Independent Peer Review Report on the South East Data and Stock Assessment Review for Gray Snapper in the Gulf of Mexico (SEDAR 51)

Tampa, Florida

20 March – 22 March 2018

Peter Stephenson

18 b La Grange St, Innaloo, WA 6018, Australia
Email peterstephenson48@gmail.com

Representing the Center of Independent Experts (CIE)

Contents

Executive Summary:

The Review Panel for South East Data, Assessment, and Review (SEDAR51) met on March 20 – March 22 to examine the stock assessment of gray snapper (*Lutjanus griseus*). The assessment prepared by the Stock Assessment group, and some additional analyses suggested by the Review Panel, were presented and discussed. All panel requests were dealt with promptly and efficiently. Useful comments were received from other participants who attended the meeting.

After reviewing and discussing many aspects of the assessment, including data employed, model structure, base model, reference points, stock status, and short-term projections, the Review Panel concluded that the assessment appeared sound and robust to a number of changes in the input. The base case assessment was accepted and endorsed by the panel. The panel began drafting the Summary Report. As the Assessment Team needed more time to complete the projections, it was agreed the Summary Report would be completed two weeks after the close of the Review Meeting.

After addressing all the terms of reference, I concluded that the assessment provided a scientifically sound basis for determining the status of the gray snapper stock in the Gulf of Mexico.

The following are my recommendations for the gray snapper assessment:

- The use of a stock recruitment steepness set at 0.99 should be re-considered in favour of considerably lower value, say 0.75.
- The use of age varying natural mortality should be investigated. If the current method of *M* declining with age does not give a significant improvement to the model fit, then the simpler assumption *M* being constant over ages should be adopted.
- The uncertainty of the stock status derived parameters (spawning stock biomass and fishing mortality) could be made more realistic by allowing natural mortality and stock recruitment steepness to have some variation in the stock assessment.
- Confidence intervals should be presented for the model estimated abundance estimates and discard proportions.
- The results for the sensitivity analyses should be presented with confidence intervals.
- The derived parameters used to determine the stock status (spawning biomass and fishing mortality), including the projections, should have confidence intervals derived from a MCMC analysis.
- Data on length composition of discards should be collected.
- Different approaches to data weighting should be considered, such as effective sample size, and the model fit used to determine the choice of method to be adopted. The reasons for choice of data weighting should be documented.
- The problems with age composition should be investigated and these data included in the assessment if possible.
- If age composition data can be included in the assessment model, then growth parameters should be estimated within the assessment model, not externally.
- The reason for the recruitment fluctuation in the 1980s should be investigated, and a solution found which makes the estimated recruitment more plausible.
- Alternate catch scenarios for the period 1945 to 1979 should be investigated.
- The change in availability in 1990 due to the introduction of legal size limits should be incorporated in the selectivity functions.
- The slope of the selectivity function for the recreational fleet should be decreased to allow for retention of fish below legal size and discard of fish just over legal size. The data confirm both are occurring.

2. Background

2.1 Overview

A Review Panel meeting for the South East Data, Assessment and Review of the gray snapper fishery in the Gulf of Mexico (SEDAR51) was held at the Gulf of Mexico Fishery Management Council Office, 2203 N Lois Ave, Suite 1100, Tampa, Florida on March 20-22, 2018. The Review Panel comprised chairperson, Dr Kai Lorenzen, Dr Luiz Barbieri, Bob Gill, and three appointees of the Centre for Independent Experts (CIE), Dr Yong Chen, Dr Laurance Kell, and Mr Peter Stephenson. The attendees of the meeting and their affiliated organizations are shown in Appendix 5. The agenda for the review is shown in Appendix 6.

Prior to the SEDAR 51 Review Meeting, the stock assessment report, supporting documents (Appendix 1), background reports (Appendix 2), and files used in the assessment model (Appendix 3) were made available to Panel members.

2.2 *Terms of Reference*

The terms of reference for the SEDAR51 stock assessment review are presented in Appendix 4. The statement of work provided by the CIE is presented in Appendix 5. This required that the SEDAR51 Review Panel members participate in an independent peer review of the assessment, assist the Review Chairman in preparing a Summary Report of the review, and also prepare an independent CIE report of the assessment and the review process. This report contains my evaluation of the assessments and the review process.

2.3 ***Acknowledgments***

My thanks are expressed to the various individuals who participated in the review meeting making the review interesting and informative. The presenters, Jeff Isley and Shannon Cass-Calay, are to be commended for the quality of their stock assessment. My thanks to the review panel members for the quality of their advice, industry participant Ed Walker, and SEDAR staff, Julie Neer and Ryan Rindone, for their useful inputs.

2.3. ***The review activities.***

Prior to the review meeting I became familiar with the background documents, the assessment report (shown in Appendix 3) as well as the SS3 data, control, and forecast files. The review committee for SEDAR51 met 22-24th March 2018 in Tampa, Florida. Those attending the review meeting were five review panel members, two members of the stock assessment group, one industry member, and three SSC personnel. The data inputs and assessment outputs were presented to the panel and several sensitivity runs requested by the panel were carried out and discussed. After consideration of the results of all the sensitivity runs, the panel suggested that the base case was the one that should be adopted. The panel began drafting the Summary Report but this was postponed for two weeks as the assessment team needed some time to improve the implementation of the projections before the Summary Report could be finalised.

4. Review relevant to the terms of reference

Here, I have presented my views of the assessment.

TOR1. Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions, and consider the following:

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

I conclude this TOR was successfully met.

Catch data. This data was not included in the datasets which are publically available on the website as there are confidentiality issues with some data. The Assessment Report documents the catch and briefly outlines the methods used to come up with catch categories. The data set contained the annual catch from six sources:

- commercial "handline" (Monroe County), denoted CHL_MC and commercial "handline" (non-Monroe County), denoted CHL_nMC, both of which pooled all commercial gears except longline,
- commercial longline, denoted CM_LL,
- recreational private, denoted REC_PR,
- recreational shoreline, denoted RECShore
- recreational headboat/charterboat, denoted REC_HB+CB.

In recent years, about 67% of the catch has been recreational private and about 12% recreational shoreline. The recreational catches prior to 1981 are not available and as these are the major component of the catch, the Data and Assessment team decided these should be filled in for the years 1945 to 1980 using the sensible approach of basing the catch on the effort expended over this period. I believe the approach used is sensible and the assumptions made are justified.

I believe there is merit in using alternate catch histories for the period 1945 to 1980 (one say 20% higher and one 20% lower) as a way of investigating the impact of variable catches on the biomass trajectories. An alternative, discussed briefly at the review, was to not use any data prior to 1981. The initial year would then be 1981, with an assumed fishing mortality in the first year. I do not recommend this approach, as the selection of this initial F would be totally arbitrary and the value chosen for this initial F would have a major impact on the comparison of the current biomass with SSB30.

Recreational fishing is the major source of the catch in this fishery and the spatial resolution of these data is poor, often reported as "Florida" or "The Gulf". If it was possible to improve the reporting of the spatial distribution of the catch, it may be possible to eventually develop historic spatial catch information based on interviews and expert opinion. A stock assessment model with a spatial component is likely to significantly improve the performance of the assessment model.

Abundance Indices.

There are eight abundance indices used in the assessment model. The data team developed standardized abundance indices for all the indices. The methods used were well documented and consistent with best practice.

Three of the indices are fishery-dependent.

- Commercial handline standardized catch-per-unit effort, denoted CHL, with data from 1993-2015. A probability threshold for inclusion of a trip was 0.34 and a delta lognormal generalized linear model was used for the analysis.
- Recreation data for private vessels, denoted MRFS_Private, with data from 1991-2015.
- Recreation shore based fishing from bridges, jetties, rocks, and shore, denoted MRFSS_Shore, with data from 1981-2015.

The diagnostics indicate that these three standardized indices were a good fit to the data.
There were five fishery independent surveys used in the assessment model.

