Center for Independent Experts (CIE) Independent Peer Review Report: SEDAR 47 Southeastern Goliath Grouper Assessment Review Workshop

JOEL RICE
July 2016

Prepared for The Center for Independent Experts

Executive Summary

This review is a result of the SEDAR 47 Southeastern U.S. Goliath Grouper Assessment Review Workshop. The State of Florida served as the lead assessment agency and SEDAR provided the mechanism for the review of the assessment. The review workshop panel consisted of an independent chair as well as six reviewers, three each from the SSC and CIE.

The review was conducted in St. Petersburg Florida during May 17th-19th. The review was conducted in an open manner with the assessment team presenting their work, responding to comments and completing some additional work at the request of the review panel. Because the standard SEDAR data and model reviews had not been held, limited background material was made available to the assessment review panel prior to the workshop. This is one of three independent peer reviews completed for the assessment, contracted through the CIE process and independent of the consensus summary being drafted (though all reviewers contributed to the consensus summary). This report should be read in conjunction with the SEDAR 47 review panel report and along with the other CIE reviewers' reports.

Goliath grouper are distributed throughout the tropical and warm temperate waters of the Gulf of Mexico and the Western Atlantic. While some tagging studies show long distance movements, genetic and other analyses support the assumption of separate stocks within the Gulf of Mexico and also between South America and the Gulf. Historically, goliath groupers were landed in southeastern United States as well as the southern Gulf states. The fishery was closed in 1990 after fisheries indicators showed an apparent decline in stock size during the 1980s. Due to perception that the stock is rebounding efforts to open the fishery lead to this review.

Information regarding the status of the stock is limited, which is partly due to the moratorium on fishing this species, but also due to a lack of reliable landings data for a fishery that had been active since at least 1918. Fishery dependent indices of relative abundance are absent for the last two decades due to the closure and what abundance information there is consists of angler survey data, diver reported sightings, and the Marine Recreational Fisheries Statistics Survey/Marine Recreational Information Program, which are intercept and phone surveys of anglers.

In general, the data for this assessment are of poor quality, and this assessment was termed 'data poor' by the assessment team. This is the fourth attempt at an assessment of the goliath grouper since 2004 (see SEDAR 3 (2004), SEDAR 6 (2006), SEDAR 23 (2010). Contrary to previous assessment attempts, this assessment did not hold a data review or assessment workshop. This assessment process would have benefited greatly from both reviews as normally carried out via the SEDAR process. The lack of preparatory review workshops created problems with evaluating assessment techniques and inputs. In general, the lack of documentation hindered the understanding of the assessment methods and their appropriateness, and the reliability of the technical conclusions regarding stock status.

Background

This review is part of the CIE process associated with SEDAR 47. The SEDAR process ordinarily consists of a data workshop and review, an assessment process that includes review of the work, application of the models and an assessment review. This document is an independent review of the assessment, based on the review workshop held in St. Petersburg (Florida), May 17th-19th 2016. The intent of this review is to ensure that the best possible assessment is provided through the SEDAR process. Prior to the meeting, the following materials were provided to the review committee: a Statement of Work (Appendix 2), including the Terms of Reference (TOR) for the assessment and the assessment report. No other documents and background material were provided prior to the meeting. At the meeting and during the drafting of the group report, a general consensus among the review panel was reached for the major findings and discussion points (with respect to the TOR). This document contains a summary of those findings as well as my own opinions about this assessment.

Description of the Individual Reviewer's Role

The background materials provided for this review consisted solely of the assessment report. Notably missing were descriptions of the data preparation and information outlining the model development. Contrary to the standard SEDAR process, no data workshop or assessment workshop took place. The assessment materials were reviewed by the reviewer prior to review workshop held in St. Petersburg, Florida, May 17th-19th 2016. During the review workshop the assessment team presented the assessments for the Goliath Grouper of the South Atlantic and Gulf of Mexico. Discussion of the assessment approach, analysis, and results occurred throughout the three-day session. Some additional analysis was requested by the review panel (RP) and presented by the assessment team (AT).

Summary of Findings

TOR 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 data providers and assessment analysts sound and robust?

Due to the lack of a data workshop, this TOR is difficult to evaluate, the analysts provided some detail in the assessment reports and during the meeting. There are salient issues with the data and its treatments which are outlined by section below. In general, this is a data poor assessment where both the quality and quantity of data are lacking.

