projections.

c. Are the methods appropriate for the available data?

Stock assessments are normally very sensitive to the natural mortality rate used, so this is a major uncertainty in the current gag and greater amberjack assessments also. Both assessments examined sensitivity to this uncertainty, but this was not a dimension explored for consideration of advice given to management. As natural mortality is often correlated with steepness, and steepness was explored in this way, it is perhaps acceptable not to have also included an examination of alternative values for *M*.

2.2.3 TOR3 Evaluate the assessment findings.

The analytical approach was appropriate and provides an acceptable basis for management advice. I agree with the comments and recommendations in the summary report. Assessment results were only available for gag, so only gag is discussed below.

a. Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?

Abundance indices

The model fit to the fishery-dependent headboat index was best overall, influenced by the relatively low CV for that index, partly contributing to the model prediction of a steep increase in biomass in recent years.

Of the fishery-independent indices, the SEAMAP video has a selectivity that should allow the steep increase in biomass in recent years to be observed. However, the observations at least in 2011 and 2012 are well below the expected values. The PC video index was not well fitted, but has a large CV and short time span. The Age 0 index includes a very below average point in 2011 followed by an additional low point in 2012. As these provide almost all information (other than discards) on very recent recruitments these produce low model estimates for recruitment in those years, with implications for projections.

CVs for abundance indices in the model were those determined by the standardization processes, and no further adjustment was made. If additional iterative reweighting, for example, had been applied, then most or all of the index CVs would have been adjusted upwards, thus giving the abundance indices less weight in the model. Recent advice on model weighting by e.g. Francis (2011) recommend that abundance indices be weighted more heavily than iterative weighting of all data sources would normally provide, so the indices as weighted in the model tend to conform to those recommendations (although simply using

observed CVs denies the existence of process error in addition to the measurement error).

Exploitation estimates

These are reliant on the exploitable biomass estimate per fleet, and the catch level per fleet annually. Exploitable biomass depends on estimated selectivity patterns and the total biomass (discussed below).

Different selectivity patterns among fleets are evident from the composition data and the Panel is satisfied with the characterization of selectivity within the current model. Given the large number of abundance indices, further work on fixing selectivity parameters that are badly estimated from the data, or highly correlated with other model parameters, is encouraged.

A considerable portion of gag catch is estimated using survey information rather than from logbook records as for the commercial fleets, so has increased uncertainty. To account for this, the Panel has recommended that the recreational catch not be fitted exactly by the assessment model, using an associated CV of 0.2. Uncertainty in the total catch for all fleets is caused by the mixing of gag catch within catch of unspecified grouper.

Biomass estimates

As the fishery underwent a long period of exploitation prior to 1963 with only catch estimates available from 1880 to 1963, the absolute virgin biomass is dependent on the productivity of the stock and therefore steepness. The steepness value for this stock is highly uncertain (see below), so also is the estimate of virgin biomass. Certainty of current biomass is dependent on the scale of the predicted recent increase due to trends in abundance indices and reduced catches.

Whether males are included in the spawning biomass affects assessment results considerably, with overall stock depletion and the sustainability of current F lower for the combined case. Without evidence to support the use of female-only SSB, the Panel agreed with the earlier assessment workshop recommendation that the combined biomass is preferred.

b. Is the stock overfished? What information helps you reach this conclusion?

This is unclear and depends on the base model configuration (male-female combined SSB worse than female only).

c. Is the stock undergoing overfishing? What information helps you reach this conclusion?

The stock is not undergoing overfishing. A combination of management interventions of lower IFQs on the commercial fishery and size and season limits on the recreational fishery appear to have successfully lowered recent F values to below F_{msy} and also F_{spr30} for the agreed base model and most sensitivities.

Overfishing measures for some sensitivities requested by the Panel were unavailable at the time of writing – particularly for steepness 0.7.

d. Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?

For the period where year class strength can be estimated by the model based on composition data (since about 1970), the stock was estimated to have generally mainly increased, except for the sharp decline attributed to the red tide event. The stock has therefore only provided one very short period of increase from which to characterize the shape of the stock-recruitment relationship, and all from after a substantial estimated initial depletion prior to 1963. Steepness is estimated by the base model to be high at >0.9. Steepness is certainly not well characterized by the estimated stock/recruitment points and therefore the Panel recommended that a fixed value of 0.85 be used for the base case, with results also to be provided for management using the estimated high value and a lower fixed value of 0.7. Acceptance that steepness cannot be estimated for this stock may also allow the estimation of an additional important uncertain parameter such as natural mortality.

e. Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?