- Southeast Area Monitoring and Assessment Program (SEAMAP) Reef Fish trawl survey denoted SEAMAP_Trawl with data from 2010-2015, although data was available in some areas back to 1987. The years from 2010-2015 were selected for the gray snapper analysis as these years covered a larger geographic area and common survey methods. The abundance index developed covered the area from Panama City to Florida Keys. The delta lognormal analysis was appropriate and the model diagnostics indicate the standardized index was a good fit to the data.
- Combined index from three stationary video surveys, denoted NMFS/FWRI_fixed with data from 1993-1994, 2004-2015. The surveys comprising this index were the NMSF Pascagoula laboratory video conducted from 1992-1997 and then 2004 to present, the NMFS Panama City laboratory video survey from 2004 to present, and the Florida Fish and Wildlife Research Institute video survey which commenced in 2008. The data were restricted to years with similar methods. A habitat index was used in the analysis. The model description and the diagnostics were well documented and the standardized index was a satisfactory fit to the data.
- Diver census survey, denoted Visual Survey, consisted of data from 1997-2015. This was restricted to Monroe County. The method of analysis was well documented and accounted for the number of zero counts. The standardized index appeared to be a rather poor fit to the data.

All of the abundance indices were scaled to the mean before they were passed to the Assessment Team. The methods used were justified and well documented, the methods of analysis were appropriate. Asymptotic confidence intervals were presented for all the indices. The data teams are to be commended for their efforts.

Discards

The recreational sector is the major component of the catch (nearly 80% of the recent catch) and these sectors fish in the shallower water where small fish are abundant. Consequently, the discards of undersized fish is a major driver of the dynamics of this fishery. A discard mortality rate of 6.5% is used in the assessment model for the recreational sector and 14% for the commercial sector.

The discards are formulated as a percentage of the catch. The observed discards for the recreational shore and recreational private were considered very uncertain, especially after 1989 and for these latter years a very large CV of 0.5 was allocated.

A commercial fisher present at the review meeting stated that the discard of gray snapper in the commercial fishery was close to zero. He said that very few undersized snapper are encountered where they fish, and gray snapper are always kept as they are always in demand at the markets and are never discarded to give preference to other species.

Length Composition data

Length data are available for the six fleets and the information is well documented and information on the location of the samples is available. The data had regional differences and the data team recommended that it be stratified at least into Monroe County and non-Monroe County. In 1991, legal sizes were introduced, with a minimum legal size of 10 inches for State waters and 12 inches for Federal waters.

Age composition

The age data were supplied to the stock assessment team but was not used in the fitting process of the stock assessment model. These data were found to have serious inconsistencies, for example, when the data were grouped by year, there was a distinct bimodal distribution of length at age. Time constraints for the assessment finalization meant that these inconsistencies could not be reconciled. The data supplied to the assessment group

apparently do not, in all cases, have the same codes for species, FL, TL, and the data sets contains sizes in inches, mm, and cm. Any inconsistency in the units of measurement could cause the problem observed in the data. Alternately, the data units and species may be completely correct, and the problem could be attributed to different growth rates from different areas of the fishery. The data from the recreational fishery often do not have good locations recorded and grouping the age-composition data into regions may not be possible for all the data.

The problem with the age data is unlikely to be due to misreading of annuli. The otoliths are read in a number of laboratories with the ageing staff meeting regularly (at least annually) to compare annuli allocation. There is an excellent ageing protocol document and a 100 item reference collection which are used by ageing staff to ensure consistency of annuli allocation. The sectioned otoliths of gray snapper are not difficult to age (my observation and the comments of SSC staff) and there is compelling data to indicate the annuli observed are annual bands. I believe the problems with the age composition data lie elsewhere than the allocation of annuli counts to the otoliths.

The age-composition data, pooled for all years, were used to generate von Bertalanffy growth parameters. Although, the bi-model pattern could be seen in the plot of the data, the assessment team believed that it was appropriate to use it to generate the growth parameters. Although there is large scatter in the lengths at age, this was not different than I have observed for many other species. I agree that the use of the pooled data for the determination of growth parameters is appropriate.

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

I conclude this TOR was successfully met.

The data uncertainties are well reported with asymptotic confidence limits provided for all datasets. The uncertainties are generally at levels that would be expected.

It was acknowledged that the age composition data had serious problems, with the obvious problem of bimodal distribution of lengths at age. Consequently, it was not used in the assessment model. This needs to be investigated so that the anomalies can be understood and hopefully corrected.

It was acknowledged that the discard data for each of the fleets had limited value as it lacked length composition, depth, and spatial location, making it difficult to apply realistically in the model. The discards data after 1989 was considered so uncertain a CV of 0.5 was allocated.

c. Are data applied properly within the assessment model?

I conclude this TOR was successfully met.

Length-composition

The model fit to the length composition data was good with the observed and model predicted distributions almost identical in most years. The Assessment Report indicates there is a lack of model fit in some fleets which could be attributed to the change in regulation in 1991 when a legal size of 10 inches was introduced in State waters and 12 inches in Federal waters. The change in availability in 1991 was not built into the assessment model, but could quite easily be accommodated in SS3 in a future assessment. It is extremely unlikely to change the spawning biomass or fishing level in the current years, but should be included to make the assessment model better represent the dynamics of the fishery.

Age-composition

The very large number of age composition samples (32000) in combination with the length composition data would greatly enhance the assessment model. The weight at age parameters could be estimated inside the model. It should be a priority to determine the cause of the unusual pattern in the data, fix the problem if possible and incorporate these data in the next iteration of the gray snapper assessment in 3-5 years

Abundance indices in the Assessment Model

The data group and the assessment group supported equal weighting of the abundance indices in the assessment. To implement this, the Assessment Team adjusted the CVs in each dataset so that the mean CV of each dataset was 0.2 but the inter-annual variation was preserved. Thus each dataset had equal weighting in the assessment model. The approach is appropriate and is well documented.

The abundance indices are rather flat but consistently show an increase in recent years. The assessment model fit to the data is good and is picking up the upward trend in recent years. This is an indication that the abundance indices are well formulated in the model. There are no asymptotic confidence intervals for the abundance indices shown in the assessment report and the confidence intervals should be included in future assessments.

The concern with all these indices is their restricted geographic range. In the assessment model, it is assumed that they represent the abundance in the whole Gulf of Mexico. The Data Team and the Assessment Team were unable to obtain spatial information on catch for the current assessment and made the pragmatic decision to build a one area model. Given the time constraints, I agree that this is an appropriate approach given the data available.

It appears that areas where the abundance indices are concentrated are areas where there is likely to be the greater recreational fishing pressure (eastern gulf and Monroe County) and this strengthens the appropriateness of this approach.

Selectivity, discard, and release mortality

The catch data and length composition data were divided into six fleets in the assessment model as each is expected to have different selectivity, discard, and mortality characteristics. The six fleets were modelled with different dome shaped selectivity functions.

The selectivity (availability) and the discard function are multiplied to give a dome-shaped retention function. The approach used is clearly documented in the SS3 Technical Manual and in the SS3 control file supplied to the Review Panel. The estimation of the parameters for a logistic and more particularly a dome-shaped functions is often very difficult, especially when the parameters are not of similar magnitude. This problem was encountered to some extent in this assessment with the parameters' magnitudes varying from 2 to 50. The assessment team put considerable effort into coming up with sensible results by estimating some parameters and fixing others in the selectivity function.

The retention parameters and discard mortality parameters were all fixed. The approach used is well documented in the SS3 control file, well considered, and appropriate. The graphs of mortality shown in Figure 4.2.3.3 and Figure 4.2.3.5 are confusing with graphs of deaths at length being fixed values of 0.65 and 0.35 respectively. Although they are correct, due to the availability in the recreational sector, I think it would be useful, if possible, to choose parameters which would make the concepts clearer to the reader.

In summary, the data were properly applied with sensible assumptions made for use of the data in the assessment model.

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

I conclude this TOR was successfully met.

Discards

- The assessment model fit to the recreational discard fraction was very poor, especially after 1989 with the model estimate being very much lower than the observed.
- For commercial handline (Monroe County, where the allocated CVs of the data were very high, the model fit was very poor with the model estimates being about 20% of the observed values.
- For Commercial handline (non Munroe County) the CVs of the observed data were low but the model fit was still very poor with the model greatly under-estimating the observed data.

The confidence intervals for the model fit were not provided, so it is not known if the data values fall within the model estimated confidence intervals.

It is not known if the model fit would improve if the CVs in the observed data were reduced. This was not investigated in the review and I feel it would not greatly improve the model fit to the discards.