Life History Information

Goliath Grouper are distributed throughout the tropics, subtropics, and warm temperate coastal waters of the Atlantic Ocean. In the United States, goliath grouper occur from North Carolina to Texas, but have patchy distribution. Genetic data indicate that the stocks on the western and eastern shores of Florida are related, and it is likely that the stock does not extend to other countries. However, within the United States the resolution of the data is not fine enough to determine spatial structure of the stock. Historical landings indicate a range from

Florida through southern Texas. Tagging data demonstrate site fidelity, but also long distance movements (~175 km). A single stock within US waters was assumed for the assessment for convenience, and the review panel considers this to be reasonable. This is one of the most fundamental questions that any assessment must deal with, and is likely correct. However, there was little consideration of changes in habitat or range over the history of the fishery. The assessment uses data from Florida only, which may or may not represent the entire population. On a finer scale juveniles are known to rear in mangrove estuaries and sub-adults/adults prefer habitat with structure such as reefs, wrecks, piers, and bridges. In general, treating the Southeastern Goliath Grouper stock as single stock is a sound decision, though more research into the connectivity outside of Florida is warranted.

Goliath grouper have been aged to 37 years old and are thought have greater (than 37 years) longevity based on the fishing pressure sustained over the last century. Growth curves that have been developed are mainly based on young fish. More ageing work is currently underway. Due to the lack of older fish in the sample, the age-length relationship is less informative at older ages. Length to weight observations show the species can attain maximum weight and length of 2 meters and 200 kilograms (KG), though there relatively few samples at the longer lengths. Historical records indicate landed fish well over 600 pounds (~270KG).

Spawning aggregations are common during the new moon from July-September, with a sex ratio approximately 1:1. Currently, the fecundity is not well studied but some information exists. New research into the reproductive characteristics may shed some light on the reproductive life history, and this could be important for management. There is some uncertainty with respect to the age at maturity, and in the model maturity is assumed to be knife edge.

Natural mortality was spoken about at length during the workshop and has been a study area for goliath grouper for at least a decade. Previous methods of estimating natural mortality resulted in lower (0.12) rates than the method recently developed, which gave a m=0.18. Questions as to which value is more appropriate are largely related to the longevity, which is an area of active research. The larger issue with the goliath grouper natural mortality is that this species experiences episodic steep increases in the natural mortality from natural phenomena such as cold kills and red tide. This increase in natural mortality is difficult to model, and resulted in some model misfit.

Catch / Removals

In any fishery assessment there is a certain amount of uncertainty that is inherent in the catch data, especially with respect to historical landings. Goliath grouper in the Gulf of Mexico have an exploitation history dating back to the late 1800s. Reported commercial landings begin in 1918, are spotty for many years thereafter and vary greatly. The adjusted statewide (FL) landings are high in the mid-1940s (>400,000 pounds), and then drop precipitously prior to 1950, thereafter staying between 50,000 and 200,000 lbs. Historical catch data are highly variable and discontinuous pre-1950. Landings after that time period are mainly restricted to the state of Florida. Adjustments were made to the reported landings for the period of 1978-1984, due to the fact that the visits by biologists to a specific dealer did not see the expected amount of goliath grouper at the dealer. The reasoning behind the exact adjustment factor was explained; however, no sensitivities were given nor were alternative catch histories given. The specific percentages as chosen by the analysts were not examined in detail and subsequent analysis could result in different percentage reductions and alternative catch trends. Recreational catch estimates are based on Marine Recreational Fishery Statistics Surveys (MRFSS) / Marine Recreational Information Program (MRIP) survey data (1981-2015). As used in the model total fishery removals consisted of commercial landings (1950–1989), reported recreational landings (1981–1989) and recreational dead discards (1990- 2014) with an assumed release mortality rate of 5%; they were considered to be known without error.

There are concerns with the removals data, and aside from the adjustment factor applied to one source, commercial discards were not estimated, which is potentially a large area of uncertainty. The report showed an approximately 7.5%-11% observed occurrence of goliath grouper in the vertical and bottom longline fisheries. The effect of these fisheries and how they have changed over the last two decades was not explored. Uncertainty with respect to discards was not explored in detail. Discard mortality was assumed to be 5% in the recreational fishery and ignored in the commercial fishery.

Recreational catch varied significantly in the average weight per fish and overall catch size by year, suggesting that the angler survey did not adequately sample the recreational fishery. Of particular concern (given the model assumptions regarding B_0) was the highly variable and discontinuous time series of catch in the first half of the 1900s. Re-examining the methods of constructing historical removals should be a research priority. Catch data are particularly hard to estimate for a species experiencing a fishing moratorium, which is definitely the case here.