The measures that rely on absolute biomass are greatly influenced by the value of steepness, and most indicators are affected by the choice of whether to include males in the spawning biomass. Dimensions of uncertainty that the Panel agreed to carry forward in management recommendations were therefore (SSBfemale only/SSBfemales+males) x (Steepness estimated/ 0.85/0.7), resulting in 6 alternatives.

2.2.4 TOR4   Evaluate the stock projections.

For gag the analytical approach was appropriate and provides an acceptable basis for management advice. For greater amberjack, the proposed approach was acceptable, but because a base model configuration was not identified during the RW, the Panel was unable to evaluate the projections for that species. The following comments apply to gag only.

a. Are the methods consistent with accepted practices and available data?

Standard and acceptable procedures were used to carry out both deterministic and bootstrap projections for the 6 scenarios recommended by the RW as representing the range of uncertainty in stock status and management recommendations due to uncertainty in model structure (3 values of steepness, male+female and female only spawning biomass). The bootstrapping provides good information of the uncertainty within an individual model configuration.

b. Are the methods appropriate for the assessment model and outputs?

Yes.

c. Are the results informative and robust, and useful to support inferences of probable future conditions?

As for all stock assessment models, the assessment and therefore the projections provide a simplified view of reality and underestimates true uncertainty, particularly about the future. However, the methods are standard when compared to other assessments. The range of outcomes provides useful insight into the true uncertainty of the results.

d. Are key uncertainties acknowledged, discussed, and reflected in the projection results?

The provision of projection results across a range of alternative values for key parameters acknowledges some of the additional between-model uncertainty. The full range of projection results was unavailable for gag at the time of writing. Results are currently available for estimated and 0.85 values for steepness and SSB including males or not (3 of the 6 recommended cases).

2.2.5 TOR5 Consider how uncertainties in the assessment, and their potential consequences, are addressed.

• Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods

Uncertainty in estimated base model parameters was evaluated both from the inverse hessian matrix and also via a parametric bootstrap procedure. This within-model uncertainty was also propagated into projections using the bootstrap procedure. However, between-model uncertainty is often greater in magnitude. Plausible alternative models are therefore often used to better indicate the true uncertainty in the assessment results. Such plausible alternative models for gag assessments could be across key fixed assumptions such as values for M and steepness, different assumptions for discard mortality, different weightings applied to abundance indices, whether to include males in the SSB and whether to include a red tide event. The Panel recommended that the uncertainty in sex structure of the spawning biomass and steepness be carried into the management advice for the current gag assessment, and at least for alternative steepness values for greater amberjack to better reflect the true range of uncertainty in assessment results.

• Ensure that the implications of uncertainty in technical conclusions are clearly stated.

Projection results that followed Panel recommendations for both gag and greater amberjack were unavailable at the time of writing.

2.2.6 TOR6 Consider the research recommendations provided by the Assessment workshop and make any additional recommendations or prioritizations warranted.

• Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.

I agree with the recommendations made in the summary report.

A summary of my own additional recommendations is as follows:

- A single document should be developed that describes reference data for catches, abundance indices and age/size composition and how they were created. Those reference sets should also be electronically archived at a single location.

- Efforts should be made/continued to collect landing statistics from other countries for important straddling stocks. For the GOM groupers and also greater amberjack, countries of importance may include Mexico and Belize (but also possibly Guatemala, Honduras, Nicaragua, Jamaica and Cuba).

• Provide recommendations on possible ways to improve the SEDAR process.

I found the SEDAR process of data workshop, assessment workshop and review workshop for gag and greater amberjack to be effective. A great deal of work has been done to examine and characterize uncertainties in source data during the Data Workshop, and considerable progress was made at the Assessment Workshop in moving stock assessments to the SS framework to better make use of all available data. The assessment teams should be congratulated on the progress made in a relatively short time, and the quality of the assessments produced.

There were some aspects of the assessment that would better have been investigated at the Assessment Workshop rather than at the Review Workshop – examples being incorporating uncertainty in catch estimates within the assessment, selection of the most appropriate starting year for recruitment deviation estimation and investigation of the causes of model instability. Local expertise with such issues particularly in relation to SS assessments is fast increasing, making it more likely that such issues will be dealt with at the AW in future. There appears to still be a need to increase the input of scientists with long experience with SS at the AW level.

I understand that the SEDAR process is aligned with management requirements, and that the focus is on management of species mostly in isolation. My following comments may not align as well with management requirements, and may therefore not be feasible, but I consider them worth making anyway. They have been made previously in a review of red and black grouper assessments.

Associated species

Snapper and gag grouper are major target species of what is most easily described as a snapper-grouper complex, and both red and black grouper are caught primarily as bycatch while targeting these other species.