The assessment team conducted a sensitivity analysis on the discard mortality, increasing the natural mortality of the recreational fleet from 6.9% to 11.9% and commercial discard mortality from 14% to 19%. The model fit to the discards improved slightly, but the review panel thought the change was not sufficient to justify changing the base case.

The model was run with the abundance indices and length composition data not used in the model fit. The fit to the discard data was not greatly improved and spawning biomass was very low after about 1987.

The discard information is currently not very informative for the assessment process. The discards need to have associated length information and discard mortality should be more detailed being related to depth and region. The discard fractions which are so uncertain that they are allocated a CV of 0.5 could possibly be omitted from the data used in the assessment model.

Indices of abundance

It is a concern that the data were collected in spatially specific areas but applied to the whole fishery. Consideration in the future should be given to running an area specific model, even if assumptions need to be made about the spatial distribution of catch.

TOR2. Evaluate and discuss the strengths and weaknesses of the methods used to assess the stock, taking into account the available data, and considering the following:

a. Are methods scientifically sound and robust?

I conclude this TOR was successfully met.

- The assessment is conducted in SS3 Version 3.24. This is a modern software package for conducting age based statistical models. It is used extensively in the United States and Australia. It has a great advantage as an assessment platform as modellers and reviewers know it is reliable and has been extensively tested, and revised. Time does not need to be spent painstakingly checking “home built” statistical model code, which all too often it found to have small bugs that have not been detected even over many years.

- The disadvantage of SS3 is that it is not obvious what values should be used in some places in the data and control files. In addition, it is sometimes difficult to get the structure of SS3 to match the peculiarities of your particular fishery. The structure you require can often be accommodated in the current model or features can be added, but this takes time to resolve. In addition, some of the parameters in the gray snapper fishery are not easily estimated, for example the double-normal selectivity function where some of the parameters are very different in magnitude (2 and 50), which sometimes makes their estimation problematic.

- The assessment model used was SS3 and all the files needed to run the model were supplied to the Review Panel with sufficient annotation to make the process used clear. The model was run successfully by Laurie Kell and the results in the assessment report were easily reproduced.

In summary, the use of the statistical model SS3 is best practice and its use in this assessment is commended.

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

I conclude this TOR was successfully met.

The files required to run the base case of the model in SS3 were supplied to the review panel (data, control, and forecast). The files were well annotated and the model was easily run and could be used to examine how sensitive the model was to changing the data, test hypotheses, and develop diagnostics.

Natural mortality

The approach in the assessment is to determine a natural mortality for the old fish using the method of Hoenig (1983) using a maximum age of 28. The oldest gray snapper observed in this fishery is 31 or 32 years old. It is quite common to take the maximum age as some proportion of the oldest observed fish (say 0.9 or 0.95). My preference is to use a conservative approach: the maximum age is the age of the oldest fish observed. The sensitivity analysis, where using M values of 0.1 and 0.2 did not make large differences to the spawning biomass in recent years, indicates that taking a maximum age as 28 or 31 is of little consequence.

More importantly, I believe that using an age-dependent M, similar to that in Lorenzen (1996) is unnecessarily complicated. If there is no significant improvement in the fit to the model using the age dependent M, I believe that the more parsimonious assumption of constant natural mortality should be used. If a constant M was adopted, it would be a simple matter to introduce variability in M using a prior (Hoenig 1983) and a small CV that would constrain M to a level consistent with expert opinion. By introducing variability in parameters in the model, a more realistic level of uncertainty can be introduced into the derived parameters, like spawning biomass and fishing mortality.

Recruitment

The SS3 model outputs of the annual recruitment (age 0 fish) showed some disturbing patterns in the 1980s. There is a pattern of extremely high recruitment in 1984, and extremely low (effectively 0) in 1985 and 1986. The pattern of recruitment jumping to extremely high value one year, and dropping to 0 the next year demonstrates that there is a problem with the estimation process or there is some signal it the data causing the chaotic behaviour. The Review Panel requested a plot of model estimated recruitment (0 year old fish) against the observed data on 0+ recruits. The fit was good in most years, but the survey data did not extend back to the years when the chaotic behaviour occurred.

The stock assessment team suggested that because the recruitment deviation estimated in the model must sum to 0, if an extremely high recruitment occurred, model may then estimate the one in the next year close to zero to compensate. I believe it is highly unlikely that SS3 would behave like this, and I have not observed this

recruitment behaviour in SS3 or "home built" statistical models. Generally, if there is no information, the recruitment deviations would be zero. I believe the model is picking up a signal in the data, possibly the length composition, resulting in this pattern of recruitment deviations.

The pattern of recruitment in the 1980s is a problem that needs to be addressed. Despite this, the recruitment appears sensible and matches the observed data in the years data are available. The recruitment anomalies in the 1980s is unlikely to affect the biomass or fishing mortality in the more recent years and would be unlikely to affect the advice on stock status.

It should be a priority to investigate the cause of the recruitment peculiarity before the next assessment in 3-5 years.

Steepness of the stock recruitment relationship

It is common that the available data do not enable the steepness to be estimated. In this assessment, steepness could not be estimated and its value is set at 0.99. This is not based on the assumption that there is no stock recruitment relationship, but rather, the assumption the value cannot be estimated. For the projections, the recruitment in the next 2 to 3 years is assumed to be similar to that in the current year.

An alternative assumption, which I think is better than that currently used in the assessment, is to fix steepness to a default value (in Australia a value of 0.75 is common). Another possibility with considerable merit, is to estimate steepness with a prior of say 0.75 and a CV of say 0.2 and a penalty that will constrain the steepness to a sensible value. By doing this, the uncertainty in the steepness estimate can be incorporated in the model.

Data Weighting

The eight standardised abundance indices were scaled to have a mean of one by the Data group. The Assessment group scaled the CVs within each dataset so that the mean CV was 0.2. When parameters were estimated in the model, the abundance indices had the same weighting.

There are six length composition data sets, one for each of the six fleets. The number of observations in a year ranges from 17 to 2908. A rule is applied to all of the abundance datasets to cap the number of observations to 100. Thus each annual data set with more than 100 observations is allocated a sample size of 100.

When the model is run, the effective sample size is calculated by SS3 using the default method of calculation. This produced six scaling factors, one per dataset, which is applied to each year of the dataset. In Francis (2011) there are eight well documented methods of calculation of effective sample size. The default method used by SS3 is TA1.8.

The effective sample size calculation is an iterative process and the model needs to be run, the newly calculated effective sample size model pasted into the data file and the model re-run. This is not too onerous as it generally takes only 4-6 iterations.

The method used for the gray snapper assessment is to calculate six values of lambda to scale the allocated sample size of 100. The calculated values of lambda, in the base case ranged from 0.14 to 0.23. When applied to the data sets, which in most cases had capped sample sizes of 100, the effective sample size, in most years ranges from 14 to 23. When the sample size is very low, say 17 fish measured, the effective sample size used in the fitting process could be as low as 2.5. I believe the method used for the gray snapper assessment is well thought out and is acceptable.

I have used the method TA1.8 (Francis 2011), but I have a preference for method from McAlaster and Ianelli (1997) denoted TA1.4 in Francis (2011) as it is easy to understand and is simple to implement. In my experience, TA1.4 generally gives a lower effective sample size than TA1.8. On suggestion from the Review

Panel, the assessment team promptly ran the SS3 using method TA1.4 using a switch "MI" to activate the calculation.

The results indicated that the spawning biomass showed little difference compared to the base case and certainly would not change the advice on stock status. The panel suggested that it would be more informative to look at the model fit, rather than just SSB, but for the purpose of this review, the Review Panel suggested that the base case should be the one used.

Selectivity

The double logistic selectivity function has an upward slope close to one. This formulation does not fit the data well as fishers are retaining fish below legal size, which was not accommodated in the model structure. In addition, the discard function was not fitting well. One of the problems here was the assumption that all fish over legal size are retained. With the bag limit of five gray snapper, fish over legal size are discarded because the bag limit is often quickly achieved (personal communication, recreational fisher). I suggest that decreasing the slope of the retention function would allow the retention of catch of fish below legal size and the discard of fish above, but close to, legal size.

c. Are the methods appropriate for the available data?

I conclude this TOR was successfully met.