Indices of Abundance

Four indices of abundance were used in the assessments. These are addressed individually below. In general model selection, development and diagnostics (apart from the final deviance table) were missing from the report. This makes it difficult to evaluate the specifics of each model. It is reasonable to believe that the majority of the decisions made by the analysts with respect to the standardization process were sound. This is not the first time that these or similar data had been standardized for use in a goliath grouper assessment. However, in practice it is customary to produce an annex or separate report that delves in to the standardization details. This would have been covered had there been a data workshop. Given the importance of the indices, especially to the catch-free assessment model, fuller clarification and explanation of the indices is needed.

Everglades National Park index: The Everglades National Park index is a fishery index that is based on a survey of anglers conducted by National Park Service biologists. This index covers important juvenile habitat, thought to be the core habitat at the beginning of the moratorium. The fish caught in this index are juveniles and sub-adults, ranging from 20-100 cm in length, which approximately corresponds to ages 2-8. This index shows low catch rates between 1981-1993, followed by a large increase in 2007, a similarly large decrease in 2008 to a low in 2010, and low minimally increasing trend from 2010-2014. The interpretation of this index is difficult because it is largely a juvenile index that tracks a changing proportion of the stock. Furthermore, this index is largely unable to take account of changes in the fishers' behavior over time (e.g. due to the moratorium; effort creep), which may or may not be significant, but has not been explored. Important aspects of this index are that it covers the period before and after the moratorium, and that this index tracks a population that was heavily affected by documented cold kills in portions of the Everglades in January of 2008 and 2010. Recently, a relative low level of increase is evident.

REEF/GGGC Dive Index: The REEF (Reef Education and Environmental Foundation) dive index is an index based on reported sightings by recreational divers who have gone through a training program in fish identification and survey techniques taught by the REEF organization. This index roughly shows an increase in sightings throughout the time series with the exception of the terminal years. A deficit of this index is that it has no rigid experimental design, and although there are numerous reports (of sightings and non-sightings), it is in general not oriented at observing goliath groupers. Data are reported as categorical variables (0, 1, 2-10, 11-100, 100+). Arbitrary criteria intended to balance the need for spatial coverage were developed by the AT to require a dive site to have at least 10 reports (0 or +) in the last 20 years, and at least one positive sighting of a goliath grouper. This data are supplemented by a targeted survey from the great goliath grouper count (GGGC) data, a targeted dive survey that is similar in method to the REEF data.

There are issues with reliability of this combined index (REEF & GGGC), hereafter the REEF index. The standardization of this data series was done with a Poisson generalized linear model, which is inappropriate for categorical ordinal data. The categorization of the counts is in logarithmic categories yet the output is in normal space. Further the binning scheme ignores changes beyond the highest categories (>100). The combination of the non-target REEF data and the targeted GGGC data is problematic for the interpretation of the index. The REEF data are non-targeted data while the GGGC are targeted survey data. The effect of combining these data is uncertain. I do not believe that the REEF/GGGC dive index is appropriate for use in the assessment for reasons of the data collection methodology, the data preparation and the standardization of the index.

MRFSS/MRIP Indices. The MRFSS/MRIP indices are based on angler intercept surveys conducted throughout Florida waters (the MRFSS program pre-dated the MRIP program). The MRFSS/MRIP data are available from 1981 through 2014. The analysts conducted refinements on the data to attempt to increase their accuracy and precision, but concluded that only data from 1997 through 2014 were adequate, partly because Goliath occurrences in the recreational catch data were sparse prior to 1997. Similar to the other indices, there was little documentation regarding data preparation and standardization. This data set was separated to create an inshore and offshore index. Both indices were highly variable in terms of the raw data, and had slightly different trends based on the coast (West vs East and the Keys). The estuarine index showed the same general trends on both coasts of Florida; however, the offshore indices from each coast showed conflicting trends in many years with the West coast increasing while the East coast is decreasing and vice versa in other areas. This index runs from 1997 – 2014. In 2006, the MRFSS program was replaced with the MRIP because the MRFSS surveys and methods were hindered by under-coverage, inefficiency, biased sampling and estimation methods. Due to the lack of information presented, we can only assume that the two surveys were correctly calibrated prior to being combined. The benefit to including these indices of abundance is that they are fishery dependent surveys that include data from the moratorium.