The data and assessments for gag, red and black grouper have largely been prepared by separating them from the other species, primarily because management action is directed towards individual species and not such complexes as a whole. However, as the species in the larger complex are closely associated (and cross-identified such as for gag and black grouper), there may be advantages and efficiency gains in the SEDAR process being directed at the complex level, rather than at particular species or sub-groups.

For example, if the data workshop was directed at the complex, then a single account could be given for how historical catch was split among the various species. It would be a simpler task to document the processes used to produce the landings, indices and length/age composition for the species as part of the complex, rather than for the separated species. Archival of the resulting data sets would also be done for the complex.

The priorities of the assessment workshop could be directed mostly towards the important (and presumably most valuable) target species, perhaps with less priority on bycatch species, and also the potential for meaningful multispecies data analyses and possibly multispecies assessments. At the very least there is scope to investigate target and bycatch analyses to highlight where TAC setting difficulties may occur, and multispecies production models could be examined as alternative stock assessments. A process that has taken steps along the multispecies path may be better positioned in the future to practically consider ecosystem-based fishery management options.

2.2.7 TOR7 Provide guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.

I agree with the specific comments in the summary report and add the following:

Stock assessment

- A search should be conducted for the most appropriate value of steepness to be used for gag and greater amberjack – either across a range of species (e.g. Ram database) or through use of a well estimated value from a closely related stock or species.

- If an appropriate fixed value for steepness is found, explore the estimation of parameters currently fixed in the model, such as natural mortality.

- Further work on fixing selectivity parameters that are badly estimated from the data, or highly correlated with other model parameters.

Simulation testing

Simulation testing can be used to verify assessments models, compare alternative assessment model structures, and to test the robustness of harvest control rules implemented by management. An often used framework for such testing is Management Strategy Evaluation (MSE). Although the implementation of a MSE system requires a fairly large resource commitment initially, once the system has been developed the ongoing maintenance can be minimal. While management benchmarks applied to gag and greater amberjack generally comply with those used in many other US fisheries, generic systems may not always work well in specific circumstances. There could be a considerable advantage in building a system to test the robustness of the current harvest strategy to the major uncertainties for gag and greater amberjack – in particular the level of discard mortality, natural mortality, and the B_{msy} estimate. A more ambitious but possibly more useful MSE for gag could be built for the snapper-grouper complex, but few multispecies MSE systems have been built thus far.

References

Francis, R.I.C.C., 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68, 1124–1138.

Appendix 1: Bibliography of materials provided for review

SEDAR 33 SAR Section II – Gag Data Workshop Report
SEDAR 33 SAR Section II – GAJ Data Workshop Report
SEDAR 33 SAR Section III – Gag Assessment Workshop Report
SEDAR 33 SAR Section III – GAJ Assessment Workshop Report
All supporting documents were also provided for these workshops via the FTP site

Appendix 2: A copy of the CIE Statement of Work

Attachment A: Statement of Work for Dr. Neil Klaer

External Independent Peer Review by the Center for Independent Experts

SEDAR 33 Gulf of Mexico Gag and Greater Amberjack Assessment Review

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

Project Description: SEDAR 33 will be a compilation of data, benchmark assessments of the stocks, and an assessment review conducted for Gulf of Mexico gag and greater amberjack. The review panel is ultimately responsible for ensuring that the best possible assessments are provided through the SEDAR process. The stocks assessed through SEDAR 33 are within the jurisdiction of the Gulf of Mexico Fishery Management Council and the state waters of Texas, Louisiana, Mississippi, Alabama, and Florida.. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements for CIE Reviewers: Three CIE reviewers shall have the necessary qualifications to complete an impartial and independent peer review in accordance with the tasks and ToRs described in the SoW herein. The CIE reviewers shall have expertise in stock assessment, statistics, fisheries science, and marine biology sufficient to complete the tasks of the scientific peer-review described herein. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall participate and conduct an independent peer review during the SEDAR 33 panel review meeting scheduled in Miami, Florida during February 24-27, 2014.

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

Tasks prior to the meeting: The contractor shall independently select qualified reviewers that do not have conflicts of interest to conduct an independent scientific peer review in accordance with the tasks and ToRs within the SoW. Upon completion of the independent reviewer

selection by the contractor's technical team, the contractor shall provide the reviewer information (full name, title, affiliation, country, address, email, and FAX number) to the contractor officer's representative (COR), who will forward this information to the NMFS Project Contact no later than the date specified in the Schedule of Milestones and Deliverables. The contractor shall be responsible for providing the SoW and stock assessment ToRs to each reviewer. The NMFS Project Contact will be responsible for providing the reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact will also be responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COR prior to the commencement of the peer review.