Stock definition

- The tagging study indicated that there is only small movement of gray snapper (generally no more than 5 miles). There is no information on otolith chemistry to complement the tagging study, but given the large number of otoliths collected it could be possible to do this at low cost.
- Information on movement of eggs and larvae is poorly known, but the large number of inlets and estuaries and the recruitment studies in Monroe County, would suggest there are many sources of recruitment to the fishery.
- Given the difficulty of obtaining spatial information on catch and different spatial extent of the abundance indices, the stock definition for the assessment must by necessity be a pragmatic compromise, and I endorse their decision to build a one area model.

Model structure

The panel discussed the possibility of using a simpler model. As discards are a driving force in this fishery, a model incorporating length-composition information and a more detailed formulation of discards and discard mortality appears necessary. A length-based Stock Synthesis model may be available shortly and its use in this fishery could be explored.

TOR3. Evaluate the assessment findings and consider the following:

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

I conclude this TOR was successfully met.

Jitter Analysis: The starting values were jittered to investigate how often the model failed to converge. The results presented to the Review Panel indicated very good model stability with convergence failure only five times in 100 trials.

The spawning biomass estimates were examined for sensitivity to changes in the input values. In none of the cases were the asymptotic confidence intervals presented.

Sensitivity analyses were conducted by the assessment group before the review panel meeting. These included

- high, and low natural mortality.
- removal of abundance indices one at a time.
- five year retrospective analysis.
- increased discard mortality.
- lambda for effective sample size set to zero.

Natural Mortality: The model was run with natural mortality M=0.15 (base case) and also M=0.1 and M=0.2. The biomass trajectories were respectively lower with low M and higher with high M as you would expect. The biomass plots had been scaled to the same point in the initial year. The change in natural mortality from the base case was considerable, the median biomass was not sufficiently different from the base case to be of concern. The asymptotic confidence limits of biomass trajectories were not presented.

Abundance indices: The abundance indices were removed one at a time from the assessment model. There was little change in the trajectory of the spawning biomass. The greatest change in the spawning biomass was when the commercial longline data were removed.

Retrospective analysis: The model was run five times, successively taking out all the data for the last year. The results showed only small variations in the spawning biomass trajectory. The Review Panel suggested that the results should be scaled to the same initial value. When this was presented, the curious pattern became clearer, where the closest agreement was between the base case and removal of 5 years of data and there was not a consistent pattern in the trajectory. No explanation was obvious, and the anomaly was not considered a serious problem. I consider the model was fairly robust to sequential removal of data.

Discard mortality: Discard mortality was increased from 6.9% to 11.9% and the spawning biomass changed very little from the base case.

Lambda=0: This case, not fitting to the age data, resulted in the greatest change in the spawning biomass with very low level in 1990 and then a rapid decrease to a level almost twice that of the other cases in 2015.

For all the sensitivity analyses conducted, except the last when lambda=0, the spawning biomass varied only slightly and the stock status conclusions were unchanged. The model was robust to the changes introduced and I endorse the use of the base case values for the assessment in the changes.

The panel suggested that they should all be scaled to same initial point. This was done promptly and efficiently. The panel suggested several other sensitivity analyses to the assessment team.

- Remove all the fishery dependent indices of abundance.
- Investigate high and low historical catch scenarios.

The analyses suggested by the panel were completed efficiently and promptly.

Abundance indices: The removal of all the fishery dependent abundance indices resulted in a slightly better model fit to the observed data (no confidence intervals presented) and a slightly lower spawning biomass in recent years. The Review Panel decided that the base case, with all indices included, should be adopted.

Historic catches: On discussion between the Review Panel and the Assessment Team, it was decided, based on experience in other fisheries, that this would probably have little impact on the model results and should not be done at this time.

The Review Panel mentioned that these sensitivity analyses were important investigations to see how robust the model was to changing the assumptions. It is generally not necessary to pass them on to policy decision makers.

After consideration of all sensitivity analyses, the Review Panel concluded that none of the cases considered made sufficient difference to the conclusions drawn from the analyses to warrant changing from the base case.

In conclusion, I consider the assessment model is robust to changes in the input parameters and the base case gives stable and robust representation of the spawning biomass and fishing mortality.

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

I conclude this TOR was successfully met.

The SSB in 2015 is 4660 mt with 95% asymptotic confidence intervals of 3323 mt to 5997 mt. Using the current definition of Minimum Standing Stock Threshold, MSST=0.5xSSB30=3330 mt, the gray snapper stock is not overfished. In fact, the lower 95% confidence interval of SSB in 2015 is just slightly below MSST.

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

I conclude this TOR was successfully met.

The fishing mortality in 2015 is 0.1347. The Maximum Fishing Mortality Threshold, MFMT=FSPR30=0.1150, thus the stock in 2015 is undergoing overfishing. The ratio of *F* to MFMT has been above 1.0 for most years since 1976 (Table 4.2.9.1 in the Assessment Report) indicating the gray snapper has been undergoing overfishing since 1976. I endorse the conclusion that stock is being overfished now and has been since 1976.

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

I conclude this TOR was successfully met.

In the assessment model, the steepness of the Beverton and Holt stock-recruitment relationship is taken as 0.99. The assessment group, on consideration of the relationship between predicted recruitment and spawning biomass (Figure 4.2.4.1), decided to run the assessment with no relationship between recruitment and stock size. The projections are constructed on the assumption that the recruitment in the next 3 years will be the same as the current recruitment, which is close to the initial recruitment.

In most assessment models, it is not possible to estimate the steepness parameter of the stock-recruitment, but fixing it at 0.99 seems to me to be unwise, as discussed above. Overall, the model predicted recruitment matches the data, and for the current assessment, I endorse the decisions made.

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?

I conclude this TOR was successfully met.

The current definition of Minimum Standing Stock Threshold (MSST) is 50% of SSB30 and the definition of Minimum Fishing Mortality Threshold, MFMT is FSPR30. These quantities are set up in the forecast file of SS3 and reliably determined by the assessment model.

TOR4. Evaluate the stock projections, including discussing strengths and weaknesses, and consider the following:

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

I conclude this TOR was successfully met.

The methods used to calculate the projections in SS3 are well documented and represent the best current practice. As the stock is not overfished, a recovery plan is not required and the projections were run with the

fishing mortality set at the FMSY proxy of FSPR30, the value specified by GMFMA. Future recruitment was set at the average recruitment in recent years (1990-2015).

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

I conclude this TOR was successfully met.

The method used to generate the projections was appropriate. The SS3 model carries through the values of the estimated parameters and their uncertainty into the projection period, with the specified fishing mortality and recruitment. The results and asymptotic confidence presented are appropriate and realistic.

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

I conclude this TOR was successfully met.

The assessment model projections are based on the assumption that recruitment, selectivity, and retention patterns will not vary from that in recent years. These assumptions are reasonable and I endorse them.

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

I conclude this TOR was successfully met.

The uncertainties of the projections were expressed as asymptotic confidence levels. In the next assessment, in 3 -5 years, when the final projections are produced, the confidence intervals should be determined from an MCMC analysis. For the current assessment, I endorse the method used in the assessment.

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

I conclude this TOR was successfully met.

Extensive exploration of the data sources and the methods used in the assessment include: jitter analysis, sensitivity analysis of natural mortality, discard mortality, abundance indices, and retrospective analysis. All these analyses did not greatly alter the current spawning biomass levels and did not alter the conclusion about the current stock status.

No different model configurations, like a multiple area model, were explored, but this is something that can be considered in the next assessment.

Some parameters were fixed in the model (natural mortality, steepness of the stock recruitment relationship, some selectivity parameters, discard mortality, discard parameters, and growth parameters). It is consistent with standard practice to fix many of these parameters but the cost is that the uncertainty in the quantities of interest, like spawning biomass and fishing mortality, will be underestimated.

Wherever possible, parameters should be estimated within the model, not estimated outside the model and brought in with an allocated CV (e.g. growth parameters, weight at age). The variation in *M* and steepness can also be easily incorporated in the assessment, particularly in the projections, to make the uncertainty more realistic.

Age composition data contained unusual features like the bimodal distribution in length at age. When attempts were made to incorporate the age data, spurious results were produced by the assessment model. The assessment group made the sensible decision to not use the age data in this assessment. When the age composition anomalies are remedied, it may be found there are spatial differences in the growth parameters which can be incorporated in the assessment.