Selectivity

The selectivity of the fisheries is separated into two time series blocks, before and after the moratorium for both the inshore and offshore fisheries. This is in general a sound decision by the analytic team. During the review workshop, the RP questioned the methods used to construct the selectivity curves because it was not entirely clear from the information provided. The selectivity associated with the index of abundance for the estuarine indices is representative of the frequency distribution of the age of fish in the estuarine catch, but because the younger fish are more abundant in the population, this 'selectivity curve' will overestimate the proportion of young fish and underestimate the proportion of older fish in the catch. Furthermore, it is questionable as to whether the largest fish (>2meters) would be retained by the majority of the fishing gear in the offshore recreational fishery. A consequence of the fact that commercial discards are ignored in the assessment is that the assumed selectivity of sub-adult and adult age class is likely misspecified. In general, the selectivity choices and development are poorly described in the assessment report.

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

The analysts reported that this was a data poor assessment and acknowledged the uncertainty, within which is at the high end of the expected level, with respect to catch. The CV's for the indices of abundance were highly variable; however, there was some confusion within the RP as to what the various CV estimates represented (how they were calculated). Modeling to standardize (remove the effect of factors other than abundance) in the index of abundance was conducted on indices prior to use as in the assessment. It would be helpful to include model diagnostics such as plots of residuals from the standardization. Plots of the nominal data with the results of standardization process and to show the effect of the modeling would be helpful to see.

Uncertainty in commercial landings due to suspected over reporting by one dealer was thoroughly discussed. The AT made a reasonable and well-explained correction, but it reduced the landings by almost 50% from the reported landings. This is a major correction and it would be helpful to make a model run using the reported landings to investigate the effect of this correction.

c) Are data applied properly within the assessment model?

The data are in general applied properly in the model. However, the information in most cases is highly susceptible to bias. For example, the indices of abundance show improbable rates of increase, and the catch estimates are adjusted based on the ad-hoc correction of reported landings. A major problem with the use of the data in the models was that all indices of abundance were input to the model. In principle, the assessment could benefit from grouping indices and fitting multiple models, using only non-conflicting indices within a model. Allowing a model to include conflicting indices breaks the assumption the indices represent the population dynamics, and results in poor fit as the model finds a non-optimal solution that is a compromise between the conflicting indices.

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

The input data series are in general not reliable to support the assessment approach and findings. Both the SSRA, and especially the catch-free model, rely heavily on the indices of abundance, which are unreliable. The SSRA model assumes that the catch is known without error. This is demonstrably not the case. There is uncertainty as to the longevity of the species and the associated natural mortality estimates. The natural mortality experienced by the population is affected by episodic high mortality events, which are not adequately reflected in any input or parameterization. The size structure and magnitude of historical landings is largely unknown, there is no independent estimate of abundance across the distribution, the changes in historical distribution are not well investigated, and the current population structure is unknown. The problems associated with the data make it difficult to consider them reliable for the assessment approach. Sufficient data did exist for the models; however, it was of insufficient quality.

TOR 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?

The assessment approach was twofold, both the "Stochastic Stock Reduction Analysis" (SSRA) (Martell et al. 2009) and the "Catch Free" model (Porch et al., 2006) were put forward as the principal assessment models presented and discussed in the assessment report, though no stock status advice or base-case model was given. Both modelling approaches have been published in the peer reviewed literature, though it is important to note that both models had been modified by persons other than the original authors or the AT to allow the inclusion of multiple survey indices and other changes. These changes to the model do not appear to have been externally reviewed. The RP was unable to fully evaluate the impact of these modifications for either model. The models share some important similarities which include:

- The underlying population model is age structured.
- A Beverton-Holt stock recruitment is assumed.
- Recruitment deviations are treated as random effects and characterise relative year class strength.
- Fishing mortality is modelled as the product of an age and year effect.

- Survey indices are treated as proportional to biomass or numbers conditioned on age specific selectivity.
- Parameters are estimated by maximizing a likelihood function.
- Penalty functions are used to constrain some of the model parameters. These are referred to as "priors" but are not true Bayesian priors and may result in improper posterior distributions.

Important differences between the models are:

- SSRA uses total fishery removals (dead catch) and treats these as known without error.
- Unlike SSRA, the Catch Free model treats selectivity, natural mortality, growth parameters and fishing mortality as parameters to be estimated.
- SSRA parameterises the stock recruitment function in terms of Fmsy and MSY, and these are the main parameters to be estimated. An important consequence of this is that the stock recruitment parameters are conditioned on the assumption of selectivity and will change if the selectivity assumption is changed.