Foreign National Security Clearance: The reviewers shall participate during a panel review meeting at a government facility, and the NMFS Project Contact will be responsible for obtaining the Foreign National Security Clearance approval for the reviewers who are non-US citizens. For this reason, the reviewers shall provide by FAX (not by email) the requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/.

Pre-review Background Documents: Approximately two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the COR the necessary background information and reports (i.e., working papers) for the reviewers to conduct the peer review, and the COR will forward these to the contractor. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the COR on where to send documents. The reviewers are responsible only for the pre-review documents that are delivered to the contractor in accordance to the SoW scheduled deadlines specified herein. The reviewers shall read all documents deemed as necessary in preparation for the peer review.

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

Tasks after the panel review meeting: Each reviewer shall prepare an independent peer review report, and the report shall be formatted as described in **Annex 1**. This report should explain whether each stock assessment ToR was or was not completed successfully during the SEDAR meeting. If any existing BRP or their proxies are considered inappropriate, each independent

report shall include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report shall indicate that the existing BRPs are the best available at this time. Additional questions and pertinent information related to the assessment review addressed during the meetings that were not in the ToRs may be included in a separate section at the end of an independent peer review report.

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

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

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting at Miami, Florida from February 24-27, 2014
3) Conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than March 14, 2013, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

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

24 January 2014	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
3 February 2014	NMFS Project Contact sends the stock assessment report and background documents to the CIE reviewers.
24-27 February 2014	Each reviewer shall conduct an independent peer review during the panel review meeting in Miami, Florida
14 March 2014	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
28 March 2014	CIE submits CIE independent peer review reports to the COR
4 April 2014	The COR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on changes. The COR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

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

Applicable Performance Standards: The contract is successfully completed when the COR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) The CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) The CIE report shall address each ToR as specified in **Annex 2**,
(3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

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

Support Personnel:

William Michaels, Program Manager, COR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov   Phone: 301-427-8155

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

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

Key Personnel:

<u>NMFS Project Contact</u>:

Ryan Rindone, SEDAR Coordinator
2203 N. Lois Avenue, Suite 1100
Tampa, Florida 33607
Ryan.Rindone@gulfcouncil.org   Phone: 813-348-1630

Annex 1: Format and Contents of CIE Independent Peer Review Report

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

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed. The CIE independent report shall be an independent peer review of each ToRs.

3. The reviewer report shall include the following appendices:

    Appendix 1: Bibliography of materials provided for review
    Appendix 2: A copy of the CIE Statement of Work

Annex 2: Tentative Terms of Reference for the Peer Review

SEDAR 33 Gulf of Mexico Gag and Greater Amberjack Assessment Review

1. Evaluate the data used in the assessment, addressing the following:

    a) Are data decisions made by the Assessment Workshop sound and robust?

    b) Are data uncertainties acknowledged, reported, and within normal or expected levels?

    c) Are data applied properly within the assessment model?

    d) Are input data series reliable and sufficient to support the assessment approach and findings?

2. Evaluate the methods used to assess the stock, taking into account the available data.

    a) Are methods scientifically sound and robust?

    b) Are assessment models configured properly and used consistent with standard practices?

    c) Are the methods appropriate for the available data?

3. Evaluate the assessment findings with respect to the following:

    a) Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?

    b) Is the stock overfished? What information helps you reach this conclusion?

    c) Is the stock undergoing overfishing? What information helps you reach this conclusion?

    d) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?

    e) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?

4. Evaluate the stock projections, addressing the following:

    a) Are the methods consistent with accepted practices and available data?

    b) Are the methods appropriate for the assessment model and outputs?

    c) Are the results informative and robust, and useful to support inferences of probable future conditions?

    d) Are key uncertainties acknowledged, discussed, and reflected in the projection results ?

5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.

    • Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods

    • Ensure that the implications of uncertainty in technical conclusions are clearly stated.

6. Consider the research recommendations provided by the Assessment workshop and make any additional recommendations or prioritizations warranted.

  • Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.

  • Provide recommendations on possible ways to improve the SEDAR process.