The confidence intervals for the spawning biomass and fishing mortality were illustrated as 95% asymptotic confidence intervals. When sensitivity analyses are being explored, this is a convenient way to express the uncertainty, but when the base case has been decided, the confidence intervals should be presented as confidence intervals from an MCMC analysis. It is expected they will not be symmetric, especially at the lower spawning stock sizes around 1990.

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

I conclude this TOR was successfully met.

Many sensitivity analyses were examined and the Review Panel concluded that the model was robust to these various alternatives, giving similar spawning biomass results. For most of these alternatives, the status of the stock still classified as "not overfished".

The confidence intervals of the model fits to abundance indices needs to be shown on the graphs. For the discard fraction, asymptotic confidence intervals in many years would be expected to be so large as to be meaningless.

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

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

- Improve the spatial resolution of the catch, especially in the recreational sector. If the catch could be split into regions in future years, it may be possible, using expert opinion, to get a sensible distribution of past recreational catches.
- Discard information needs to contain length composition and spatial information, especially depth.
- The release mortality rates should be more detailed, extending over a broader depth range.
- The abundance indices should be extended to cover more of the fishery.
- Resolve the age composition data problems and investigate the spatial variability of weight at age.
- If spatial catch and age composition data can be developed, an area and age based statistical model should be developed.

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

- It would be useful to have an introductory session on the biology of the species and the operation of fishers in its capture.
- It would be useful for someone to be available during the Review who participated in the Data Workshop and Assessment Workshop so that Review Panel questions can be more easily addressed.
- The files used in the assessment model should be supplied as they are used in the model, not as pdf files.
- It would be useful for the reviewers and also for the assessment presentation, to have two assessment documents, one with the methods and explanation and another with the figures so they can be viewed simultaneously.
- Improve communication between various groups within the SEFSC to improve data and assessment products.

TOR7. Consider whether the stock assessment constitutes the best scientific information available using the following criteria as appropriate: relevance, inclusiveness, objectivity, transparency, timeliness, verification, validation, and peer review of fishery management information.

- All of the review documents were relevant to the fishery and contained sufficient information to be informative. The data supplied for the modelling process were complete and well annotated. The age composition data was not supplied and it was clear early in the process that this data would not have been useful for the review process.
- The assessment team were transparent and candid about problems with the data and with the model fitting process. After discussion, the Review Panel agreed that sensible and pragmatics decisions were made on how the data should be used in the assessment.
- The background documents were supplied to the Review Panel early in the process but the Assessment Document and the assessment model files were rather late. There were good reasons for this and it did not detract from the smooth running of the review.
- The SS3 model could be run easily from the supplied files and the model outputs, which matched those of the base case supplied by the assessment team, were easily produced.
- It would be useful for the Review Panel to have a description of the SEDAR process for data analysis, assessments, and the post assessment Council Process.

I conclude that the gray snapper assessment constitutes the best scientific information available on the stock's status. The spawning stock biomass and fishing mortality were clearly presented with confidence intervals, reference points were clearly articulated, and stock status was determined. Based on this stock status, projections were generated.

TOR8. Provide suggestions on key improvements in data or modelling approaches that should be considered when scheduling the next assessment.

- Currently the discard data are just the proportion of the catch discarded. The discards data should be improved so that it includes the length composition of the discard for the various fleets. This information would be a major improvement in the model structure and would be expected to greatly improve the model fit to the discard data.
- The age-composition data need to be thoroughly investigated and the problems resolved. Hopefully the age-composition can be incorporated into the model for the next assessment. Fitting to length composition and age composition data would be a major improvement in the assessment. In addition, the weight-at-age can be estimated inside the model with the associated uncertainty flowing into the estimates of spawning biomass and fishing mortality.
- The discard mortality rates should be investigated to see if they can be applied by depth rather than by just recreational and commercial.
- The selectivity function should be revised so they incorporate the change in the legal size in 1991.
- I consider fixed natural mortality to be a simpler and preferable approach to an age dependent natural mortality.
- The fixing of stock-recruitment steepness to a lower value should be considered.
- The recruitment in the 1980s should be investigated to see if it is being driven from information in the data, for example the length composition data.

TOR9. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

The Peer Review Summary Report has been prepared.

5. References

Francis, R.I.C.C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences 120*: 221-230.

McAllister M. K., Ianelli J. N. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm, *Canadian Journal of Fisheries and Aquatic Sciences*,1997, vol. 54 (pg. 284-300).

Lorenzen, K (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. *Journal of Fish Biology 8: 627-647.*

Appendix 1. Review Workshop Terms of Reference

1. Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions, and consider the following:
   a. Are data decisions made by the DW and AW 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 and discuss the strengths and weaknesses of the methods used to assess the stock, taking into account the available data, and considering the following:
   a. Are methods scientifically sound and robust?
   b. Are assessment models configured properly and consistent with standard practices?
   c. Are the methods appropriate for the available data?
3. Evaluate the assessment findings and consider 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, including discussing strengths and weaknesses, and consider the following:
   a) Are the methods consistent with accepted practices and available data?
   b) Are the methods appropriate for the assessment model and outputs?
   a. Are the results informative and robust, and useful to support inferences of probable future conditions?
   c) 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 Data and Assessment workshops 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.

8. Consider whether the stock assessment constitutes the best scientific information available using the following criteria as appropriate: relevance, inclusiveness, objectivity, transparency, timeliness, verification, validation, and peer review of fishery management information.

8. Provide suggestions on key improvements in data or modeling approaches that should be considered when scheduling the next assessment.

9. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

*This assessment will follow a Benchmark-track approach.

Appendix 2: Documents provided for the Review

SEDAR51-DW-01 Brief Summary of FWRI-FDM Tag-Recapture Program and Brief Summary of FWRI-FIM Tag-Recapture Data. Rachel Germeroth, Kerry Flaherty- Walia, Beverly Sauls and Ted Switzer. Submitted 4 Nov 2016.

SEDAR51-DW-02 Summary of length and weight data for gray snapper (Lutjanus griseus) collected during NMFS and SEAMAP fishery independent surveys in the Gulf of Mexico. David S. Hanisko and Adam Pollack. Submitted 20 March 2017.

SEDAR51-DW-03 Gray Snapper Abundance Indices from SEAMAP Groundfish Surveys in the Northern Gulf of Mexico Adam G. Pollack, David S. Hanisko and G. Walter Ingram, Jr. Submitted 7 April 2017 and updated: 12 June 2017.

SEDAR51-DW-04 Length frequency distributions for gray snapper length and age samples collected from the Gulf of Mexico Ching-Ping Chih. Submitted 9 April 2017.

SEDAR51-DW-05 Gray snapper *Lutjanus griseus* Findings from the NMFS Panama City Laboratory Camera Fishery-Independent Survey 2005-2015 C.L. Gardner, D.A. DeVries, K.E. Overly, and A.G. Pollack. Submitted 7 April 2017.

SEDAR51-DW-06 Reproductive parameters for the Gulf of Mexico gray snapper, Lutjanus griseus, 1991-2015 G.R. Fitzhugh, V.C. Beech, H.M. Lyon, and P. Colson. Submitted 13 April 2017.

SEDAR51-DW-07 SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Grey Snapper. Matthew D. Campbell, Kevin R. Rademacher, Michael Hendon, Paul Felts, Brandi Noble, Ryan Caillouet, Joseph Salisbury, and John Moser. Submitted 10 April 2017.

SEDAR51-DW-08 Description of age data and estimated growth of Gray Snapper from the northern Gulf of Mexico: 1982-1983 and 1990-2015 L.A. Thornton, L.A. Lombardi, and R.J. Allman. Submitted 14 April 2017.

SEDAR51-DW-09 SEDAR 51 Stock ID Working Paper S51 Stock ID Working Group. SubmittedFebruary 2017

SEDAR51-DW-10 Indices of abundance for Gray Snapper (Lutjanus griseus) from the Florida Fish and Wildlife Research Institute (FWRI) video survey on the West Florida Shelf Kevin A. Thompson, Theodore S. Switzer, and Sean F. Keenan. Submitted 21 April 2017.

SEDAR51-DW-11 Gray Snapper Abundance Indices from Inshore Surveys of Northeastern Gulf of Mexico estuaries. Kerry E. Flaherty-Walia, Theodore S. Switzer, and Amanda J. Tyler-Jedlund. Submitted 24 April 2017. Updated: 27 April 2017.