The models are well known variants of age-structured production models and can be regarded as scientifically sound in the form that they were published. Whether they are considered "robust" depends heavily on the data used. Here "data" may include constants, such as the fishery removals, age at maturity, selectivity, M, etc., that do not enter the likelihood as well as observations, such as survey indices that do. Where data enter the model as constants, it is particularly important that they are accurate to avoid cumulative errors. Fishery removals and selectivity, for example, can be critical in determining the model outcome, yet there is considerable uncertainty surrounding the values used in these assessments.

It is not possible to conclude that the methods are robust and various analyses reported in the assessment document (e.g., the MCMC runs for the catch free model, and the sensitivities run during the assessment review) demonstrate that the results are highly dependent on the specific configuration of the model presented. As such the methods are not robust. Where priors are used, as is the case in these assessments, it is particularly important to examine whether these are updated by the observations, and to examine the sensitivity of model estimates to the priors. These diagnostics were not done, which prevented a RP assessment of robustness.

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

A major failing of both models is that they included all the indices at once. This is a failing because the indices have different trends (most notably the reef index increases while the others drop off. The catch free model was for the most part configured properly. Only two configurations were presented, with higher and lower natural mortality. These seem to be parameterized correctly; however, more analysis of the prior assumptions is needed along with diagnostics showing the priors and posteriors. The SSRA model shared the same failings as the catch–free model (inclusion of conflicting indices, overlapping priors); however, in addition the SSRA model (as implemented here) assumes that the population is at the virgin level in the first year of the model. This is an improper configuration given the history of exploitation for this species. Furthermore, with highly uncertain catch the use of a production model that assumes known catch is not consistent with standard practice.

c) Are the methods appropriate for the available data?

The catch-free model is an appropriate tool given the data; however, the SSRA is highly dependent on the catch being correct. The SSRA could be used to explore the impact of alternative catch series given the data but this was not done. Both models could be used as part of a suite of techniques to characterize stock status and the associated uncertainty. The available data could potentially be analysed using a variety of models, including

surplus production models and other data-poor approaches, provided it was presented in a framework that showed the relative strengths of the data components. Much more thought needs to be given to the implication of handling the fishery removals as known constants and developing changes to the SSRA/catch free models so that this issue can be explored.

TOR 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?

The RP felt that it was unable to evaluate the abundance, exploitation, and biomass estimates from both the catch free and SSRA models. Given the uncertainties in the data and the extreme sensitivity of the SSRA to the choice of the start year (because it assumes virgin stock size at the start year) and the inclusion of the REEF index (which is in conflict with the others), it does not support inferences on stock status. Given the uncertainties in the data, the reliance of the SSRA on the assumption that catch is assumed to be known without error, the SSRA model outcome does not support inferences on stock status.

Both models provided insight into the possible population dynamics over the last few decades, but neither is sufficient for management advice and thus not useful to support status inferences.

By definition, the presented catch-free model can only provide estimates of relative abundance, exploitation, and biomass so there is no information provided to the model to allow scaling to absolute values. The Catch-Free model has previously been used in SEDAR 6 and SEDAR 23 where they were adopted to provide relative estimates and to provide guidance on the possible recovery time of goliath grouper. At SEDAR 23, the Catch-Free model was employed again, but the context was changed with a management need to provide OFL and ABC recommendations. However, the Catch-Free model cannot provide this information as it does not use data on removals to scale necessary estimates, and because it cannot take account of possible future exploitation patterns. Thus, for SEDAR 47 this model is again not appropriate for stock status determination.
At the SEDAR 23 Review Workshop (RW), a SSRA model was presented for exploratory purposes, but the review panel did not use it to make inferences about stock status, as it had not been previously considered by the SEDAR 23 Assessment Workshop. In principle, with better quantification of removals and conducting various sensitivity runs, the SSRA could be used to provide more relevant information for management. However, the SEDAR 47 RP concluded that the SSRA model critically depends on credible inputs of removals, which were deemed too uncertain in the current iteration. Furthermore, during the course of the review, the review panel discovered that the model initializes virgin biomass at the beginning of 1975. Thus for SEDAR 47 the Review Panel does not consider the SSRA model appropriate for stock status determination.

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?
Neither model was sufficient to infer stock status and support management decisions.

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

The stock recruitment curve was estimated internal to the model and was assumed to be a Beverton-Holt relationship. The robustness of the chosen stock recruitment relationship was not explored. Therefore, it is not known how informative the presented stock recruitment relationship is.