7. Provide guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.

Annex 3: Tentative Agenda for

SEDAR 33 Gulf of Mexico Gag and Greater Amberjack Assessment Review
February 24-27, 2014
Miami, FL USA

Monday
1:00 p.m.	**Convene**	
1:00 – 1:30	**Introductions and Opening Remarks**	**Rindone**
	- Agenda Review, TOR, Task Assignments	
1:30 – 5:00	**Assessment Presentations and Discussions**	**SEFSC**
5:00 p.m. - 6:00 p.m.	**Panel Work Session**	**Powers**

Tuesday
8:00 a.m. – 11:30 a.m.	**Assessment Presentations and Discussions**	**SEFSC**
11:30 a.m. – 1:00 p.m.	**Lunch Break**	
1:00 p.m. – 3:30 p.m.	**Panel Discussion**	**Powers**
	- *Assessment Data & Methods*	
	- Identify additional analyses, sensitivities, corrections	
3:30 p.m. – 3:45 p.m.	**Break**	
3:45 p.m. – 5:00 p.m.	**Panel Discussion**	**Powers**
	- Continue deliberations	
	- Review additional analyses	
5:00 p.m. - 6:00 p.m.	**Panel Work Session**	**Powers**

*Tuesday Goals***:** Initial presentations completed, sensitivities and modifications identified.

Wednesday
8:00 a.m. – 11:30 a.m.	**Panel Discussion**	**Powers**
	- Review additional analyses, sensitivities	
	- Consensus recommendations and comments	
11:30 a.m. – 1:00 p.m.	**Lunch Break**	
1:00 p.m. – 3:30 p.m.	**Panel Discussion**	**Powers**
	- Final sensitivities reviewed.	
	- Projections reviewed.	
3:30 p.m. – 3:45 p.m.	**Break**	
3:45 p.m. – 5:00 p.m.	**Panel Discussion/Work Session**	**Powers**
	- Review Consensus Reports	
5:00 p.m. - 6:00 p.m.	**Panel Work Session**	**Powers**

Wednesday Goals: Final sensitivities identified, preferred models selected, projection approaches approved, final results made available. Summary report drafts begun.

Thursday
8:00 a.m. – 12:00 p.m.	**Panel Work Session**	**Powers**

12:00 p.m.   ADJOURN

Appendix 3: List of participants

Monday, February 24, 2014

Mike Armstrong-CIE
Benjamin Blount-SSC
Meaghan Bryan- SEFSC
Shannon Cass-Calay-SEFSC
Nancie Cummings- SEFSC
John Froeschke- Council Staff
Jeff Iseley- SEFSC
Neil Klaer-CIE
Michael Murphy-FWC
Anders Nielsen-CIE
Sean Powers-SSC
Ryan Rindone-Council Staff
Adyan Rios- SEFSC
Skyler Sagarese- SEFSC
Charlotte Schiaffo-Council Staff
Jessica Stephen- SEFSC
Greg Stunz-SSC
Jake Tetzlaff- SEFSC
Katyana Vert-Pre-UF

Tuesday, February 25, 2014

Mike Armstrong-CIE
Benjamin Blount-SSC
Meaghan Bryan- SEFSC
Shannon Cass-Calay-SEFSC
Nancie Cummings- SEFSC
John Froeschke- Council Staff
Jeff Iseley- SEFSC
Neil Klaer-CIE
Michael Murphy-FWC
Anders Nielsen-CIE
Sean Powers-SSC
Ryan Rindone-Council Staff
Adyan Rios- SEFSC
Skyler Sagarese- SEFSC
Charlotte Schiaffo-Council Staff
Jessica Stephen- SEFSC
Greg Stunz-SSC
Jake Tetzlaff- SEFSC
Katyana Vert-Pre-UF

Wednesday, February 26, 2014

Mike Armstrong-CIE
Benjamin Blount-SSC
Meaghan Bryan- SEFSC
Shannon Cass-Calay-SEFSC
Nancie Cummings- SEFSC
John Froeschke -Council Staff
Justin Grubiche-PEW
Jeff Iseley- SEFSC
Neil Klaer-CIE
Michael Murphy-FWC
Anders Nielsen-CIE
Sean Powers-SSC
Ryan Rindone-Council Staff
Adyan Rios- SEFSC
Skyler Sagarese- SEFSC
Charlotte Schiaffo-Council Staff
Jessica Stephen- SEFSC
Greg Stunz-SSC
Jake Tetzlaff- SEFSC
Katyana Vert-Pre-UF

Thursday, February 27, 2014

Mike Armstrong-CIE
Benjamin Blount-SSC
Meaghan Bryan- SEFSC
Shannon Cass-Calay-SEFSC
Nancie Cummings- SEFSC
John Froeschke -Council Staff
Neil Klaer-CIE
Anders Nielsen-CIE
Sean Powers-SSC
Ryan Rindone-Council Staff
Adyan Rios- SEFSC
Skyler Sagarese- SEFSC
Charlotte Schiaffo-Council Staff
Greg Stunz-SSC
Jake Tetzlaff- SEFSC