SEDAR51-DW-12 Standardized Catch-Per-Unit Effort Index for Gulf of Mexico Gray Snapper *Lutjanus griseus* Commercial Handline Fishery (1993 – 2015). Matthew W. Smith. Submitted 26 April 2017.

SEDAR51-DW-13 Commercial Landings of Gray or Mangrove Snapper (*Lutjanus griseus*) from the Gulf of Mexico. Refik Orhun and Beth Wrege.

SEDAR51-DW-14 Standardized Reef Fish Visual Census index for Gray Snapper, *Lutjanus griseus*, for the Florida reef track from Biscayne Bay through the Florida Keys for 1997- 2016. Robert G. Muller. Submitted 6 June 2017.

SEDAR51-DW-15 Indices of abundance for Gray Snapper (*Lutjanus griseus*) using combined data from three independent video surveys Kevin A. Thompson, Theodore S. Switzer, Mary C. Christman, Sean F. Keenan, Christopher Gardner, Matt Campbell, Adam Pollack. Submitted 15 June 2017.

Appendix 3: Reference Documents

SEDAR51-RD01 Short-Term Discard Mortality Estimates for Gray Snapper in a West-Central Florida Estuary and Adjacent Nearshore Gulf of Mexico Waters Kerry E. Flaherty-Walia, Brent L. Winner, Amanda J. Tyler-Jedlund & John P. Davis.

SEDAR51-RD02 Regional Correspondence in Habitat Occupancy by Gray Snapper (*Lutjanus griseus*) in estuaries of the Southeastern Kerry E. Flaherty & Theodore S. Switzer & Brent L. Winner & Sean F. Keenan.

SEDAR51-RD03 Improved Ability to Characterize Recruitment of Gray Snapper in Three Florida Estuaries long the Gulf of Mexico through Targeted Sampling of Polyhaline Seagrass Beds Kerry E. Flaherty-Walia, Theodore S. Switzer, Brent L. Winner, Amanda J. Tyler-Jedlund & Sean F. Keenan.

SEDAR51-RD04 Conservation Genetics of Gray Snapper (*Lutjanus griseus*) in U.S. Waters of the Northern Gulf of Mexico and Western Atlantic Ocean John R. Gold, Eric Saillant, N. Danielle Ebelt, and Siya Lem.

SEDAR51-RD05 Developmental patterns within a multispecies reef fishery: management applications for essential fish habitats and protected areas Kenyon C. Lindeman, Roger Pugliese, Gregg T. Waugh, and Jerald S. Ault.

SEDAR51-RD06 Age, growth, and mortality of gray snapper, *Lutjanus griseus*, from the east coast of Florida Michael L. Burton.

SEDAR51-RD07 Ingress of transformation stage gray snapper, *Lutjanus griseus (*Pisces: Lutjanidae) through Beaufort Inlet, North Carolina Mimi W. Tzeng, Jonathan A. Hare, and David G. Lindquist.

SEDAR51-RD08 Biological response to changes in climate patterns: population increases of gray snapper (*Lutjanus griseus*) in Texas bays and estuaries James M. Tolan and Mark Fisher.

SEDAR51-RD09 Returns from the 1965 Schlitz tagging program including a cumulative analysis of previous results. Dale S. Beaumariage.

SEDAR51-RD10 Recruitment dynamics and otolith chemical signatures of juvenile gray snapper, *Lutjanus griseus,* among West Florida estuarine and coastal marine ecosystems Cecelia Lounder.

SEDAR51-RD11 Reproductive biology of gray snapper (*Lutjanus griseus*), with notes on spawning for other Western Atlantic snappers (Lutjanidae). M.L. Domeier, C. Koenig, and F. Colman.

SEDAR51-RD12 Climate-related, decadal-scale assemblage changes of seagrass-associated fishes in the northern Gulf of Mexico F. Joel Fodrie, Kenneth L. Heck, Jr., Sean P. Powers, William M. Graham, and Kelly L. Robinson

SEDAR51-RD13 Response of coastal fishes to the Gulf of Mexico oil disaster F. Joel Fodrie and Kenneth L. Heck Jr.

SEDAR51-RD14 Variation in the isotopic signatures of juvenile gray snapper (*Lutjanus griseus*) from five southern Florida regions Trika Gerard and Barbara Muhling.

SEDAR51-RD15 Temporal and spatial dynamics of spawning, settlement, and growth of gray snapper (*Lutjanus griseus*) from the West Florida shelf as determined from otolith microstructures. Robert J. Allman and Churchill B. Grimes.

SEDAR51-RD16 Regional variation in the population structure of gray snapper, *Lutjanus griseus,* along the West Florida shelf R.J. Allman and L.A. Goetz.

SEDAR51-RD17 Evaluating juvenile thermal tolerance as a constraint on adult range of gray snapper (Lutjanus griseus): A combined laboratory, field and modeling approach Mark J. Wuenschel, Jonathan A. Hare, Matthew E. Kimball, and Kenneth W. Able.

SEDAR51-RD18 Growth variation, settlement, and spawning of gray snapper across a latitudinal gradient Kelly Denit and Su Sponaugle.

SEDAR51-RD19 Age, growth, mortality, and radiometric age validation of gray snapper (*Lutjanus griseus*) from Louisiana. Andrew J. Fischer, M. Scott Baker, Jr., Charles A. Wilson, and David L. Nieland.

SEDAR51-RD20 Southeast Florida reef fish abundance and biology: Five year performance report Luiz R. Barbieri and James A. Colvocoresses.

SEDAR51-RD21 Larval ecology of a suite of snappers (family: Lutjanidae) in the Straits of Florida, western Atlantic Ocean E. K. D'Alessandro, S. Sponaugle, and J. E. Serafy.

SEDAR51-RD-22 Multidecadal otolith growth histories for red and gray snapper (Lutjanus spp.) in the northern Gulf of Mexico, USA. Bryan A. Black, Robert J. Allman, Isaac D. Schroeder, and Michael J. Schirripa.

SEDAR51-RD-23 Investigations on the Gray Snapper, *Lutjanus griseus*. Walter A. Starck II and Robert E. Schroede.

SEDAR51-RD-24 Age-size Structure of Gray Snapper from the Southeastern United States: A Comparison of Two Methods of Backcalculating Size at Age from Otolith Data A.G. Johnson, L.A. Collins, and C.P. Keim.

Appendix 4: Documents related to the Assessment

S51_Assessment_Report_FOR_RW.pdf

Appendix A_data_SS.dat

Appendix B_ctrl_SS.dat

Appendix C starter_SS.dat

Appendix D forcast_SS.dat

BASE_Gray Snapper_Run 42_estFs.zip

Appendix 5: Statement of Work

National Oceanic and Atmospheric Administration (NOAA) National Marine Fisheries Service (NMFS)

Center for Independent Experts (CIE) Program
External Independent Peer Review

SEDAR 51 Gulf of Mexico Gray Snapper Assessment Review

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf). Further information on the CIE program may be obtained from www.ciereviews.org.

Scope

SEDAR 51 will be a compilation of data, an assessment of the stock, and CIE assessment review conducted for Gulf of Mexico Gray Snapper. The review workshop provides an independent peer review of SEDAR stock assessments. The term review is applied broadly, as the review panel may request additional analyses, error corrections and sensitivity runs of the assessment models provided by the assessment panel. The review panel is ultimately responsible for ensuring that the best possible assessment is provided through the SEDAR process. The stock assessed through SEDAR 51 is within the jurisdiction of the Gulf of Mexico Fisheries Management Council and the states of Florida, Mississippi, Alabama,

Requirements

NMFS requires three (3) CIE reviewers to conduct an impartial and independent peer review in accordance with the SoW, OMB guidelines, and the TORs below. The CIE reviewers shall have expertise in stock assessment, statistics, fisheries science, and marine biology sufficient to complete the primary task of

providing peer-review advice in compliance with the workshop Terms of Reference.

Tasks for Reviewers

1) Review the following background materials and reports prior to the review meeting.
    a. Working papers, reference documents, and the Data Workshop and Assessment Process Reports will be available on the SEDAR website: http://sedarweb.org/sedar-51
2) Attend and participate in the panel review meeting. The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to answer any questions from the reviewers, and to provide any additional information required by the reviewers.
3) After the review meeting, reviewers shall conduct an independent peer review report in accordance with the requirements specified in this SoW, OMB guidelines, and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.
4) Each reviewer should assist the Chair of the meeting with contributions to the summary report.
5) Deliver their reports to the Government according to the specified milestones dates.