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

The RP felt that the quantitative estimates produced by both the catch free and SSRA models were not reliable. During the RW the Panel requested various sensitivity runs be produced from the SSRA, including starting the data at 1975, dropping the REEF diver data index, and including each index in isolation. The stock status determinations produced from these various sensitivity runs varied greatly and contributed to the lack of confidence that the RP had in the model's ability to accurately estimate stock status for the goliath grouper population.

TOR 4 – Evaluate the stock projections, including discussing strength 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?

No projections were presented for SSRA. The methods used to project the future relative biomass based on the results of the catch free model are consistent with accepted practices. Note that due to the characteristics of this type of assessment model, the projections are not appropriate (or useful) to support inference as to absolute future conditions, or harvest levels. Uncertainty in the catch free model is reflected in the projection results through the use of MCMC as well as by completing sensitivity runs with two levels of natural mortality and modifications to which selectivities were estimated by the model.

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

a) Comment on the degree to which methods used to evaluate the uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods.
b) Ensure the implications of uncertainty in technical conclusions are clearly stated.

The AT approached the characterization of uncertainty in the individual assessment models via measures of precision (standard errors, CVs, and posterior distributions) associated with model outputs, and via sensitivity analyses to key model assumptions (e.g. natural mortality). In general, the methods used to characterize the uncertainty in the assessment were acceptable; however, the estimates of the uncertainty associated with parameter estimates cannot be provided because the model results are not accepted. The results of the sensitivity analyses conducted prior to and during the review workshop implied that the uncertainty in the technical conclusions was high.

The largest uncertainty in the assessment was the catch. The uncertainty associated with the catch was not addressed in the SSRA and by definition the catch-free model did not address this. The potential consequences

stemming from the misspecification of the catch are significant, not reflected in the assessment, and would affect the technical conclusions of the population status estimates. One effect of assuming that the catch is known without error is that uncertainty reflected in the 95% Bayesian central interval was so tight it showed almost no uncertainty in the population trajectory.

Uncertainty with respect to the longevity and the associated natural mortality (*m*) was addressed by running variants of the catch-free model at two levels of *m*. This was useful to determine the effect of alternative estimates of natural mortality which showed a large change in the status of the stock based on the choice of *m.*

Additional sensitivity analyses were performed for both models at the request of the reviewers. For the catch-free model, the team conducted model runs fixing the selectivity curves vs estimating the selectivities based on priors. Model outputs under the different assumptions were provided as a means of assessing consistency of the conclusions.

The SSRA model was run with and without various indices of abundance and for different time periods (1975-2014 against 1950-2014). The results indicated that biomass trends were strongly impacted by the changes in start date. In addition, the sensitivity runs with different indices led to further changes in the model fit and predictions. Together these provide good indications of the high degree of uncertainty in model results.
The RP was unable to fully evaluate the requested sensitivity runs (with a start year of 1975), because the models assumed a virgin stock biomass in 1975, which is likely unrealistic. As a consequence, the RP was unable to fully compare the model in the report with these sensitivity runs. The report did include retrospective analysis for the SSRA; however, their utility was diminished by the lack of stability of the assessment model.

Examination of the uncertainty in the assessment was facilitated by the fact that the original models had been coded in ADMB, which provides as part of its standard output standard errors for estimated parameters and derive variables, and can also produce posterior distributions for estimated and derived parameters via MCMC. In general, this is an appropriate method to address the uncertainty. Of particular concern is the fact that the MCMC output from the catch-free model was interpreted despite the results implying the model was misspecified (2 of the 8 chains did not converge, and were deleted). In general, with MCMC it is helpful to present the priors as well as the estimated posterior distributions of estimated parameters, which was not done for either model.

TOR 6 – Consider the research recommendations provided and make any additional recommendations or prioritizations warranted.

The AT listed a series of research recommendations in the assessment report, all of which would provide valuable information that would aid in the understanding of some aspects of the life history or the population dynamics of goliath grouper but not necessarily improve a future assessment. Having a reliable time series of removals is crucial to the type of modeling attempted in this assessment. Research regarding a reliable and comprehensive catch series is recommended as a priority. If the reliability of a single catch series is low, then construction of multiple plausible time series should be undertaken to address the uncertainty with respect to catch. Additional recommendations with respect to estimating the catch would be to estimate the interaction and discard mortality associated with the commercial fishery. Standardizing the method to estimate the recreational removals could also improve the confidence in the catch series. Note that the SEDAR 23 RW concluded that “The next benchmark assessment cannot be successfully completed without data from the research recommended by the Data, Assessment, and Review Panels.” The outcome of (this) SEDAR 47

benchmark assessment process indicates that much of this information is still needed in order to successfully complete an assessment for goliath grouper. Specific research and monitoring efforts that could improve future assessments for goliath grouper include:

Developing additional data from outside of Florida. That there is limited data from other Gulf states may be indicative of low population size, a change in habitat, or poor sampling. This research topic is related to the geographical validity / usefulness of the assessment for regional management.