Foreign National Security Clearance

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

Place of Performance

The place of performance shall be at the contractor's facilities, and at Tampa, FL.

Period of Performance

The period of performance shall be from the time of award through March 31, 2018. The CIE chair and each reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Within two weeks of award	Contractor selects and confirms reviewers
Approximately two weeks later	Contractor provides the pre-review documents to the reviewers
March 20-22,2018	Panel review meeting
Approximately two weeks later	Contractor receives draft reports
Within two weeks of receiving draft reports	Contractor submits final reports to the Government

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards:

(1) The reports shall be completed in accordance with the required formatting and content; (2) The reports shall address each TOR as specified; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel

All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $10,000.

Restricted or Limited Use of Data

The contractors may be required to sign and adhere to a non-disclosure agreement.

Project Contact:

Julie A. Neer, Ph.D.
SEDAR Coordinator Science and Statistics Program
South Atlantic Fishery Management Council
4055 Faber Place Drive, Suite 201, North Charleston, SC 29405
(843) 571-4366 *Julie.neer@safmc.net*

Annex 1: Peer Review Report Requirements

1. The report must 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 report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR, in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs.

   a. Reviewers must describe in their own words the review activities completed during the panel review meeting, including a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, but especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the summary report that they believe might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The report shall represent the peer review of each TOR, and shall not simply repeat the contents of the summary report.

3. The report shall include the following appendices:

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

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

Annex II: Review Workshop Terms of Reference.

1. Evaluate the data used in the assessment, including discussion of the strengths and
    a. weaknesses of data sources and decisions, and consider the following:
    b. Are data decisions made by the DW and AW sound and robust?
    c. Are data uncertainties acknowledged, reported, and within normal or expected levels?
    d. Are data applied properly within the assessment model?
2. Are input data series reliable and sufficient to support the assessment approach and findings?
    a. Evaluate and discuss the strengths and weaknesses of the methods used to assess the stock, taking into account the available data, and considering the following:
    b. Are methods scientifically sound and robust?
    c. Are assessment models configured properly and consistent with standard practices?
    d. Are the methods appropriate for the available data?
3. Evaluate the assessment findings and consider 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, including discussing strengths and weaknesses, and consider 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.
    a. 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
    b. Ensure that the implications of uncertainty in technical conclusions are clearly stated
6. Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted.
    a. Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.
    b. Provide recommendations on possible ways to improve the SEDAR process.

8. Consider whether the stock assessment constitutes the best scientific information available using the following criteria as appropriate: relevance, inclusiveness, objectivity, transparency, timeliness, verification, validation, and peer review of fishery management information.

8. Provide suggestions on key improvements in data or modelling approaches that should be considered when scheduling the next assessment.

9. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

*This assessment will follow a Benchmark-track approach.

Appendix 6. SEDAR Review Workshop Panelist Instructions

SEDAR Review Workshop Overview

SEDAR Review Workshops provide independent peer review of stock assessments prepared through SEDAR data and assessment workshops. The goal of the review is to ensure that the assessment and results presented are scientifically sound and that managers are provided adequate advice regarding stock status, management benchmarks, and the general nature of appropriate future management actions. The Review Panel has limited authority to request additional analyses, corrections of existing analyses and sensitivity runs.

An analytical and presentation team, composed of a subset of the Assessment Workshop panel and representing the primary analysts for each assessment, will be present at the workshop to present assessment findings, provide an overview of assessment data, provide additional results or model information, and prepare any additional analyses requested by the Review Panel. Although many individuals contribute to a SEDAR assessment, the Review Panel is ultimately responsible for ensuring that the best possible assessment is provided through the SEDAR process.

The review panel shall not provide specific management advice. Such advice will be provided by existing Council Committees, such as the Science and Statistical Committee and Advisory Panels, following completion of the assessment. SEDAR review workshop panels are typically composed of a Chair appointed by the Cooperator (usually an SSC member), three reviewers appointed by the CIE (Center for Independent Experts), and up to two additional SSC reviewers appointed by each SEDAR Cooperator having jurisdiction over the stocks under review.

All reviewers are independent, meaning that they should not have contributed to the assessment under review and should not have a role in any management deliberations or actions that may stem from the assessment. All SEDAR workshops, including the Review Workshop, are open, transparent, public processes administered according to the rules and regulations governing Federal Fishery Management Council operations.

All SEDAR workshops are recorded and transcripts of workshop discussions may be prepared upon request through the SEDAR Steering Committee. The names and affiliations of reviewers will be disclosed in the review workshop documents. The Review Workshop Report will be publicly distributed along with the other SEDAR Workshop working papers and workshop reports. The public may submit written comments in accordance with Council guidelines once the report is disseminated to the relevant Council.

Review workshop panelists receive the Data and Assessment Workshop Reports, supplemental analytical materials including all working papers and reference documents from prior workshops; and general information regarding the Review Workshop, including the agenda, report outlines, terms of reference, and participant list. Review panelists are expected to read and review the provided materials to become familiar with the assessment.

The charge to each SEDAR Review Workshop is specified in Terms of Reference approved by the SEDAR Cooperator(s) having jurisdiction over the stocks under review. During the review the Review Workshop panel will prepare a Review Panel Report for each stock assessed addressing each of the Terms of Reference. The report should represent the views of the group as a whole, but shall also include any dissenting views of individual panelists if appropriate. Outlines and example documents will be provided by SEDAR staff.

Review Workshop Panel General Instructions

The Review Panel Chair is responsible for compiling, editing, and submitting the Review Panel Report to the SEDAR Coordinator by a deadline specified in the assessment schedule. At the start of the workshop the Chair will assign each panelist specific duties, such as drafting specific report sections. The Chair may select one panelist to serve as assessment leader for each stock assessment under review. The assessment leader is

responsible for preparing initial drafts of text addressing Terms of Reference for the assigned assessment. Such duties may be further subdivided if workshop manpower allows. The SEFSC will provide a rapporteur to take notes on the discussions so that panellists can more fully participate in discussions and assist the analytical team in documenting panel recommendations.

The Review Panel's primary responsibility is to ensure that assessment results are based on sound science, appropriate methods, and appropriate data. During the course of review, the panel is allowed limited flexibility to deviate from the assessment provided by the Assessment Workshop. This flexibility may include modifying the assessment configuration and assumptions, requesting a reasonable number of sensitivity runs, requesting additional details and results of the existing assessments, or requesting correction of any errors identified. However, the allowance for flexibility is limited, and the Review Panel is not authorized to conduct an alternative assessment or to request an alternative assessment from the technical staff present. The SEDAR Steering Committee recognizes that determining when modifications constitute an 'alternative' assessment is a subjective decision, and has therefore determined that the Review Panel is responsible for applying its collective judgment in determining whether proposed changes and corrections to the presented assessment are sufficient to constitute an alternative assessment. The Review Panel Chair will coordinate with the SEDAR Coordinator and technical staff present to determine which requests can be accomplished and prioritize desired analyses.

Any changes in assessment results stemming from modifications or corrections solicited by the review panel will be documented by the lead analysts in an addendum included in the final Stock Assessment Report. If updated estimates are not available for review by the conclusion of the workshop, the review panel shall agree to a process for reviewing the final results. Any additional or supplemental analyses requested by the Review Panel and completed by the analytical team may, at the discretion of the chair and panel, be documented either through a supplemental report or included in the Review Panel Report.

If the Review Panel finds an assessment deficient to the extent that technical staff present cannot correct the deficiencies during the course of the workshop, or the Panel deems that desired modifications would result in an alternative assessment, then the Review Panel shall provide in writing the required remedial measures and suggest an appropriate approach for correcting the assessment and subsequently reviewing the corrected assessment.

Review Workshop Panel Participant Information

Serving as a review workshop panellist is a considerable time commitment that requires more than simply the daily sessions of the review workshop. Panellists will need to set aside time in the weeks prior to the workshop to review data and assessment documents. During the workshop, time beyond that of the scheduled daily sessions may be required to complete workshop tasks and reports. Time is required following the workshop to review and finalize panel reports.

Review panellists are expected to author components of the workshop report and may conduct supplementary analyses or data summaries. Panellists should come prepared with a laptop computer for these tasks.

The SEDAR Steering Committee and SEDAR Coordinator establish deadlines for document submission. SEDAR staff distributes working documents and support materials (agenda, participant instructions) to workshop participants, typically two weeks prior to the workshop.

SEDAR Workshop Panellist Code of Conduct

SEDAR workshop panel decisions are based on science. Discussions and deliberations shall not consider possible management actions, agency financial concerns, or social and economic consequences. Such items will be dealt with by appropriate groups at appropriate times.

SEDAR workshop decisions are based on consensus. There is no formal voting in SEDAR. SEDAR Review Workshop Panels are encouraged to reach a group consensus that all participants can accept, which may include agreeing to acknowledge and present multiple possibilities. If this is not feasible, then each reviewer may state his or her individual opinion with regard to the Terms of Reference and is responsible for providing appropriate text that captures their opinion for the Review Workshop Report.

Personal attacks are not tolerated. Advancement in science is based on disagreement and healthy, spirited discourse is encouraged. However, professionalism must be upheld and those who descend into personal attacks will be asked to leave. Those criticizing the work and recommendations of others are expected to do so constructively and to offer reasonable solutions to go along with any criticisms.

Agendas provided for SEDAR workshops are approximate. Published starting and ending times for the Workshop will be adhered to so that participants may make appropriate travel plans, but starting and ending times of daily workshop sessions will be dictated by progress toward assigned tasks. Certain deadlines must be met during the workshop to ensure that all decisions are made and the group has a chance to review draft documentation, thus panellists should expected extended work days and evening working sessions. Webinar ending times are approximate.

Appointed participants are discouraged from leaving early. Considerable time and financial expense is involved in organizing and convening SEDAR workshops and participants should expect the workshop to last the allotted time. Many important decisions may be finalized in the closing hours. Every effort is made to accommodate travel requirements within the standard work week and avoid the need for weekend travel for most participants. Participants are responsible for notifying the Co-operator which appointed them to the workshop if they cannot meet this obligation. There are no guarantees that decisions made early in a workshop will not change as new information comes to light later in the workshop.

SEDAR data and assessment workshops are working sessions, not expanded reviews. Those appointed to serve as SEDAR workshop panellists are expected to contribute by participating in panel discussions, serving in working groups, drafting text, conducting analyses, reviewing workshop reports, and ensuring that the highest quality standards are met. When discussing issues, panellists are expected to provide constructive suggestions and alternative solutions, not simply criticism and critique.

Workshop reports are prepared by workshop participants. Each participant who accepts an appointment to a workshop is also accepting responsibility to ensure discussions and recommendations are thoroughly and accurately reflected in the workshop report. Each panellist is individually responsible for ensuring that their points and recommendations are addressed in workshop reports; they should not rely on others to address their concerns.

Review Workshop Networking and IT

Wireless access will be available at each SEDAR workshop to provide internet access. All reports and documents pertaining to the review will be available on the South Atlantic Council's FTP server. IT staff will be available during the first day of the review workshop to aid each participant in securing network access.
Review Workshop Technical Reviewer Responsibilities

1. Approximately two weeks prior to the meeting, the reviewers shall be provided with the stock assessment reports, associated supporting documents, and review workshop instructions including the Terms of Reference. Reviewers shall read these documents to gain an in-depth understanding of the stock assessment, the resources and information considered in the assessment, and their responsibilities as reviewers.

2. During the Review Workshop, reviewers shall participate in panel discussions on assessment methods, data, validity, results, recommendations, and conclusions as guided by the Terms of Reference. The reviewers shall

develop a Review Panel Report for each assessment reviewed. Reviewers may be asked to serve as an assessment leader during the review to facilitate preparing first drafts of review reports.

3. Following the Review Workshop, reviewers shall work with the chair to complete and review the Review Workshop Reports. Reports shall be completed, reviewed by all panellists, and comments submitted to the Chair within two weeks of the conclusion of the workshop.

4. Additional obligation of CIE-appointed reviewers: Following the Review Workshop, each reviewer appointed by the CIE shall prepare an individual CIE Reviewer Report and submit it in accordance with specifications provided in the Statement of Work.

SEDAR Review Workshop Report Outline

The Review workshop report is Section V of the Stock Assessment report and should be prepared cooperatively by the Review panel and the Coordinator in the following format.

Review Workshop Report

Cover Page [SEDAR]

Table of Contents [SEDAR]

1. Introduction [SEDAR]

1.1 Workshop Time and Place

1.2 Terms of Reference 1.3 List of Participants

1.4 List of Review Workshop Working Papers and Documents

2. Review Panel Report [Review Panel]

Executive Summary

2.1 Terms of Reference

List each Term of Reference, and include a summary of the Panel discussion regarding the particular item. Include a clear statement indicating whether or not the criteria in the Term of Reference are satisfied.

2.2 Summary Results of Analytical Requests - *sensitivities, corrections, additional analyses, etc.*

2.3 Additional Comments *(if necessary, to address issues or discussions not encompassed above)*

3. Submitted Comment - *any submitted, written comment or opinion statements [Review Workshop participants or observers]*

Appendix 7 List of Participants

Review Panel Members

Dr Kai Lorenzen (Chair), School of Forest Resources and Conservation, University of Florida.
Dr Luiz Barbieri, Program Administrator, Fish and Wildlife Research Institute.
Dr Laurence Kell (CIE), Hempstead, UK.
Peter Stephenson (CIE), Innaloo, Australia.
Dr Yong Chen (CIE), Professor of Fisheries Science, University of Maine.
Bob Gill, Scientific and Statistical Committee, Fishery Management Council.

Presenters

Dr Jeff Isley, Assessment Scientist, South East Fisheries Science Centre.
Dr Shannon Cass-Calay, Stock Assessment Team Leader, South East Fisheries Science Centre.

Other attendees

Dr Julie Neer, SEDAR Coordinator, South Atlantic Fishery Management Council.
Ryan Rindone, SEDAR Liason, Gulf of Mexico Fisheries Management Council.
Ed Walker, Industry Representative.
Beth Wrege, South East Fisheries Science Centre.
Charlotte Schiaffo, Gulf of Mexico Fisheries Management Council.
Karen Hoak, Gulf of Mexico Fisheries Management Council.

Appendix 8 Meeting Agenda

SEDAR 51 Gulf of Mexico Gray Snapper Assessment Review
March 20-22, 2018. Tampa, FL

Tuesday

9:00 a.m.	Introductions and Opening Remarks - *Agenda Review, TOR, Task Assignments*	Coordinator
9:30 a.m. – 11:30 a.m.	Assessment Presentations - *Assessment Data & Methods* - *Identify additional analyses, sensitivities, corrections*	Analytic Team
11:30 a.m. – 1:00 p.m.	Lunch Break	
1:00 p.m. – 6:00 p.m.	Assessment Presentations (continued) - *Assessment Data & Methods* - *Identify additional analyses, sensitivities, corrections*	Analytic Team
6:00 p.m. – 6:30 p.m.	Public comment	Chair

Tuesday Goals: Initial presentations completed, sensitivity and base model discussion begun

Wednesday

8:00 a.m. – 11:30 a.m.	Panel Discussion - *Assessment Data & Methods* - *Identify additional analyses, sensitivities, corrections*	Chair
11:30 a.m. – 1:00 p.m.	Lunch Break	
1:00 p.m. – 6:00 p.m.	Panel Discussion/Panel Work Session - *Continue deliberations* - *Review additional analyses* - *Recommendations and comments*	Chair

Wednesday Goals: sensitivities and modifications identified, preferred models selected, projection approaches approved, Report drafts begun

Thursday

8:00 a.m. – 11:30 a.m.	Panel Discussion - *Final sensitivities reviewed.* - *Projections reviewed.*	Chair Chair
11:30 a.m. – 1:00 p.m.	Lunch Break	
1:00 p.m. – 5:30 p.m.	Panel Discussion or Work Session - *Review Reports*	Chair
5:30 p.m. – 6:00 p.m.	Public comment	Chair

6:00 p.m. ADJOURN

Thursday Goals*:* Complete assessment work and discussions, final results available. Draft Reports reviewed.