Improving the indices of abundance, for example thoroughly examining the influence of the GGGC survey on the REEF index, fully examining the standardization process for this survey.

Develop and/or explore methods to take into account episodic mortality events. In general, this should be quite possible to do with most assessment models.

Research projects oriented at improving the life history information for goliath grouper are recommended. Specific needs are basic reproductive information including: size and age at maturity for each sex, sexual sequence with size and age for each sex, estimates of longevity and fecundity. There are some ongoing studies in this area.

Additional research is needed on the age structure of the catch, especially in the offshore recreational fishery, and the commercial bycatch.

Discard mortality estimates are needed across the species distribution and fisheries. For the SEDAR 47 assessment, a fixed discard mortality rate estimate was applied to the post-moratorium harvest. However, the uncertainty around this estimate is unknown.

The previous assessment (SEDAR 23) recommended that goliath grouper should be genetically sampled from areas across the stock range in the South Atlantic and Gulf of Mexico to allow for a more thorough examination of the current single stock definition. The SEDAR 47 RW was presented with a brief summary of these efforts, which seem to support that single stock definition. Like many other sources of information informing the SEDAR 47 assessment, this information remains in progress or is incomplete and has not yet been vetted by peer review. Further genetic sampling of the goliath grouper population across the species distribution in order to complete this analysis should be completed prior to the next benchmark assessment. Examination of spawning aggregations over the entire distribution range should include seasonality, sex ratios, and individual fidelity.

TOR 7 – 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.

The nature of the data, data choices, and choice of modeling framework did not provide results that can be considered best available scientific information. Details are provided under various TORs above.

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

A major hindrance to the review of this assessment is the lack of clarity regarding the available data and its treatment (including the collection, data grooming, modeling). The inclusion of a complete data workshop report would remedy part of this issue. Along the same lines, a complete assessment workshop report would be helpful for assessing the model development, as well as facilitate the daylighting of some aspects of the analysis that were not included in the assessment report. An assessment workshop would also help assure the RP that any changes to the code by the AT were appropriately implemented and function as intended. Research into the catch estimates (including commercial discards and historical catch), as well as research into the CPUE series, and the longevity (and the associated natural mortality) should be prioritized.

Appendix 1: Bibliography of materials provided for review

All materials are available on the SEDAR website, http://sedarweb.org/sedar-47-review-workshop. The materials listed below were provided prior to or during the review.

SEDAR 47 Stock Assessment Report Goliath Grouper_Updated.pdf

SEDAR 47 RW Presentation I - Data and Methods

SEDAR 47 RW Presentation II - Catch-free Model

SEDAR 47 RW Presentation III - SRA Model

SEDAR 47-RW-01: The tpl file, data file, and control file for a Stochastic Stock Reduction Analysis (SSRA) program

Appendix 2: A copy of the CIE Statement of Work

Statement of Work

External Independent Peer Review by the Center for Independent Experts

SEDAR 47 Southeastern Goliath Grouper Assessment Review Workshop

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

Project Description: SEDAR 47 will be a compilation of data, an assessment of the stock, and CIE assessment review conducted for Southeastern Goliath Grouper. 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 stocks assessed through SEDAR 47 are within the jurisdiction of the South Atlantic and Gulf of Mexico Fisheries Management Council and the states of Florida, Georgia, South Carolina, and North Carolina, Mississippi, Alabama, Louisiana, and Texas. The Terms of Reference (ToRs) of the peer review are attached in Annex 2. The tentative agenda of the panel review meeting is attached in Annex 3.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge 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. Experience with data-limited or catch-free assessment methods would be preferred. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in St. Petersburg, FL during May 17-19, 2016.

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

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

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

http://deemedexports.noaa.gov/

http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background

information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

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

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

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

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

1)Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting tentatively scheduled in St. Petersburg, FL during May17-19, 2016.

3) Tentatively in St. Petersburg, FL during May 17-19, 2016 as specified herein, and conducts an independent peer review in accordance with the ToRs (Annex 2).
4) No later than REPORT SUBMISSION DATE, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Dr. Manoj Shivlani, CIE Lead Coordinator, via email to mshivlani@ntvifederal.com, and Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in Annex 2.

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

March 29, 2016	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
April 29, 2016	NMFS Project Contact sends the CIE Reviewers the pre-review documents
May17-19, 2016	Each reviewer participates and conducts an independent peer review during the panel review meeting
June 9, 2016	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
June 23, 2016	CIE submits CIE independent peer review reports to the COTR
June 30, 2016	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

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

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified

in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

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

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

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

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

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

Support Personnel:

Allen Shimada
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
Allen Shimada@noaa.gov Phone: 301-427-8174

Manoj Shivlani, CIE Lead Coordinator
NTVI Communications, Inc.
10600 SW 131st Court, Miami, FL 33186
mshivlani@ntvifederal.com Phone: 305-968-7136

Key Personnel:
NMFS Project Contact:
Julie A Neer
SEDAR Coordinator
4055 Faber Place Drive, Suite 201
North Charleston, SC 29405
(843) 571-4366
julie.neer@safmc.net

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

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

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

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

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

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

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

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

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review

   Appendix 2: A copy of the CIE Statement of Work

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

Annex 2: Terms of Reference
SEDAR 47 Southeastern Goliath Grouper Assessment Review Workshop

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 used 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?
   c) Are the results informative and robust, and useful to support inferences of probable future conditions?
   d) Are key uncertainties acknowledged, discussed, and reflected in the projection results?
5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.
   - Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods.

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

6. Consider the research recommendations provided by the 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.
7. 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 guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.
9. Ensure that stock assessment results are clearly and accurately presented in the Stock Assessment Report and that reported results are consistent with Review Panel recommendations. If there are differences between the AW and RW due to the reviewer's request for changes and/or additional model runs, etc. describe those reasons and results.
10. CIE Reviewer may contribute to a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

Annex 3: Agenda

SEDAR 47 Southeastern Goliath Grouper Review Workshop

Saint Petersburg, Florida

17-19 May 2016

Tuesday

9:00 a.m. Introductions and Opening Remarks Coordinator

- Agenda Review, TOR, Task Assignments

9:30 a.m. – 11:30 a.m. Assessment Presentations Analytic Team

- Assessment Data & Methods
- Identify additional analyses, sensitivities, corrections

11:30 a.m. – 1:00 p.m. Lunch Break

1:00 p.m. – 6:00 p.m. Assessment Presentations (continued) Analytic Team

- Assessment Data & Methods
- Identify additional analyses, sensitivities, corrections

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 Chair

- Assessment Data & Methods
- Identify additional analyses, sensitivities, corrections

11:30 a.m. – 1:00 p.m. Lunch Break

1:00 p.m. – 6:00 p.m. Panel Discussion/Panel Work Session Chair

- Continue deliberations
- Review additional analyses
- Recommendations and comments

6:00 p.m. – 6:30 p.m. Public comment 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 Chair

- Final sensitivities reviewed.
- *Projections reviewed.* **Chair**

11:30 a.m. – 1:00 p.m. Lunch Break

1:00 p.m. – 5:30 p.m. Panel Discussion or Work Session Chair

- Review Reports

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.

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

LIST OF PARTICIPANTS

Workshop Panel

Marcel Reichert, Chair ..... Chair, SSC
Carolyn Belcher ..... SSC
Mary Christman ..... SSC
Robin Cook ..... CIE Reviewer
Bob Ellis ..... SSC
Desmond Kahn ..... CIE Reviewer
Joel Rice ..... CIE Reviewer

Analytic Representation

Joe O'Hop ..... FWRI, St. Petersburg
Joseph Munyandorero ..... FWRI, St. Petersburg

Observers

Dustin Addis ..... FWRI, St. Petersburg
Steven Atran ..... GMFMC
Shanae Allen ..... FWRI, St. Petersburg
Michael Drexler ..... Ocean Conservancy
Doug Gregory ..... GMFMC
Elizabeth Herdter ..... FWRI, St. Petersburg
Michelle Masi ..... FWRI, St. Petersburg
Bob Muller ..... FWRI, St. Petersburg
Mike Murphy ..... FWRI, St. Petersburg
Brian Schoonard ..... GMFMC

Staff

Julie Neer ..... SEDAR
Ryan Rindone ..... GMFMC Staff
Charlotte Schiaffo ..... GMFMC Staff